TEACHERS' TOPICS

Using Answer-Until-Correct Examinations to Provide Immediate Feedback to Students in a Pharmacokinetics Course

Adam M. Persky, PhD, and Gary M. Pollack, PhD

School of Pharmacy, University of North Carolina at Chapel Hill

Objectives. To implement an answer-until-correct examination format for a pharmacokinetics course and determine whether this format assessed pharmacy students' mastery of the desired learning outcomes as well as a mixed format examination (eg, one with a combination of open-ended and fill-in-the-blank questions).

Methods. Students in a core pharmacokinetics course were given 3 examinations in answer-until-correct format. The format allowed students multiple attempts at answering each question, with points allocated based on the number of attempts required to correctly answer the question. Examination scores were compared to those of students in the previous year as a control.

Results. The grades of students who were given the immediate feedback examination format were equivalent to those of students in the previous year. The students preferred the testing format because it allowed multiple attempts to answer questions and provided immediate feedback. Some students reported increased anxiety because of the new examination format.

Discussion. The immediate feedback format assessed students' mastery of course outcomes, provided immediate feedback to encourage deep learning and critical-thinking skills, and was preferred by students over the traditional examination format.

Keywords: critical thinking, assessment, survey, anxiety, answer-until-correct, pharmacokinetics, examination

INTRODUCTION

Pharmacokinetic principles underlie many aspects of pharmacotherapeutic decision-making. Clinicians use their understanding of pharmacokinetics to choose appropriate agents for therapy, appropriate dosing schedules, and appropriate monitoring strategies. Because pharmacokinetics combines physiologic, pharmacologic, and mathematical principles, it is challenging for instructors to facilitate pharmacy students in acquiring sufficient knowledge and skills to make dosing decisions. Evaluating the skill development in pharmacokinetics also is challenging because the evaluation ranges from basic concepts and definitions to calculations or case studies. Because of the range of questions with respect to levels of learning, examination format is often heavily weighted towards open-ended questions. This open-ended format requires significant grading time, from a couple of days to a week. By the time examination results are made avail-

able to the students, new material is being discussed in class, limiting the impact of the examination as a learning opportunity. A multiple-choice examination reduces the time required to grade examinations, but does not test the student in the same manner as an open-ended question by allowing points to be earned (ie, partial credit) for a student's proximate knowledge. Moreover, feedback on performance is still delayed, albeit to a lesser degree. In some instances, feedback can be given promptly by providing answer keys or reviewing the examination immediately after completion of the examination. The potential limitations to these approaches are unwillingness of faculty members to post examination keys (potentially to allow use of examinations in future years) or lack of sufficient faculty time to review the examinations immediately after the examination. An immediate-feedback examination that can assess the student's depth of knowledge (full or proximate) as well as an open-ended question examination is a possible solution to the limitations of traditional examination styles.

Immediate feedback is considered 1 of the 7 principles for good practice in education.[1] One reason immediate feedback is a good educational practice is because it improves learning in various situations compared to delayed feedback.[2-5] Formative assessment, ie, assessment

Corresponding Author: Adam M. Persky, PhD, Clinical Assistant Professor, 109G Beard Hall, CB#7360, Division of Pharmacotherapy and Experimental Therapeutics, School of Pharmacy, University of North Carolina at Chapel Hill, Chapel Hill, NC 27599-7360, (phone) 919-966.9104, (fax) 919.966-0197, E-mail: apersky@unc.edu

that provides feedback on a student's work, is a key element in deep learning.[6] Examinations should evaluate student knowledge, correct any misinformation the student has, and produce new knowledge in the process.[7] When using typical examinations, students do not know the correct solutions until hours, days, or in some cases weeks later; thus, a traditional examination (ie, summative assessment) may not facilitate deep learning as well as a feedback-oriented approach (ie, formative assessment).

An additional issue with open-ended problems is the relationship between 1 or more questions. If the answer to 1 question serves as the basis for 1 or more additional questions, a mistake on 1 problem may cause subsequent problems to be answered incorrectly. The immediate-feedback approach prevents this from occurring; students will know the correct response after each question and will be less likely to have 1 mistake compounded into larger or repetitive errors. An immediate feedback assessment has been developed and validated to accomplish these purposes.[2,7] This examination format can behave like an essay examination in some ways by offering partial credit and allowing questions to build upon each other. In addition, there is a high correlation between scores on essay examinations and behaviors on the immediate feedback format.[8]
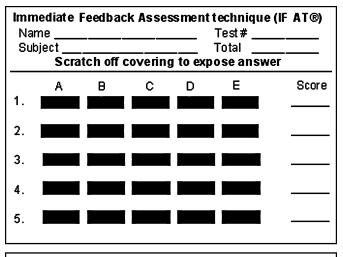
This manuscript communicates initial experiences with immediate feedback examinations in a core, foundational course in pharmacokinetics in a professional pharmacy program. The hypothesis underlying this study was that the immediate feedback format would yield equivalent ability to assess student mastery of the desired course outcomes as compared to mixed-question format examinations (eg, open-ended problems, multiple choice, fill-in-the-blank, matching) administered the previous year as measured by examination scores. A secondary hypothesis was that students would favor the immediate feedback format because it would allow students to assess their own strengths and provide them with the opportunity to re-think and re-work problems in the answer-until-correct format. The primary goal of this study was to evaluate whether the answer-until-correct format represented an assessment tool that was comparable in scope (ie, content), level (ie, level of learning), and scoring, to the mixed format examinations used previously in this course.

DESIGN

Pharmacokinetics instruction in the School of Pharmacy at The University of North Carolina at Chapel Hill consists of 3 courses starting in the spring of the first year and continuing to the spring of the second year. Pharmacy 413 (PHCY 413) occurs in the fall semester of the second year and serves as the foundational pharmacokinetics

course of the professional pharmacy curriculum. PHCY 413 is a 3-credit course that meets 3 times a week for 50 minutes per class. All classes are taught using synchronous video teleconferencing. Three examinations were administered in this course during designated examination weeks. The first and second examinations were worth 100 points, and each contributed 25% to the final grade in the course. The cumulative final examination was worth 150 points and contributed 38% to the final grade in the course. The remaining points for the course (50 points or 12%) were earned from weekly reflective writing assignments. Examinations were multiple-choice, answer-until-correct format only using immediate feedback forms (IF-AT, Epstein Educational Enterprises, New Jersey). This was in contrast to the previous year, in which the test format was approximately 50% open-ended problems, with the remainder of items being multiple-choice, true/false or short answer questions.

Examinations were constructed based on content to sufficiently assess the learning objectives for that section of material (approximately 2 to 4 questions per class period/learning outcome). Questions also were constructed to assess level of learning according to Bloom's taxonomy, consistent with the stated learning objectives. The levels of learning were separated into 3 categories to facilitate classification of questions, and were consistent with the school's curriculum mapping efforts: Level 1 (knowledge and comprehension), Level 2 (application), and Level 3 (analysis, synthesis and evaluation); the questions were sequentially arranged within the examination according to the level of the question (eg, Level 1 in the beginning, Level 2 in the middle, Level 3 at the end of the examination). The number of questions at each level was determined by the desired outcomes of the course (or section material) with the majority of questions focusing on the application of knowledge (ie, Level 2); if the students acquired sufficient skills development on all the desired outcomes, they would score in the range of 88% to 91%. Each examination question had 5 answer choices; in cases when no reasonable fifth choice existed, 4 choices were used.

During the first 4 weeks of class leading up to the first examination, students were acclimatized to the examination format through weekly preparatory quizzes. These quizzes did not count towards the overall grade in the course. At the end of each week, approximately 3 to 5 questions were posted at the end of class, and students worked through the problems utilizing the immediate-feedback answer form. For each question on this form (Figure 1) there were 5 scratch-off blocks labeled A through E. Students were required to scratch off the block associated with their answer selection. If their selection
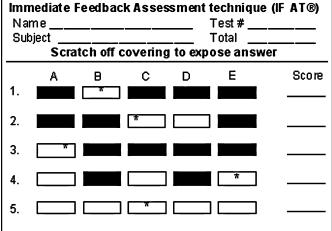
Figure 1. Depiction of examination form in its original form (top) and after answers have been scratched (bottom).

correct on the second attempt. For example, on a 36-question examination, a student could earn a 90% by answering 32 questions correctly on the first attempt and receive zero credit for the remaining 4 questions; alternatively, the student could score 90% by receiving full credit on 27 questions (ie, 5 points) and partial credit (ie, 3 points) on the remaining 9 questions. Each examination also included a scaling factor to convert the raw score to a percentage so students would know exactly how they faired on the examination in relation to the course grading scheme. The average number of questions per examination was approximately 41 on mixed-format examinations and 36 on the immediate-feedback format examinations.

Scores from the 2006-2007 academic year were compared to those from the 2005-2006 academic year using a Mann-Whitney test, with $p<0.05$ set as the criterion for statistical significance. A nonparametric test was used because the left-hand skewness of the distribution of scores and failure of the Kolmogorov-Smirnov normality test for the grade distribution. In addition, student performance was assessed based on level of the questions related to Bloom's Taxonomy. A repeated-measures ANOVA on ranks with question level as the repeated measure was used because of the failure of the normality test on score distribution and the unequal variance across the question levels. A Tukey post-hoc test was used to detect the location of any statistical differences with $p < 0.05$ set as the criterion for statistical significance.

Attitudinal survey instruments were administered at the end of each of the first 2 examinations, and included questions to probe a potential increase in anxiety due to the novel format, impact of the format on examination preparation, and fairness of the grading scale.

ASSESSMENT

In 2006 (ie, the experimental year) there was 144 students spread over 2 campuses; the Chapel Hill campus (CH) had an enrollment of 130 students and Elizabeth City State University (ECSU), a satellite campus, had 14 students. The control class (ie, the 2005), which was located in Chapel Hill and not video-teleconferenced, had an enrollment of 132 students. Based on admission profiles from the 2005 and 2006 cohorts, these classes were equivalent in composition for sex (approximately 33% vs 36%), age (19-45 years vs 19-56 years), number of students with previous degrees (approximately 65%), average Pharmacy College Admissions Test (PCAT) score (392 vs 400) and average entering grade point average (3.5 vs 3.5).

The approximate breakdown of each examination relative to level of learning is described in Table 1. Examination 1 assessed more basic knowledge, including

was the correct answer, scratching would reveal a star; an incorrect answer was indicated by a blank block.

Grading of each examination question was based on number of attempts required to arrive at the correct answer or solution for that question. Questions that were answered correctly on the first attempt (ie, one block scratched) earned full credit (5 points); questions that were answered correctly on the second attempt (2 blocks scratched) earned 3 points, and questions that were answered correctly on the third attempt (3 blocks scratched) earned 1 point. Students were encouraged to scratch as many blocks as necessary to discover the correct answer, although questions requiring more than 3 attempts earned no points for the student. With this grading scheme, students would need to answer a minimum of 75% of the questions correctly on the first attempt to obtain an A (ie, 90%), a minimum of 50% of the questions correctly on the first attempt to obtain a B (ie, >80%), and a minimum of 13% of the questions correctly on the first attempt to obtain a C (ie, >65%) with the remaining questions being

Table 1. Composition of Each Examination Based on Bloom's Taxonomy[a]

Bloom's Taxonomy	First Examination		Second Examination		Final Examination	
	No (%)	Median (IQR)	No (%)	Median (IQR)	No (%)	Median (IQR)
Level 1: Knowledge and Comprehension	16 (44)	4.7 (0.41)	8 (24)	5.0 (0.25)	9 (24)	4.6 (0.56)
Level 2: Application	16 (44)	4.8 (0.25)[b]	21 (62)	4.4 (0.71)[b]	23 (62)	4.4 (0.39)
Level 3: Analysis, Synthesis and Evaluation	4 (11)	4.0 (1.0)[b,c]	5 (15)	4.2 (0.8)[b,c]	5 (15)	3.4 (1.4)[b,c]

[a]Data are presented as number of questions and percent of total questions; question scores are presented as median and interquartile range (IQR) with each question worth 5 points
[b]$p < 0.05$ compared to Level 1 by one-way repeated measures ANOVA on ranks with Tukey post-hoc
[c]$p < 0.05$ compared to Level 2 by one-way repeated measures ANOVA on ranks with Tukey post-hoc

definitions of terms and fundamental calculations. Examination 2 focused on more conceptual aspects of the material, and applying those concepts to hypothetical real-world situations. The final examination directly assessed content from the final third of the course, but also included holistic application questions encompassing multiple aspects of the course.

Examination 1 resulted in a significantly higher performance ($p<0.001$) for the answer-until-correct format compared to performance on the mixed format examination (Table 2), but the difference (2.6 points), which was less than 1 letter grade, was not highly significant from an academic standpoint. On average, students earned 5, 8, and 14 points of partial credit on the first, second, and final examination, respectively. The medians for each level of learning indicated, in all cases, that students performed significantly lower on Level 3 questions compared to either Level 1 or Level 2 questions (Table 1). There were no differences between the 2 campuses with respect to examination grades on any examination (data not shown). In the 2005-2006 academic year, examinations were not mapped to level of learning; thus, comparisons could not be made. During the control year (mixed format questions used),

students' were provided 2 hours to complete each examination. The same amount of time was used for the first examination in the answer-until-correct format; however, instructors found that 2 hours was an insufficient amount of time to complete the examination in this format, so students were allowed 3 hours for each of the 3 examinations.

Approximately one third of the class indicated that answer-until-correct format increased their level of anxiety for the first examination 1 (Table 3). After the second examination, that percentage decreased to approximately 12% (Table 3). The majority of students indicated that they liked the immediate knowledge of whether they had arrived at the correct answer or had made a mistake (73%), and the opportunity to have multiple attempts at answering each question correctly (approximately 98%). These responses were consistent with student comments on the open-ended survey questions. Students felt they were adequately prepared for the format, and agreed that the scoring of the examinations were fair (Table 3). Eighty percent of students felt the format did not hinder their performance. When asked via open-ended questions to identify aspects of the format that they did not like, the most frequent responses were "increased anxiety" and

Table 2. Comparison of Scores Between Students in a Pharmacokinetics Course Who Took Immediate Feedback, Answer-Until-Correct Format Examinations and Students Who Took Mixed Format Examinations

Course	Examination	Median (IQR)[a]	Median w/o Partial Credit (IQR)
Students Taking Immediate	First	93.5 (5.6)[d]	89.6 (8.4)
Feedback, Answer Until-Correct	Second	90.9 (10.3)	82.6 (16.2)
Format Examinations[b] (n = 144)	Final	134 (11.4)	117 (16.2)
Students Taking Mixed Format	First	90.9 (7.9)	—
Examinations[c] (n = 132)	Second	90.0 (10.8)	—
	Final	135.0 (13.0)	—

Abbreviations: IQR = interquartile range
[a]Examinations were out of 100 points except the final which was out of 150 points
[b]Study group, students enrolled in the 2006-2007 academic year
[c]Control group, students enrolled in the 2005-2006 academic year
[d]$p < 0.001$ compared to the previous year's respective examination (Mann-Whitney test)

Table 3. Survey of Student Attitudes After Completing Immediate Feedback, Answer-Until-Correct Format Examinations in a Pharmacokinetics Course

Survey Questions	Mean
Survey Administered Following the First Examination	
I would describe my anxiety as high for this exam	3.2 (1.2)
I had more anxiety about the material for this exam than the testing method itself	3.5 (1.2)
The testing method increased my level of anxiety	2.8 (1.3)
I like that I knew whether I got the answer right or wrong immediately	4.1 (1.1)
I like that I had multiple chances to answer the questions	4.9 (0.8)
I think the scoring of questions reflects my level of understanding of the material	4.4 (0.9)
I think the scoring of questions was fair	4.7 (0.6)
I prefer to have the more difficult questions worth more points than the easier questions.	1.9 (1.1)
The in-class preparation was adequate in getting me acquainted to this type of testing	4.4 (0.8)
I would have liked more practice with this method of testing prior to the exam	2.5 (1.1)
I felt this method of testing hindered my performance	1.7 (1.0)
This exam agreed with the learning objectives for this material	4.5 (0.7)
The format of the exam changed the way I prepared/studied for the exam	2.2 (1.2)
I think I will remember what I missed on this exam and will remember it for any future cases.	3.7 (1.0)
Getting answers wrong on the exam forced me to rethink the problem and identify my errors	4.5 (0.8)
I prefer the format of this exam over typical multiple choice (ScanTron) exams	4.1 (0.8)
I prefer the format of this exam over open-ended questions and problems	4.3 (1.0)
Survey administered following second examination	
I would describe my anxiety as high for this exam	3.1 (1.2)
I had more anxiety about the material for this exam than the testing method itself	4.0 (1.1)
The testing method increased my level of anxiety	2.0 (1.2)
Compared to Exam 1, the material on Exam 2 was more challenging	3.9 (0.9)
The multiple choice format is an appropriate way to evaluate your learning about pharmacokinetics	4.4 (0.7)
How would you rate your learning in this course to this point? (5 = very high; 1 = very low)	4.4 (0.6)
I am learning at a higher level in this course compared to others (low level = memorization; high level = analyzing & evaluating information or synthesizing ideas)	4.3 (0.7)

Responses were based on a 5-point Likert scale on which 1 = strongly disagree and 5 = strongly agree.

"the mess as a result of scratching the answers." When the students were asked their preferred examination format, answer-until-correct format was their first choice (80% of responders), with their second choice being either a combination of answer-until-correct format and open-ended questions, or standard multiple-choice questions (29% versus 38% of responders, respectively). The lowest rateed options were group tests (14%) and oral examinations (66%).

DISCUSSION

This study investigated the use of immediate feedback, answer-until-correct examinations to provide a method for assessing student mastery of material, as well as an opportunity to correct any misinformation that students had. In terms of an assessment tool, this examination format appeared to be equivalent to more traditional approaches, as performance on the answer-until-correct examination was equivalent to examination scores from the previous year. The added benefit of this approach in assessing mastery of material may be the iterative nature of the format. While open-ended questions allow the instructor to assign partial credit, this can be subjective without appropriate grading rubrics, and partial credit in and of itself does not indicate whether the student can think through a particular issue or correct mistakes to solve the problem at hand. This format also allows a student to assess his or her own mastery of the material, indicates to the student areas of potential misconception, and allows the student to think about and rework problems. Each of these elements potentially increases deep learning.[9]

Because, examination scores have been historically high for the course for which this examination format was developed, there were not significant improvements in test performance (grades) to be gained. Since the use of

this format seeks to improve deeper learning and thinking skills (eg, critical thinking), improvements in long-term retention and ability to apply information to complex problems may not be reflected well in short-term assessments of student knowledge. Despite the examinations in the study year and control year being criterion-based (ie, based on stated learning outcomes), there may have been some variability in item difficulty. Given that examination scores were similar between years, there is no information suggesting that one year's examination was more difficult than another, but future use of this format will allow easier tracking of item difficulty as the results will be recorded on a question-by-question basis versus a more cumulative method (eg, points per page) used for the mixed format. The preparatory quizzes, which were not used in the control year, could possibly impact examination performance as they were a form of assessment, but it is unlikely they would have a strong impact given there were other means (ie, problem sets, old examinations) that allowed students to self-assess their knowledge base.

The students preferred this format of assessment despite the initial increase in anxiety surrounding preparation for the first assessment. When the examination format was explained to the students on the first day of class, there was a sense of trepidation. The use of preparatory quizzes acquainted the students with the format and was an important part in developing a comfort level that ultimately allowed the examinations to focus on assessing the mastery of outcomes rather than the newness of the method of evaluation. Some students reported that they would have benefited from additional practice with the format prior to the first examination, although the vast majority of students felt that the preparation was adequate. By the second examination, however, the novel format did not seem to present further test anxiety for most students. Anxiety level can impact test performance,[10-12] with test performance typically being inversely related to levels of test anxiety.[13-15] Despite potential anxiety, the students in this cohort did not report that the format of the examination hindered their performance. This is confirmed by consistent examination scores compared to historical controls. One potential reason a performance decrement was not found in this study was because the examination was formatted in terms of increasing question difficulty (ie, low-level questions in the beginning, high-level questions at the end). Wise et al noted that immediate feedback on examinations with randomly arranged questions had a negative effect on test performance and increased anxiety, whereas when the same items when arranged in terms of increasing difficulty, showed no performance decrement and less anxiety.[12]

A subgroup of individuals who experience anxiety associated with new tasks or approaches may be at a disadvantage with this type of assessment.[16] Indeed, negative feedback (ie, incorrect answers) tends to be most detrimental to performance by individuals who evidence or report a higher level of anxiety[17,18] and some studies have shown that the immediate-feedback format increases anxiety in some students.[10-12] Such students may be less comfortable with the immediate feedback approach, and their scores may be lower than those resulting from a more traditional multiple-choice answer format.[3] In contrast, other studies have indicated that immediate feedback is associated with decreases in test anxiety.[19-21] In terms of student ability, performance for low-ability students may not be negatively impacted by immediate feedback examinations, whereas feedback may impair performance of high-ability students,[22] this would be difficult to assess in this cohort of students as most would be considered high-ability students based on course performance and their entrance criteria (eg, PCAT and grade point average).

Overall students reported that they preferred the immediate feedback format, and that they preferred it over almost every other testing method to which they had been exposed previously. The 2 main reasons cited for this preference was the immediate knowledge of success or failure and the ability to make multiple attempts to ascertain the correct answer. The preference for this type of assessment is consistent with other studies in which the majority of undergraduate students preferred this method of testing in a variety of courses.[9,23] DiBattista and colleagues also have found the preference for this format to be independent of both test performance and a variety of personal characteristics.[23] As per the objectives for use of the immediate feedback format, the method resulted in students leaving the examination setting knowing not only the correct answers, but how they scored on the examination as a whole. The students also indicated that the format of the examination forced them to rethink some problems and therefore increased their understanding of topics for which they did not demonstrate initial mastery, that is, corrected any misinformation.

The largest challenge associated with this format of testing was translating formerly open-ended types of questions into multiple-choice questions. This translation required a significant amount of time to write from 2 perspectives: probing higher-level learning (according to Bloom's Taxonomy), and creating answer choices that would represent common mistakes associated with conceptual misunderstanding. This latter point is a potentially added benefit to the approach in that one can quickly scan the answer sheets to identify frequent incorrect choices

and use that information for future class discussions. More work is required, to develop better approaches to relating examination performance to content mastery and competency, and to differentiate mistakes made due to misinformation (ie, the student was confident an incorrect response was actually correct) from those associated with a lack of information (eg, a simple guess).

The use of this format for assessment will continue for several reasons. One reason is that it gives the instructor the ability to map components for topic area and learning taxonomy. This mapping can facilitate the assessment of other innovative approaches which may impact learning because grading is on a more continuous scale (ie, students can score 5, 3, 1, or 0 points) compared with the standard dichotomy of a "correct" or "incorrect" scale typical of multiple-choice formats. A second reason is the potential for this type of assessment to promote deep learning which will hopefully have a positive impact on patient care. The final reason is students' in their first year hear from students who have used this assessment method and are, in some ways, looking forward to the assessment method.

Limitations of this format are the inability to track the order of answer selection and inability to determine which incorrect answers were selected in an easy manner While this latter problem is easily resolved with ScanTron sheets used in the traditional multiple-choice format, the recording of answer selection in the immediate-feedback format requires a copious amount of data entry. A possible solution is a computer-based version of the examination that tracks students' responses. In future offerings of the course, the answer-until-correct format will be incorporated into quizzes imbedded in a multimedia-enriched environment that will allow tracking of the order of answer selection.[24] The financial cost is minimal (approximately $200 for 1000 sheets) and the instructor/faculty time required is no greater than that with any other examination formats except when writing multiple-choice questions for the higher levels of learning.

CONCLUSION

The answer-until-correct examination format allows students to re-work or re-think their mistakes, potentially resulting in deeper learning. This format also allows mapping of examination questions to learning outcomes. Students tend to enjoy this examination format, though some students experience some anxiety.

REFERENCES

1. Chickering AW, Gamson ZF, American Association for Higher Education Washington DC. Seven Principles for Good Practice in Undergraduate Education, Education Commission of the States, Denver, CO: 6. ED282491
2. Dihoff RE, Brosvic GM, Epstein ML, et al. Provision of Feedback during Preparation for Academic Testing: Learning Is Enhanced by Immediate but Not Delayed Feedback. *Psycholog Rec.* 2004;54:207-331.
3. DiBattista D, Gosse L. Test Anxiety and the Immediate Feedback Assessment Technique. *J Exp Educ.* 2006;74:311-27.
4. Kluger AN, DeNisi A. Effects of feedback intervention on performance: A historical review, a meta-analysis, and a preliminary feedback intervention theory. *Psycholog Bull.* 1996;119:254-84.
5. Dihoff RE, Dihoff RE, Brosvic GM, et al. The Role Of Feedback During Academic Testing: The Delay Retention Effect Revisited. *Psychological Record.* 2003;53:533-48.
6. Rushton A. Formative assessment: a key to deep learning? *Med Teach.* 2005;27:509-13.
7. Epstein ML, Epstein BB, Brosvic GM. Immediate feedback during academic testing. *Psycholog Rep* 2001;83:889-94.
8. Epstein ML, Brosvic GM. Immediate Feedback Assessment Technique: Multiple-choice test that 'behaves' like an essay examination. *Psychologic Rep.* 2002;90:226
9. Epstein ML, Brosvic GM. Students prefer the immediate feedback assessment technique. *Psycholog Rep.* 2002;90:1136-8.
10. Delgado AR, Delgado AR, Prieto G. The effect of item feedback on multiple-choice test responses. *British J Psychol.* 2003;94:73-85.
11. Kluger AN, Kluger AN, DeNisi A. Feedback interventions: Toward the understanding of a double-edged sword. *Current Directions Psychol Sci.* 1998;7:67-72.
12. Wise SL, Wise SL, Plake BS, et al. The effects of item feedback and examinee control on test performance and anxiety in a computer-administered test. *Computers Human Behavior.* 1986;2:21-9.
13. Clark JW, II, Clark JW, II, Fox PA, et al. Feedback, test anxiety and performance in a college course. *Pychol Rep.* 1998;82:203-8.
14. Hembree R. Correlates, causes, effects, and treatment of test anxiety. *Rev Educ Res.* 1988;58:47-77.
15. Musch J, Musch J, Broder A. Test anxiety versus academic skills: A comparison of two alternative models for predicting performance in a statistics exam. *Br J Educ Psychol.* 1999;69:105-6.
16. Hill KT, Wigfield A. Test anxiety: a major educational problem and what can be done about it. *Elem School J.* 1984;85:106-26.
17. Auerbach SM and Auerbach SM. Effects of orienting instructions, feedback-information, and trait-anxiety level on state-anxiety. *Psychol Rep.* 1973;33:779-86.
18. Hill KT, Eaton WO. The interaction of test anxiety and success-failure experiences in determining children's arithmetic performance. *Dev Psychol.* 1977;13:205-11.
19. Arkin RM, Schumann DW. Effects of corrective testing: an extension. *J Educ Psychol* 1984;76:835-43.
20. Morris LW, Fulmer RS. Test anxiety (worry and emotionality) changes during academic testing as a function of feedback and test importance. *J Educ Psychol* 1976;68:817-24.
21. Rocklin T, Rocklin T, Thompson JM. Interactive effects of test anxiety, test difficulty, and feedback. *J Educ Psychol.* 1985;77:368-72.
22. Wise SL, Wise SL, Plake BS, et al. Providing item feedback in computer-based tests: effects of initial success and failure. *Educ Psychol Meas.* 1989;49:479-86.
23. DiBattista D, Mitterer JO, Gosse L. Acceptance by undergraduates of the immediate feedback assessment technique for multiple-choice testing. *Teach Higher Educ.* 2004;9:17-28.
24. Persky AM. Multi-faceted approach to improve learning in pharmacokinetics. *Am J Pharm Educ.* 2008;72:36.