3.7

USING AN ACTIVE-LEARNING QUIZ SERIES IN AN INTRODUCTORY METEOROLOGY COLLEGE COURSE

Neil F. Laird[1], Laura A. Golson[1], Kathryn E. Wood[1], and Karen Feasel[2]

[1] Department of Geoscience, Hobart & William Smith Colleges, Geneva, NY
[2] Department of Psychology, Hobart & William Smith Colleges

1. INTRODUCTION

The purpose of this investigation is to present and assess the utility of an active-learning activity which educators can use as a simple and engaging instructional tool at the college introductory course level. It is important that students learn foundational concepts and materials that can be applied to answering "knowledge questions". These types of questions, which include multiple choice and true/false questions, are often used to evaluate the recall ability of students and assess their understanding of basic scientific relationships (McConnell et al. 2003).

The instructional tool examined in this study is a semester-long series of quizzes used during four offerings of an introductory-level meteorology course. The data were collected at the University of Illinois at Urbana-Champaign and Hobart and William Smith Colleges. The two primary objectives of the quiz series were to (a) allow students an opportunity to self-assess their understanding of course material, and (b) provide an opportunity for the instructor to assess student comprehension in an interactive manner and reinforce material discussed during previous course meetings. The current investigation focuses on the examination of the quiz series as a student self-assessment tool and whether the students' ability (or choice) to use the quiz series as a measure of their understanding was related to performance on subsequent exams.

2. METHODOLOGY

This study examines four sets of quiz data. Two data sets come from an introductory geoscience course offered during the fall 2004 and spring 2005 semesters at Hobart and William Smith Colleges, a small liberal arts institution. Each course had approximately 30-40 students. The other two sets of data were collected from an introductory atmospheric sciences course taught in the fall semesters of 2002 and 2003 at the University of Illinois at Urbana-Champaign. The introductory course taught at the University of Illinois had approximately 170-190 students during each semester. The organization of the course and topics taught remained similar across all semesters.

Author address: Neil Laird, Department of Geoscience, Hobart & William Smith Colleges, 4002 Scandling Center, Geneva, NY 14456. email: laird@hws.edu

The quizzes were administered at ten random class meetings during each semester. The structure of each quiz was based on the procedures used for the popular television game show, "Who Wants to be a Millionaire?" Each quiz contained six multiple choice questions, with individual questions worth the point value of the question number. For example, the second question was worth 2 points. Twenty-one points were possible on each quiz when all six questions were answered correctly. If the student incorrectly answered questions 1, 2, or 3, they received a score of 0 points. If the first three questions were answered correctly, the student established a threshold score of 6 points. This threshold score was received with an incorrect answer to questions 4, 5 or 6.

Similar to the game show, the unique quiz structure allows students to decide at which question to stop answering, without penalty. Additionally, the risk of losing points by incorrectly answering questions was included in an attempt to promote student self-assessment beyond the standard knowledge question approach. Table 1 shows the number of points which a student risked losing in association with an incorrect answer, provided all previous questions had been answered correctly. Notice that the points risked are always less than or equal to the potential points gained for questions 1 – 5. For question 6, the risk of answering incorrectly greatly increases the potential of losing points. This aspect is a significant deviation from the standard approach of multiple choice assessment where students *only accumulate* points with correctly answered questions and are typically not penalized for incorrect answers.

Question Number & Point Value	Points Risked
1	0
2	1
3	3
4	0
5	4
6	9

Table 1. Quiz point structure and points risked with each question assuming all previous questions were answered correctly.

Each quiz was considered a bonus opportunity and performance was not directly factored into a students' semester grade. Instead, the cumulative score of all quizzes was determined at the end of the semester and top students received their choice of two awards. The top students were identified by either placing in the highest 10% of quiz score totals or reaching a previously established threshold of points (e.g., 149 points). The awards presented at the end of the semester included the addition of 20 bonus points to any exam grade or a grade of an 'A' on the third non-cumulative exam without attendance. The awards were significant to provide a substantial incentive for students to maintain active participation in course discussion and review of materials.

Three non-cumulative exams were given during each semester to evaluate student understanding of the course material. Table 2 shows the number of quizzes given each semester and their distribution in relation to each exam. The groupings prior to each exam (A, B, and C represented by colors in Table 2) were used for derivation of comparative variables and conducting statistical analyses. In order to provide a representative sample for investigation of the quiz series as a student self-assessment tool, students who did not complete at least five quizzes (half of the quizzes offered) were excluded from the analyses. This resulted in a utilized sample size of 365 with the exclusion of 68 from our total sample of 433 students.

The two main questions addressed in this study are (a) Does this unique quiz structure promote self-assessment by students? and (b) Is self-assessment used during the quizzes related to performance on subsequent exams?

	FA2002	FA2003	FA2004	SP2005
Q1	9/12	9/15	9/15	2/3
Q2	9/17	9/25	9/24	2/8
Q3	10/1	10/7	9/27	2/17
Q4	10/22	10/23	10/8	3/24
Q5	10/29	10/28	10/18	3/31
Q6	10/31	11/4	10/25	4/12
Q7	11/19	11/18	10/27	4/14
Q8	11/21	11/20	11/8	4/19
Q9	12/5	12/4	11/15	4/26
Q10	12/10	12/9	11/22	4/28

Table 2. Distribution of quizzes during each semester. Colors represent similar quiz grouping prior to each of the three exams.

3. RESULTS & ANALYSES

The analyses presented in this section investigate the level of self-assessment which students used during quizzes and whether these opportunities for self-evaluation of their knowledge translated into enhanced performance on exams. Two important variables collected from each quiz were the number of questions answered and the number of questions answered correctly. Examination of these variables will set the foundation for the evaluation of student usage of the quizzes. An additional variable used to determine the level of student self-assessment is the Accuracy of Self-Assessment (ASA), defined as the proportion of the number of questions a student answered correctly divided by the number of questions answered.

Figures 1a, 1b, and 1c show the frequency distributions of the number of questions answered on quizzes during the semester. The distributions are presented for each of the three quiz groupings which were differentiated using exams (see Fig. 1). The mean of questions answered increased steadily through the first (mean = 5.04), second (mean = 5.20), and third groupings of the semester (mean = 5.58). The mean values and distributions show that a large number of students typically answered at least 5 questions on each quiz during the semester. These results suggest that the increased amount of points at risk from questions 5 to 6 may have been a greater limiting factor towards the beginning of the semester. As the semester progressed, students either (a) became more familiarity with the quiz structure, (b) were willing to accept more risk in an attempt to obtain the bonus award, or (c) had a greater understanding of complex material and concepts.

The number of questions answered correctly was used as a measure of student performance. Figures 2a, 2b, and 2c show the frequency distribution of the number of questions answered correctly on quizzes. The mean increased through the first (mean = 3.58), second (mean = 3.77), and third groupings of the semester (mean = 4.44). The distribution in the third grouping shows a noticeable shift towards more students correctly answering 5 to 6 questions. The percentage of students in that range increasing from 9.3% and 17.5% for groupings A and B, respectively, to 47.2% for grouping C. These results show that student performance improved during the semester; however this improvement may have been a result of several factors, such as (a) students' increased familiarity with the quiz structure during the semester, (b) an increase in the number of questions attempted, and (c) greater understanding of difficult material and concepts.

As already discussed, a main factor which adds to the uniqueness of the quizzes is the opportunity for students to assess their own knowledge. Unlike other quizzes, students may be more likely to use a specific approach for answering questions to avoid the risk of losing points already gained. The approach to taking

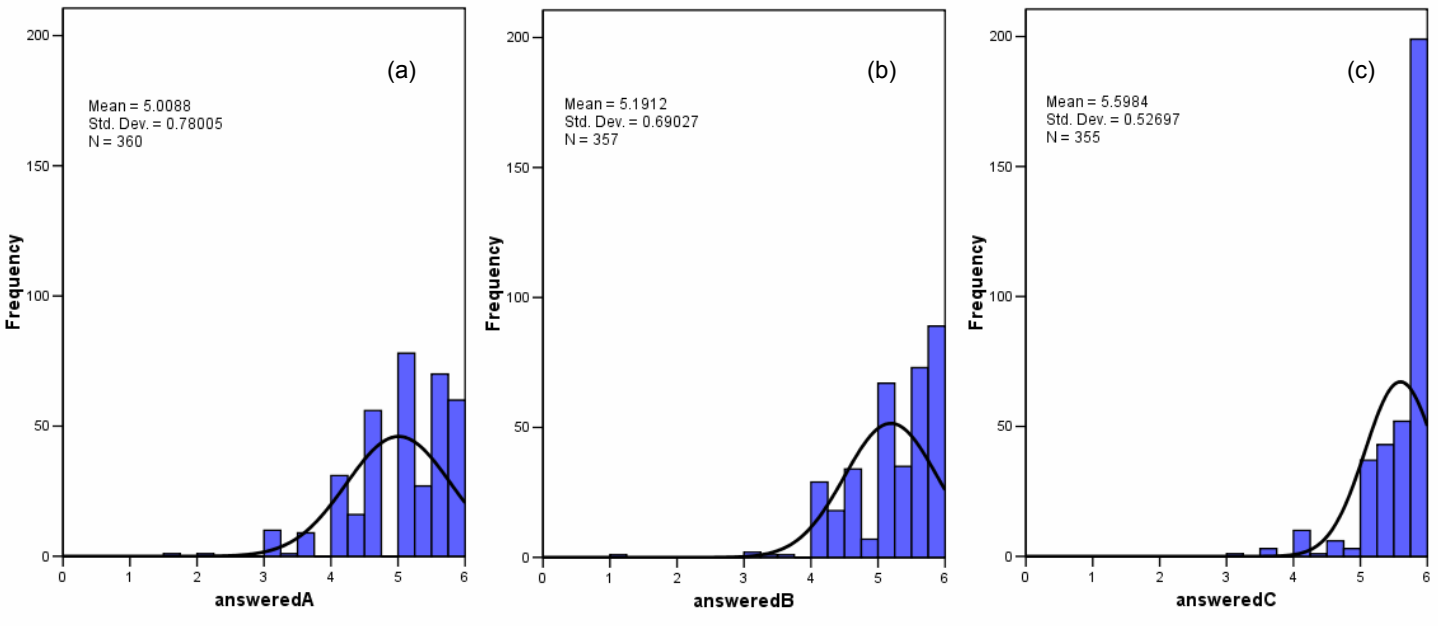

Figure 1. Distribution of the average (mean) number of quiz questions answered by each student for quizzes prior to exam 1 (grouping A), exam 2 (grouping B), and exam 3 (grouping C). The mean and normal distribution (black line) of each grouping is shown.

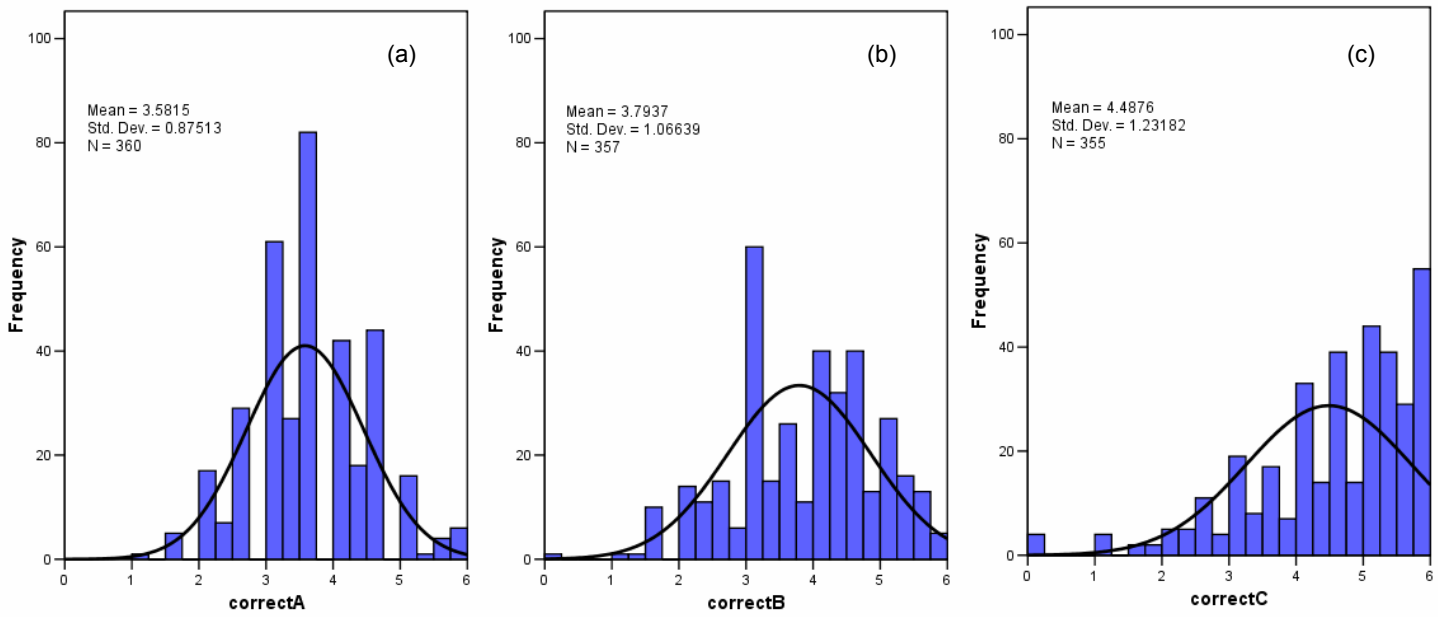

Figure 2. Distribution of the average (mean) number of quiz questions answered correctly by each student for quizzes prior to exam 1 (grouping A), exam 2 (grouping B), and exam 3 (grouping C). The mean and normal distribution (black line) of each grouping is shown.

the quiz may be different for each student, although two pronounced approaches can readily be identified with our data; (1) students who stop answering questions when they cease to know the information, and (2) students who answer questions regardless of their understanding of the material (i.e., guessers). These groups can be identified by examining values of ASA, a representative measure of student self-assessment. The values of ASA range from 1.0, all attempted questions were correctly answered, to 0.0, all attempted questions were incorrectly answered.

Figures 3a, 3b, and 3c show the frequency distribution of ASA for quiz groupings A, B, and C. The results clearly demonstrate that most students are achieving an ASA value of 0.6 by successfully attempting questions 1, 2, and 3, since the average number of question answered was ≥ 5 (Fig. 1). An ASA value of 0.83 is equivalent to a student answering 5 of 6 questions correctly. As evident in Figs. 3a, 3b, and 3c, the percentage of students successfully answering 5 of 6 questions increases markedly from 23.1% and 34.1% for groupings A and B, respectively, to 57.2% for grouping C. This improvement suggests that students more accurately assessed how much information they knew as the semester progressed.

Students armed with a better assessment of their knowledge of course material prior to exams should be in a better position to study more effectively and score higher on an exam than they would have been without this understanding of their limitations. Figures 4a, 4b, and 4c compare ASA derived from quiz information to subsequent exam scores for each student during quiz groupings A, B, and C. It is evident that as the semester progressed, both ASA and exam scores increased. The Pearson Correlation, a bivariate correlation measured from -1.0 to 1.0, was calculated using student ASA values and exam scores within each quiz grouping. The correlation was positive and statistically significant at the 0.05 level for grouping A and at the 0.01 level for groupings B and C. The correlations between ASA and exam score for groupings A, B, and C were 0.112, 0.154, and 0.312, respectively. These results indicate that students who more accurately assessed their knowledge during the quizzes (i.e., large ASA value) tended to score higher on exams, and that this relationship strengthened during the semester.

4. CONCLUSIONS

The unique quiz, fashioned after the television game show “Who Wants to be a Millionaire?”, appears to successfully function as a self-assessment tool for students in an introductory meteorology course. Most students answer fewer than 6 questions, suggesting that they are actively assessing the limits of their knowledge when they take the quizzes. On average, students tended to both answer and correctly answer more quiz questions as the semester progresses.

ASA also improved over the course of the semester, such that students on average got more of the questions they answered correct. These results suggest the plausibility of a causal link between the quiz structure (i.e., the structural features that encourage self-assessment) and accuracy of self-assessment. It is not possible on the basis of these correlations and analyses, however, to rule out other factors (e.g., acquiring a better fundamental knowledge of the atmosphere over the course of the semester, and "practice effects" associated with increasing experience with the quiz structure) that may have contributed to improved ASA.

The modest, but statistically significant, positive correlations between ASA and exam performance suggest that improved ASA is an advantageous outcome, since accuracy of self-assessment predicted subsequent indicators of learning. Causal links also cannot be inferred from these relationships (e.g., it may simply have been the case that students who more accurately assessed the limits of their knowledge also knew the course material better). These correlational results most likely underestimate the strength of the relationship between ASA and exam performance; as computed, the ASA variables are imperfect measures of self-assessment accuracy, since the risk of losing points for correct answers is probably not great enough to discourage guessing until questions 5 or 6 on the quizzes.

The effectiveness of this quiz method as a tool that encourages students to evaluate the limits of their knowledge might be improved by eliminating or lowering the “safe” 6 point (i.e., question 3) threshold, although the potential benefits of doing so must be weighed against the potential costs associated with structuring the quiz in a manner that encourages students to answer fewer items.

Acknowledgment: This research was completed by the second and third authors as an undergraduate independent study project during the fall 2005 semester at Hobart and William Smith Colleges. Support for conference registration and attendance was provided by the Department of Geoscience and the William Smith Deans office.

5. REFERENCES

McConnel et. Al., 2003, Assesment and active learning strategies for introductory geology courses, *J. Geoscience Edu.* **51**, 205-216.

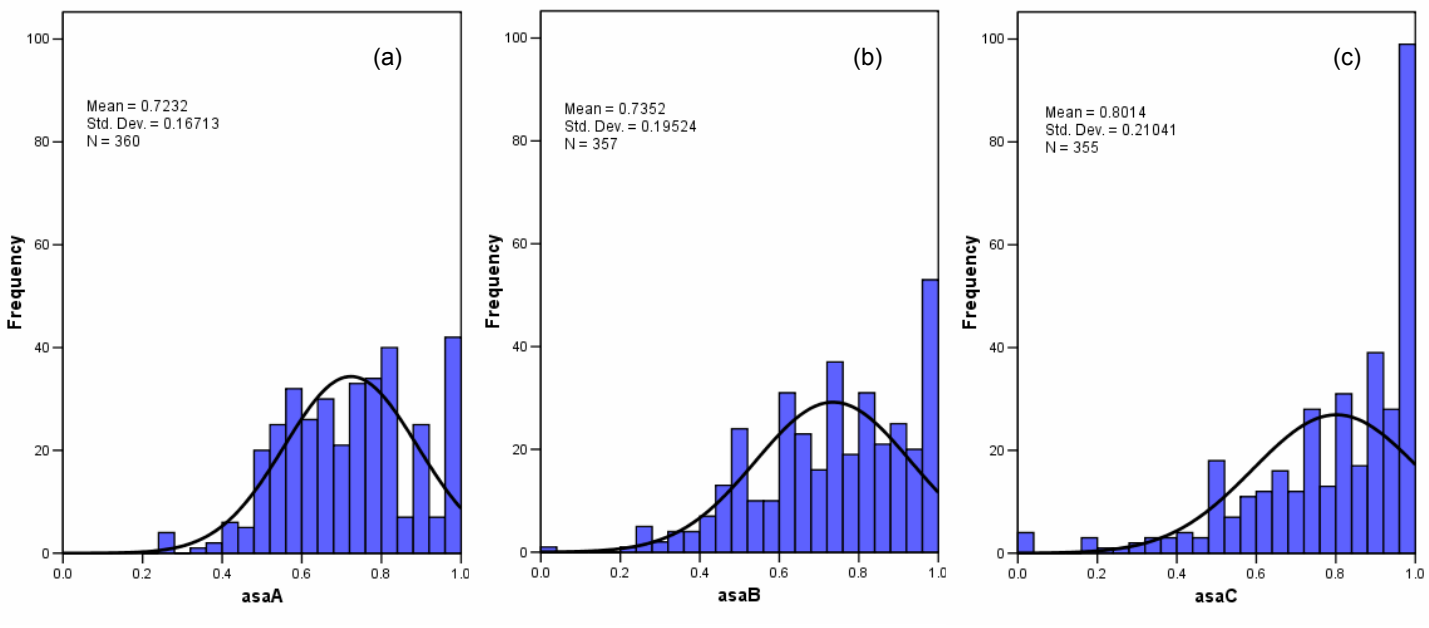

Figure 3. Distribution of the average (mean) accuracy of self-assessment (ASA = # correct / # answered) for each student for grouping A, grouping B, and grouping C. The mean and normal distribution (black line) of each grouping is shown.

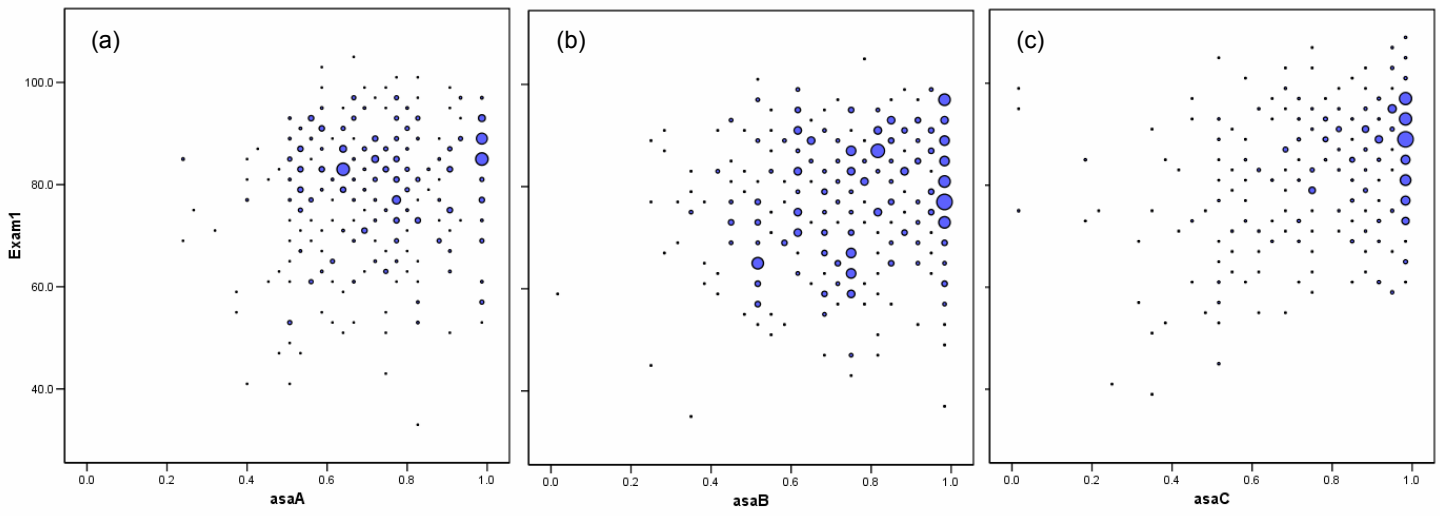

Figure 4. The average (mean) accuracy of self-assessment (ASA) for grouping A, grouping B, and grouping C versus the corresponding subsequent exam score. The size of the plotted point represents the number of students within a bin (i.e., larger plotted point = more students).