Presented at the Research Congress 2013
De La Salle University Manila
March 7-9, 2013

THE QUESTIONS FOR ASSESSMENT (QfA) SCHEME IN A PHYSICS CLASS: AN ALTERNATIVE ASSESSMENT METHOD FOR IMPROVED METACOGNITIVE AWARENESS, CONCEPT RETENTION AND TRANSFER?

Resty C. Collado[1], Lydia S. Roleda[2]
[1]Department of Natural and Engineering Sciences, Miriam College High School
[2]Science Education Department, De La Salle University

Abstract. This study explored the effects of writing multiple choice questions (MCQs) to students' performance in physics tests, retention and transfer of physics concepts, and metacognitive awareness. In lieu of teacher-made seatwork, homework and quizzes, the students in an experimental group (N=36) wrote MCQs (N=1,942) with accompanying justification for their alternatives, a feature of the Questions for Assessment (QfA) scheme. Their performance in three separate achievement, retention and transfer tests were compared against a control group (N=34). One-tailed z-tests at 0.05 confidence level confirm significant difference in the performance of the groups in two achievement tests (lowest p = 0.04), with the experimental group posting the higher average. One retention test (p = 0.03) suggests that QfA has an impact to retention of concepts, although studies done in a longer period could rule out the adjustment period of the students in adapting to the method. None of the transfer tests show higher performance of the experimental group, but as time progresses, a trend of decreasing *p*-value can be observed and reaching significant level. Schraw and Dennison's (1994) Metacognitive Awareness Inventory (MAI) was also administered to all students before and after the study. A significant increase in MAI scores was observed in the experimental group, suggesting that the writing of MCQs help students become more metacognitively aware. In conclusion, QfA improves student performance in physics tests, retention of physics concepts, and metacognitive awareness, but not transfer.

Key Words: multiple-choice questions; metacognition

1. INTRODUCTION

Low comprehension impedes higher cognitive skills (Newman, 1977). Rosenshire et. al. (1996) meta-analyzed more than 17 reading comprehension studies and found out that students who wrote questions about a reading selection have higher comprehension that those who do not, which serves as a primary basis for this study. Instead of simply writing questions, some studies required the students to write MCQs for a pharmacy class (Pittinger & Lounsberry, 2011) and an organizational behaviour class (Fellenz, 2004). Other studies required students to write the justifications for the correctness or wrongness of the alternatives of an MCQ (Dodd & Leal, 1988; Nield & Wintre, 1986; Faize, Dahar & Niwaz, 2010). This study required students to write

MCQs with accompanying justifications, which were treated as their homework, seatwork and quiz. As an abstruse subject, physics requires a high level of comprehension of its concepts.

2. METHODOLOGY

All participants of the study are physics students of class 2013 of Miriam College High School (MCHS). Special attention was given to controlling other variables and reducing researcher bias. Both control (N=34) and experimental (N=36) groups have the same group average science grade from first year to third year, received the same content by means of a unified learning plan and discussion resources, and the same amount of time on a particular topic.

The researcher assigned another teacher to teach the groups (hereto referred as *QfA Implementing Teacher*), and another teacher to write the tests (hereto referred as *QfA Achievement Test Writer*). This procedure ensures minimal interaction between the groups and the researcher. An orientation as regards QfA scheme was held first in the experimental group. This orientation covers the discussion of Bloom's taxonomy, guidelines for writing MCQs and the marking system of QfA seatwork, homework and quizzes. While the experimental group was being oriented to the scheme, no course content was discussed in the control group. The study lasted for two months.

The QfA Scheme and the Experimental Group. Students in the experimental group made MCQs after a particular topic has been discussed. These were treated as seatwork and were called *Challenge the Teacher* seatwork. This routine follows a *reciprocal teaching* style, where the students write the questions (no justifications and indicator of the correct response) which the teacher answers. A volunteer student presents her MCQ to be answered by the teacher, or the teacher selects from the submitted MCQs if there is no volunteer. The MCQ is then projected in front of the class. A digital scanner was set up for this purpose. While answering, the teacher justifies each option as to why it is correct or wrong, and writes the justifications on the board while thinking aloud. All student-written MCQ homeworks are written on the *Aha! Notebook*. The alternatives of MCQ-homework must be justified and the correct response indentified by the student. Each seatwork and homework is worth 5 points, where knowledge and comprehension questions were given 2.5 points, and analysis and applications questions were given 5 points. For QfA Quizzes, the students are allowed to write a maximum of 1 knowledge question, 2 comprehension questions, and 4 application or analysis questions, which correspond to 1, 2 and 3 points, respectively. In quizzes, alternatives must be justified and the correct response must be identified as well. Every quiz is worth 17 points. Students may prepare for a QfA quiz by writing questions the day before the quiz, but *forced content* in QfA quizzes allows some important chapter competencies to be tested by the teacher and also a motivation for students to study since they cannot fully prepare for quiz due to teacher placed forced content.

The Control Group. While the experimental group was being oriented to the scheme, no course content was discussed in the control group to guarantee that the groups will start receiving physics lessons at the same time. Activities such as film viewing, laboratory groupings and science project brainstorming were used to fill-in the control group's time. The same teacher who taught the experimental group also taught the control group, where students receive teacher-made seatwork, homework and quizzes. To ensure that the control group receives the same course content and discussion strategy, the researcher developed learning plans and PowerPoint presentations that were used by the QfA Implementing Teacher and the QfA Achievement Test Writer in their classes. These resources were aligned to the course syllabus for the school year which was jointly developed by all the physics teachers in MCHS. Consistency of teaching methods and course contents were further elaborated in the weekly meetings of the researcher with all the physics teachers.

The Tests. After chapters of thematic unity were discussed, an achievement test (AT) was administered. These ATs are in MCQ format. To reduce researcher bias, these were written by the QfA Achievement Test Writer who did not have any knowledge of the study. A total of three ATs were taken by the students in the duration of research. The scores in these ATs were used to gauge the effect of writing MCQs to student performance in physics tests.

Each AT had a corresponding transfer and retention test (T&RT). T&RTs are surprise tests to effectively measure retention. These were simultaneously given to the groups. The first T&RT was administered 18 days after AT1; T&RT2 twenty-two days after AT2; and T&RT3 eleven days after AT3. T&RTs have two parts. The first section is a test of retention. The items in this part are in MCQ format. The second section is a test of transfer which contains highly contextualized items which are in essay or problem-solving format.

To gauge how much the students were practicing self-regulatory habits, Schraw and Dennison's MAI was also administered to the students following a pre- and post-test design.

3. RESULTS AND DISCUSSION

Effect of Writing MCQs to Student Performance in Physics Tests. The effect of QfA to student performance in physics tests was measured using the three ATs. The results of the three ATs are summarized below:

Table 1. Summary of test results for Achievement Tests (ATs)

	Control (C)		**Experimental (E)**	
Achievement Test (max score)	N_C	$\bar{X}_C \pm s_C$	N_E	$\bar{X}_E \pm s_E$
1 (49)	33	29.73 ± 6.24	33	29.48 ± 5.05
2 (38)	32	22.75 ± 5.15	35	24.89 ± 4.84
3 (38)	34	24.79 ± 5.75	36	26.92 ± 4.16

The experimental group significantly outperformed the control group in two out of three ATs. The effect size for the second and third AT is 0.4 which means that 66% of the control group is below the average student in the experimental group. This implies that the practice of writing MCQs has a positive effect on the performance of students in a physics tests.

Effect of Writing MCQs to Student Retention of Physics Concepts. The first part of the T&RTs, which is a test of retention, was used to gauge how much MCQ-writing had affected the forgetting curve of students. The following table summarizes the retention test results:

Table 2. Summary of test results for Retention Tests

	Control (C)		Experimental (E)	
Retention Test (max score)	N_C	$\bar{X}_C \pm s_C$	N_E	$\bar{X}_E \pm s_E$
1 (19)	33	9.36 ± 2.92	34	8.71 ± 2.90
2 (20)	32	10.78 ± 2.42	36	11.08 ± 3.34
3 (20)	33	8.19 ± 3.03	36	9.56 ± 3.07

The experimental group outperformed the control group in terms of the average score in two out of three tests. The effect size for the second T&RT, part I, is 0.12 which means that 54% of the students in the control group are below the average student in the experimental group; while T&RT3 posted a 0.5 effect size, implying that the average student in the experimental group outperformed 69% of the control group. Based on statistics, however, only one retention test shows a significant difference in the performance of the groups. This could be attributed to the number of days elapsed before the T&RT was administered. The writing of MCQs has a positive effect on student retention, and the forgetting curve difference between the groups is significant only up to about 11 days.

Effect of Writing MCQs to Student Transfer of Physics Concepts. The second part of the T&RTs was used to gauge how much concepts the students were able to transfer to respond to test items that are not in MCQ and are highly contextualized. For example, instead of simply solving uniform motion problems, a transfer problem asks the students to synchronize a clock on the Earth and on Mars using uniform motion; or make a reaction time strip using principles of free-fall. A well-defined rubric was developed to mark student responses to transfer tests. The following table shows the results of the transfer tests:

Table 3. Summary of test results for Transfer Tests

	Control (C)		Experimental (E)	
Transfer Test (max score = 10)	N_C	$\bar{X}_C \pm s_C$	N_E	$\bar{X}_E \pm s_E$
1	33	3.12 ± 2.23	34	3.76 ± 2.34
2	32	3.94 ± 2.14	36	4.72 ± 2.53
3	32	2.09 ± 1.09	36	2.64 ± 1.61

Average-wise, the experimental group outperformed the control group in all transfer tests. The effect size for the first, second and third transfer test is 0.3, 0.4 and 0.5, respectively. This translates to more than half of the control group (62%, 62% and 66%, respectively) performing below the average student in the experimental group in any of the transfer test. However, the difference between the group's performances is statistically insignificant.

Nonetheless, it can be observed that the difference is approaching significant levels as time progressed.

***Adjustment Effect*.** It can be observed that in all the types of tests, the experimental group had a performance that is about the same as the control group in the first administration of that type of test. This is attributed to the fact that writing MCQs is new to the students and it took time for the students to adjust to the procedure. The students in the experimental group reflected this on their feedback as regards the scheme which was collected at the end of the study.

Effect of Writing MCQs to Metacognitive Awareness. Schraw and Dennison's Metacognitive Awareness Inventory (MAI) was used in this study to measure how self-regulating the students' habits are. From the original 100-mm continuous scale, the study adopted an 8-point Likert-type scale, "1" which means *"Absolutely false of myself"* and "8" which means *"Absolutely true of myself."* Fifty-two statements in the inventory reflect self-regulating practices; thus a student with a high MAI score is more metacognitively-aware than a student with a low MAI score. The following table shows the results of the pre- and post-MAI.

Table 4. Summary of test results for MAIs

	Control (C)		**Experimental (E)**	
MAI (max score = 8)	N_C	$\bar{X}_C \pm s_C$	N_E	$\bar{X}_E \pm s_E$
Pre	28	5.84 ± 0.57	34	6.09 ± 0.58
Post	28	5.90 ± 0.64	34	6.53 ± 0.54
Difference		0.07 ± 0.46		0.44 ± 0.55

The students in the experimental group significantly improved their metacognitive awareness after two months of writing MCQs, unlike the students in the control group who had teacher-made seatwork, homework and quizzes.

Schraw and Dennison (1994) also categorized MAI items as either *knowledge of cognition* or *regulation of cognition*. The former category implies that a person knows how, when, and why to use a metacognitive skill, while the latter category implies that a person knows how to manage learning. Each category has further classifications. Out of the 52 MAI items, the experimental group posted the highest mean differences in five items categorized as *monitoring, declarative knowledge, evaluation* or *planning* metacognitive skills.

Students who wrote MCQs increased a lot of their *planning* metacognitive skills. MAI items categorized as such are those which reflect methods in preparation for learning. As evaluated by the students themselves, there is an increased level in the time spent for studying due to QfA scheme. By writing MCQs about the topic learned, the students anticipate questions and errors as well. Furthermore, as Fellenz concluded in his study, the students can develop a better test preparation strategy and a firm grasp of how an MCQ works (Fellenz, 2004).

The MAI average differences of the two academic groups are equal to each other. This implies that writing MCQs will help increase the metacognitive awareness of any student regardless of academic standing.

Seven Students' Misconceptions. At the end of two months, not less than 1,492 MCQs were written by the students; each was checked and analyzed by the researcher for qualitative data. Seven science misconceptions were identified in the student-written MCQs.

1. *Inability to Properly Estimate the Magnitude of a Dimension.* Students provide numerical values depicting dimensions that are unrealistic. This includes a "26.3 m high apple tree," "a volleyball [that] has a mass of 7 kg," and "a fast plane that has a speed of 1.5 m/s."
2. *Misinterpreting a Displacement vs. Time Graph.* Students' written MCQs show inability in identifying the displacement and time axes of a graph. This led to erroneous analysis of graphical data. Moreover, students interpret a displacement vs. time graph with a constant slope as the movement of an object that is accelerating.
3. *Identifying a Force as Something an Object Has.* Students wrote MCQs about forces which imply that a force is something that an object owns. For example, "*find the ball's static friction," "An airplane… was giving off 22 N [of] force,"* and *"…the force of gravity of the hanging shorts."*
4. *Failing to Anticipate Other Possible Equations which Yield Different Answers.* Several students gave empirical values in the MCQ stem which do not agree with the indicated key response due to two applicable equations which result to different answers.
5. *Considering the Mass of a Falling Object to Explain its Motion.* Most students in the experimental group are still Aristotelian.
6. *Failing to Cite Velocity Against Acceleration.* Student-written MCQs ask for the direction of the net force even without stating if the object is speeding up or slowing down, or asking for the sign of acceleration without stating if the velocity is constant or changing, or where the velocity is directed.
7. *Addressing Physics Concepts with Improper Terms.* The misuse of the word *weight* when in fact the student was pertaining to *mass* is one of the physics concepts. Another example is when an empirical value has been incorrectly described by a physical event (or vice-versa, relative to the intention of the student). In one instance, a student wrote *"A piece of gum rolled in paper was thrown downwards…"* while the student is thinking of a zero initial velocity.

4. CONCLUSIONS

With the quantitative and qualitative data gathered from the study, QfA is a pedagogical tool with four uses. First, it could be used to enhance cognitive skills, starting from comprehension, which leads to better performance in physics tests. Second, QfA could be used to improve the metacognitive skills of students regardless of their academic standing. Third, student-written MCQs could be used to identify student misconceptions. And fourth, student-written MCQs could be used as formative or summative assessment.

When students write MCQs about topics discussed in a physics class, the students perform better in physics tests and retain the concepts better than those who receive teacher-made assessment tasks. QfA helps improve student performance in physics tests, enhance concept retention and increase metacognitive awareness, but not concept transfer.

5. ACKNOWLEDGEMENT

This paper would not be completed without the assistance of the Department of Science and Technology – Science Education Institute.

6. BIBLIOGRAPHY

Dodd, D., & Leal, L. (1988). Answer justification: removing the "trick" from multiple-choice questions. Teaching of Psychology, 15, 37-38.

Faize, F., Dahar, M., & Niwaz, A. (2010). Improving students' performance on multiple-choice questions in physics - practice in justifying the options. International Journal of Academic Research, 26, 204-207.

Fellenz, M. (2004). Using assessment to support higher level learning: the multiple choice item development assignment. Assessment & Evaluation in Higher Education, 29, 703-719.

Nield, A. F., Wintre, M. G. (1986). Multiple-choice questions with an option to comment: student attitudes and use. Teaching of Psychology, 13, 196-199.

Pittinger, A., Lounsberry J. (2011). Student-generated questions to assess learning in an online orientation to pharmacy course. American Journal of Pharmaceutical Education, 75, 1-8.

Rosenshine, B., Meister, C., & Chapman, S. (1996). Teaching students to generate questions: a review of the intervention studies. Review of Educational Research, 181.

Schraw, G. & Dennison, R. S. (1994) Assessing metacognitive awareness. Contemporary Educational Psychology, 19, 460-475.