An Examination of the Relationship between Active Participation in Test Development (APTD), Student Performance and Student Attitudes

Robert B. Sweeney, Jr.†
Kelly Mosteller, Masters Graduate Student
Roy J. Daigle‡

University of South Alabama
School of Computer and Information Sciences
Mobile, Alabama 36688
(334) 460-6390

Abstract

Many advanced courses in computing curricula seek to combine theory and skills through complex projects. Courses of this type may include applications development, database programming, systems analysis and design, senior project, and applied software engineering. These courses present challenges to both student and instructor for maintaining a global view the project while working at the detail level and for developing a higher level of understanding of the project. Previously, one of the authors used an approach called Active Participation in Test Development (APTD) as an attempt to address these challenges. The basic belief underlying APTD is that by providing students with the opportunity to participate in the examination generation process, they are given the chance to reflect on the meaning of in academic terms, to discover a standard by which their understanding might be measured, and to apply that standard in self-assessment. The objective of this paper is to report the results of a study of the influence of the approach on student attitudes and performance in the course.

Keywords: Active participation in test development, student performance, student attitudes, depth of knowledge, complexity

1. BACKGROUND

Recently professional computing organizations have prepared curricula documents [ACM 1991, IS 1997] that emphasize the need for higher levels of knowledge and skills in problem solving, teaming, and all forms of communication from our computing graduates. Many approaches have been proposed to the educational community: employment in the curriculum of compatible individual and group problem-solving models, use of cooperative and collaborative in-class assignments, adoption of active learning approaches, implementation of approaches that emphasize higher levels of development of written and oral communication skills [Murry 1990]. Another approach, Active Participation in Test Development (APTD) [Daigle, 1998], seeks to contribute to student development in all of the emphasis areas by using a collaborative approach among students and instructor for test development in project-oriented courses.

The most elementary view of testing is as a means of assessing student performance. However many instructors view testing in a broader context: testing is seen as an extension of the teaching process itself. Some testing methods involve giving students a second opportunity to complete exam questions, thus allowing them a fresh start in which learning may occur as well as an opportunity to improve grades. Other methods allow students to work in teams, which encourages them to simultaneously review and teach each other, to gain experience in debate, and to improve grades [Murry 1990]. These and other testing approaches are designed to be a part of the teaching process.

† sweeney@cis.usouthal.edu
‡ daigle@cis.usouthal.edu

Educators often observe poor student preparation for exams. Even when students have prepared for an exam in advance, there is always some uncertainty that reviewed material will appear on the exam. This uncertainty increases for courses involving projects from which students are additionally expected to observe project generalizations. The APTD approach was used to encourage early preparation for exams and practice in articulating generalizations independent of a specific project. The approach also provides feedback to the instructor regarding the concepts students consider important and their understanding of the higher-level implications of a project. This feedback affords the instructor an opportunity to provide remediation and clarification.

As stated in [Daigle 1998], Active Participation in Test Development involves a four-phase process: Call of test item submission, Class cooperation for integration and review to prepare a test, Instructor review of prepared test, and Instructor preparation of the examination. The perceived benefits for the approach were based only upon instructor observations. The goal of this research is to study the influence of the APTD approach on student attitudes and performance.

2. RESEARCH OBJECTIVES

The goal of this research is to determine if there is a relationship between the use of APTD and student performance on examinations. APTD involves allowing students in a class to prepare questions and answers suitable for inclusion on a class examination periodically throughout the duration of the class. The use of APTD should result in a better understanding of the course material and a deeper understanding of a project, which in turn should result in higher examination scores. Students who prepare more questions are expected to perform better than those who prepare fewer questions. Students who prepare questions of greater knowledge complexity are expected to perform better than those who prepare lower knowledge level questions.

The purpose of APTD is to involve students in the creation of course examination material. The reason for this involvement is to increase the student's knowledge, understanding, and retention of the course material, which will then hopefully result in better student performance in the course. The theory we are studying is that the use of APTD is related to student test performance.

Three hypotheses related to student performance were developed:
H1: The quantity of student questions submitted as part of APTD will not affect the student’s test performance.
H2: The complexity of student questions submitted as part of APTD will not affect the student’s test performance.
H3: The interaction of quantity and complexity of the student questions submitted as part of APTD will not affect the student’s test performance.

In addition, the survey collected student attitudes regarding their preferences for subject areas covered in the course and as to their overall impression on the use of APTD in the course. Two additional hypotheses related to student attitudes were developed:
H4: Students will not have a preference for particular subject areas in the course.
H5: Students will be indifferent as to their opinions on the use of APTD in the course.

Some key benefits of APTD have been noted which support the validity of these hypotheses.

1) Students reflect on the material in the course to extract appropriate questions. This requires more than memorizing facts, it requires reflective thought on the material.
2) The students give feedback to the instructor before the test. Thus, it is more apparent when the students do not understand the instructor intention for topic importance.
3) The instructor may use this feedback fine-tune lectures for appropriate emphasis of key topics.
4) The students carry knowledge and understanding of material to future classes thus better preparing the student for the materials in those classes. The APTD teaches a methodology for preparation for test review in this and future classes.

The Depth of Knowledge scale from ISECON ’97 was used to operationalize complexity because it was developed specifically for use in Information Systems-related curricula. The Depth of Knowledge scale is based upon a Bloom’s Taxonomy of Knowledge [Bloom 1956], a widely used theory regarding how knowledge can be divided into different levels of complexity. In his taxonomy, Bloom identifies six levels of comprehension in increasing order of complexity ranging from a simple recitation of facts (Level 1) and a usage of facts (Level 2) up through the highest level (Level 6) where new facts are evaluated in comparison to extant knowledge. Table 1 illustrates how Bloom’s Taxonomy is related to the IS ’97 Model Curriculum [Davis, et al. 1997].

3. METHODOLOGY

The study involved students in a database-programming course who are preparing for entry into a masters program in Computer Science or Information Systems; they possess an undergraduate degree but they lack a foundation in computing. Prerequisites to this course consist of accelerated courses in object-oriented programming (Java), data and file structures (C++), architecture and operating systems, and networks and communications. The course is an accelerated course covering topics typically found in two courses in the Information System undergraduate curriculum. The course examines implementation and access of databases via event-driven applications developed with visual programming tools and covers other topics such as elementary E-R modeling, data integrity, referential integrity, report development, and interface design. For our purposes, this course was divided into approximately 10 learning units that grouped similar course material together. After each learning unit had been covered, homework was administered requesting 2 to 3 test-quality submissions from each student. These submissions were to consist of test-quality questions and appropriate answers. Although students were aware of depth of knowledge levels from the accelerated object-oriented programming course (two semesters earlier), there was no review of these levels provided nor were there example questions. Collected submissions were separated into categories based on the IS ’97 Depth of Knowledge guidelines outlined in Table 1.

Questions were divided into three categories of complexity based on the three lowest levels of the IS ’97 Depth of Knowledge categorization. No questions that would appear to fall into the fourth or fifth IS ’97 Depth of Knowledge categories were submitted. The first of these categories involves simple questions based upon “awareness” information. These are the most basic type of questions requiring the lowest level of knowledge. The second category is described as “literacy" and has a higher level of knowledge and complexity than questions from the first category. The third category is the most complex; it includes "concept/use" type questions that require high levels of comprehension and the ability to use the knowledge described in the first two categories.

Table 2 gives a summary of the questions submitted by the students. Column 1 identifies each subject. Columns 2, 3, and 4 identify the number of questions classified into one of the first three IS '97 Depth of Knowledge categories, respectively. Column 5 is the total number of questions submitted by each subject over the course of the semester. Column 6 is the percentage of IS '97 Level 2 and Level 3 questions submitted by each subject relative to the total number of questions submitted. Column 7 is the total number of IS '97 Level 2 and Level 3 questions submitted. Columns 8 and 9 show the midterm and final exam scores for each subject and column 10 is the final weighted average of the midterm and final examinations.

4. RESULTS

In order to test hypothesis 1, that a larger quantity of questions submitted will not result in higher scores in the course, we performed various correlation analyses comparing the total number of questions submitted (column 5) with the midterm (column 8), final (column 9), and weighted average (column 10) scores, respectively. The respective correlation coefficients for these variables were 0.009483, -0.0544846, and -0.031755424, which indicates that there was not a significant correlation between the quantity variable and any of the three score variables. Consequently H1 cannot be rejected.

In order to test hypothesis 2, that students submitting more complex questions will not obtain higher scores in the course, various correlation analyses were performed comparing the percentage of IS '97 Level 2 and Level 3 questions submitted (column 6) with the midterm (column 8), final (column 9), and weighted average (column 10) scores. The respective correlation coefficients for these variables were 0.380946, 0.44494123, and 0.491031, which again indicates no a significant correlation existed between the quality variable and any of the three score variables. As a result of this analysis, H2 is not rejected.

Finally, we examined the relationship between both quantity and complexity of individual student question submission and the course scores. We performed a correlation analyses between the total number of Level 2 and Level 3 questions submitted by each student (column 7) and the student’s midterm, final exam, and weighted average score (columns 8, 9, and 10, respectively). The respective results of those correlation analyses were 0.335366, 0.420902, and 0.451859, each of which indicate that no significant correlation exists between these variables. Consequently, we conclude that we cannot reject H3. It is apparent then that there is no relationship between student performance and either the level of complexity of questions submitted or the quantity of questions submitted as part of the APTD process.

We conducted a survey in order to assess student attitudes regarding various areas of course materials. The course material was broken into 14 broad categories. Then students were asked to anonymously report their interest level for each category. The results of the survey show that, as a whole, the students reported high interest in most categories. Students reported highest interest those areas that appear to be of most direct use to them: the use of Microsoft Access and Visual Basic, and the use of relational DBMS models in general. The areas of highest interest among those were the most elementary and most common aspects (forms, bounds, data control, etc). The least interest was reported for DBMS models other than relational, and especially for normalization. These results suggest that we can reject H4.

Speculation dictates that, as a whole, the class will perform better on Access, Visual Basic, and relational DBMS model aspects of the exam, as opposed to other areas especially involving normalization. Table 3 contains the summarized results of the survey. The results are based on an 11-point scale from 0 (no interest) to 10 (high interest).

We survey students to assess their views on the effectiveness of APTD. Table 4 contains a summary of those responses. All areas assessed were generally positive, with the overall means of the different assessments areas outline above as 1.86, 2.12, and 1.82 on a five point Likert scale where 1=strongly agree, 2=agree, 3=neutral, 4=disagree, and 5=strongly disagree. These overall positive results indicate that we can reject H5: Students will be indifferent as to their opinions on the use of APTD in the course.

5. DISCUSSION

In general, students preferred project-related course material to theory-related and they were overall positive in their assessment of the use of APTD. Student performance and its relationship to the quantity and complexity of their APTD submissions were also assessed using correlation analysis. No significant correlations were found between student test performance and the quantity or complexity of their APTD submissions. There are several possible explanations for the results of the study.

The lack of a significant correlation between the quantity of questions submitted by students in APTD and the various examination scores might be explained by the idea that creating even a large number of very low knowledge level questions does not improve the student’s ability to perform on examinations that included primarily higher-knowledge-level questions. Although the questions submitted by the students were used as the basis for the creation of the class examinations, particularly in determining the subject areas to emphasize on the examination, the instructor did reword some of them to test the students on a higher knowledge level.

An explanation for the lack of a significant correlation finding between the complexity of questions submitted by students and student examination performance is that we did not provide specific instruction on the different levels of knowledge complexity to the subjects before beginning the APTD process. The great majority of questions submitted were at a very low level of knowledge complexity. We could classify only 6 of 252 (2.3%) total questions submitted as IS ’97 Level 3 Depth of Knowledge and only 44 of 252 (17.5%) as either Level 2 or Level 3. Only 4 of the 20 students submitted a question that could be categorized as Level 3. An increase in the participation rate of students submitting higher-level depth of knowledge questions may reveal here-to-fore undiscovered relationships between the variables in question.

Yet another factor to consider is the higher frequency of requests for submission. During its initial use, Daigle and Doran made test item submission voluntary and requested submissions twice during the semester: for the midterm and final examinations. In this study, students were asked to submit test items on a weekly basis. The task of test item submission may have become a routine chore to complete rather than an opportunity to prepare multi-concept test items.

Future studies might include a review session for the Depth of Knowledge levels according to IS ’97, reducing the number of requested submissions to four or fewer per semester, and emphasize complexity rather than quantity.

6. FUTURE DIRECTIONS

There are several implications of this research. Many educators are searching for more effective methods of delivering course material than the typical lecture approach. Furthermore, many educators are also attempting to increase students' depth of knowledge related to these course materials and therefore build a foundation of knowledge that students can base further learning on. The ATPD approach as noted by Daigle and Doran [Daigle, 1998] "provides a framework that can be reused for identifying a standard of assessment and for self-assessment against the standard, critical skills for self-management for life-long learning."

7. REFERENCES

The Association for Computing Machinery, ACM/IEEE-CS Joint Task Force, Computing Curriculum, 1991.

Bloom, Benjamin S. (Ed.) et al, The Taxonomy of Educational Objectives: Classification of Educational Goals. Handbook 1: The Cognitive Domain. New York: McKay Press, 1956.

Daigle, Roy J., Michael V. Doran. “Cultivating Life-Long Learning Through Student Participation in Exam Development”, Association for Information Systems Conference, 1998.

Davis, Gordon B., Gorgone, John T., Couger, J. Daniel, Feinstein, David, L., Longenecker, Herbert E, IS ’97: Model Curriculum and Guidelines for Undergraduate Degree Programs in Information Systems. Association of Information Technology Professionals, 1997.

Murry, John P., 1990. “Better Testing for Better Learning College Teaching”. Fa11 1990, Vol. 38 Issue 4, p148, 5p.

IS ’90, ’94, ’95 Depth of Knowledge	Bloom Levels of Knowledge	Template for Writing Behavioral Objectives Students completing …will be able to	Meaning of Depth of Knowledge Level and Activities Associated with Attaining that Level
1 Awareness	1 Knowledge Recognition	Define… List characteristics of… Name components of… Diagram… List advantages/disadvantages of…	Introductory Recall and Recognition Class presentations, discussion groups, reading, watching videos, structured laboratories. Involves only recognition, but with little ability to differentiate. Does not involve use.
2 Literacy	1 Differentiation	Compare and contrast… Explain… Write/execute simple… Define functional capabilities that are… Describe interrelations of…to related objects	Knowledge of Framework and Contents, Differential Knowledge Continued lecture and participative discussion, reading, team work and projects, structured labs. Requires recognition knowledge as a prerequisite. Requires practice. Does not involve use.
3 Concept/Use	2 Comprehension Transition/Extrapolation Use of Knowledge	Use… Communicate the idea of… Form and relate the abstraction of… as… Given a set of…, interpolate/extrapolate to… List concepts/major steps in…	Comprehension and Ability to Use Knowledge *when Asked* Requires continued lab and project participation, presentation involving giving explanations and demonstrations, accepting criticism; may require developing skills in directed labs.
4 Detailed Understanding Application	3 Application Knowledge	Search for correct solution to… and apply it to… Design and implement a… for… Write syntactically correct… and/or debug… Apply the principles of… to… Implement a… and maintain it	Selection of the Right Thing and Using it *without hints* Semi-structured team-oriented labs where students generate their own solutions, make their own decisions, commit to and complete assignments, and present and explain solutions.
5 Skilled Use	4 Analysis 5 Synthesis 6 Evaluation	Develop/originate/institute… Construct/adapt… Generate novel solutions to… Come up with new knowledge regarding… Evaluate/judge the relative value of… with respect to…	Identification, Use and Evaluation of New Knowledge An advanced level of knowledge for those very capable of applying existing knowledge in which *denovo* solutions are found and utilized in solving and evaluating the proposed new knowledge.

Table 1 - Knowledge Levels, Templates for Objective Writing, and Meaning of the Depth Levels with Associated Learning Activities [Davis, et al. 1997]

1	2	3	4	5	6	7	8	9	10
Subject ID	**IS '97 Level 1 Questions Submitted**	**IS '97 Level 2 Questions Submitted**	**IS '97 Level 3 Questions Submitted**	**Total Questions Submitted**	**% of Level 2 & Level 3 Questions Submitted**	**Level 2 & Level 3 Total Questions Submitted**	**Midterm exam score**	**Final exam score**	**Final weighted average**
1	10	1	1	12	17%	2	72.3	71.7	72.0
2	9	4	0	13	31%	4	91.0	78.3	83.8
3	6	7		13	54%	7	74.7	95.0	86.3
4	8	2	2	12	33%	4	80.3	91.7	86.8
5	9	3	1	13	31%	4	98.0	100.0	99.1
6	13		2	15	13%	2	79.3	81.7	80.7
7	12	4		16	25%	4	77.7	76.7	77.1
8	13			13	0%	0	78.0	85.0	82.0
9	6			6	0%	0	55.0	85.0	72.1
10	14	1		15	7%	1	60.7	76.7	69.8
11	13	2		15	13%	2	81.3	100.0	92.0
12	9			9	0%	0	81.0	66.7	72.8
13	17			17	0%	0	71.3	63.3	66.8
14	8	1		9	11%	1	87.0	76.7	81.1
15	13	1		14	7%	1	74.0	91.7	84.1
16	12	3		15	20%	3	92.0	93.3	92.8
17	8	1		9	11%	1	91.0	95.0	93.3
18	9	5		14	36%	5	89.0	95.0	92.4
19	8	2		10	20%	2	97.0	95.0	95.9
20	11	1		12	8%	1	83.3	93.3	89.1
Total	**208**	**38**	**6**	**252**		**44**			
Average							**80.70**	**85.58**	**83.49**

Table 2 - Summary and Classification of Questions Submitted by Student and Student Performance Measures

Area of Interest	Mean	Standard Deviation
Data Modeling (reality, ANSI-SPARC)	7.39	1.75
Hierarchical DBMS model	7.68	1.86
Network DBMS model	7.58	1.80
Relational DBMS model	9.05	1.39
Data structures of DBMS Models	8.11	1.70
Database terminology (key, foreign key, referential integrity, …)	8.68	1.53
Database Environment and DBMS functionality.	8.42	1.61
E-R modeling	7.89	1.91
ACCESS projects	9.00	1.41
Relational operators and SQL commands	8.63	1.80
Elementary Visual Basic Projects (forms, bounds, data control)	9.11	1.33
Intermediate Visual Basic Projects (database access, reporting)	9.00	1.49
Advanced Visual Basic Projects (database maintenance, control array, multiple data controls, grid controls)	8.79	2.10
Normal Forms (1NF, 2NF, 3NF, BCNF)	6.89	2.00
Mean of Means / Std. Dev.	**8.30**	**1.69**

Table 3 - Summary of Question Types Submitted Survey

Strongly Agree	Agree	Neutral	Disagree	Strongly Disagree		
1	2	3	4	5	**Mean**	**Standard Deviation**
Regarding the question and answer submissions…					**1.86**	**0.80**
Guidelines regarding the format of the questions were clear.					2.21	0.92
Two questions per week are adequate for topic coverage.					1.74	0.81
My submissions were representative of the topic coverage.					1.74	0.65
I participated fully in the submissions.					1.74	0.81

I feel that preparation of the submissions…	**2.12**	**0.89**
Was beneficial to me for exam preparation.	2.05	1.08
Encouraged a greater depth of review than I would have typically performed.	2.37	0.90
Increased my knowledge and comprehension of the material.	1.84	0.76
Improved my grade in the course.	2.21	0.85
Was a worthwhile activity.	2.11	0.88

I feel that the class submissions, as a whole…	**1.82**	**0.68**
Should well represent the material covered.	1.74	0.56
Reflect the emphasis placed on the material.	1.74	0.73
Would create a challenging, yet realistic, exam.	2.05	0.78
Reflect the comprehensibility of the lectures.	1.74	0.65

Mean of Means / Std. Dev.	**1.94**	**0.80**
Std. Dev. of Means / Std. Dev.	**0.23**	**0.13**

Table 4 - Summary of APTD Effectiveness Survey