Engaging Students to Work with Self-Assessment Questions: A Study of Two Approaches

Peter Brusilovsky, Sergey Sosnovsky
School of Information Sciences, University of Pittsburgh,
135, North Bellefield ave.
Pittsburgh, PA, 15260, USA
{peterb, sas15} @ pitt.edu

ABSTRACT

We explored two approaches for encouraging introductory programming students to use the web-based, self-assessment system, QuizPACK. An "organizational" approach applied specially constructed classroom quizzes, while the "technical" approach introduced adaptive guidance. Our studies demonstrated that each of these caused a dramatic increase in system use. This approach could be useful in many other contexts, when an educationally beneficial system is underused by students.

Categories and Subject Descriptors

K.3.1 Computer Uses in Education: Distance learning. K.3.2 Computer and Information Science Education: Self-assessment.

General Terms

Performance, Human Factors, Languages.

Keywords

Self-assessment, Personalization, Engagement, Introductory Programming.

1. INTRODUCTION

Researchers in the field of Computer and Information Science (CIS) Education have developed a wealth of computer tools to support the various aspects of teaching and learning CIS courses. Many developed tools have been evaluated in the classroom and proven to be useful. A good number of these tools have been made available to the community through dedicated web sites and educational digital libraries [12]. However, we have now learned that the mere availability of a good tool, although proven beneficial for students, is not enough to ensure its broad educational impact. Special provisions have to be made to encourage instructors to adopt the tool and to engage learners to use it. Research into this area has already been started. The report of the Working Group on Improving the Educational Impact of Visualization [9] set a precedent by discussing some obstacles that prevent instructors from adopting and students from using a technology that "can greatly benefit learners and instructors alike." The following report [8] suggests several ways to ease the instructor-adoption barrier. Our work attempts to solve this problem from two opposite directions, by exploring two ways to engage students in a broader use of an educationally effective web-based tool (previously adopted by the instructor and available through the course web site).

Our concern is the broader use of a specific category of educational tools, which we call student-driven tools. Student-driven tools are created to assist student learning, yet their use is not required and does not count towards the student's course grade. A large fraction of visualization and simulation tools are student-driven. Unlike a variety of assessment-driven tools such as CourseMarker [5] that the students must use in order to complete their assignments, it is up to the students to decide to what degree and how frequently they choose to use the student-driven tools. An instructor might work hard to provide a good set of educational tools of known benefit to the students, only to discover that these tools are dramatically underused.

This is exactly the situation we found ourselves in when we began the project reported on in this paper. Over the last four years, our team has been working on individualized self-assessment questions. We developed QuizPACK [10], a system that can generate and evaluate parameterized questions in the domain of C-programming. We also used this system in a self-assessment mode in several programming-related classes. While the system was demonstrated and recommended by the instructor, no special provisions were made to encourage the students to work with it. The use of QuizPACK benefited the students. Our studies demonstrated significant correlation between the use of the system by individual students and their course performance as measured by the exams and the final grade [11]. Yet, the protocol analysis demonstrated that the students used QuizPACK much less than it deserved.

A known way to encourage the use of educationally beneficial tools is to convert them from student-driven to assessment-driven tools. TRAKLA [7] and WebtoTest [1] provide good examples of the assessment-driven use of traditionally student-driven technologies. However, this conversion is not always possible. For example, our self-assessment questions cannot be used as a reliable homework assessment tool. The student can simply compile the code of the question and get the answer without thinking. Using them for classroom assessment in controlled

ITiCSE'05, June 27–29, 2005, Monte de Caparica, Portugal.

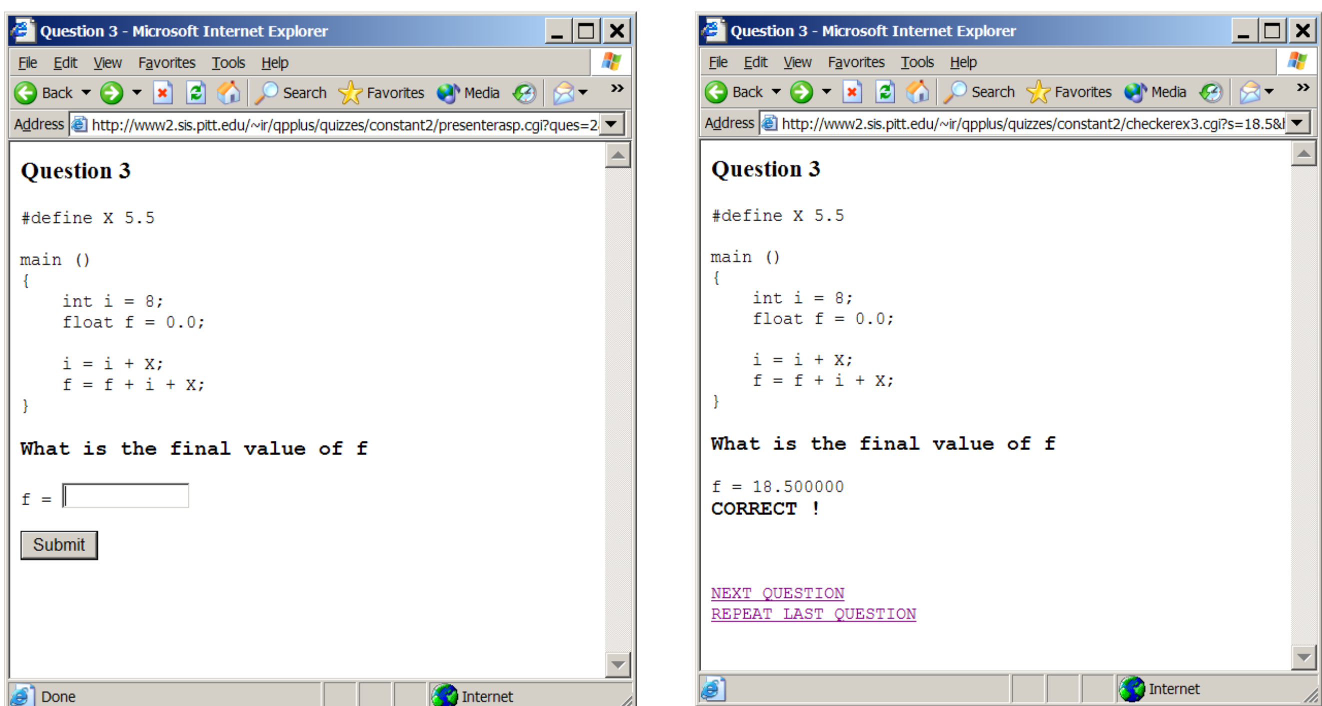
settings is possible, and we have tried that. However we have found out that students strongly prefer traditional paper-based quizzes. As a result self-assessment became the primary use of QuizPACK. Over the last three semesters, we have been exploring two alternative approaches to engaging students in self-assessment questions. Both of these approaches brought sizeable results. This paper reports our work. After a brief presentation of the QuizPACK system, we will introduce the "engagement approaches" we explored and report on their impact on both student engagement with QuizPACK and student course performance.

2. INDIVIDUALIZED QUESTIONS IN QUIZPACK

QuizPACK focuses on a special category of questions that are used in programming courses – code-execution questions. In code-execution questions, the student is given a fragment of a program and is asked to predict the value of a particular variable or a string to be printed when this fragment has completed its execution. This kind of question is very important, because they enable the teacher to check the student's understanding of the semantics of programming-language constructs. Figure 1 demonstrates a sample code-execution question presented by QuizPACK. Unlike code-execution questions in traditional "static" quizzes, QuizPACK questions are parameterized. This means that one or more constants viewed in the body of the question are actually instantiated parameters. They produce different random numbers for each student who takes the quiz, as well as for the same student attempting the question several times. The student has to fill in the answer and hit the "submit" button. In response, the system generates an evaluation screen for the student (Figure 1, right). This screen lets the student rethink the question and his/her answer. In particular, the student may want to attempt the same question again by using the "repeat last question" link.

The parameterized nature of the QuizPACK questions is the source of its power. As an assessment tool, a reasonably small number of question patterns can be used to produce individualized assessments for even large classes. In the self-assessment context, the same question can be used again and again with different parameters, allowing every student to achieve understanding and mastery without boredom. The benefits of using parameterized questions for assessment have been confirmed by a number of studies. For example, long-term studies of the well-known CAPA system [6] demonstrated that individualized exercises can significantly reduce cheating, while at the same time improving student understanding and exam performance. Our work has focused on exploring the value of parameterized questions in a self-assessment context.

QuizPACK supports the authoring and delivery of parameterized questions. An author provides the core content of a question, which is a parameterized fragment of code to be executed and an expression (usually a variable) that has to be evaluated by the student at the end of the fragment execution. The system does the rest: randomly generating the question parameter, creating a web-quiz presentation of the parameterized question, receiving the student's input, comparing the student’s answer with the actual result of running the parameterized code "behind the stage," reporting scores to the student, and recording the results. Information about QuizPACK implementation and authoring support can be found in [10]. The system is freely available for external users. All readers are welcome to try it at http://www2.sis.pitt.edu/~taler/QuizPACK.html.



a) b)

Figure 1. Student interface of QuizPACK

3. ENCOURAGING STUDENTS TO USE QUIZPACK

The first comprehensive evaluation of QuizPACK was performed in 2002 in the context of an introductory programming class. For every lecture of the course the instructor provided one or two self-assessment quizzes of 5 questions each. The quizzes were recommended in class, but their use was completely voluntary. As we pointed out in the introduction, our first evaluation of QuizPACK [11] revealed a significant (p=0.0016) and close to significant (p=0.0589 and 0.0618) correlation between the amount and quality of work with QuizPACK and the different course performance parameters. At the same time, we discovered that the system was quite underused. During two semesters of 2002, QuizPACK was made available to 81 (39 + 42) undergraduate students, but only 49 of them tried QuizPACK (i.e., attempted at least one question). While a few students used the system heavily (the largest number of questions attempted by a single student over a semester was 319), most students attempted just a few questions. Only 22 students (less than a third!) used QuizPACK relatively actively (attempting at least 30 questions). Given the clear positive value of working with QuizPACK, we decided to provide some additional motivation for students to use it.

3.1 The "Organizational" Approach: Classroom Quizzes

The first approach that we explored was purely organizational. Knowing the role of classroom quizzes in motivating student learning, we decided to encourage the use of QuizPACK by changing the format of classroom quizzes. Classroom paper-and-pencil quizzes were regularly used in our introductory programming classes. Typically, each semester an instructor administers 8 to 10 classroom quizzes. During a quiz the students were allowed 10 minutes to answer 5 questions related to topics from the two most recent lectures. Until 2003, we used rather standard multiple-choice quizzes, unrelated to QuizPACK. In 2003 we introduced fill-in-the-blank paper quizzes produced with QuizPACK. For each of these quizzes, we selected five randomly instantiated QuizPACK questions out of the 15-20 questions offered for self-assessment on material from two of the most recent lectures. The students were informed about this. Thus, the students who worked with QuizPACK after each lecture had a chance to practice each of the forthcoming classroom questions, though with different parameters. We used QuizPACK in this context during Spring and Fall 2003. During this time, 73 students had access to QuizPACK, 60 (82%) of them tried it at least once, and 51 (70%) worked with the system regularly.

As we can see from Figure 2, the use of paper-based QuizPACK quizzes for regular classroom assessments dramatically affected the use of QuizPACK by the students. We measured the quantity of student work being done with QuizPACK by following several factors. The average student *activity* (average number of questions attempted by students who used the system at least once) measures the plain volume of student work. The average *number of sessions* (number of independent sessions of work with the system) and the average *course coverage* (the ratio of questions attempted at least once to the total number of questions) measure the distribution of this work over the duration of the course and course topics. The *percentage of active students* (students who attempted more than 30 questions) measures the distribution of this work over all the students of the class. Most of these performance variables *tripled or nearly tripled* in 2003 in comparison with 2002. (It is actually surprising that the growth rates for these different measures are so close). Other specific changes occurred too: The maximum number of questions attempted by a single student rose to 510. At the same time the average success, measuring the quality of student work, slightly decreased in 2003. The students started to use more complicated quizzes from the advanced sections of the course on a regular basis. These same quizzes had been ignored by most of the students in 2002.

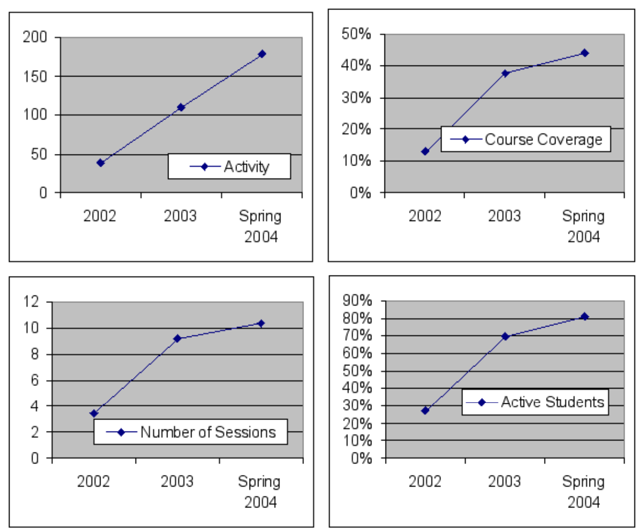


Figure 2. System use in three classroom studies

Table 1. A comparison of three studies of self-assessment quizzes

Year	Total students	Students ever used the tool	Active students	Average activity	Max activity	Average success	Average course coverage	Average num. of sessions	Attempts per question
2002	81	49 (60%)	22 (27%)	37.93	319	49.34%	12.91%	3.44	1.39
2003	73	60 (82%)	51 (70%)	110.18	510	43.55%	37.56%	9.18	1.69
2004	27	24(89%)	22(81%)	178.44	753	39.18%	44.02%	10.33	2.59

To confirm the correlation between student usage of QuizPACK and class performance, we introduced a new performance variable called *knowledge gain,* in 2003. To isolate the past experience factor, we administered a pre-test (before the first lecture on C) and a post-test (during the Final Exam). The paper-and-pencil pre-test and post-test featured the same ten QuizPACK fill-in-the-blank questions, but with different parameters. The *knowledge gain,* the measure for growth of student knowledge of semantics, was calculated as the difference between post-test and pre-test scores. The regression analysis confirmed significant dependency between usage of QuizPACK and student quiz performance (p=0.0020). Moreover, the nearly significant dependency between success and use of QuizPACK and student course performance, as measured by the Final Exam and course grades, became significant (p=0.0006 and 0.0007). In addition, we found a significant dependency (p=0.0163) between usage of QuizPACK and student knowledge gain.

3.2 The "Technical" Approach: Personalized Guidance

Encouraged by these results, we decided to explore a second approach to draw students toward use of the system. This time we focused on a technology-based approach. Following our past work on adaptive hypermedia [4], we attempted to implement adaptive guidance for students who were trying to decide which quiz to take. Adaptive guidance in adaptive educational hypermedia systems is most often implemented with so-called adaptive annotation technology [2]. With adaptive annotation, each link to an educational object is annotated with adaptive visual cues that inform the user about important properties of this object. In the past we explored two kinds of visual cues – one kind that shows whether specific content is ready to be learned and another kind that expresses the student's relative mastery level of a specific concept or topic. In several studies we demonstrated the benefits of this approach [4].

To evaluate whether it was beneficial to add adaptive annotation to our self-assessment section, we developed the QuizGuide system. QuizGuide was implemented as a value-added service for our eLearning architecture KnowledgeTree [3]. It is an interface between the user and the original QuizPACK system, helping the student select the most relevant quizzes. Figure 3 shows the interface of QuizGuide. The *quiz presentation area* (right on Figure 3) shows the familiar QuizPACK interface. The *quiz navigation area* (left on Figure 3) uses adaptive annotation to help students select a topic for self-assessment. To reflect both the goal and knowledge relevance of each topic in one icon, QuizGuide uses the "target-arrow" abstraction. Each topic is annotated with a target icon. The number of arrows in the target reflects the level of student knowledge for that topic: the more arrows shown, the higher the student's knowledge level. The intensity of the target's color shows the relevance of the topic to the current learning goal: the more intense the color, the more relevant the topic. Topics that are not ready to be studied are annotated with a barred target. Since student goals and knowledge are constantly changing, different icons will be shown almost every time the student accesses QuizGuide. To see changes in the user model that occurred during the current session, the student clicks on the refresh icon. Thus at any time during his/her independent work with self-assessment quizzes, a student can clearly see which topics are most important to practice or which topics require the student to do more work.

We explored QuizGuide in the Spring 2004 semester, in the same introductory programming course. QuizGuide was introduced to the students after the Midterm exam. For the first part of the course, the students were able to access standard QuizPACK quizzes. But during the second half, access to both adaptive and non-adaptive quizzes were made available to the students. Note that both QuizGuide and KnowledgeTree provide access to *the same set of quizzes*. The QuizPACK performance recording mechanism does not depend on the way of access, so data obtained from both QuizPACK and QuizGuide are considered equally when building the user model.

During the semester time, 27 students had access to quizzes, 24 (89%) of them tried at least one question, and 22 (81%) students worked with the system regularly. Of the students who used quizzes regularly, 10 students (46%) stayed with QuizPACK (three of them did not try QuizGuide at all, the other seven tried it, but most continued to use QuizPACK). Six students (27%) completely moved to QuizGuide after its introduction, though they had actively used QuizPACK before. The remaining six students (27%) started to use quizzes actively only after QuizGuide was presented and at that point began to use QuizGuide almost exclusively.

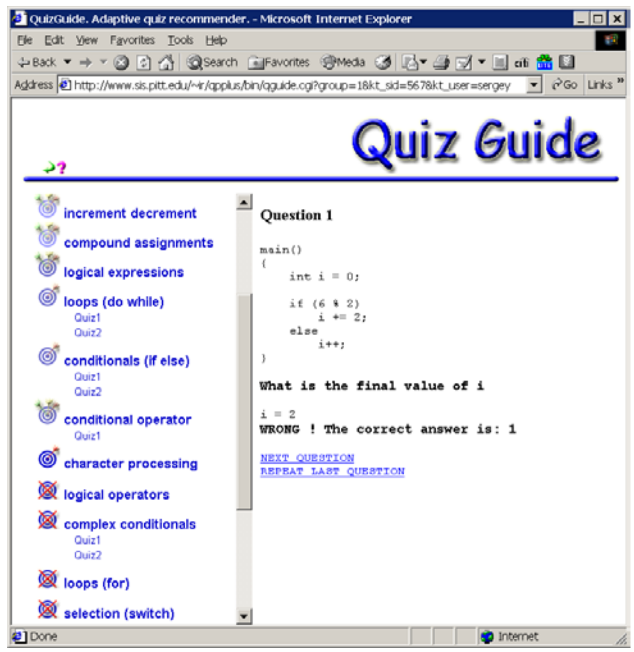


Figure 3. Student interface of QuizGuide

As we can see from Figure 2 and Table 1, the introduction of QuizGuide resulted in further increase of all system use variables. The average number of questions attempted by students rose most dramatically reaching nearly 180. Since we knew previously that the use of self-assessment quizzes affects learning, it was not surprising to find that the average knowledge gain in Spring 2004 rose from 5.1 to 6.5. To better understand the value of adaptive annotation we attempted to compare QuizGuide and QuizPACK sessions within the same class. The comparison confirmed that adaptive annotation provided an additional motivation for the students to use quizzes: for students actively working with

QuizGuide the average session length was 24 question attempts, while for those working with QuizPACK, it had been only 14 attempts. In addition, we discovered that the percentage of correctly answered questions in the QuizGuide sessions is also larger: 44.3% versus only 35.6% in the QuizPACK sessions.

4. CONCLUSIONS

In our classroom study of the QuizPACK system, we discovered that work with self-assessment quizzes strongly benefits students. Despite that fact, the majority of students resisted use of the system. We explored two different ways to engage the students to begin to work with self-assessment quizzes. We attempted to achieve this through a change in course organization and through adaptive annotation technology. Our studies demonstrated that both approaches were able to achieve a dramatic increase of most system use variables. We think that our approach could be useful in other contexts where students underuse an educationally beneficial system. While the organizational approach may not be directly applicable to essentially different kinds of educational software, the adaptive annotation approach looks very promising and highly reusable.

5. ACKNOWLEDGEMENT

This material is based upon work supported by the National Science Foundation under Grant No. 0310576.

6. REFERENCES

[1] Arnow, D. and Barshay, O. WebToTest: On-line programming examination using WebToTeach. SIGCSE Bulletin - Inroads. 31, 3 (1999), 21-24.

[2] Brusilovsky, P. Adaptive hypermedia. User Modeling and User Adapted Interaction, 11, 1/2 (2001), 87-110, available online at http://www.wkap.nl/oasis.htm/270983.

[3] Brusilovsky, P. KnowledgeTree: A distributed architecture for adaptive e-learning. In: Proc. of The Thirteenth International World Wide Web Conference, WWW 2004 (Alternate track papers and posters), (New York, NY, 17-22 May, 2004), ACM Press, 104-113.

[4] Brusilovsky, P. and Eklund, J. A study of user-model based link annotation in educational hypermedia. Journal of Universal Computer Science, 4, 4 (1998), 429-448, available online at http://www.iicm.edu/jucs_4_4/a_study_of_user.

[5] Higgins, C., Hegazy, T., Symeonidis, P., and Tsintsifas, A. The CourseMarker CBA System: Improvements over Ceilidh. Education and Information Technologies, 8, 3 (2003), 287-304.

[6] Kashy, E., Tsai, Y., Davis, N.E., and Wolfe, S.L. Using Networked Tools to Promote Student Success in Large Classes. Journal of Engineering Education, 87 (1998), 385 - 390, available online at http://www.nscl.msu.edu/~thoennes/personal/papers/jee87_385.pdf.

[7] Korhonen, A. and Malmi, L. Algorithm simulation with automatic assessment. In: Proc. of 5th annual SIGCSE/SIGCUE ITiCSE conference on Innovation and technology in computer science education, (Helsinki, Finland, ACM Press, 160-163.

[8] Naps, T., Rößling, G., Anderson, J., Cooper, S., Dann, W., Fleischer, R., Koldehofe, B., Korhonen, A., Kuittinen, M., Leska, C., McNally, M., Malmi, L., Rantakokko, J., and Ross, R.J. Evaluating the educational impact of visualization. ACM SIGCSE bulletin, 35, 4 (2003), 124–136, available online at http://doi.acm.org/10.1145/960875.960540.

[9] Naps, T.L., Rößling, G., Almstrum, V., Dann, W., Fleischer, R., Hundhausen, C., Korhonen, A., Malmi, L., McNally, M., Rodger, S., and Velázquez-Iturbide, J.Á. Exploring the role of visualization and engagement in computer science education. ACM SIGCSE bulletin, 35 (2003), 131–152, available online at http://doi.acm.org/10.1145/960568.782998.

[10] Pathak, S. and Brusilovsky, P. Assessing Student Programming Knowledge with Web-based Dynamic Parameterized Quizzes. In: Barker, P. and Rebelsky, S. (eds.) Proc. of ED-MEDIA'2002 - World Conference on Educational Multimedia, Hypermedia and Telecommunications, (Denver, CO, June 24-29, 2002), AACE, 1548-1553.

[11] Sosnovsky, S., Shcherbinina, O., and Brusilovsky, P. Web-based parameterized questions as a tool for learning. In: Rossett, A. (ed.) Proc. of World Conference on E-Learning, E-Learn 2003, (Phoenix, AZ, USA, November 7-11, 2003), AACE, 309-316.

[12] Zia, L.L. The NSF National Science, Technology, Engineering, and Mathematics Education Digital Library (NSDL) Program: New Projects in Fiscal Year 2003. D-Lib Magazine, 10, 3 (2004).