Web-based Parameterized Questions as a Tool for Learning

Sergey Sosnovsky, Olena Shcherbinina, Peter Brusilovsky
School of Information Sciences
University of Pittsburgh
Pittsburgh PA 15260
sas15@pitt.edu, ols1@pitt.edu, peterb@mail.sis.pitt.edu

Abstract: Web-based parameterized questions is one of the most promising kinds of Web knowledge assessment. In an assessment context a reasonably small number of parameterized questions can be used to produce assessments for large classes. In self-assessment context, the same question can be used again and again with different parameters allowing every student to achieve mastery. This paper reports the results of our objective and subjective evaluation of parameterized questions as a learning tool in programming courses. The results show that parameterized questions as served by our QuizPACK system provide an exceptional learning tool. There is also evidence that this tool was especially helpful for female students who usually lag behind male students in programming courses.

Introduction

Web-based quizzes have emerged into the primary tool for evaluation and self-evaluation of student knowledge in the context of Web-based education [Brusilovsky & Miller 2001]. All major Web courseware management systems support authoring and delivery of on-line quizzes made from static questions. Researchers working in the field of E-learning attempt to extend the power of this tool in several directions. One of the most promising directions is known as *individualized* or *parameterized* questions.

A parameterized question is essentially a pattern of a question created by an author. At the presentation time, the pattern is instantiated with randomly generated parameters from a particular set. Every question pattern is able to produce a large or even unlimited number of different questions. This, in an assessment context a reasonably small number of question patterns can be used to produce individualized assessments even for large classes. Moreover, same patterns can be used in different versions of the same course, different semesters, and even different courses. In self-assessment context, the same question can be used again and again with different parameters allowing every student to achieve mastery. Cheating-proof question patterns become re-usable non-devaluating assets and can be collected over courses and over years in re-usable assessment libraries.

Individualized questions have been explored in the field of educational technology and came back into focus recently as one of the promising research directions in the field of Web-enhanced education. A number of pioneer systems such as CAPA [Kashy et al. 1997], WebAssign [Titus, Martin & Beichner 1998], EEAP282 [Merat & Chung 1997], or Mallard [Graham, Swafford & Brown 1997] have explored the use of individualized questions in different contexts. A very encouraging model has been provided by CAPA system [Kashy et al. 2001; Kashy et al. 1997]. The CAPA researchers have not limited their effort to developing a new tool for authoring and administering individualized exercises, they also have performed a number of careful studies of this new technology. The reported results [Kashy et al. 2001; Kashy et al. 1997] provide very clear evidence that individualized exercises can significantly reduce cheating, improve student understanding and exam performance.

Individualized questions have become a working assessment technology in such domains as physics, chemistry, or mathematics where an answer can be calculated using a parameterized formula. The general focus of our work in this direction is to apply the ideas of parameterized questions in a non-traditional and challenging application area of programming languages. Assessing programming knowledge is quite different from knowledge assessment in "formula-based" subjects and individualized questions have not yet been previously explored in this field. Recently we have created a proof-of-concept system QuizPACK that is capable to deliver Web-based dynamic individualized exercises for programming-related classes that use C language. An earlier version of QuizPACK has been presented at ED-MEDIA'02 conference [Pathak & Brusilovsky 2002].

QuizPACK presents a parameterized question to the student just like a typical static fill-in question (Figure 1, left). One or more constants that the student can see in the body of the question are, actually, instantiated parameters. They are different for different students taking the quiz as well as for the same student attempting the

question several times. The students has to fill-in the answer and to hit "submit" button. In response, the system generates an evaluation screen for the student (Figure 1, right). This screen lets the student to re-think the question and the answer. In particular, the student may want to attempt the same question again by using the Back button and reloading the question screen. The student can attempt the same question many times. We welcome the readers to try a sample quiz from QuizPACK system at http://www2.sis.pitt.edu/~taler/QuizPACK.html

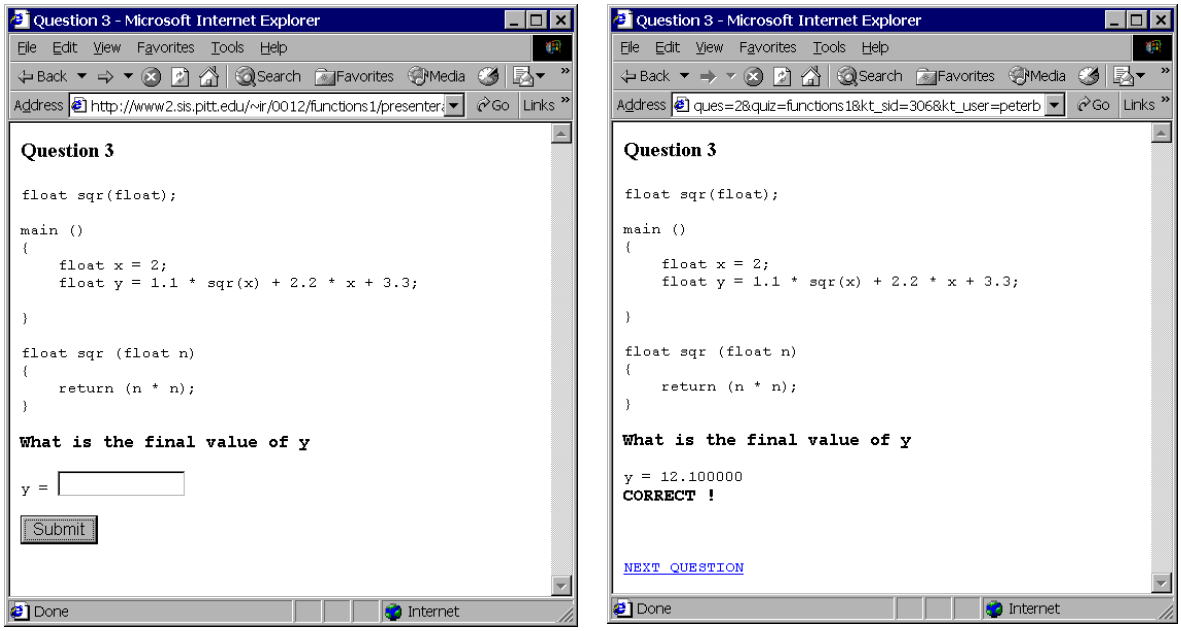

Figure 1: The user interface of QuizPACK

We consider QuizPACK as both – a practical educational tool and a research tool. We have already used QuizPACK in the context of several programming classes in 2001-2003. Our current research goal is to explore systematically the use of parameterized questions in programming-related subjects. We want to find the ways of using this technology in programming-related courses that maximize its strong sides and minimize known problems. In the past we already run several small scale formative studies of QuizPACK system during each of the courses where we have used the system [Pathak & Brusilovsky 2002]. These studies enabled us to develop a better version of the system and prepare a large number of parameterized questions. This paper presents our most recent comprehensive study that we have run during the Spring Semester of 2003 at the University of Pittsburgh. Next section of the paper presents the focus of our study and the following sections present and discuss the results. The technical and implementation side of QuizPACK project are not presented here due to the lack of space; they can be found in [Pathak & Brusilovsky 2002].

Evaluating Parameterized Questions as a Learning Tool

Currently parameterized questions are used predominantly in an assessment context. It is quite natural since in this context parameterized questions provide a great benefit by beating cheating and allowing teachers to accumulate banks of re-usable questions. However, our early experiments with QuizPACK demonstrated that the students themselves are more interested in using parameterized questions in a self-assessment context. Indeed, if positioned properly, parameterized questions may become a very useful learning tool. While classic self-assessment quizzes simply allow the students to check their overall level of knowledge of a particular topic to decide whether further reading or studying is required, parameterized quizzes allow the students to achieve mastery on a topic. An incorrect answer to a parameterized question opens a focused learning opportunity. A student can take a focused study of the topics involved and check the new level of understanding by attempting to answer the same question again – now with a different parameter. This loop can be repeated several times until understanding is achieved. Our recent study attempted to answer the question whether this style of using parameterized questions as a learning tool is attractive to the students and whether it benefits the student knowledge.

The study has been performed in the context of a Spring 2003 undergraduate course "Introduction to Programming" taught to University of Pittsburgh students who are interested to obtain a BS degree in Information Science. QuizPACK was one of the learning tools available to all students taking this course for the whole duration of the course. For each of 16 lectures in the course devoted to C programming QuizPACK offered at least two 5-question quizzes. While our past experience with QuizPACK in the same context demonstrated that many students use this tool on a regular basis, we have attempted to provide an additional motivation for using the system. To do that, we have changed the format of weekly classroom quizzes. Instead of relying on traditional multiple-choice questions we have used fill-in-the-blank questions taken directly from QuizPACK database. Once a week the students took a 10-minutes quiz composed of 5 QuizPACK questions (that the students may have already seen - with different parameters - when working with QuizPACK). It does provided extra motivation while avoiding the cheating issue. Comparing this arrangement with the situation in theprevious Fall 2002 semester (when we were using regular multiple-choice quizzes), we have found that about 30% more students used QuizPACK self-assessment quizzes regularly. The details and results of our study performed in this context are reported in the next section.

QuizPACK evaluation

The goal of the QuizPACK evaluation was to measure the objective and the subjective value of QuizPACK as a learning tool. To determine the objective value we tried to find relationships between students' work with QuizPACK quizzes and their course performance. The student course performance was measured by two dependent variables that were both components of their final grade: their total score on weekly classroom quizzes and the final exam score. Note that these parameters are quite different. While the quizzes were measuring students' knowledge of C language and its semantics, the exam assessed mainly their programming skills - the ability to understand, modify, and write programs.

As with any introductory course, the student starting knowledge level of C programming and programming in general varied significantly - from complete novices to students who were able to write solid programs in the past (the latter argued that they need this course as a refresher). It was natural to expect that the studentquiz and exam scores depend not only on their work over the duration of the course (including QuizPACK work), but also on the starting level of knowledge. To isolate the past experience factor we have administered a pre-test (before the first lecture on C) and post-test (during the Final exam). The pre-test and post-test featured the same ten QuizPACK fill-in-the-blank questions (with different parameters). We have calculated the *knowledge gain*, the third dependent variable as the difference between post-test and pre-test scores.

To evaluate the subjective effect of the QuizPACK, we have asked all students who used QuizPACK *enough to qualify* to fill in a questionnaire. A student was considered qualified if he or she worked with at least 10 or more 5-question quizzes associated with 6 or more different lectures and used this feature during several sessions for at least 20 days before taking the questionnaire. To motivate the students 3 extra credit points were offered for taking the questionnaire.

Objective Performance Evaluation

We have performed a Linear Regression Analysis (using SPSS software) to find out how the student work with QuizPACK measured by *explanatory variables* influenced *explained (dependent) variables*. This type of analysis is typical to study problems with cause-effect relationships between variables. As mentioned above, our explained variables were the student *grade on final exam* measured in percents, *their quiz performance* measured in points (maximum number of points gained on 9 quizzes was 45) and their *knowledge gain* also measured in points. Explanatory variables wereQuizPACK *activity* measured as the total number of questions attempted by the student in QuizPACK and *success* - the percentage of correctly answered questions calculated as the total number of questions attempted divided by total score (number of correctly answered question). Note that each parameterized question may be attempted several times. Each of these attempts (correct or incorrect) is counted in the activity and success parameters. Most typically, students were working with the same question until the very first successful attempt, however, a number of students kept working with the questions even after the first success that usually resulted in several registered successful attempts.

The results of the regression analysis for the quiz performance are shown in the Table 1. As we can see, the explanatory variable *success* does have strong influence (with a high significance = 0.029) on the student *quiz performance*. Note that R Square is equal to 0.335 - it means that QuizPACK usage could explain 34% of the in-class quizzes grade. Significance of full equation is high, (in ANOVA table Sig.=0.038 which is less than 0.05).

Model Summary

R	R Square	Adjusted R Square	Std. Error of the Estimate
.578(a)	.335	.251	18.15815

a Predictors: (Constant), ACTIVITY, SUCCESS

ANOVA(b)

	Sum of Squares	df	Mean Square	F	Sig.
Regression	2653.030	2	1326.5	4.023	.038(a)
Residual	5275.496	16	329.7		
Total	7928.526	18			

a Predictors: (Constant), ACTIVITY, SUCCESS

Coefficients (a)

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	25.635	13.098		1.957	.068
	SUCCESS	60.573	25.329	.517	2.391	.029
	ACTIVITY	.027	.041	.140	.648	.526

Table 1: The influence of QuizPACK work on the quiz grade

Similar results were obtained by performing the regression analysis for the final exam grade. However, the influence of the *success* parameter was less evident. R Square for the final exam was 0.19, it means that QuizPACK usage could explain only 19% of the final exam grade. Significance of full equation also low, in ANOVA table Sig.=.192 which is more than .05. However, we can conclude that the variable success has some influence on the final exam grade, because level of significance for this variable was 0.088 that is close to 0.05 and the variable points to right direction in has + sign in equation.

In both cases, there was no visible influence of the *activity* variable. It means that the student successful work with QuizPACK had a strong influence on their knowledge of semantics as measured by the quizzes and also have some influence on their programming skills as measured by the final exam. At the same time, it was shown that simply playing with QuizPACK is not sufficient: to gain knowledge, the students should strive to answer QuizPACK questions correctly.

The above analysis alone does not provide sufficient evidence of the role of QuizPACK since it does not isolate the reason why the students were achieving a good success grade with QuizPACK. Naturally, some of them just had a reasonable level of starting knowledge. These students were using QuizPACK more as a self-checking tool most often giving the correct questions with the first attempt. In contrast, novice students had to work with QuizPACK more persistently attempting each question several times until one or more correct answers were achieved. This is where we need to see the influence of QuizPACK on the *knowledge gain.*

The results of the regression analysis for the *knowledge gain* are shown in the Table 2. As we can see, the explanatory variable *activity* has influence on the student knowledge gain with significance = .023. The value of the R Square (0.309) tells that QuizPACK activity can explain 31% of knowledge gain by student. Significance of full equation is high, in ANOVA table Sig.=0.052 which is almost 0.05. At the same time, the variable *success* has no visible influence on the knowledge gain. Our analysis of student profiles tells us that low knowledge gain has been achieved by two groups of students - the students who had some good past experience and had high pre-test and post-test score and the students who were not working hard on the course despite of their low starting knowledge level. These students had low pre-test and post-test score. Both groups worked with QuizPACK less than average, but had very difference success rates. At the same time, those novice students who worked a lot with QuizPACK have also benefited a lot achieving higher knowledge gain.

This is a very encouraging result showing that QuizPACK can serve not as a mere self-assessment tool, but also as a powerful learning tool.

Model Summary

R	R Square	Adjusted R Square	Std. Error of the Estimate
.556(a)	.309	.222	1.60708

a Predictors: (Constant), ACTIVITY, SUCCESS

ANOVA(b)

	Sum of Squares	df	Mean Square	F	Sig.
Regression	18.466	2	9.233	3.6	.052(a)
Residual	41.323	16	2.583		
Total	59.789	18			

a Predictors: (Constant), ACTIVITY, SUCCESS

Coefficients (a)

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	5.655	1.159		4.878	.000
	SUCCESS	-3.797	2.242	-.373	-1.694	.110
	ACTIVITY	.009	.004	.553	2.513	.023

Table 2: The influence of QuizPACK work on the student *knowledge gain*

On the qualitative level the value of QuizPACK was also quite visible. Of 46 students in our class only 11 students have ignored QuizPACK (have not used it at all or used it very little attempting less than 15 questions). Interesting is that 6 of these 11 have failed the course completely obtaining D or F grades (on the contrary, all students who used QuizPACK have got at least C and none of them failed the course). Of the remaining 5 students four performed worse on the final exam than on the midterm exam. Midterm exam have been mostly devoted to Karel the Robot and evaluated more student potential programming skills while the final exam evaluated their knowledge of C language.We think that ignoring QuizPACK as a learning tool these students lost theiradvantage and got a grade lower than they deserved. The only student of these"ignorant" 13 who demonstrated good level of performance despite of working very little with QuizPACK was a very strong student (top grade on the final exam) with reasonable programming experience.

Student Subjective Evaluation

In our class of 46 we have collected 31 filled questionnaires from students who have gained enough QuizPACK experience to provide reliable evaluation of the system (out of 35 who were qualified to take it). A discriminant analysis that we have conducted to evaluate the quality of collected data has shown that 95.2% of respondents provided valid data.

Due to the lack of space, the following analysis focuses four of twelve questions. The overall results are summarized on Figure 2. The first question measured the student general attitude to the system. The majority of them (87.10%) believe that self-assessment quizzes "can significantly help them during the course", 6.45% answer that self-assessment quizzes "can help them during the course" and last 6.45% say that this tool "can sometimes be of help". None of the students has answered that self-assessment quizzes are "useless for their course".

The second questions asked the student opinion about the system's ability to generate the same question with different data and providing thereby a chance to work with the same question again and again. The dominated majority (80.65%) of respondents answered that this feature is "very useful", 16.13% answered that it is "useful" and only 3.23% – that it "could be useful, but in very few cases". None of the students answered that it is "useless".

The third question asked about the type and the content of question material. 29.03% of the students answered that they are "exactly right to be most helpful", 64.52% – that they are "good and helpful overall", 6.45% – that they are "sometimes helpful, but could be much better". None answered that the type and the content of quizzes are "not helpful at all".

The question about system interface have generated the following data. 45.16% of the students felt that it is "very good", 38.71% – that it is "good" and 16.13% – that it "has some problems or lacks some features". Again, none of the students answered that QuizPACK student interface "has some major problems".

As this analysis shows, the students were very positive about the system. More than 90% of students have provided at least light-positive answer to the three first questions. Moreover, for the first two questions, more than 80% provided a strong-positive answer. We have been evaluating different systems in the classroom and this is the strongest results we have ever achieved. QuizPACK was a clear champion of our class. Even the relatively plain system's interface - arguably the least appreciated feature of the system - has received more than 80% of light-positive feedback. We also think that it is remarkable that answers on the question about system's ability to generate the same question with different data are very close to the data, collected for the first question reflecting the general attitude to the system. None of other questions, concerning students' evaluation of other system's features shows such similarity of answer profiles. We may hypothesize that the students highly positive attitude to the system is strongly connected with the system ability to create parameterized quizzes.

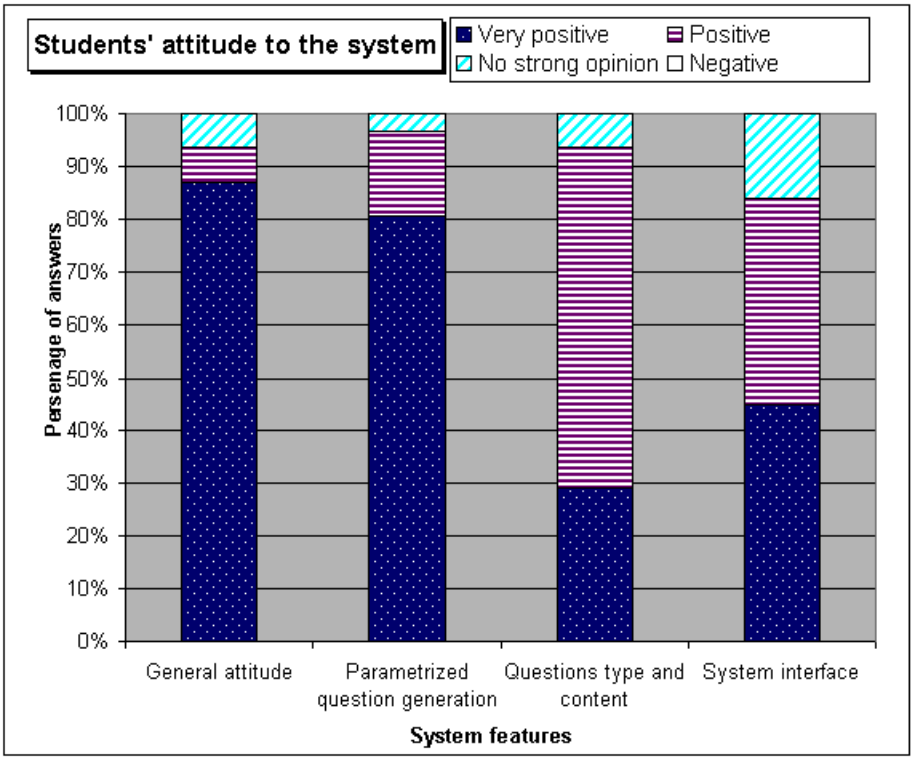

Figure 2: Students' attitude to the different features of QuizPACK

Next question asked the students if they agree with the statement "I will recommend my friends taking this course next semester to use this tool". 70.97% of the students choose "strongly agree", 12.90% – "agree", and 16.13% – "no strong opinion". None of the student choose "disagree" or "strongly disagree". As it we can see these numbers are also very close to the data, reflecting general attitude to the system, that could indirectly confirm the validity of students' evaluation.

We have also used two questions to examine the students attitude to the fact that self-assessment quizzes have been using as a source for in-class quizzes. 67.74% of the students thought that "it is a very good arrangement", 25.81% believed that "it is quite good", 6.45% told that "it makes some sense, though it is far from perfect", finally, none of the students thought that "it is a completely wrong arrangement". Concerning consideration of the QuizPACK as a tool for preparation for classroom quizzes, 58.06% thought that it "helps

them a lot", 32.26% thought that it is "quite" helpful, 9.68% – that it helps "a little" and none thought that it "didn't help at all".

QuizPACK usage encourages students to more actively use such programming tools as debugger. More then 51% of respondents have answered that they "have used debugger to find out where they are wrong when giving incorrect answer" "almost every time" or "often". 35.48% have said that they have used it sometimes and only 12.9% have said that they never have used a debugger.

To complement the general analysis we have attempted to compare the attitude of different groups of student to QuizPACK. So far we have completed the gender analysis that have brought some remarkable results. Figure 3 shows the profile of answers, given by female students, male students and students in general on the questions, reflecting students' attitude to the system. To compose the profile of the answers, we have calculated the average answer for each question and expressed it in percents. Here 100% mean that all students in this category give the strong-positive answer and 0% that all students give strong-negative answer. It is remarkable, that the line, corresponding to female student answers dominates the line showing male students answers. The difference is essential; it is about 10% for several questions. It particular, the graph shows clearly that *all* female students who took part in the study gave strong-positive answer to the first two questions! Thisis a very remarkable result. We were evaluating several educational tools in the past, but we have never observed such a high level of praise for any other educational tool from female students. Note that we can't attribute this extremely positive attitude of female students to QuizPACK to the fact that they have been working with the system more actively or were more successful. In fact female students have not been working significantly more with QuizPACK than male students (average numbers of questions attempted by female and male students are 187 and 178 correspondingly). Their average success was not higher either (37.12% of correct answers for female students and 37.26% for male students). We may hypothesize that an exceptionally high attitude of female students to QuizPACK expresses the fact that QuizPACK was a really critical tool for female students over the duration of the course. It is remarkable that in our course female students, traditionally lagging behind male students in technical subjects, were able to achieve a higher average grade than male students. Interesting is that even female students, who have got "C" or "B-" or lower grade have given strong-positive answers in our questionnaire.

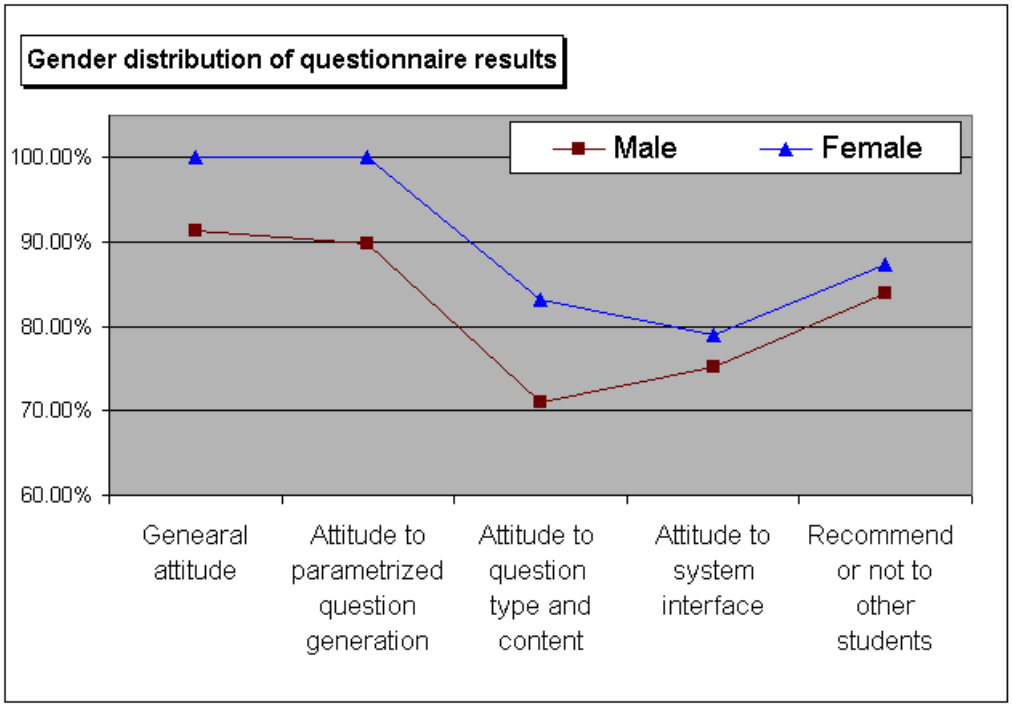

Figure 3: Gender distribution of questionnaire results

Summary

We have reported the results of our study that attempted to measure the role of parameterized questions as a learning tool in the context of a programming course. The study shows that the student successful work with this tool highly correlates with their success on classroom quizzes and also correlates with their final exam performance. The amount of student work with parameterized questions is also a significant predictor of their course knowledge gain. The students have regarded the QuizPACK system exceptionally high. Most highly was rated the system ability to serve the same question several times with different parameters. Female students evaluated different features of QuizPACK more positively than male students. The percentage of female students selected strong-positive answers in the system assessment questionnaire reached 100% on several questions. These results show that parameterized questions as served by QuizPACK system provide an exceptional learning tool in the context of a programming course. There is also evidence that this tool was especially helpful for female students who usually lag behind male students in programming courses. We intend to perform further analysis to determine more specifically how QuizPACK can help different groups of students.

References

[Brusilovsky & Miller 2001] Brusilovsky, P., & Miller, P. (2001). Course Delivery Systems for the Virtual University. In Tschang, T., & Della Senta, T. (Eds.), Access to Knowledge: New Information Technologies and the Emergence of the Virtual University. Amsterdam: Elsevier Science. 167-206.

[Graham, Swafford & Brown 1997] Graham, C.R., Swafford, M.L., & Brown, D.J. (1997). Mallard: A Java Enhanced Learning Environment. WebNet'97, World Conference of the WWW, Internet and Intranet, AACE. 634-636.

[Kashy et al. 2001] Kashy, D.A. et al. (2001). Individualized interactive exercises: A promising role for network technology. 31st ASEE/IEEE Frontiers in Education Conference, IEEE.

[Kashy et al. 1997] Kashy, E. et al. (1997). Using networked tools to enhanse student success rates in large classes. 27th ASEE/IEEE Frontiers in Education Conference, Stipes Publishing L.L.C. 233-237.

[Merat & Chung 1997] Merat, F.L., & Chung, D. (1997). World Wide Web approach to teaching microprocessors. FIE'97, Frontiers in Education Conference, Stipes Publishing L.L.C. 838-841.

[Pathak & Brusilovsky 2002] Pathak, S., & Brusilovsky, P. (2002). Assessing Student Programming Knowledge with Web-based Dynamic Parameterized Quizzes. ED-MEDIA'2002 - World Conference on Educational Multimedia, Hypermedia and Telecommunications, AACE. 1548-1553.

[Titus, Martin & Beichner 1998] Titus, A.P., Martin, L.W., & Beichner, R.J. (1998). Web-based testing in physics education: Methods and opportunities. Computers in Physics, 12 (Mar/Apr), 117-123.