

Improving Learning by Collaborative Testing

Sergio Ochoa, Luis A. Guerrero, José A. Pino & César A. Collazos

Universidad de Chile

David Fuller

Pontificia Universidad Católica de Chile

We have conducted experiments to improve examinations in Computer Science courses. We present a strategy that promotes learning during the examination process. This strategy has been implemented through two techniques. ETT (Examining Together Technique) and CET (Collaborative Evaluation Technique) are the techniques we have used.

According to Hay, the objective of an examination can be (a) to measure the real level of the students' knowledge, and (b) to measure the students' capacity to relate and integrate the acquired knowledge (Hay, 1996). On the other hand, Ebel and Frisbie maintain that the objective of examinations is to measure the students' achievements, to motivate and direct their process of learning (Ebel & Frisbie, 1986). These are some of the most accepted interpretations of the main objectives of examinations. However, we think that examinations can also be utilized for learning, especially if collaborative learning activities can be integrated into them.

The task of designing and grading examinations is additionally complex in knowledge areas in which there are no unique solutions. In these areas, knowledge is highly interrelated and its validity depends on its context. This is why there are few absolute truths, but many guidelines or suggestions of paths to follow. An example of such area is Software Engineering (part of Computer Science and Engineering).

On the other hand, in these traditional exami-

nation scenarios, various problems associated with the evaluation of knowledge can be observed. Some of the most relevant problems that appear are:

1. *Ambiguity*: The student's and the examiner's interpretation of the same exam question may differ.
2. *Anxiety*: The students show anxiety before and after the exams. The pretest anxiety is due to the student's uncertainty about the knowledge that will be evaluated. On the other hand, the post-test anxiety is due to the fact that examination results take too much time to be corrected, so many students are anxious to know their results. This anxiety can interfere with the student's learning process (University of Wisconsin at Milwaukee, 1990).
3. *Lack of uniformity in grading criteria*: Usually, courses have more than one Teaching Assistant (TA) in charge of evaluating tests. Even though they receive a grading outline, a uniform criterion is not typically used. In many cases, an important part of the final grade depends on the teacher's subjective judgement.

4. *Sense of injustice*: Many students do not dare to ask the instructor or the correctors about dubious correction of their exams. It is also possible that the teacher does not give a convincing answer to a student's protest. Indeed, there might be more than one correct answer for one exam item, and these alternatives sometimes are not considered by the corrector.

As a form of attacking these problems and using the examination as a learning scenario, two evaluation techniques were designed. The work hypotheses were the following: (a) the exam is the appropriate moment to support the student's learning, and (b) a collaborative exam strategy provides a tool for both measuring knowledge acquired by the students and promoting learning. Clearly, hypothesis *b* depends on *a*.

To prove these hypotheses, in 1998 we began to experiment with a group evaluation technique in college level courses in Computer Science. This technique was called *Examining Together Technique (ETT)*. Subsequently, we designed a new evaluation technique, based on the previous one. This new technique was called *Collaborative Evaluation Technique (CET)*. Although CET has not undergone much experimentation, the results obtained thus far are quite encouraging. In the present article, the two evaluation techniques are described and the experiments results are presented. These results show the impact that collaborative activities have on the testing scenario in college courses, in Computer Science.

Related Work

In the established scenario, it is possible to use a self evaluation scheme to reduce the effect of the post-test anxiety (part of problem 2), preferably in the lesson following the exam. In this technique, the student corrects his/her own exam, based on outlines given by the instructor, and assigns a grade to each question. This auto-correction can be seen as a problem of matching the student's answers and what the instructor considers to be the correct answers. In areas with little exactitude, where there are generally no unique answers, this technique must be used with care. When the self-evaluation ends, the examination can be revised by an assistant or by the instructor,

who can modify the grade that the student assigned himself. Even though self-examination reduces the effect of post-test anxiety, the student is still susceptible to the rest of the problems described in the previous section. On the other hand, the amount of learning that this technique provides is low, because students do not have the possibility of discussing and negotiating answers.

Co-evaluation or evaluation of peers is another technique that could be used to solve some of the problems. This technique is similar to self-examination, and consists of each student correcting the exam of another student, based on an outline given by the instructor. At the end of the evaluation process, the instructor revises all corrections and modifies what she deems necessary. This type of evaluation has the same strengths and weaknesses as self-examination. Even more, these evaluation techniques do not reward or punish for good or bad corrections, so students have little motivation to do this task.

Another alternative that can be used is multiple choice. This technique also reduces post-test anxiety as well as the effects associated to the lack of uniformity in the correction criteria (problem 3). This technique requests that the student should choose the right answer from a provided list of optional answers to each question. After the examination is finished, the instructor could indicate the right answers, thereby eliminating the post-test anxiety. The process of correction is reduced to the problem of matching the student's answers to the ones indicated by the instructor.

There are evaluation techniques that are a blend or adaptations of those just mentioned, like for example, those proposed by Shen *et al.* at the New Jersey Institute of Technology (Shen, Cheng, Cho, Hiltz & Bieber, 2000; Shen, Hiltz, Cheng, Cho & Bieber, 2001). To be used, these techniques have to be adapted to each scenario. The technique proposed by Shen *et al.* involves four steps. First, each student proposes an exam question. Second, each student selects one question and answers it. Third, the student who created the question grades answers and provides justifications of the grading. Finally, advanced students (e.g., Ph.D. students) do a review of the grading and lastly, the instructor provides the final grade. Clearly, this strategy contributes to the reduction of the ef-

fects of problem 2, but nothing can be assured with respect to the rest of the problems.

None of the evaluation techniques presented above are capable of reducing or neutralizing all the problems, and promote learning during the evaluation process at the same time. Therefore this work proposes a new strategy, which has demonstrated to be useful in overcoming these challenges. The strategy consists of introducing group work as part of the examination process.

Examining Together Technique (ETT)

In 1998, we designed ETT as a form of solving the typical problems of traditional examinations, and also to test our hypotheses (*a* and *b* defined in Sect. 1). This technique involves three phases, *pre-test*, *test* and *post-test*. Since each of these phases is carried out during class time¹, the evaluation process involves three consecutive classes (in different days).

During the pre-test, concepts and the relationships among concepts are clarified, to facilitate their assimilation and improve their application. This stage seeks to reduce the pre-test anxiety (part of problem 2) and is not graded, but constitutes an important aid when it comes to answering the exam.

The test phase responds to the objectives proposed by Hay (Hay, 1996) and is done individually. The work done by the student during this stage is graded in the same way as a traditional examination.

Finally, the post-test is carried out. The objective is to give the student the possibility to learn independently of the grade of his exam. It also contributes to the reduction or elimination of the four problems indicated in the introduction. Basically, this phase asks the students to generate an outline of the exam solution, based on a discussion and negotiation among themselves. Afterwards, each student evaluates his original exam² according to this outline. The correctors assigned by the instructor grade the exam, and they do not know the grade proposed by the students. Finally, the good corrections done

by the students are rewarded. Each of these phases is described in detail below.

Pre-test Phase

During the ordinary classes, the instructor transmits the knowledge and, surely, the concepts in an isolated way. Even though knowledge and concepts can be understood at that time, if they are not organized in an adequate way, a great part of this information could be lost. According to Novak, human beings need to relate their knowledge to make it more understandable and in that way, more usable (Novak, 1998). The human being thinks better in terms of structures (Novack & Sano, 1995; Novak, 1998), so structuring and organizing concepts (for example: relate, classify, establish hierarchies, etc.), helps better understand the knowledge acquired. On the other hand, Constructivism establishes as one of its postulates the construction of new knowledge over the base of the actual structure of the student's knowledge. This activity requires, on one side, the understanding of the new knowledge, and on the other, that the knowledge be inserted into the student's mental scheme (Novak, 1998).

The main objective of the pre-test is thus to help the student organize and relate the knowledge, to make it easy to understand and apply it. Moreover, this activity gives the student a slot of time in which to mature his knowledge before the examination. The student's performance during the examination depends fundamentally on the quantity and quality of the assimilated knowledge.

The pre-test is done during the class before the examination and involves only one activity. There are many types of activities intended to achieve the objectives of the pre-test, but the most successful have been the discussions and the solution of problems in groups. The discussions are spaces for interaction between students and instructor. The instructor usually brings to class a set of 8 to 10 questions concerning the knowledge that will be evaluated in the test. These questions should be similar in complexity to those in the test, and must not point to unique or trivial answers. Ideally, questions with open answers should be chosen. The answers to such questions need to be related to other concepts or scenarios to prove their validity. This generates discussion among students

¹ In Chile, the class lasts 90 minutes

² The original exam is photocopied to prevent students changing answers.

and in consequence, typical activities of this scenario appear, for example, self-explanation, idea appropriation and negotiation (Dillenbourg, Baker, Blake, & O'Malley, 1995). When this phase ends, the students should have a clear understanding of the most important concepts, and of the relationships among them.

Another activity with good results in the pre-test phase is the solution of problems in groups. The instructor poses a problem with similar complexity to that which could be asked in the test, and whose solution involves an important part of the knowledge to be examined. In this scenario, the students propose, discuss and justify their approaches to problem solutions. When this activity finalizes, a consensus solution to the proposed problem must be obtained. The role of the instructor during this activity is to coordinate the students' interventions, and to channel them towards the solution of the problem. This activity is somewhat similar to the discussion, and therefore mechanisms like self-explanation, idea appropriation and negotiation seem appropriate (Dillenbourg, Baker, Blake, & O'Malley, 1995).

It is recommended to use discussions in those cases in which the test is based on questions. If the test is centered on the solution of problems, it is convenient to do a pre-test using the second proposed activity. Since both discussions and group solution of problems show the main characteristics and the complexity of the test, the students reduce their level of anxiety. This reduction of anxiety permits them to do better on their examination (University of Wisconsin at Madison, 1990).

During the pretest the instructor takes the role of moderator. Thus, he is responsible for coordinating and channeling the students' interventions. The students assume the role of constructors of the solution. The instructor assigns to one of the students the role of secretary. He shall be in charge of taking note of the solutions or agreements reached. A copy of the notes made by the secretary shall be given to each student. The correctors need not participate in this phase.

Test Phase

Our educational system is accustomed to individual evaluations of the students. Therefore the test

phase is conducted in the traditional way, that is face-to-face written test.

During the design of the test, it must be remembered that the test will be evaluated using the proposed technique. Hence, it is recommended that, of the total time required to answer a question, 70% of the time is for thinking, analyzing and outlining solutions, and the remaining 30% is to answer the question. The main objective of the exam should be to state challenges that require the application and relation of concepts in order to give a solution. The examinations should not reward writing speed or memorization capacity, but creativity and ability to apply the acquired knowledge.

During the test process the students can consult the bibliography if they wish, because this is a scenario similar to what they will find in real life. If the student solved the exam successfully, it means that he was capable of understanding and applying the course knowledge in a domain that was initially unfamiliar to him.

Post-test Phase

The objective of this phase is to give the student the possibility of learning from his mistakes, and also to establish clear guidelines for the correction process. The post-test phase is divided in three parts: *search for solutions*, *self-evaluation* and *exam grading*. The first two activities are done during the class after the examination, and the students, assistants and instructor participate. The student's participation in these activities is voluntary. If he decides not to participate, his exam will be corrected in the traditional form.

For the third part, called exam grading, only the assistants and partially the instructor partake. This activity should be finished after the class in which the search for solutions and self-evaluation were done but before the next class. Each of these three parts is described below.

Search for Solutions

This activity consists of the construction of the test solution, according to the discussion among the students. The instructor presents the first question of the exam so that the students can discuss and negotiate the solution among them. The students can not

ask questions during this process, they can only make statements of negations, with the corresponding justification.

The instructor assigns a time slot in which the students must obtain the solution. If this time slot ends and the students have not reached an agreement, each student must expound his own solution based on his knowledge and the previous discussion. Afterwards, each student corrects the previously analyzed item in his exam. To do this, the self-evaluation process described below is used, considering the results of the previous discussion. This process is then applied to each and every one of the exam questions. Finally, the students grade the exam as a whole, in accordance with the individual grades assigned to each of the questions.

During this activity, the instructor plays the role of moderator. She should assure the fulfillment of the preestablished rules of the execution of this activity, and the times assigned for the discussion of each question. The students take the role of constructors of the solution. Unlike the pretest case, there is no secretary now.

This search for solutions gives the student the possibility of expounding, arguing and validating his viewpoint before his peers. This generally produces a greater sensation of justice with regard to the evaluation of his work. When this phase is over, the student has quite a clear idea of the result of his test, which helps reduce the post-test anxiety.

On the other hand, this activity also helps the instructor know how ambiguous her questions were, how clearly the concepts were handed, how balanced the test was, and other benefits. The instructor's frankness regarding her work is perceived by the students, and generally is expressed as an increase in the empathy between the instructor and the students.

Self-evaluation

Self-evaluation lets each student grade his answers on the photocopy of his test. If an answer does not obtain the maximum (or the minimum) grade, the student must justify in writing the assigned grade. This provides the correctors the basis for each assigned grade.

It is customary that the students place their justifications on the left margin of the page, meanwhile

the correctors do so on the right margin. Both use differently colored pens. To facilitate the correction process and to make it comparable, it is stipulated that all questions of a test must be qualified with a grade between 1.0 (minimum) and 7.0 (maximum).³ The instructor can assign a different weight to each question.

When all questions are graded, the students compute the total grade of their tests and hand in the graded tests to the instructor. Then, the instructor hands them over to her assistants, so that they can carry out the *exam grading*. This task is described in the following section.

Exam Grading

The assistants evaluate the copies of the students' exams, which have been graded by them. The students' answers are examined and graded. If the final grade assigned by the corrector differs in no more than 0.4 points from the grade assigned by the student, the grade for that question will be equal to the corrector's grade, plus a bonus of 0.5 points. Otherwise, the grade will be the same. The assignment of the bonus in the correction of the item is to reward the student for having understood what the correct answer was. Even though the student did not answer the test well, the post-test phase gives him the chance to learn and improve his grade.

This bonus shall not be assigned to those grades that required a justification but do not have one. Nor will it be assigned to those answers that agree with the grade, but have a wrong or insufficient justification.

Results Obtained with ETT

This technique was tested with various Computer Science courses in three different universities (Table 1): Pontificia Universidad Católica de Chile (PUC), Universidad de Chile (UCH), and Universidad Nacional de San Juan, Argentina (UNSJ). These courses correspond to college and graduate levels. For some of these courses, the proposed technique was evaluated more than once (several semesters).

³ The Chilean grading range is from 1.0 to 7.0. minimum passing grade is 4.0

During the experimentation process four exams were taken in each course: two traditional exams and two exams taken using the proposed technique (see Figure 1). All were done in the same order and following the same patterns.

At the end of each course, an opinion poll was conducted, to get the students' opinions. For this purpose, the participants had to qualify with a value between 1.0 and 7.0 (from poor to excellent) each of the questions of a questionnaire. Some of the more relevant questions were the following:

1. How satisfied are you with the ETT used?
2. How fair were the grades assigned to the exams?
3. How much did the collaborative evaluation technique help in the identification of ambiguities, both in the formulation of questions, and in the correction of answers?
4. If you were in charge of teaching a course in the area of Computer Sciences, how willing would you be to apply the proposed evaluation technique?

5. How uniform were the correction criteria for the exams?
6. Would it be suitable to penalize bad corrections on the part of the students?
7. Do you feel this form of evaluation is better than traditional exams?
8. How much did this evaluation technique help you organize the knowledge delivered by the course?
9. Do these type of examinations help evaluate the instructor's work?

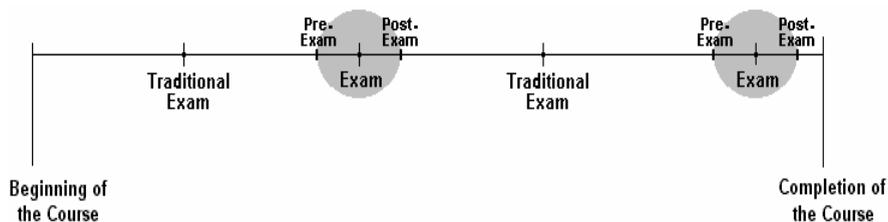
Instructors, assistants (correctors) and students answered the questionnaire. Apart from the previous questions, there were specific questions for each category. For example, students also answered question 10 below, assistants answered question 11 and instructors answered question 12.

10. How much did the evaluation technique help you learn during the test?

Table 1: Courses in which the ETT technique has been applied

Course name	University	Semester/Year	Level
Introduction to Programming	PUC	II/2000	College
Soft. Engineering Laboratory	PUC	I/99, II/99, I/00, II/00	College
Groupware Laboratory	PUC	I/2000, II/2000	Graduate
Software Engineering	UCH	I/2001, II/2001	College
Development of Soft. App.	UCH	II/2001	College
Software Engineering Lab.	UNSJ	I/2000	Graduate
Groupware Laboratory	UNSJ	II/2000	Graduate

Figure 1: Exams Taken in the Courses



11. How much did the evaluation technique help you improve the test correction?
12. How much did the evaluation technique help you improve the evaluation method?

The results obtained in the questionnaires were averaged by role (instructor, corrector, student). In total, 349 students, 5 instructors and 12 correctors participated. Table 2 shows a summary of these results.

From the results shown in Table 2, it is clear that all the process participants were more than satisfied with the evaluation strategy (in the 1.0 to 7.0 scale, 4 is considered fair, 5 is good, 6 is very good and 7 is excellent). The answers for question 10 show that hypothesis *b* (section 1) could be true. Moreover, exam grades obtained using ETT were between 10% and 20% higher than grades obtained using the traditional evaluation. Also, 80% of the students obtained the bonus in their final grade, which means that they learned during the test. These results show clear signs in favor of hypothesis *b*. Hypothesis *a* proof will require more work, but the results obtained regarding hypothesis *b* gives hope of its validity.

Regarding the typical problems of traditional examinations, it can be said that ETT does not avoid ambiguous writing, but it makes it explicit, showing clearly where the ambiguity is and which are the valid interpretations. This prevents the appearance of different interpretations among students and correctors. The results obtained in question 3 corroborate this statement.

The pre and post-test anxiety have been attacked through pre and post-test activities. During the pre-test the student gets an idea of the test type and complexity, and thus, he reduces his level of anxiety. Afterwards, in the post-test, during the self-evaluation, he can obtain quite a good idea of his test results. Although it is difficult to prove a reduction in anxiety

levels, the answers to questions 4 and 7 show the students have a good predisposition to use this technique.

The answers for question 5 do not indicate problems in the uniformity in correction criteria. On the other hand, if these problems were obvious, they would be reflected in the sensation of injustice that the students would have regarding this evaluation process. Answers to question 2 show a high level of satisfaction regarding the fairness with which the test was graded. The students, assistants and instructors share this opinion. In the courses in which this process was used to experiment, the percentage of complaints associated to the traditional examination correction was between 2% and 17% of the total of students. In the situations where ETT was applied, the percentage of complaints never surpassed 5%.

On the other hand, the attendance to pretests was between 80% and 100%. This percentage is high taken into account that in these universities, the average student attendance is roughly 50%. The post-test attendance was between 95% and 100%. This shows a great interest on the part of the students in participating in this evaluation process.

Table 3 shows the variance values for the information obtained from the questionnaires, categorized by question and by role. Clearly, a high level of agreement with each question qualifications can be assessed.

Some positive effects that were not foreseen during the design stage of this technique were the following ones:

- The aid provided by ETT in the improvement of the evaluation process due to feedback from the students.
- The students have the chance of evaluating the instructor's work, with respect to the information communication and the exam design.

Table 2: Questionnaire results for ETT

Role	1	2	3	4	5	6	7	8	9	10	11	12
Instructors	6.5	6.8	6.5	6.8	6.1	3.1	6.6	5.0	6.1	N/A	N/A	6.7
Correctors	6.7	6.7	6.8	7.0	6.3	2.7	6.8	6.1	5.9	N/A	6.8	N/A
Students	6.1	6.5	6.9	6.5	6.0	0.5	6.7	6.3	6.0	6.5	N/A	N/A

ETT Gains and Losses

Like all evaluation techniques, ETT has advantages and disadvantages. The most relevant advantages observed during experimentation are:

- It promotes and incites learning. The student can learn even after the test, independently of the obtained result.
- It reduces or neutralizes many negative effects associated with the typical problems of traditional examinations.
- It allows correctors to get a clear idea of each question answer, before beginning to correct.
- It promotes discussions and other social activities. It increases the students' emotional quotient (empathy).
- It allows the student to learn how to self-evaluate.
- It aids the student in the organization of his information, concept connection and application.
- It allows the instructor to improve her work, using the post-test data.
- It creates a natural relationship of collaboration among students, instructors and assistants.

The main disadvantages observed in the application of this technique were the following ones:

- The examination is longer than usual (three classes in total).
- If the coordinator does not control and coordinate the students' interventions well, then the post-test might not serve its purpose.
- A test suitable for this type of evaluation must be designed. This means there is additional design work for each exam.
- During the pretest and post-test, negative effects may appear for the students, like: *free riding*, *evaluation apprehension*, *air time fragmentation* and *domination* (Nunamaker, Dennis, Valacich, Vogel, & George, 1991).

Collaborative Evaluation Technique (CET)

After applying ETT for a number of years, and with very good results, it was decided that some improvements could be introduced. These improvements involve the creation of collaborative learning environments in the pre and post-test phases, with the objective of reducing the negative effects of free riding, evaluation apprehension, air time fragmentation and domination (Nunamaker, Dennis, Valacich, Vogel, & George, 1991).

Following a collaborative way of work is not simply making the students work in a group fashion; it is required that each group also pursue a common goal. It is required that the participants assume roles that are related, that complement each other and are different in order to achieve this goal. Therefore, the participants must work in small groups and with a common goal. This goal is tightly connected to the participants' individual goals, in such a way that each participant can achieve his individual learning objectives, only if the rest accomplish theirs.

Developing a collaborative activity requires the definition of a series of characteristics, like: the nature of the collaborative task, the number of participants, and the collaboration environment (Bannon, 1989). Kagan (Kagan, 1992) has defined four basic principles that must be present in any collaborative activity: *simultaneous interaction*, *equal participation*, *positive interdependence*, and *individual responsibility*. These elements not always can be acquired in the ETT because the teacher can not simply ask students to start projects and encourage peers to learn together. Collaborative learning is more effective if individuals and groups have to evolve within well-specified scenarios. Just assembling students in groups and telling them to work together may not be

Table 3. Data Variance

Role	1	2	3	4	5	6	7	8	9	10	11	12
Instructors	0.12	0.12	0.10	0.14	0.16	0.20	0.12	0.15	0.17	N/A	N/A	0.18
Correctors	0.16	0.11	0.12	0.13	0.18	0.17	0.15	0.12	0.19	N/A	0.15	N/A
Students	0.15	0.13	0.18	0.17	0.16	0.21	0.14	0.13	0.22	0.13	N/A	N/A

sufficient. For this reason, it is necessary to structure collaborative activities in order to promote positive interdependence among members of the group.

Keeping in mind the aspects defined by Kagan, and the ones proposed by Collazos *et al.* for the implementation of collaborative learning models in the classroom (Collazos, Guerrero, & Vergara, 2002), a better technique than ETT is described. This new technique, called Collaborative Evaluation Technique (CET), introduces collaborative learning activities in the pre and post-test phases. Unlike ETT, the execution of the three phases (pretest, test and post-test) are compulsory in CET. Each phase is carried out during a class. These phases are described below.

Pre-test Phase

The students work in collaborative learning groups during the pre-test, doing an activity related to the subject studied. This activity has been designed based upon the collaborative learning technique called JIGSAW (Aronson, Blaney, Stephin, Sikes, & Snapp, 1978), (see also <http://www.jigsaw.org>). This technique requires that students work in small groups, solving a set of questions asked by the instructor. The group objective is that all the members of the group learn how to solve each of the questions, reaching a consensus solution. To this effect, the proposed activity for the pre-test has four sub-phases: *collaborative group formation, solution discussion in pairs or expert teams, revisions in pairs* and *information delivery*.

The instructor assigns the work groups, trying to obtain heterogeneous groups, in terms of levels of knowledge. For this purpose, she can base her decisions on previous evaluations or the students' histories. If that information is not available, groups should be formed at random. Each group should have four or five students (Kagan, 1994). The set of questions handed by the instructor should be the same as the number of group members. Once the groups are formed, the instructor numbers the groups and hands over the questions. Then, each member of the group is assigned a question, and his responsibility is to become an expert on that subject (individual responsibility). The role of the instructor during the CET is to promote that members of the group really work in a

collaborative fashion, trying to intervene when someone is not participating, but the teacher does not answer the questions about the problematic situation.

In the solution discussion in pairs or expert teams, each individual leaves his group and looks for another student who was assigned the same question (it is recommended that it be done in pairs, but it can be done in groups of three, and maximum four). These students have to interact and negotiate their viewpoints, with the objective of solving the assigned problem. They can aid themselves with any didactic material (books, notes, etc.) while doing this activity. The instructor decides the length of this activity and sees to that the preestablished times are globally fulfilled. The revision of the problematic situation in pairs can be repeated more than once, until all solutions have reached all groups.

During the revisions in pairs, the students with the same assigned questions group again in pairs, but the pairs must be different than in the previous activity. Then, the solution that was found in the previous phase is revised, and the presentation of this solution to the group is posed. Finally, during the information presentation, the students return to their original groups, where each member informs his fellow group members about the solution to the problem that had been assigned to him.

The positive interdependence present in this phase is of the "reward" type since one of the analyzed questions will appear in the test. The instructor must guarantee that this will happen. Consequently, if each member of the group does his part of the work well, everybody wins, because they have solved one of the test questions beforehand.

Test phase

Each student answers his test individually, in a similar fashion as that proposed for ETT. The objective is to solve the questions posed by the instructor.

Post-Test phase

During the class after the test, groups of four or five students are formed so as to answer the test questions. Then, individually, each student grades his own exam. Each group must hand the solutions to the instructor. For the correction of their own exams, the students use the solution outline constructed by the

group. In a similar way as ETT, if the grade that a student assigns to an exam item differs in no more than 0.3 points (i.e., plus or minus 5% of the grade range) from the grade assigned by the corrector, the final grade for that question will be equal to the corrector's grade, plus a 0.5 point bonus. A member of the group must keep track of the time, because at the end of the activity, all questions must have been corrected. The positive interdependence stems from the fact that good solutions to the questions, constructed with the aid of all students, help everyone to be more objective in his own evaluation.

Results obtained with CET

This technique began to be applied in the first semester of 2002 for the Software Engineering course in which ETT had already been applied. Even though these are the first results in the application of CET, the data obtained is quite promising. In this experiment, 16 students, two assistants and one instructor participated. At the end of the test, the same questionnaire as ETT was filled out. The results are presented in Table 4.

The results from the questionnaire were similar to those for ETT, but there are two interesting aspects. The first is that most students agreed on penalizing bad corrections (question 6). This may be a consequence of the new design of the examination activity. All students have now the opportunity to present their opinions and influence others' opinions. The new design not only increases cooperation among participants but it broadcasts well-founded opinions.

The second interesting aspect is that students not only rated CET better than ETT but the variability of opinions was significantly decreased (see Table 5). This could be an indication of students' preference for CET. Rates from instructors and assistants were also better, but the sample size is too small to reach any conclusion from that. All 16 evaluated students voluntarily did the three CET phases. The exam average grade was 6.0 (the range is 1.0 to 7.0). This means a high degree of achievement. Even though the range of agreement of self-correction grade and assistant's grade (for bonus) was reduced from 0.4 to 0.3, 87% of the students got the bonus for at least one question.

Concerning variance on the poll results, only student were included in the computation because the sample size for instructors and assistants was too small. Results are similar to those obtained for ETT (Table 5).

The sample size is small, but results show CET is better than ETT. With this new technique the negative aspects mentioned earlier could be avoided. There will not be "free-riders", because the efforts of all the group members are required, and this is necessary for successfully reaching the common goal. Each group member has a unique contribution to make to the joint effort because of his resources and/or role and responsibilities, therefore there can not be social loafing.

Also, during last year we have applied CET in CC20A (Computer Science II). This course has approximately of 90 students. Some of the most important results are shown in the Table 6.

Table 4. Questionnaire results for CET

Role	1	2	3	4	5	6	7	8	9	10	11	12
Instructors	6.4	6.7	6.9	7.0	6.9	6.5	6.7	5.2	6.0	N/A	N/A	6.5
Correctors	6.5	6.8	6.8	7.0	6.9	6.3	6.8	5.7	6.2	N/A	6.4	N/A
Students	6.5	6.7	6.9	6.8	6.8	6.0	7.0	6.5	6.0	6.7	N/A	N/A

Table 5. Data variance for CET

Role	1	2	3	4	5	6	7	8	9	10	11	12
Students	0.12	0.12	0.10	0.13	0.10	0.12	0.10	0.11	0.19	0.09	N/A	N/A

C1 and C2 correspond to the average grades obtained in the first test, and second test. We applied the CET to the C2 during last year (2002/1 and 2002/2). If we analyze in a detailed way, we can observe that the grades have increased during the academic semester where we have applied the technique: the C2 grades when applying CET (2002/1 and 2002/2) are better than the grades of corresponding C2 tests of previous years and grades of the C1 test in the same year.

Since CET groups are smaller than ETT groups, they are less prone to evaluation apprehension, air fragmentation and domination. Besides, all students must participate, preventing the appearance of the negative effects. None of the four negative effects occurred during the experiment, but valid conclusions can only be obtained after additional experimentation.

Discussion

The following suggestions will increase the likelihood of success of collaborative activities in a classroom (Collazos, Guerrero & Vergaro, 2002):
Programming

- Evaluate each participant's performance individually.
- Keep the size of the groups small.
- Select the members of the group that have to do monitoring activities at random.
- Assign roles to the participants in such a way that, during the activity, they can be rotated.
- To train students before the pretest in collaborative techniques.

Table 6. Average grades for CC20A course

Semester	C1	C2
2000/1	3.0	3.41
2000/2	3.99	4.2
2001/1	4.0	3.45
2001/2	4.2	4.28
2002/1	3.3	5.0
2002/2	3.87	4.9

The instructor must do these activities taking various roles, such as educational designer, cognitive mediator and instructor (Collazos, Guerrero, & Vergara, 2001). This work is important to get the best performance from the group members.

It is important to highlight that these collaborative learning activities require a great deal of preparation and practice from the students, who are generally not prepared to do this type of tasks. In the courses taught at our universities, usually there is one assistant instructor that teaches one class a week, where problems are solved and questions addressed. It is proposed that in all these laboratory classes, collaborative learning activities similar to pre and post-test activities be done. In this way, the students not only get involved in a greater number of collaborative learning activities, but they also can practice these methods for actual pre and post-test activities.

The hardest problem we have found with the CET is that teachers need to do additional work, because they need to organize two kinds of test, during the pretest and test phases. They also need to copy the test of the students. On the other hand, there is a significant reduction in the number of complaints after the test results are given. In the CC20A course, for example, the number of complaints has been reduced in more than 80%.

Conclusions

In university courses and in areas with inexact knowledge like Software Engineering, most of the examinations are done following the FTF (face-to-face) model (Mason & Woit, 1998). These examinations are done in a written fashion. This model has various problems, like ambiguous questions, anxiety, lack of uniformity in correction criteria and sense of injustice.

At the same time, these techniques point to (a) measuring the students' real level of knowledge, and (b) measuring the students' capacity to relate and integrate acquired knowledge (Hay, 1996). None of them take advantage of the examination scenario to promote learning. Therefore, this article presents a strategy to promote learning through the examination process. This strategy has been implemented through two student knowledge evaluation tech-

niques, which have been called ETT (Examining Together Technique) and CET (Collaborative Examining Technique). Both solve the typical problems to which the traditional exams are prone. These techniques have been used in the Computer Science area, with very good results.

The ETT technique uses group activities to measure the level of knowledge and to promote learning. It takes advantage of the evaluations so as to “improve” learning, making the students participate in collaborative learning activities before and after the test. However, in ETT, problems like free-riding, evaluation apprehension, air time fragmentation and domination (Nunamaker, Dennis, Valacich, Vogel & George, 1999) can appear sometimes. Therefore, the CET (Collaborative Examination Technique) evaluation technique was designed, based on ETT. This new technique increases the level of interaction among the participants, during the pre and post-test phases, by applying collaborative learning techniques. Only one sample has been obtained thus far from the experimentation of this technique. The obtained sample reflects the same result patterns as ETT. For this single experiment, there was no record of the four problems that ETT shows.

According to the results obtained, it can be said that the examination process is a good opportunity to introduce collaborative learning activities into the classrooms. Testing is incorporated as an additional learning activity.

Acknowledgement

This work was supported by Fondecyt (Chile) grants N° 1000870 / 1030959, and grant thesis scholarship N° 49 from Departamento de Postgrado y Postitulos, Universidad de Chile (2002)

References

Aronson, E., N. Blaney, C. Stephin, J. Sikes and M. Snapp. *The Jigsaw Classroom*. Beverly Hills, CA. Sage Publishing Company, 1978.

Bannon, L. *Issues in Computer-Supported Collaborative Learning*. Procs. of NATO Advanced Workshop on Computer-Supported Collaborative Learning (Claire O'Malley Ed.), Maratea, Italy, 1989.

Center for the Improvement of Instruction. *Teaching at the University of Wisconsin-Madison: A Sourcebook*. Chapter 16: Exams & Quizzes. Madison, WI: University of Wisconsin at Milwaukee, 1990.

Collazos C., Guerrero L., Vergara, A. Collaborative Learning: A change in the role of the teacher (In Spanish). III Chilean Congress on Higher Education in Computing, Punta Arenas, Chile, 2001.

Collazos C., Guerrero L., & Vergara, A. A guide for the implementation of a collaborative learning model in the classroom (In Spanish). *ITINERANTES Journal*, 1, 2002.

Crawford, K and Fekete, A. *What do exam results really measure?* Procs. of the Second Australian Conference on Computer Science Education, July 1997.

Dillenbourg, P., Baker, M., Blake, A. & O'Malley, C. The Evolution of Research on Collaborative Learning (Spada, H. and Reimann, P. Eds.), *Learning in Humans and Machines: Towards an Interdisciplinary Learning Science*, 189-211, Oxford, Elsevier, 1995.

Ebel, R. and Frisbie, D. *Essential of Education Measurement*. Fourth Edition. Prentice-Hall. Englewood Cliffs, NJ, 1986.

Hay, I.: Examinations: Achievement Tests. *Journal of Geography in Higher Education*, 20(1) 137-142. 1996.

Jacobs, G.M. *Cooperative Learning or Just Grouping Students: The Difference Makes a Difference*. RELC Seminar, Singapore, 1997.

Johnson D., Johnson R. *Learning Together and Alone* (4th Ed.), Needham Heights, MA. Allyn and Bacon, 1994.

Johnson D., Johnson R., & Holubec, E.: *Cooperation in the classroom*. Interaction Book Company, Edina, MN, 1998.

Kagan, S.: *Cooperative Learning*. San Juan Capistrano, CA. Kagan Cooperative Learning, CA, USA, 1992.

Kagan, S. and Kagan, M. The Structural Approach: Six Keys To Cooperative Learning. *Handbook of Cooperative Learning Methods*. (S. Sharon, Ed.) 115-133. Westport, CT. Greenwood Press. 1994.

Kumar, Amruth N. *The Design of on-line tests for Computer Science I and their effectiveness*. Procs. of 29th ASEE/IEEE Frontiers in Education Conf. San Juan, Puerto Rico, Nov. 10 - 13, 1999.

Mason, D. and Woit, D. *Integrating technology into Computer Science examinations*. Proc. of the 29th SIGCSE Tech. Symp. on Computer Science Education. Atlanta, USA. pp. 25-28. Feb. 26 to Mar. 1, 1998.

Mullet K. and Sano, D. *Designing Visual Interfaces: Communication Oriented Techniques*. SunSoft Press. Prentice-Hall, 1995.

Novak, J. *Learning, creating, and using knowledge: Concept maps as facilitative tools in schools and corporations*. Lawrence Erlbaum Associates. 1998.

Nunamaker, J., Dennis, A., Valacich, J., Vogel, D. and George, J. Electronic meeting systems to support group work. *Communications of the ACM*, 34(7) 40-61. July 1991.

Shen, J., Cheng, K., Cho, Y., Hiltz, R. and Bieber, M. *Evaluation of an on-line collaborative examination process*. Proc. of the Americas Conference on Information Systems (AMCIS 2000), Long Beach, California, USA. 1791-1797. August 10-13, 2000.

Shen, J., Hiltz, R., Cheng, K., Cho, Y., and Beiber, M. *Collaborative examinations for asynchronous learning networks: Evaluation results*. Proc. of the 34th Hawaii Int. Conf. on Systems Sciences, IEEE Computer Society Press, Los Alamitos, CA. Maui. Jan. 3-6, 2001.

Sergio F. Ochoa is an Assistant Professor of Computer Science at the University of Chile. He received his PhD in Computer Science from Catholic University of Chile. His research interests include Computer-Supported Collaborative Work, Software Engineering, Educational Technology. Dr. Ochoa is a member of ACM, IEEE and Chilean Computer Society. Currently, he is IT Consultant of some Chilean organizations.

Luis A. Guerrero is an Assistant Professor of Computer Science at the Universidad de Chile. He received his PhD in Computer Science at the Pontificia Universidad Católica de Chile, a M.Sc. degree in Computer Science from Instituto Tecnológico de Costa Rica, and a Bachelor degree in Computer Science from Universidad de Costa Rica. His research interests include the modeling, design and building processes of collaborative applications, object-oriented technologies and collaborative learning.

José A. Pino is Associate Professor of Computer Science and Director of the PhD program in Com-

puter Science at the Universidad de Chile. His research interests include Computer-Supported Collaborative Work, Human-Computer Interaction and Software Industry Studies (SCCC) and President of Information Technology). He has co-authored six books and published research papers in international conferences and journals, including Journal of the ACM, Communications of the ACM. Decision Support Systems, Interacting with Computers and Information Technology and People.

César A. Collazos is an Assistant Professor of Computer Science at the Universidad del Cauca. He is pursuing his doctoral studies in Computer Science at the Universidad de Chile. He received his Bachelor degree in Computer Science from Universidad de los Andes in Bogota (Columbia). His research interest include Computer-Supported Collaborative Learning and Human-Computer Interaction, especially the modeling, design and building process of educative software applications.

David Fuller is an Associate Professor of Computer Science at the Pontificia Universidad Católica de Chile. He received his PhD in computer Science from imperial College of Science, Technology and Medicine at the london univeristy, England, a MS in computer Science from the University of California, Los Angeles, and a BS in Electrical Engineering from Pontificia Universidad Catolica de Chile. His research interests include collaborative systems, systems development, object-oriented frameworks and visual programming.