Experiments in evaluation: towards an eXtreme Learning method

Juan C. Dueñas, Félix Cuadrado
ETSI Telecomunicación, Avenida Complutense nº 30, Ciudad Univesitaria, E-28040 Madrid
Universidad Politécnica de Madrid
Madrid, España
{jcduenas, fcuadrado}@dit.upm.es

***Abstract*— In the last years, the field of education has advanced boosted by the introduction of Information Technologies. However, these advances have occurred to a lesser extent in the support of evaluation and assessment of students' learning, even when the improvements in evaluation would enhance the whole learning process. In this paper, the processes involved in assessment in engineering university learning are identified, described, reviewed and analyzed in order to ensure fast feedback to the students. A method for evaluation called eXtreme Learning is introduced which builds upon highly granular and continuous evaluation, integration of evaluation within the general learning process and fast feedback of evaluation results to students. At the same time, the method keeps some principles of traditional learning considered mandatory for its application in university education, as they certify the level of knowledge of the student.**

***Index Terms*— e-Assessment, educational methods, computer assisted assessment.**

I. Introduction

The integration of Information Technologies has contributed greatly to the advancement of the majority of educational processes such as information gathering, cooperative learning, distance education, housekeeping, etc. However, these advances have occurred to a lesser extent in the support of evaluation and assessment of students' learning, even when the improvements in evaluation would enhance the whole learning process.

In the rest of the paper, a model of the activities typically involved in university learning evaluation are presented; then we perform an analysis on such process in order to improve the feedback to students. A series of experiments in real university settings has been done, whose results are presented and discussed. As a consequence, an evaluation process called eXtreme Learning is introduced, which allows applying fast feedback loops into the traditional university learning process. At the same time, the method keeps some principles of traditional learning considered mandatory for its application in university education, as they certify the level of knowledge of the student.

The method relies on the application of well-known principles of pedagogy, such as choosing the proper granularity for each learning activity in contents and students' effort, ordering the knowledge acquisition following a logical path, enforcing the gathering of key knowledge, putting the knowledge into practice (particularly adapted to engineering education), evaluating the core concepts, converting the evaluation into a learning experience, and performing additive marking. Also, the role of e-learning technology is discussed, as the method, having a cornerstone on fast evaluation, cannot be applied without the proper technical support.

This article describes the results obtained in the application of the proposed method to a first course in computer programming in engineering syllabus over two different semesters. The specific characteristics of the subject impose additional requirements as students must face different studying conditions and context, while keeping at the same time the motivation for the kind of studies they have chosen. The application of the method has been maturing for a number of years as the selection of the proper granularity and sequence of learning contents was guided by a "trial and error" strategy. Then, we applied the method in the computer programming in the first course of Telecommunication Engineering in Universidad Politécnica de Madrid (UPM), for two different semesters, over a large set of students on a real setting. The control group exhibited better levels of success (percentage pass/total of students) and better learning results (marks over the whole subject) in the two semesters.

Outcomes from the experiences provide hints about the usefulness of the approach. Also the potential problems and drawbacks in the application of the approach are discussed.

II. Educational Assessment

In order to track the outcome of the learning process it is necessary to evaluate student's process. This is achieved using a wide range of methods for evaluating students performance and attainment including formal testing and examinations, practical and oral assessment and classroom assessment among others [1][2].

The role of evaluation activities in the university education has been analyzed as far as the second half of XIX century, by the philosophical school of K. F. Krause. At that time, throughout all Europe there is a strong current about pedagogy at university that identifies the concept of traditional exam as one of the key elements to be replaced or complemented by other evaluation methods (such as continual evaluation), in order to improve students motivation and get closer to the actual knowledge [3].

Manuscript received November 6, 2009. This work was supported in part by the Vicerrectorado de Ordenación Académica y Planificación Estratégica from Universidad Politécnica de Madrid (UPM) under grant "XL-TIC".

As regards evaluation, pedagogues usually distinguish between summative and formative assessment [4]. The first is usually applied once the learning period has been completed, is driven by the academic staff, can be verified and marked, validating that the knowledge /objectives have been obtained. On the other side, formative assessment is applied continuously, and intends to help the student improve the learning, providing feedback and recommendations on his own taken process.

Some tendencies try to conciliate both approaches; among them, "Authentic Assessment" [5] focuses on the acquisition of competencies and the evaluation of this achievement. Thus, evaluation must be a planned, integrated process related to professional competences. However, up to our knowledge, no formal efforts have been performed in order to combine summative and formative assessment into a single method that gets the best of both and avoids the problems they face.

In traditional evaluation process –at least in the university education-, whose goal is to qualify or certify the level of knowledge of the student at the end of a learning time, only summative assessment is applied, consisting of a final exam at the end of the teaching semester. This evaluation method has been used extensively due to some of its characteristics: it allows comparison in the group of students as well as the checking of a knowledge threshold, it is cheap in organizational terms; it is predictable in methods, secure and fair. It is also mandatory by legal conditions in some universities.

The traditional evaluation process has been improved in the last years by the introduction of the "continuous evaluation" approach, in which the evaluation activities are performed throughout the learning period (not only at its end). These processes bring the advantages of formative assessment for improving the quality of student learning. Its application allows for: increased feedback to the students about their learning process, increased motivation, feedback to the staff about the quality of teaching, consideration of evaluation activities as learning (it is also formative) [6]. It is a difficult method, as many times the evaluation gets uncoupled from teaching and from learning, it is time consuming, it may be unplanned, increases bureaucracy overhead in the learning process, and it may put strong pressure to students, which is clearly undesirable to retain them inside the academic environment (the "retention problem" refers as the capabilities to keep students focused on the syllabus they have chosen), as stated in [7].

However, the incorporation of formative assessment into course assessment presents several challenges to improve the formative capabilities of the process, such as the adaptation of learning objects to the characteristics and context of each student; or the consideration of learning paths, defined as the sequence of consumption of learning objects by the students. This sequence defines a path or trajectory of navigation on the set which can be defined in the learning planning activities, chosen by the student, or explored by social algorithms. In any case, there is an optimal trajectory, and the actual one that must be compared with the ideal in order to know the quality of the learning process, qualify the student, and suggest corrective actions. All this can be considered as a control loop, and as such, a fast feedback and an adequate response to deviations make the system stable. On the technical side, some of the outcomes related to this theory are: the definition of learning paths, the algorithms to adapt these paths to each student, and the resources for the evaluation of the knowledge.

The usage of Information and Communication Technologies (ICT) on evaluation is increasingly growing, leading to the concept of e-assessment as the study of the usage of computers and related technology within the assessment process. It has many advantages over traditional (paper-based) assessment, as it automates most of the repeatable tasks non depending on human judgment, such as getting the answers provided from students, storing them, marking or communicating results [8]. There are some successful systems supporting e-assessment based on forms or questionnaires, such as WebCT.

Particularly useful are those systems which integrate evaluation into the general learning process: learning Management Systems (LMS), also known as learning platforms. They are based on a Web server providing software modules for administrative and tracking processes on a learning system. These kinds of systems also ease distributed collaborative learning from prepared activities and contents; both in synchronous and asynchronous way, based on Internet services. Students follow the lessons, perform the scheduled activities, communicate with professors and other students, and produce statistical and marking records. Some LMS are Blackboard, Moodle and Claroline .

However, being powerful tools, LMS are usually devised for its usage in distributed settings with students spread across a region. Computer security measures, such as the usage of password authentication are in place. But ensuring the student gave answers to evaluation questions without help is still a process requiring human intervention (traditional examination in a certain place in which the student gets reduced communication capabilities is the best example). There are some evaluations that provide legal effects and therefore security and reliability must be preserved; so far these are performed purely in the summative mode without the help of ICT infrastructures.

III. Evaluation process

Therefore, we focus our efforts on the integration of formative and summative evaluation in the context of university education; particularly for engineering degrees which are regulated by law. So, for university education that follows the French-European model (university degree habilitates for a professional title or profession), the student must fulfill a series of competences; and this fulfillment must be assessed by certified staff –a kind of governmental agents. The administrative and legal effects of university education expand those of simply learning: knowledge must be demonstrated and evaluated in practice. In Spain, some branches of engineering follow this model, so the professional activity is regulated, and its access is also specified and controlled.

In the presented context evaluation must meet several requirements:

- Formal education (we define it as any kind of education part of a certified program that produces legal effects, ranging from kindergarten to university), requires validation and verification activities. To produce a legal effect (passing the course) it is necessary to get a proof that the required knowledge and competence have been acquired.
- Evaluation must be normative, as the levels of competence (learning objectives) the student may get should be described in a clear manner, also specifying the different degrees in which the objectives can be covered.
- Evaluation must be precise so it is possible to infer the coverage of these levels of competence by the student from the results of evaluation (examinations, assignments). For this purpose it must also be objective, meaning that no external conditions may affect the results of the evaluation, such as affective bias from staff towards students.
- Evaluation must be secure. This is the main detrimental factor for a greater adoption of e-assessment tools and techniques. Security relates to the authentication of students, non cheating, non plaguing, control of the processes, etc. For this purpose, face-to-face examinations are still a powerful evaluation tool.
- Evaluation must be fair for comparison between students. Even in the engineering university studies we are presenting, and despite the fact that the key point for students in each subject is the pass/no pass decision (all mandatory subjects must be passed to get the degree); each student is given a final mark in each subject. Individual marks or the aggregate results are used for comparison of egregated students by some engineering companies.
- Evaluation must also help in the learning process, helping each student to get knowledge about his/her level of competence and guiding the selection of optional subjects. During the learning period in each subject, availability of in-the-middle, continuous or formative evaluation results helps the student to focus on the issues he/she would require more effort. It is also a motivational agent.
- Finally, the academic and physical conditions in which the evaluation is performed, as regards the evaluation technique, the time for evaluation, the place and space it is performed on, the availability of consulting material, the text of questions, etc, must be handled with care to get the aforementioned goals but at the same time be realistic within the restrictions in staff and costs universities face.

Face-to-face examination must still be part of the process. Nonetheless, evaluation will be more efficient the more continuous it is; besides, it will be more focused the narrower the context of application. In continuous evaluation schemes, a formative approach can be supported We believe this infrastructure can play a fundamental role reducing cycle execution total time, and improving learning through a fast feedback, and adaptation to student pace.

Traditional university education has been mostly driven by these activities, performed at the end of a long learning time period (year, semester). This is a one-shot process, focused on the legal effects (pass, fail). It has been shown that evaluation and assessment activities have a strong impact in the overall learning process. The main goals of these activities are: give students a mark and ranking, help and motivate them by a quick and continual feedback, drive the syllabus development and also to contribute to analyze the teacher performance.

However, we think traditional methods provide poor feedback to students. It has already been empirically demonstrated that, in ICT (technological disciplines related with communications and computers) teaching knowledge is mainly acquired by means or practical work. The main drivers for student's motivation are learning-by-doing, and the formative role of mistakes [9] [10].

Because of that, in the last years traditional education has evolved from the "one shot process", quite similar to the traditional waterfall software development process, to "continuous evaluation" models. These models perform several iterations in the course period, including learning and evaluation in each cycle. However, the number of cycles is limited by the complexity of the complete assessment process. The process is composed by many activities, and involves professors, teachers, and university officers. Each step must also be performed with the required security, controlling agents' communication, detecting forgery and copies, as well as storing the documental proof of the evaluation.

Next, a generic process for evaluation is presented. It has been obtained from the observation of the different examination methods that are followed by the university staff; while it leaves apart the evaluation based on the observation of the student. Figure 1 shows the five activities currently performed in legally binding (paper based) university examinations. They have been identified from our experience and conversations with other university staff. There are five tracks, of which third and fourth affect to the quick feedback to the student.Other activities, such as the evaluation based on the delivery of homework or project results (in Project Based Learning approaches) could also be accommodated to this model.

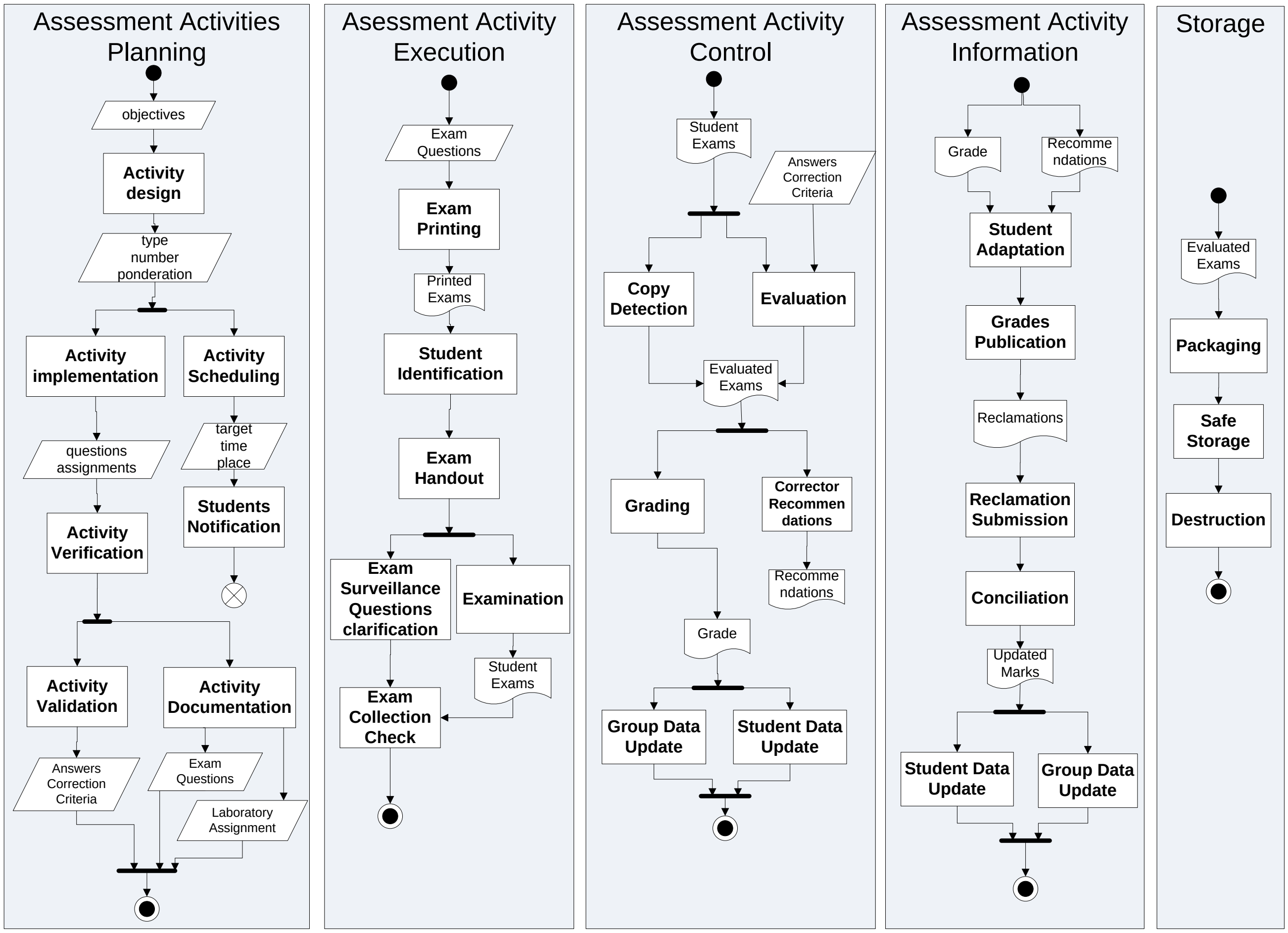

Figure 1 The Assessment Process

1. Development of the evaluation activity. Deals with all the actions required to prepare the examinations ensuring the questions or assignments are accurate and suitable. These activities are performed by the academic staff in a secure mode. They can be made in advance even before the subject period so they do not put an additional burden on academic staff.
2. Execution of the activity. Active participation of each student is required at this point, as he/she performs the examination under objective conditions.
3. Marking, which implies evaluation of answers and grading, and constitutes the largest effort to be performed and therefore the highest delay in the feedback to students.
4. Information to students of their performance levels and grades. Also conciliation activities may take place at this point if the student feels in disagreement with the final mark. Review and conciliation activities are regulated by internal norms in our university, but these regulations are only applied for final examinations.
5. Storage of results as any other legally binding document should be safely stored for 5 years.

It is interesting to note that all the activities affect the overall qualities of security, reliability and validity –we call these cross-cutting concerns. It is also worth noticing that so far marking and feedback to students are operations whose effort depends at least linearly to the number of students, hindering the application of methods that imply several examinations, and producing in general poor results as regards the time of feedback.

As it has been mentioned before, these activities are suitable to be automated, but on the other hand a purely computer based approach may not meet the security requirements that legally binding evaluations hold. That is why we approach the problem of evaluation as a mix of purely asynchronous computer based evaluations (in its application to Computer Science I these are programming assignments), with synchronous paper base exams, performed several times during the semester.

Assessment activity development is the first subprocess. Only the teacher participates in it, and its overall intent is to completely define the assessment activities that will allow measuring the students' fulfillment of the course objectives. The process starts by analyzing these objectives. Depending on the characteristics of the subject, the students' profile, the organizational context and the estimated relevance and criticality of each specific part, the teacher will prepare a high-level design of the subject assessment. By analyzing the objectives, a preliminary definition of the amount of assessment activities, its relative weight to the final mark, and the nature of each type of assessment (written exams – with problems, short and /or long questions - individual or group assignments and laboratory practices) will be performed. Once these activities have been defined, they can be integrated to the overall course planning, scheduling them over the semester, and, if necessary, performing room reservations for its execution. At this point, they can be notified to the students in order to allow a reasonable preparation interval.

In parallel with the organizational tasks, the academic staff will “implement” each activity, defining the specific questions or assignment specifications that will constitute each assessment activity. Once defined, these questions must be reviewed, in order to ensure they are appropriate to the objectives they are intended to assess. This verification should if possible be applied by a different teacher than the creator, in order to provide a broader perspective. Finally, the activity must be properly documented, obtaining the final enunciate of the exam questions, assignment specifications and support material. At the same time, the activity must be validated, in order to define the expected answers and the associated correction criteria. This way, the activities have been completely defined.

The next sub-process is the activity execution by the students, under the teacher surveillance. At this point there are some variations depending on the activity kind. We will describe a written examination, as it is the reference model for any other assessment activity. Prior to the examination time, the teacher must print the physical exams. At the examination room, after verifying the identity of the students, they will be handed out the paper, and the students will start the examination. Over this process, teachers watch over the students and clarify the exam questions. After the scheduled time, the teacher collects the student exams and checks they are formally correct. On an e-Learning context, these activities can also be carried out, as long as the identity of the students is assured.

On a later stage, the teacher proceeds to the marking sub-process, where student answers are evaluated with the reference of the correction criteria, detecting potentially copied answers from different students. Once the evaluation is complete, the teacher will obtain two results from it, first, it is possible to define recommendations to the student, based on the wrong answers and the nature of its mistakes. Grades must be obtained from the correctness of the questions. With that information, both individual student data and group data will be updated by the teacher.

Once the marks have been assigned, the teacher will adapt the grade and the specific recommendations to the student, in order to improve the quality of the provided feedback. Once prepared, the complete information will be published to the students, which will have an open time frame to submit any reclamation they consider appropriate to the notified marks. For each reclamation, there will be a conciliation process where the teacher and student will review the exam evaluation, and, if the reclamation is deemed right, will update the student mark. Finally, individual and group data will be updated with the changes from the accepted reclamations.

After the assessment has been completed, it is the teacher’s duty to ensure that a proper storage is applied to the student proofs. This way, the student exam answers or assignments (either physical or digital) will be safely stored over a set amount of time (stated by the legal normative) and will be destroyed after that interval has passed.

On the light of our requirements for university evaluation, and the observation and formulation of current evaluation processes, we concluded that: it would be feasible to integrate summative and formative evaluations in our environment, this integration would need for several short cycles of assessment in which the feedback of results to students should be very fast, additionally e-learning technologies would help in supporting automatable tasks as much as possible. We performed some experiments in order to check these hypotheses.

IV. Experiments

We have applied the principles of frequent and rapid feedback to students through continuous, formative assessment to a particular subject from the Telecommunications engineering degree offered at UPM. For those countries such as USA where this engineering discipline is lesser known, it can be viewed as the combination of Electrical Engineering with Computer Science. The specific subject where the experimental techniques were applied is “Fundamentos de programación”, a mandatory part of the first semester of the first course of Telecommunications studies which is very similar to “Computer Science I” in other countries. The subject is allocated 47 hours teaching spread along the semester. The subject neither has formal prerequisites nor depends on other subjects –whereas several subjects in the following courses depend on it. This subject has been chosen for experimentation because it offers specific characteristics that make it suitable for improvements in the evaluation process:

- The subject holds both theoretical (knowledge) and practical (operational) competences. Its learning process requires a proactive and practice-oriented effort from the student. As there are few computer programming subjects in the degree, students must obtain a minimum level of proficiency so they can apply programming concepts in later subjects with no more formal learning, despite it is in the first course. Operational contents require a higher level of interaction between student and teacher than pure cognitive contents.
- It builds design competences; students must be able to analyze computer programs, but they must also be able to design or create small programs. This is the first real experience in engineering they get, which requires the student to create his/her personal problem solving process.
- It is offered to freshmen students; so the typical problems of retention of students, wrong selection of studies, adaptation of their study mechanism, and anxiety appear among them. The subject must cope with those factors offering additional motivation as being in the very first semester most of the subjects are on basic (non applied) disciplines such as Maths or Physics.

For several years the academic staff has been adjusting both the amount of lessons and its sequence, organizing the subject following an “objects first” approach [11]. Table 1 specifies the topics dealt with, the order in which they are presented, and the time (in minutes) allocated to each of them for the lectures. It may seem controversial issues in this table are: The time allocation for the introduction and basic definitions may seem controversial as these definitions do not provide operational capabilities, but they are intended to serve as a “training time”, allowing students to get a feeling on the subject –it is the very first semester in the university-.

TABLE 1 SUBJECT LECTURES TIME ALLOCATION

Lesson	Contents	Time(min)
1	Introduction: basic definitions	240
2	Simple classes, objects, references	240
3	Primitive types, operations, arrays	360
4	Methods in classes	240
5	Sentences and simple algorithms	300
6	Classes	240
7	Inheritance and interfaces	360
8	Collections (simple data types)	240

Another key factor which impacts the course organization is the large number of students, ranging from 350 to 500, split into 3 to 5 separate groups. This presents practical difficulties in order to organize the contents of the subject, perform teaching, and include the proper evaluation mechanisms. With these factors in mind the teaching staff agreed to follow a hybrid approach for the summative assessment, shared by all the groups. This way, a quarter of the final mark was assessed from the results of continuous evaluation (performed independently at each group), with the remaining 75% being determined by the final exam, common to every group.This ensured the formality of the process.

In the first semester of course 2007-2008 the first experiment on evaluation improvement was performed. In order to check our hypotheses, we made an effort to apply a large number of assessment cycles over the teaching period, and provide the feedback to students from each of them in a short period (less than 48 hours). This way, we performed a total of 12 assessment activities over 15 weeks (almost one evaluation a week), switching between programming assignments (that could be submitted through the Web) and written exams with short questions (which were performed in the university rooms following the evaluation process presented in Figure 1). The specific contents for each test were very limited, whereas the degree of difficulty was growing incrementally. For all the cases we published the reports before 48 hours from the evaluation.

At the end of the four-month course we validated our approach with the students by means of surveys and by checking marking results against the students from the remaining groups, which were not involved in the experiment. As the marking for each student was primarily determined by the final exam, common to all of the groups, the comparison of the results allows checking the usefulness of this approach to continuous evaluation. The analyzed group had 97 students, whereas the remaining four containing 328 (group composition being assigned at random). Marking results comparison shows that the pass rate was 5% better than the average overall mark, with the absolute grades improving an 8%.

On top of the raw results, further reflection was done over the experience, including the results of the student survey performed at the end of the semester. While the experience was overall positive, there was room for improvement over the implementation. Student feedback informed that the amount of continuous work felt very pressing, and was not directly tied to the final qualifications. In addition to that, it was not only costly for the teacher but also in our impression detracted time for explanation because of allocating so much time to pure evaluation. We believe that the number of continuous assessment cycles was excessive hence hampering the usefulness of this approach, and experiencing diminishing returns when compared with the required effort. These factors guided the evolution of the process for the second experiment.

In the second year of application (first semester of the course 2008-2009) the number of students in the control group was slightly smaller (89), but the same number of groups (5) remained. Based on the previous conclusions, when planning this new execution of the experiment we opted for reducing by 25% the number of continuous assessment activities (from 12 to 9). Table 2 summarizes the assessment activities carried out over this experiment, including information about the weight these activities had on the final grade, the kind of activity (written-programming lab), and the examination contents. In addition to the assessment definition, estimated effort for the execution, preparation and correction-notification time, are presented (in minutes). As it was shown in Figure 1, the preparation time is the time spent by the staff preparing the evaluation and correctors; execution time is the time actually taken by the students in performing the evaluation activity; and correction and publication includes the time to scan and upload written examinations, and for practical assignments the time spent by professors in building and publishing the listings, as well as performing manual inspection of the solutions given by students that did not perform correctly by the automated evaluation system (in these kinds of activities, time spent on erroneous answers turns to be the dominant factor).

TABLE 2 ASESSMENT ACTIVITIES FOR THE SECOND EXPERIMENT

Activity	Week	Weight	Contents	Execution time	Preparation time	Marking time
Exam 1	6	0.25	Short questions: objects, references, expressions, arrays	30	30	240
Lab 1	7	0.18	Simple classes, basic types, operations	60	30	240
Exam 2	9	0.85	Short questions: Methods, classes	45	30	240
Lab 2	10	0.18	Sentences, simple algorithms	120	30	240
Lab 3	11	0.18	Arrays	120	30	360
Exam 3	12	0.25	Regular exercises: classes, interfaces, arrays, sentences	60	30	240
Lab 4	13	0.18	Interfaces, inheritance	180	30	360
Exam 4	14	0.25	Question marks, regular exercises: interfaces, collections(data structures), algorithms	60	30	240
Lab 5 (opt)	15	0.18	Design	240	30	240
Final exam	16	7.5	Full contents	120	360	360
Total		10.0		915+120	270+360	2400+ 360

After applying the same validation in this second iteration we noticed an improvement in the results, which are shown at Table 3. Each column lists the statistics from one of the groups, with A being the subject of the experiment. For each of them group size, average mark and percentage of students passing the course is shown. It can be seen how both the average mark (over 10.0) and especially the pass rate (with 5.0 being the threshold) were significantly improved over the rest of the groups as well as the increased difference when compared to the previous experiment.

TABLE 3 RESULTS FROM THE SECOND EXPERIMENT

GROUP	**A**	B	C	D	E
SUBMITTED	**89**	58	47	49	67
AVG MARK	**6.06**	5.53	4.98	5.68	5.65
PASS/SUBMIT	**74**	47	53	49	60

V. DISCUSSION: THE XL METHOD

The presented experiments show the results of applying a set of improvements to traditional evaluation activities which we propose to name “eXtreme Learning” (XL), in analogy to the software development methodology eXtreme Programming [12]. We have found this name used by some UK pedagogic communities [13]. It is also a commercial name that offers services to K12 (pre university) students and organizations. In both usages of the term, it seems to be oriented to the primary and secondary school context. They move the principles and values of XP (eXtreme Programming) to the domain of learning and teaching, but so far they seem to be focused on the first value of XP: communication. For us, XL is a wider method, consisting primarily on the adoption of multiple fine granularity learning cycles in university education. It serves students to control their learning trajectories, providing incremental, fast feedback, integrating summative and formative evaluation.

From the results of its two initial applications we have obtained a set of conclusions which we believe can be extended to other implementations of this approach.

The first key aspect for a successful application is to find the right level of granularity for the activities. There must be a sufficient number of them, in order to provide a focused and frequent feedback to the students. On the other hand, a too large number of tests can lessen the effectiveness of the process, as students can feel overwhelmed by the load of examinations, the time for introducing the new concepts is reduced in concordance, as well as imposing unnecessary load to the teaching staff.

Applying this technique does not only affect the design of the assessment activities, but also influences the course planning. Subject contents must be structured in a sequential development, with each new concept building only over the previous ones. Course contents should be well defined and individually testable. The combination of these factors, will allow to focus only on the tested concept in each of the assessment activities, improving the quality of the feedback provided to students.

As regards the design of the evaluation activities, there are two factors which should be taken into account. Continuous evaluation activities should assess the different types of competences to be obtained from the subject (base knowledge, skills, design capabilities...). In addition to that, it is advisable to select types of activities which can be efficiently evaluated (e.g. short answers / questionnaires instead of long problems), in order to enable a quick feedback process. From the data received by the surveys, we believe that less than 48h is a good threshold value to optimize the usefulness of the provided feedback.

Clearly, from the previous point it can be concluded that e-assessment tools are a great fit to this approach, as they help reduce the required effort for marking, although as it was mentioned before, they do not completely remove the need for manual checking. In addition to that, the security and legal concerns impede to adopt them as the only mechanism, so face to face examination must still be the predominant summative assessment technique.

Finally, we will briefly discuss the relationship between the specific characteristics of the course subject to the experiments and the XL approach.

As regards the suitability for freshmen students, we believe the formative assessment capabilities of this approach can be critical to help them have a smoother transition from their previous learning habits to the steep requirements of the university education. This is not as critical for students from higher courses, as they have already developed the required learning skills.

The other key factor from those experiments is the high number of students per class, which did not allow providing personalized feedback and learning paths for specific students. However, as rapid feedback informs the students on their continuous progress, it does enable them to take a proactive role where, after looking at the results of his personal progress he can request individual lessons (tutorial) to his teachers. Regarding group size it must also be mentioned that the required effort from the teacher increases linearly with the number of students (and is multiplied by the number of cycles), so it can turn out to be very costly in the presented context.

VI. CONCLUSIONS

Over this article we have presented the XL method, applying the eXtreme Learning approach to a traditional university context. Bare bones, our interpretation relies on the execution of lots of assessment tests (both written examinations and practical exercises), corrected and marked quickly, so the student can get a fast feedback on his/her learning process.

Before applying this technique to our context of work we have first analyzed how the domain-specific requirements impact the assessment process. This way, both the formal and security constraints of engineering degree education and the challenges of adapting the subject to freshmen students have been identified. In addition to that, the general assessment process has been described, with an emphasis on the potential techniques for minimizing the complete time required for its execution, which is the only way to provide rapid and frequent feedback to students.

We have defined and implemented this approach for two consecutive years to a first year subject from the telecommunications engineering degree, with several improvements for the later iteration. In our experiments we have found that this fast feedback improves students’ motivation, improves the quality of learning and reinforces the active role of students in learning.

However, it must be noted that there are several factors which have to be considered before applying this approach to other contexts. First of all, on a traditional university context, the method cannot substitute, but must complement instead the traditional, legal evaluation methods. In addition to the organizational context, it is vital to find the right degree of granularity for the evaluation, and ensure the distribution of course contents allows applying highly focused, low-coupled tests.

As future work we plan to increase the role of e-assessment tools in the process. From our experience the availability of an IT infrastructure that eases the correction, communication and publication activities can substantially reduce the teacher effort, greatly increasing the applicability of this technique. In addition to that, we also plan to improve the quality of the feedback generated by the automated correction system for non-correct submissions, in order to provide personalized feedback to the students.

References

[1] C.V. Gipps, “Beyond Testing. Towards a Theory of Educational Assessment” Oxon, UK, Ed. Routledge Farmer, 1994.

[2] J. Tagg. “The learning paradigm college.” Bolton, MA, Anker Publishing Company 2003.

[3] J.P. Ramsden. “Learning to teach in Higher Education.” Routledge, London. 1992.

[4] T. Crooks, T. “The Validity of Formative Assessments”, Assessment in Education: Principles, Policy & Practice, Nov. 1997, Vol. 4, Issue 3.

[5] S. Gielen, F. Dochy, S. Dierik, “Evaluating the Consequential Validity of New Modes of Assessment: The Influence of Assessment on Learning, Including Pre-, Post-, and True Assessment Effects” from Optimising New Modes of Assessment: In Search of Qualities and Standards, Springer, 2003, pp-37-54.

[6] M. Yorke, “Formative assessment in higher education: Moves towards theory and the enhancement of pedagogic practice”, Higher education vol.:45 issue:4 p.p.477-501.

[7] C. Connolly, E. Murphy, S. Moore. “Programming Anxiety Amongst Computing Students-A Key in the Retention Debate ?” IEEE Transactions on Education, 52-1. Feb 2009.

[8] D. O'Hare, D.M. Mackenzie, “Advances in Computer Aided Assessment”. Staff and Educational Development Association Ltd, Birmingham, UK., 2004.

[9] M. C. Linn “The cognitive consecuences of programming instruction in classrooms”. Educational Research, Vol.14, n 15, 1985.

[10] Goold and R. Rimmer, “Factors affecting performance in first year programming,” SIGCSE Bull., vol. 32, pp. 39-43, 2000.

[11] S. Cooper, W. Dann and R. Pausch, “Teaching Objects First in Introductory Computer Science” SIGSE Bull. Vol 35, Issue 1, 2003., pp-191-195.

[12] K. Beck. “Extreme Programming Explained: Embrace Change.” Addison-Wesley, 2000

[13] K. J. Babbage. “Extreme Learning.” Scarecrow Education, 2004.