Chapter XIII

Iterative Design and Evaluation of a Web-Based Experimentation Environment

Anh Vu Nguyen-Ngoc, Ecole Polytechnique Fédérale de Lausanne (EPFL), Switzerland

Yassin Rekik, Ecole Polytechnique Fédérale de Lausanne (EPFL), Switzerland

Denis Gillet, Ecole Polytechnique Fédérale de Lausanne (EPFL), Switzerland

Abstract

Nowadays, Web-based experimentation environments provide an excellent instrument to add flexibility in traditional engineering curricula. This chapter presents a model for the evaluation of such environments. The proposed model relies on an iterative evaluation paradigm. It allows the integration of different analysis methods including quantitative and qualitative analysis, and social network analysis. The chapter also describes the iterative user-centered design and development of the *eMersion environment developed at the Ecole Polytechnique Fédérale de Lausanne (EPFL), as well as the results and analyses of the evaluation process carried out in the automatic control laboratory courses using the eMersion environment from the 2002 winter to the 2005 summer semesters at the EPFL. The evaluation was performed to study different aspects relevant for an online learning community in engineering education, such as participation, flexibility, learning performance, collaboration, and community social structure.*

Introduction

Automatic control is a mandatory course offered to various engineering degree programs including electrical, mechanical, and micro-engineering curricula at the Ecole Polytechnique Fédérale de Lausanne (EPFL). In automatic control, as in other engineering domains, laboratory activities—or hands-on activities in general—play an essential role in theoretical knowledge reinforcing and know-how acquisition. Hands-on activities also help in increasing students' motivation.

For about a decade, academic institutions have tried to meet the increasing student needs for professional competencies, personal development, and career planning, including the necessary skills for teamwork and lifelong learning. Furthermore, engineering departments have had to solve the logistical dilemma of educating more students with fewer resources while maintaining the quality of education. Within this challenging context, the so-called *flexible* learning paradigm (Gillet, 2003; Kazmer & Haythornthwaite, 2005; Mosterman et al., 1994) happened to be helpful. This paradigm is leading towards the development of a hybrid-learning scheme in which the traditional courses are combined with online activities that can be carried out at anytime and from anywhere. In addition to providing students with new online resources, the flexible learning paradigm also sustains the development of a learning community. All people involved in a course, including the educators, the tutors, the teaching assistants (TAs), and the students, who synchronously and asynchronously interact with each other and with laboratory resources, form what is called an online learning community.

Web-based experimentation is one of the online activities that plays a key role in the development and deployment of the flexible education paradigm in engineering education. Web-based experimentation stands for hands-on activities carried out online using either simulators (virtual experimentation) or remote connection to real laboratory equipment (remote experimentation). Typical Web-based experimentation sessions are mediated by tutors and TAs. There might be some face-to-face (f2f) sessions in which the students work in the laboratory with the presence of the tutor and/or TA (see Figure 1 as an example), but most of the learning activities take place online. This bimodal context requires special features to effectively support the online learning community.

First of all, the content delivered in online engineering courses includes not only static documents, textual presentations, or video presentations, but also computation, graphics generated on-the-fly, real devices measurements, etc. Hence, the environments supporting Web-based experimentation must provide necessary functionalities to enable monitoring, measuring, and manipulating the virtual or real experimentation resources. They also require additional software components supporting the organizational and the collaborative tasks associated with the hands-on activities.

Secondly, Web-based experimentation environments should encourage students to carry out experimentation in a flexible way. In other words, students are allowed to perform multi-session experiments. For instance, they can do the first part of the experiment at school, and pursue the rest of it at home thanks to the remote access to the laboratory equipment.

Thirdly, Web-based experimentation environments should provide shared spaces, as well as online collaboration facilities with which students can find, share, and co-construct knowledge. These components help the students actively create their own contextual meaning, rather than passively acquiring knowledge structures created by others. In an active learning perspective, students need to interact with their peers, collaborate, discuss their positions, form arguments, reevaluate their initial positions, and negotiate meaning.

Last but not least, Web-based experimentation environments should support awareness. In learning and especially in flexible learning, awareness (Dourish & Bellotti,

Figure 1. Hands-on activities in f2f learning modality

1992) plays a very important role for every member of the community. Tutors need awareness to have a general perception of the class activities, to monitor the class progress, and to detect problems in order to intervene in time. Students need awareness to have a perception about their progress compared with other groups. Awareness is also necessary for students to find potential collaborators for exchanging documents and ideas, or to ask for help.

As a summary, in order effectively and efficiently to support online communities in engineering education, Web-based experimentation environments have to integrate the components supporting multiple interaction dimensions, including not only the interaction with the experimentation resources, but also collaboration (interaction between students), tutoring (interaction between students and TAs), and data exchange (interaction among the Web components themselves). Furthermore, awareness features should be provided explicitly. Although several institutions have recently developed Web-based experimentation environments (Atkan, Bohus, Crowl, & Shor, 1996; Faltin, Böhne, Tuttas, & Wagner, 2002; Ogot, Elliott, & Glumac, 2003; Schmid, 1998; Tzafestas, Alifragis, & Palaiologou, 2005), not one satisfies all these requirements. Such environments have mostly focused on the interaction between the students and the experimentation resources. In some cases (e.g., Faltin et al., 2002), students have been provided with a shared workspace such as BSCW (http://bscw.gmd.de). However, the collaboration, the tutoring, and the data exchange in the context of flexible engineering education are still very limited or not supported.

Flexible learning and Web-based experimentation resources have been integrated progressively within the automatic control course in the engineering curricula at the EPFL. This chapter describes the valuation scheme and results obtained between the 2000 winter and the 2005 summer semesters regarding the deployment of the flexible scenario and the associated Web-based experimentation environment called eMersion for the course mentioned previously. The next section deals with some evaluation issues concerning Web-based experimentation environments. Then the model proposed for the evaluation of such online learning environments is detailed. A section is also dedicated to the presentation of the successive designs and refinements implemented. The following section is about the evaluation instruments and results. Finally, the chapter ends with some concluding remarks.

Evaluation Issues of Web-Based Experimentation Environments

User-centered evaluation is a newly emerging facet of the Web-based experimentation environment development. Evaluation is one of the main challenges as well as a prerequisite that could allow students to profitably exploit the environment. In single-user applications, it is already difficult to test the perceptual, cognitive, motor variables (Card, Moran, & Newell, 1983). It is however extremely difficult to evaluate multi-user applications (Grudin, 1988), especially to evaluate Web-based experimentation environments that support collaborative hands-on activities where many interactions take place at both a technical and a social level. Another very important point that needs to be evaluated is the learning performance of students participating in such an online course. In the traditional classroom, there are several methods that the tutor can use to evaluate students' learning process and to know about the students' progress. In an online environment, the tutor can mainly evaluate what he/she has access to.

Some initial attempts to evaluate Web-based learning environments in engineering education have been reported in Faltin et al. (2002), Ogot et al. (2003), Roppel, Hung, Wentworth, and Hodel (2000), and Tzafestas et al. (2005). These works have considered employing existing usability engineering methods applied to a small population of students. The favorite methods employed were empirical ones (Rosson & Carroll, 2002) such as field study, usability testing in a laboratory, or controlled experiments. In fact, various important aspects related to the online learning community in Web-based experimentation environments have been neglected. Actually, the evaluation should provide answers to questions about participation, learning performance, flexibility, collaboration, and social structure of the online learning community. The variety and complexity of the interaction processes and the need to consider the system from both social and technical points of view (Nguyen-Ngoc, Rekik, & Gillet, 2005b) require mixed and integrated evaluation methods that combine different sources of data and different analysis techniques applied at different phases from the analysis to the design, and up to the exploitation stages of the environment. By using different sources and methods at various points in the evaluation process, the evaluator can build on the strength of each type of data collection and minimize the weakness of any single approach. A multi-method approach to evaluation can increase both the validity and reliability of evaluation data (Frechtling & Sharp, 1997).

The eMersion environment (Gillet et al., 2003; Gillet, Nguyen-Ngoc, & Rekik, 2005) has been iteratively designed, developed, and deployed since the year 2000 on a semester basis. A model for the evaluation of Web-based experimentation environments has emerged from this iterative process. Then it has been generalized with the aim of providing a new structured framework to cope with the specific requirements of evaluating online learning environments in engineering education. This evaluation model, the instrumentation feedback model for evaluation, is detailed in the next section.

Instrumentation Feedback Model for Evaluation

The term *instrumentation feedback model* was coined in the work of Leifer (1997). This term is used in the sense of observing both independent and dependent variables in an automatic feedback control environment.

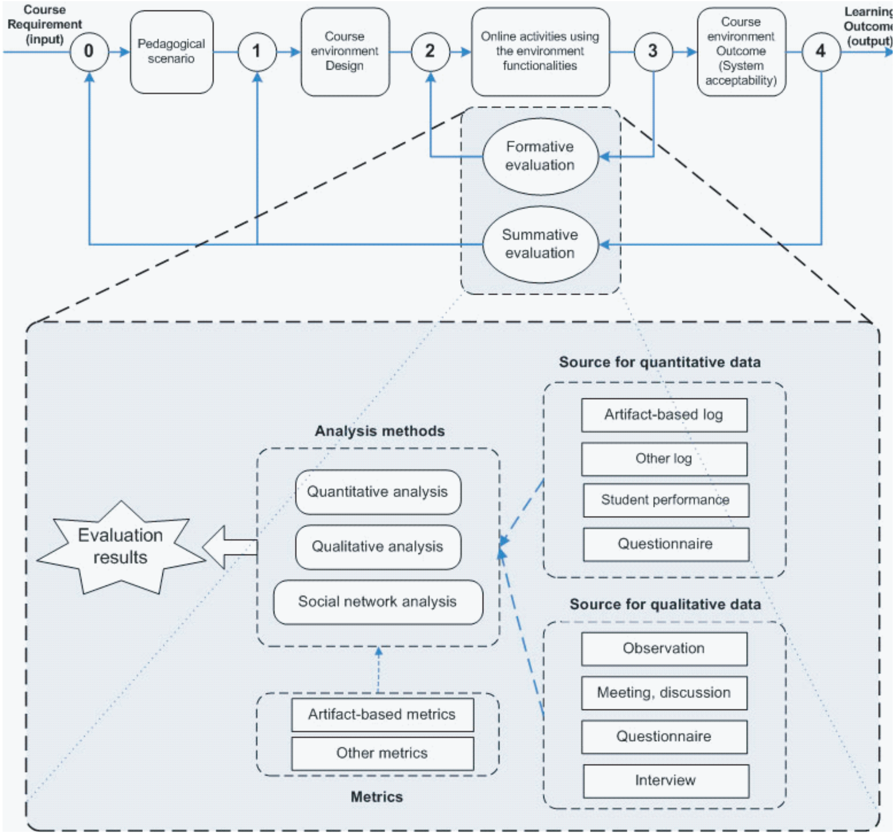
Our model includes five instrumentation nodes (see Figure 2). Each one represents a phase in the process of learning using the online environment. The outcomes are differentiated into levels, and each of them is evaluated and validated through a feedback path. The output of the evaluation process at one node could provide feedback and influence the input of another node.

The input of the whole process is the online course requirements. From these requirements, the pedagogical scenario can be designed. It is important to integrate the design and the development process around scenarios. Scenarios have people built-in, they are specific, they are grounded in the real world, and they describe an existing or envisioned system from the perspective of participative and non-participative users, including a narration of their goals, plans, and reactions (Rosson & Carroll, 2002). At Node 1, the pedagogical objectives and the course requirements are already defined. Based on these definitions, the course environment is designed or redesigned. By redesigned, we mean that some fundamental concepts of the environment need to be modified or replaced. At Node 2, the tutors and the students' requirements are defined in greater detail. The system functionalities that facilitate the online learning activities are also specified at this node.

The evaluation is carried out at Node 3 and Node 4, for the innermost, formative evaluation loop from Node 3 to Node 2, or in other words, the formative evaluation process takes place during the course. The goal of the formative evaluation is to identify the aspects of the system that can be improved, and to provide guidance on what to change in the design. One big constraint in applying formative evaluation is that it must not disturb the students who are currently using the system. Thus, in general, only minor modifications of the system functionalities are allowed. The summative evaluation loop at Node 4 is aimed at measuring the acceptability of the system (Nielsen, 1993). The summative evaluation loop may lead to the modification of the pedagogical scenario (the loop from Node 4 to Node 0) or to the redesign of the environment (the loop from Node 4 to Node 1).

In the proposed model, all the analysis methods are fed with data coming from different sources, meeting the need for capturing different forms of interaction in an online engineering learning community. The basic instruments providing quantitative data are automatic data coming from the log, questionnaires, and the student's learning performance. In a Web-based experimentation environment like eMersion, the *artifact-based log* constitutes an interesting support reflecting the student hands-on

Figure 2. Instrumentation Feedback Model for Evaluation



activities and interactions within the online community. The concept of artifact is used to represent any kind of data that could be saved, extracted, and analyzed during hands-on activities. It can be shared and can facilitate the interaction among members of the learning community. Because of the important role of an artifact-based log, it is separated from other forms of log. The instruments providing qualitative data are observations, interviews, and discussions directly with students and TAs.

The analysis methods include quantitative and qualitative analysis, and social network analysis. Social network analysis (SNA) methods are applied to construct the social structure and to find the interaction patterns in the learning community. SNA (Wasserman & Faust, 1994) is an approach that focuses on the study of patterns of relationships between members in a community.

Evidently, the choice of the evaluation methods may be changed from one course to the other. It depends on the pedagogical scenario as well as the evaluation objectives. Basically, the evaluation analyses are carried out to estimate predefined *metrics*. We have proposed a set of candidate metrics that could be useful to measure the usability and the utility of the environment supporting the online engineering learning community. These metrics are briefly presented as follows:

- **Metrics for User Learnability** (Shneiderman, 1998)**:** To measure the time and effort students spend to be able to use the environment and the resources provided to achieve specific tasks accurately and completely.
- **Metrics for User Acceptability, Participation, and Satisfaction:** To see if students accept and participate in the new learning paradigm, and how satisfied they are.
- **Metrics for Learning Performance:** To see if there is any difference in learning performance when students carry out experimentation remotely compared with when they carry out experimentation locally.
- **Metrics for Learning Pattern:** To measure the possible patterns students prefer to follow in their online courses.
- **Metrics for Environment Comprehensiveness, Effectiveness, and Efficiency:** To measure if the environment provides all necessary information and functionalities to respond to the users' needs.
- **Metrics for Flexibility:** To measure how often students participate in flexible sessions, how they divide tasks among members in the same group, and so forth.
- **Metrics for Interaction in the Community:** To measure the interaction patterns in the online engineering learning community.
- **Metrics for Social Structure in the Community:** To measure the social relationships, the activeness, the knowledge distribution, and the mediation role of members in the online engineering learning community.

The proposed metrics are defined at a fairly high abstraction level. They can be somewhat considered as important features that need to be considered in order to evaluate an online learning environment, and more precisely speaking, a Web-based experimentation environment and the online learning community using that environment. Most of these metrics are based on the artifact analysis and calculation. Hence, they are called *artifact-based metrics*. Not all of these metrics need to be calculated. Again, the appropriate choice depends on the evaluation phase as well as on the evaluation objectives.

The following sections will be used to illustrate how this model has been applied to evaluate the automatic control laboratory course. First, we will present the course setting, and then discuss the iterative design of the eMersion environment that has been carried out for this course. Finally, the evaluation results will be presented.

The eMerson Design History

The Automatic Control Laboratory Course Setting

Traditional Automatic Control Course

The academic year at EPFL is divided into a winter and a summer semester. There is a strict separation between lectures, exercise sessions, and laboratory assignments set by the study programs and the course schedule. Every week, two hours of lectures are taught to the students enrolled, followed by one hour of in-class exercise supervised by a TA. The laboratory assignments, which can last for two or four hours depending on the degree program, are also completed under supervision of a TA.

Flexible Automatic Control Course

Flexible learning deployment implies some changes in pedagogical methods such as the structure, the presentation, and the organization of information. The pedagogical scenarios have been established and evaluated progressively from the year 2000. All laboratory assignments have been reorganized into two-hour modules. They are structured into three parts: introduction, experimentation, and examination. The introduction part is dedicated to the presentation of the learning objectives, the freedom offered by the flexible learning, and the learning tools. The experimentation part is split into three to seven hands-on modules depending on the degree programs in which the students are enrolled. The examination part is carried out as a laboratory test.

The hands-on modules are composed of two parts. The first one is dedicated to a preliminary analysis and design activity called the *prelab,* which has been introduced to ensure that students have the prior knowledge necessary to benefit from the hands-on experiment, and to motivate them to do preparatory work on their own. Students need to submit a prelab document to the TA to be granted the right to proceed further with the actual experimentation, called the *labwork.* The labwork consists of carrying out a real experiment and of validating the preliminary design of the physical device. No fixed schedule is imposed on the students; only the sequence of modules has to be followed. The laboratory test consists of performing a random module and then presenting the associated results to the tutor. The course lasts for 14 weeks.

The students enrolled in the course have the possibility of following different learning modalities. The modalities vary according to the presence of a TA, and according to the students' location. When group members work together in the presence of a TA, they are in f2f condition. Students can also work in flexible sessions and remotely access the physical laboratory devices and/or computer simulation tools. In whichever learning modality, the students use the same Web-based experimentation environment, the eMersion environment.

The evolution history of the eMersion environment can be divided into four major periods, which started with the 2000 winter semester. These periods will be presented in the following sections.

The eMersion Evolution History

First Period, 2000 Winter Semester: Observation and Analysis

We proceeded with a classical f2f setting during the first year of the project. The students had regular f2f sessions with two TAs in a laboratory room. The laboratory workbenches were equipped with either an electrical drive or a thermal process trainer connected to a Macintosh computer through an analog/digital converter board. Several software applications were available on the computer: LabVIEW for controlling the connected device and acquiring sample data points, and SysQuake (http://www.calerga.com), which executes Matlab-compatible scripts for analysis and design.

The experts in education science observed a total of six hands-on sessions. Two hands-on sessions were slightly modified to conduct a controlled experiment for understanding the effect of distance in getting the TAs' help. For that purpose, the TAs were not present in the laboratory room, but they were accessible by telephone.

The observations have shown a cognitive overload for the students to master at the same time as several user interfaces, mathematical analysis and design concepts, and the experiment itself. The students' working method was to save data produced by the LabVIEW application and/or snapshots of mathematical plots to local files that they could take home on a floppy disk and/or to print their results. The sessions with simulated distance showed that students did not use the telephone and preferred to get assistance from their co-located colleagues. They exchanged data using floppy disks and printed documents.

Second Period, 2001 Winter and 2002 Summer Semesters: The eMersion Version "Niceberg"

The main challenge of the second year was to experiment with a new organization of work. That organization was based on a mix of flexible sessions with planned f2f sessions. In flexible sessions students work without the presence of TAs, who were reachable asynchronously by telephone or by e-mail.

The eMersion environment was changed from a collection of standalone applications into a Web-based experimentation environment. The LabVIEW application was replaced with a Java applet, and the SysQuake application was replaced with SysQuake Remote, which is a thin-client consisting of a Web form for editing and submitting scripts to a SysQuake engine located on a server. In addition, online manuals, online experimentation protocol, bibliography, and reference documents were also available in the environment. Figure 3 shows the environment portal (available only in French) from which students can perform experiments and can use the different facilities provided.

During these semesters we introduced two preliminary versions of shared workspaces for students working online. The first one called *Niceberg* was based on a Web-based content management system. The second one called the *Lab Journal* was a Web-based shared workspace that provided various editing functionalities. Niceberg integrated a desktop with a forum, a space for accessing the submitted laboratory reports, and various facilities for supporting students working online. The TA had access to the laboratory reports submitted and could annotate these reports with structured notes. The Lab Journal provided several workspaces for structured text fragments (in forms of XML fragments) imported from other documents, for manual notes, images, and electronic messages (see Figure 4). All these documents could be combined together to form a report. Both Niceberg and the Lab Journal

Figure 3. The eMersion portal of a group

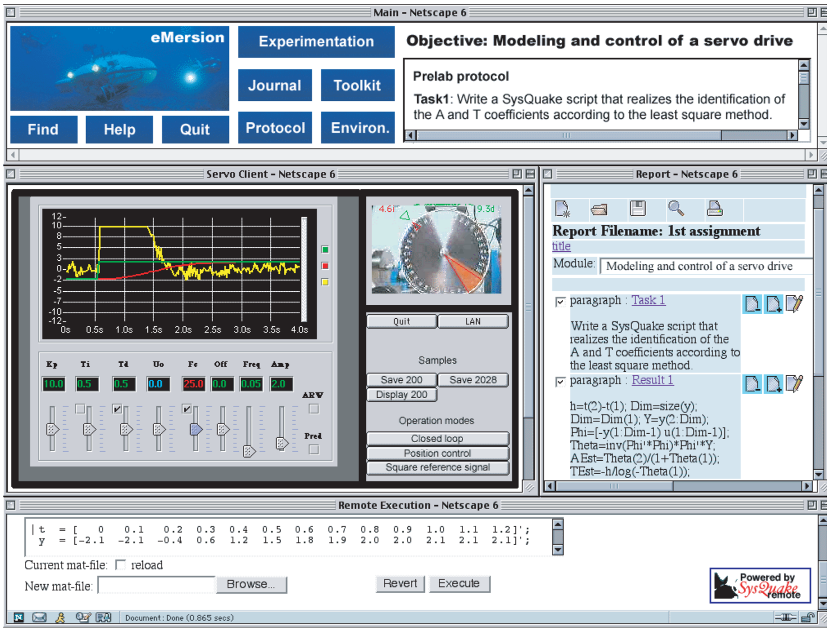
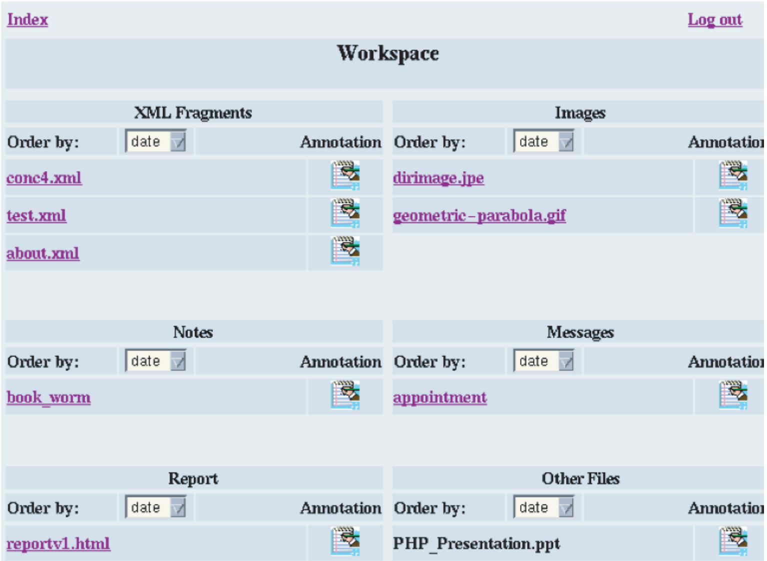


Figure 4. The Lab Journal user interface



had functionalities that allowed students to submit their reports to the TA. The access to these journals was based on password identification, but everybody could see the files in other students' journals except for those that were marked as hidden by the owners.

In fact, the Lab Journal has played the role of an electronic laboratory journal for *each* group. Laboratory journals take a privileged place in engineering education (McCormack, Morrow, Bar, Burns, & Rasmussen, 1991; Myers et al., 1991). They serve as chronological repositories for experimentation resources, planning, and realization. Laboratory journals, as a special kind of document archive, are used extensively by students in the execution of their own work and to share information with others. The activity history, the details, the results of a series of experiments, and the knowledge developed can be captured in a laboratory journal and then be reused in the same or in another session by the same or by another student. The metaphor of laboratory journals can acquire the collaboration support property of paper and paper-like instruments within a community, which has been demonstrated through many empirical studies (e.g., Schmidt & Bannon, 1992; Sellen & Harper, 2002). To summarize, an electronic laboratory journal that combines the peculiarities of a paper laboratory journal with the features of database systems and Web access is an appropriate instrument for sustaining collaboration and interaction in a Web-based experimentation environment.

The observations and the focus groups gave rise to a lot of criticism on the environment. The forum that had not been used in the 2001 winter was removed for the summer semester and replaced with a messaging system embedded in the students' workspaces. However this messaging system was also not used; students preferred e-mail as a means for communication within the community. In both prototypes the structured editing functionalities were not used as they were complicated, and in addition, students preferred to create reports with a real-text editor such as MS Word. For data collection, students would cut and paste information from the Experimentation applet's output console to a text editor and save it to a local file. In fact, the students used the journals only for submitting reports to the TA. As a result, the main goal of the journal, which was for collecting data and for supporting data sharing and exchange among students in the community, was not fulfilled at all.

We attributed the failure of the journals to a wrong choice of functionalities and to a poor design of the user interface. First, the structured notes editing functions were not appropriate. Second, it was too difficult and required many extra steps to import data into the journals from the other components such as the Experimentation applet. The difficulty of importing data into the journals and the flexible context were the source of the discontinuity of interaction (Nguyen-Ngoc, Rekik, & Gillet, 2005b), which clearly prevented the collaboration and interaction in the online engineering learning community, and also complicated the student hands-on tasks.

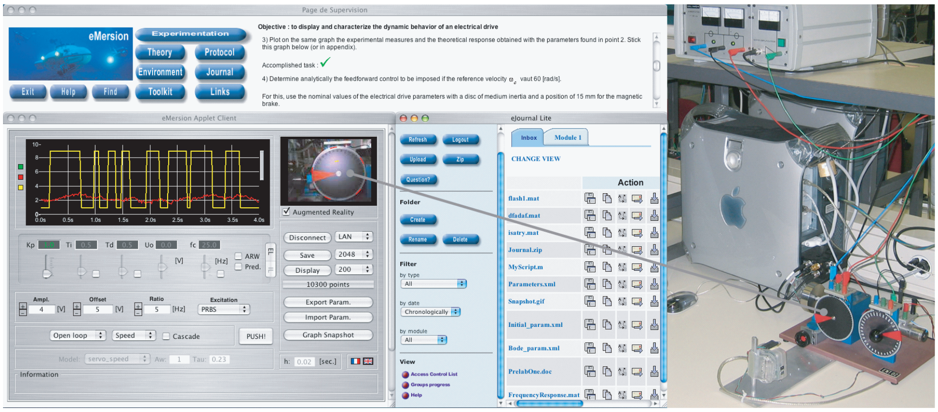
Third Period, 2002 Winter to 2004 Summer Semester: eMersion 1

The lessons learned from the first two periods led us to redesign the eMersion environment. The eMersion 1 environment included three main components: the Experimentation Console for experimentation activities; the Lab Journal, which was renamed *eJournal*, as a collaboration space; and the Toolkit Console, which was the SysQuake Remote component for mathematical analysis and design. In the Experimentation Console, the equipment was visualized in real time using a Web cam. The image quality was improved using virtual reality techniques that gave students more feeling of *reality*. Students could choose between different modes of connection such as LAN or ADSL. Using the eJournal, students could import/export a set of parameters, as well as save the experimentation results and snapshots displayed on the Experimentation Console. The experimentation results stored in the eJournal could then be processed using SysQuake Remote. This point was quite important since it facilitated the continuity of interaction within the community while carrying out the experiments (Nguyen-Ngoc et al., 2005b).

The interface of the eMersion 1 environment is illustrated in Figure 5.

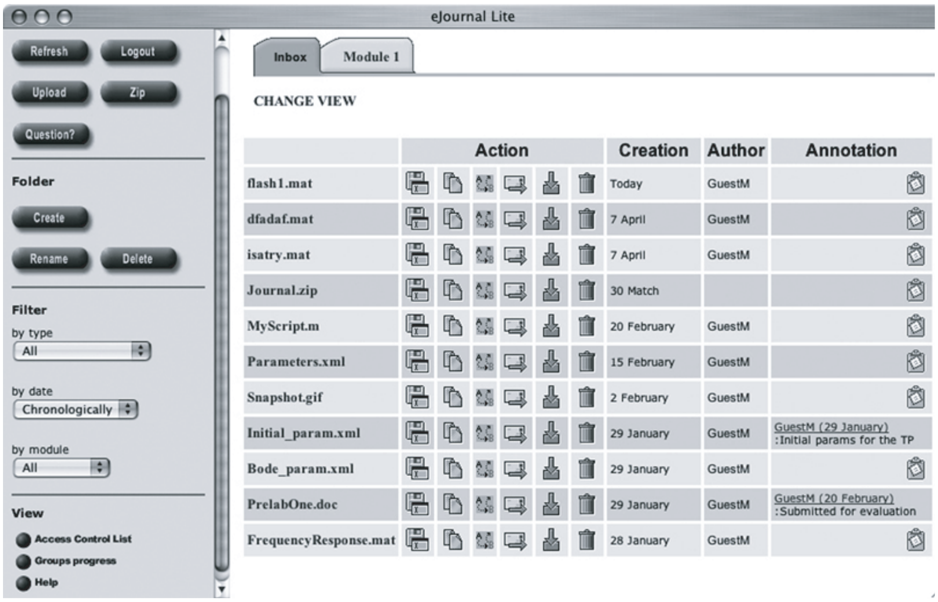
The eJournal was completely redesigned. All complex structured text editing and asynchronous messaging functionalities were removed. Its role of supporting interaction and collaboration in the online engineering learning community was stressed. The eJournal main space looked like the mailbox of an e-mail client, except that it did not contain e-mail but rich-type documents (see Figure 6), namely *fragments*.

Figure 5. The eMersion 1 environment as used to remotely control an electrical drive



In fact the concept of fragments also plays the role of artifacts as presented in the instrumentation Feedback model for evaluation. Any fragment was typed, representing different kinds of data. The fragments with different types were handled differently. Tags could be assigned to fragments when they were created in order to ease their processing and sharing. A list of tags corresponding to the assigned tasks was automatically generated from the experimentation protocol.

Figure 6. The eJournal interface in the eMersion 1 environment



Using the eJournal, the members of the online engineering learning community were provided with many different ways to collaborate with one another. Students could submit their fragments to the TA. The fragments could be annotated. In the 2002 winter semester, two different annotation systems were provided: one was Wiki based, which allowed students freely to create and edit Web page content linked to the fragment, and another was based on a simple HTML form. Students could directly send fragments with associated annotations, or send questions with attached fragments to other groups or to TA via an integrated e-mail system. This mechanism was used for prelab submission, and it could be used to get contextualized support. Students could also copy/move fragments from one eJournal to another. The fragment was at the same time an instrument and a result of the interaction and collaboration process. As an example, the experimental results of a student are saved in his eJournal when he has finished his assignment, and shared with his group colleagues for further processing in the next assignment.

The eJournal enabled many services that generate awareness information. Besides the availability awareness such as the user presence and the user location, many other kinds of group awareness based on the fragment activities analysis and calculation, called *fragment-based awareness,* were also provided in an external page. Such awareness provided information about group activities, group progress, and the social structure of the community (Nguyen-Ngoc, Gillet, & Sire, 2004b).

Fourth Period, 2005 Summer Semester: eMersion 2

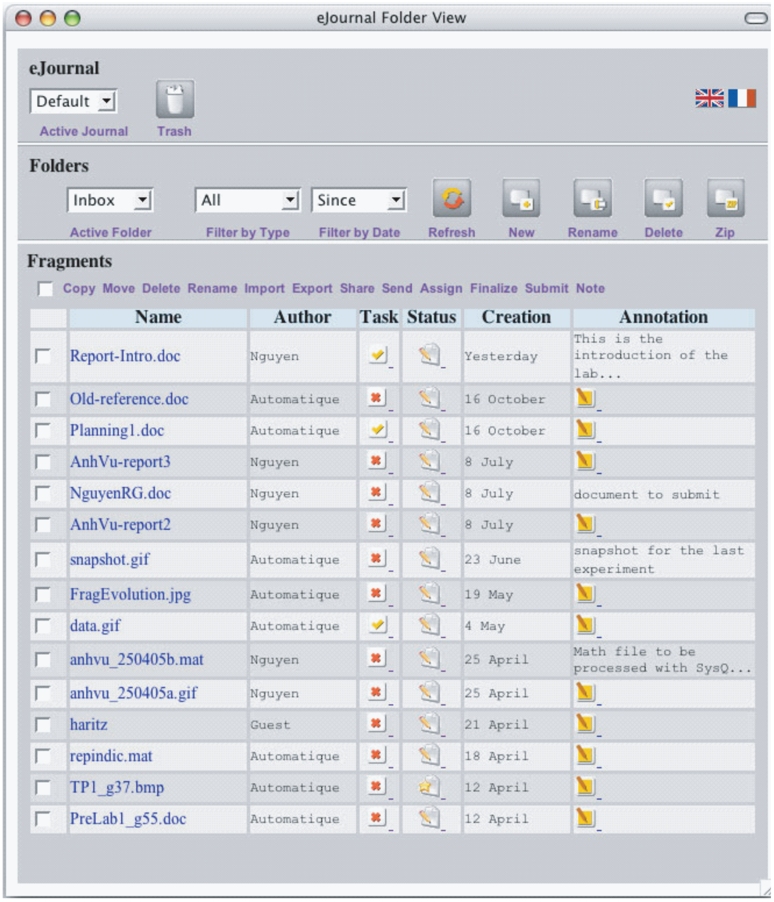
The eMersion 1 environment almost fulfilled all the designers' and the students' expectations. However, the incremental adaptations carried out during the course of its utilization made the code not as clean as it should have been. In addition, partner institutions mentioned their interest for using the environment for their own courses. Hence, it was decided to completely rewrite the code to make it more modular for further adaptations and for release under an open source scheme. The functionalities provided by the environment were regrouped as services, and the possibility of integrating new tools supporting the online community as plug-ins was implemented.

The resulting eMersion 2 also better integrates awareness features. Relevant information for the group and the class progresses are displayed in real time. Hence, it better supports students' self-motivation and autonomy development while using the online environment. The experimentation protocol was also redefined so that each task requires a *deliverable,* which is what the students are supposed to achieve after finishing a task. Basically, the student needs to respond and/or submit a deliverable in order to pass to the next task. Different kinds of deliverables could be defined. However, for this version a deliverable can only be a fragment. Depending on the experimentation modules, the deliverables for a task could be mandatory or elective.

This means that, for some tasks, the students just work for themselves. In such a case they can simply *finalize* the current task by *tagging* the fragment in an appropriate way. For this purpose, a status flag has also been added in the eJournal (which is another form of awareness). When a fragment is finalized, the flag is changed and the progress indicators are updated. When a fragment is submitted, the flag is changed, the progress indicators are updated, and the fragment is sent to the TA.

Figure 7 illustrates this new user interface of the eJournal. The two visible flags enable one to change the language of the GUI *on-the-fly*.

Figure 7. The eJournal interface in eMersion 2



Evaluation of the Automatic Control Course

Evaluation Instruments

This section presents the results of a comparative evaluation study carried out from the 2002 winter to the 2005 summer semesters. The evaluation took place in an iterative process through the different loops presented with the purpose of studying the participation, learning performance, flexibility, collaboration, and social structure aspects of an online engineering learning community. Another objective was to improve the user interface design.

During the course, the developer and the evaluator were present in the laboratory with TA and students (f2f modality). By observing the behavior of the students and the TA, and by talking with them whenever they faced problems in using the environment, the evaluator could find the potential bugs of the system as well as different minor aspects of the system that could be improved. The log data also helped to facilitate this formative evaluation process. This evaluation loop (from Node 3 to Node 2 in the Evaluation Model) iterated during the whole semester.

At the end of the semester, questionnaires were distributed to the students. Our questionnaires were based on the IBM CSUQ Questionnaire (Lewis, 1993) with some extensions (Nguyen-Ngoc, Gillet, & Sire, 2004a). The questionnaires were used to measure the metrics for user acceptability, participation, and satisfaction.

The fragment-based log was also analyzed. Fragments that originated from components of the Web-based environment and which were directly imported to the eJournal were called *intra-fragments.* Fragments that were uploaded from a local user's computer were called *extra-fragments.* These were created using external applications. Fragments that were created during f2f sessions were called *f2f-fragments,* while those created during flexible learning modalities were called *flexible-fragments.* The intra-fragments helped to observe the amount of student work that took place within the Web-based environment. This measure reflected the metrics of environment comprehensiveness, effectiveness, and efficiency. The flexible-fragments measure was linked to the importance of f2f learning modalities compared with flexible learning modalities—that is, the metrics for flexibility.

The volunteer students were interviewed. The tutor also organized a meeting in which all TAs of the course could express their ideas and their comments about the environment.

One should bear in mind that the result of the summative evaluation loop could cause major modifications and improvements of the environment for the following semesters. For each evaluation loop, different analysis methods were carried out.

The next section shows some of the results from the evaluation process carried out in the automatic control laboratory courses during these five semesters at the EPFL.

Evaluation Population

- In the 2002 winter semester, 30 students enrolled in the fourth year of the mechanical engineering degree program participated in the course. For the sake of simplicity, this sample was called Group Winter 2002.
- In the 2003 summer semester, 96 students enrolled in the third year of the micro-engineering degree program participated in the course. This represented the Group Summer 2003.
- In the 2003 winter semester, 49 students from mechanical engineering and 6 students from electrical engineering enrolled in the course. They were fourth-year students. This represented the Group Winter 2003.
- In the 2004 summer semester, 47 students from electrical engineering, 97 students from micro-engineering, and 12 students from mechanical engineering participated in the course. They were all third-year students. This represented the Group Summer 2004.
- In the 2004 winter semester, there was no course.
- In the 2005 summer semester, 39 students from electrical engineering, 69 students from micro-engineering, and 9 students from mechanical engineering participated in the course. They were all third-year students. This represented the Group Summer 2005.

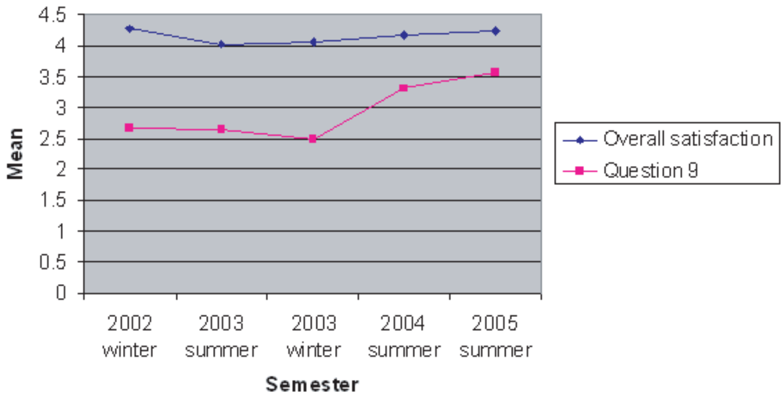
In total, during this period of five semesters, 454 students used the eMersion environment to perform hands-on activities. The evaluation results have been reported elsewhere (Fakas, Nguyen-Ngoc, & Gillet, 2005; Gillet et al., 2005; Nguyen-Ngoc et al., 2004a; Nguyen-Ngoc et al., 2005a). For the sake of simplicity, only *representative* results will be presented and discussed here.

Evaluation Results

Metrics for User Satisfaction

Among the 181 students enrolled in the course from the 2002 winter to the 2003 winter semester, 129 returned the questionnaires distributed (71.3%). In these three semesters, we encouraged students to spend some time to fill in the questionnaires and return them right after the laboratory test. In the 2004 summer semester, stu-

Figure 8. Mean of satisfaction (2002 winter to 2005 summer semesters)



dents could return the questionnaires approximately one month after the test. In fact, this was an examination period at the EPFL, and only 22 questionnaires were returned (14%). From the experience obtained from the 2004 summer semester, we also prepared an electronic version of the questionnaire accessible to all enrolled students in the 2005 summer semester. For this semester, 74 questionnaires were filled in and returned (62.2%). Figure 8 shows the mean of overall satisfaction, and that for question 9: "The system provides error messages that clearly help me to resolve problems." This question received the worst ranking and greatly reduced the general satisfaction. In fact, as implementing a help system had been quite time consuming and it was not the priority of the development team, only basic features were provided. Although this bad score was not a surprise, it was an example of the difficulty of providing an efficient help system for an online community. It is interesting to underscore that despite the fact that no help system was introduced, the 2004 and 2005 results are significantly better. This shows that a well-designed environment does not necessarily need a help system to be understood and used, while a bad one requires additional support resources.

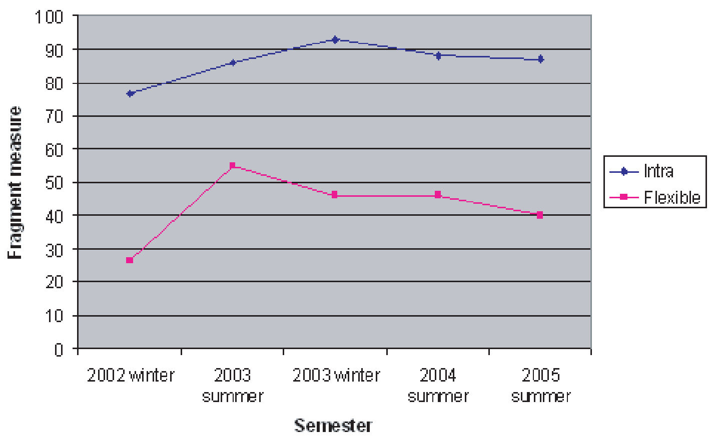
Students were also asked to provide the three most positive and three most negative aspects (in order of importance) at the end of the questionnaires concerning the usage of the environment and the environment itself. The most frequent positive comment of the system was its flexibility. The integration of all the necessary tools in one integrated environment also appears to be important in the students' positive comments. Students also enjoyed different interactive and collaborative features provided by the eJournal. They also liked the hands-on activities that reinforced their theoretical knowledge. The majority of negative comments concerned technical problems (e.g., server and client crashes) and the complexity of the interface (many windows for many tools).

Metrics for Environment Comprehensiveness, Metrics for Flexibility

We carried out the analysis of fragment logs for all five semesters. On average, about 86% of the fragments were created within the environment with the Experimentation component and the SysQuake Remote component; the other 14% were fragments created with external applications and then uploaded to the environment (e.g., MS Word documents). The number of fragments created in flexible sessions corresponded to 42.6%. The intra-fragment and flexible-fragment measures of each semester are shown in Figure 9.

One should recall that the summative evaluation loop (from Node 4 to Node 1 in the evaluation model) at the end of the semester provided feedback for the system design for the next semester. The summative evaluation results may lead to fundamental modifications of the environment. During the 2002 winter semester, we proposed two annotation mechanisms; one was based on the Wiki mechanism. However, very few students used this annotation mechanism. Thus in the next version for the 2003 summer semester, this mechanism was dropped. Since the 2003 summer semester version, the possibility of sustaining the continuity of interaction has been improved. As a consequence, the intra-fragments and the flexible-fragments have increased greatly from 76.67% and 26.29% in the 2002 winter semester, to 86% and 55% in the 2003 summer semester. Since then, the flexible-fragment ratio has slightly decreased. This might be explained by the fact that more teaching assistants were available in f2f sessions. Thus students benefited more in working directly with them in the laboratory. In addition, in 2004 and 2005, enough workbenches were available for all the students to work simultaneously. This was not the case

Figure 9. Fragment measures (2002 winter to 2005 summer semesters)



in 2002 and 2003. It was in fact a logistical constraint that was initially the main motivation for the development of the eMersion environment. Later, the pedagogical motivations became more important.

To have a clear view about these fragments, one should see the examples in Figures 10 and 11. In these figures, each column represents the number of created fragments by a micro-engineering group of the Group Summer 2004. In each column, the white part represents the intra-fragments. The black part represents the extra-fragments. Figure 11 represents the same data but from another perspective. The black part shows the fragments that were created in f2f sessions, while the white part is the number of fragments created in flexible sessions.

One should not forget that we applied more or less the same evaluation methods for the evaluation loops. However, the evaluation variables and parameters for the next loop (or next semester) may be modified depending on the result and on the requirements.

Metrics for Learning Performance

Since the 2003 summer semester, we started considering the group performance (via the grade of the group members). Analysis in the Group Winter 2003 and Group Summer 2004 showed that there was a statistically significant correlation between the number of created fragments and the group performance (obtained via the groups'

Figure 10. Intra- and extra-fragments produced by micro-engineering groups during 2004 summer semester

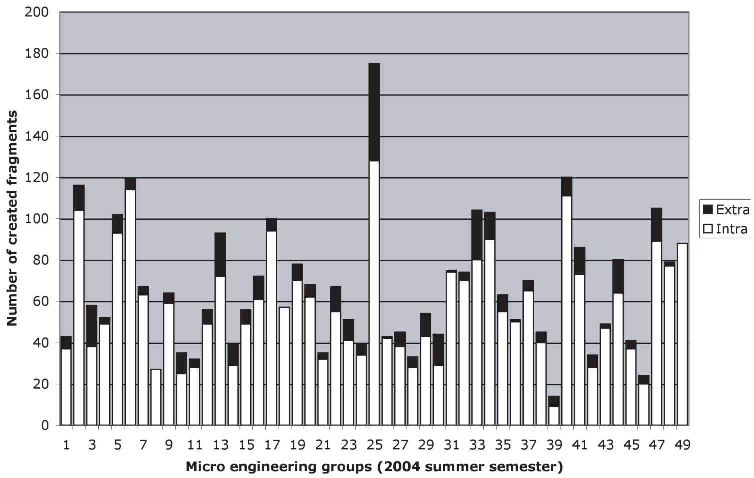
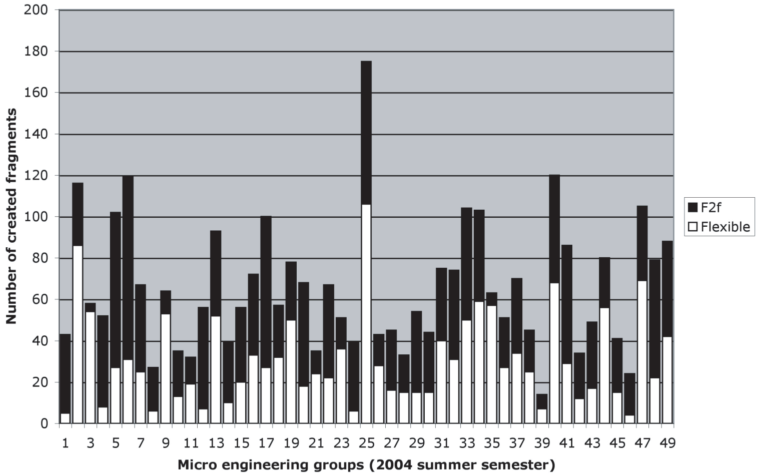


Figure 11. F2f- and flexible-fragments produced by micro-engineering groups during 2004 summer semester



grades). The Pearson product-moment coefficient correlation between these two variables was 0.522 ($p<0.01$) for the Group Winter 2003, 0.296 ($p<0.05$) for the Group Summer 2004, and 0.3 ($p<0.05$) for the Group Summer 2005. We have found no statistical correlation between these two variables in the Group Summer 2003.

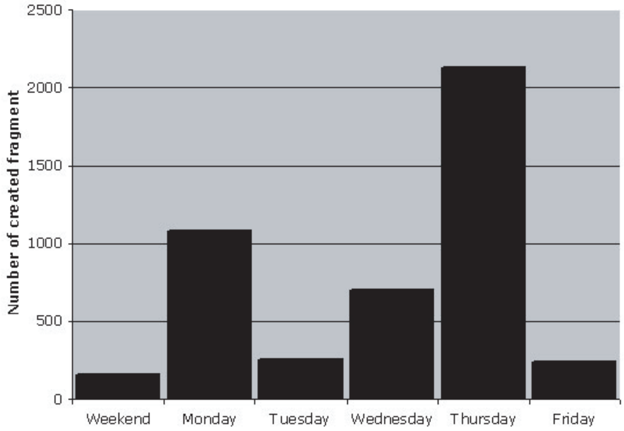
We divided all groups into two sub-groups: the first one preferred working in flexible modalities (high flexibility groups), the second one worked mostly in f2f modalities (low flexibility groups). This classification was based on the flexible-fragments of all groups. A group was classified as high flexibility if its flexible-fragments were more than or equal to 50%. Actually, for the Group Summer 2003, the grade mean of high and low flexibility groups was 5.04 over 6 (S.D.=0.58) and 5.07 (S.D.=0.6), respectively; for the Group Winter 2003, these were 5.05 (S.D.=0.69) and 5.12 (S.D.=0.56); for the Group Summer 2004, both sub-groups received the same grade mean of 4.3 (S.D.=1.05); and finally for the Group Summer 2005, these were 4.69 (S.D.= 1.1) and 4.65 (S.D.=1.12).

The results showed that there was no significant difference between the educational outcomes from students who performed the experiment remotely compared with those who preferred carrying out the experiments in f2f sessions.

Metrics for Learning Pattern

Since the 2003 summer semester, we have considered the learning pattern of students in the online engineering learning community. In the 2003 summer semester,

Figure 12. Cumulative number of fragments created each day of the week during the 2004 summer semester



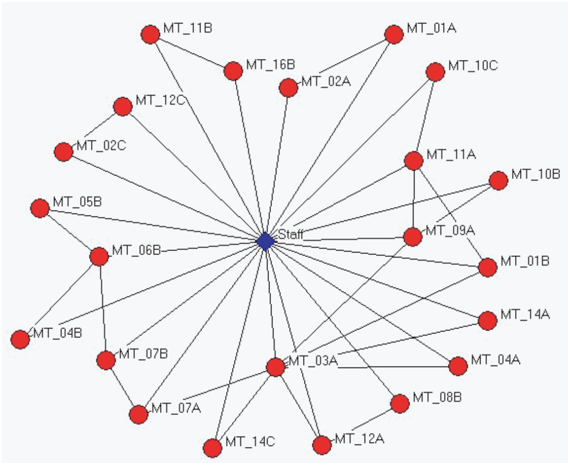
1.4% of fragments were created during weekends, and 2.5% of fragments created in the evening and at night—that is, from 6:00 p.m. to 7:00 a.m. the next day. These numbers were 6.6%-4.4% and 3.5%-17.4% in the 2003 winter and 2004 summer semesters, respectively.

We noted that students worked most actively on the days in which there were laboratory sessions. Figure 12 shows a histogram illustrating the cumulative total number of fragments created each day of the week during the 2004 summer semester. In this semester, there was one f2f session every Thursday (from 10:15 a.m. to 12:00 noon) for groups from micro-engineering degree programs, and every Monday (from 5:15-7:00 p.m.) for groups from mechanical and electrical engineering degree programs.

Metrics for Interaction and Social Structure

Last but not least, we performed different SNA methods to find the interaction patterns between different groups, as well as the social structure in the community. The SNA methods have been carried out since the 2003 summer semester. For establishing the community structure and interaction patterns, we were interested in those techniques giving information about structural properties of the network as a whole, and particularly those related to cohesion (Woodreff, 1999) such as sociogram, clique, and Freeman's centrality degree (Wasserman & Faust, 1994). These methods were applied to each semester to provide so-called *social structure awareness* for tutors and TAs (Nguyen-Ngoc et al., 2004b). As an example, Figure

Figure 13. Sociograms of the interactions found during the 2004 summer semester



13 shows a sociogram representing the social structure established in the Group Summer 2004 community.

In a sociogram, nodes (circles) represent groups, and lines represent the interaction between groups. Different shapes and colors are used to refer to some special groups. For example, the Staff group (tutors and TAs) is represented by the central diamond.

Discussion

The metrics calculated previously help to answer most of our evaluation objectives—that is, to study various aspects of an online engineering community. We find the results satisfactory concerning the "acceptability goal" as shown by the metrics for user satisfaction. However, the mean satisfaction is not much higher than the neutral scale point, thus suggesting much room for improvement.

The *participation goal* is also reached as all the groups created a significant number of fragments. As a corollary, we believe that the "participation goal" contributes to the "acceptability goal" as evidence of the use of the environment.

The metrics for environment completeness and metrics for flexibility show that the students took advantage of different learning modalities. These metrics also show that the system functionalities satisfy the needs of students while performing online hands-on activities.

SNA contributed to identifying the interaction patterns at different levels: the community, the group, and the individual. It also shows the interaction in time—that is, the interaction between students from different semesters. In fact the metrics for interaction and social structure show that staff members still play the most important role in the knowledge distribution within the community. The SNA measure can be used not only at the evaluation phase, but also during the learning process to provide awareness information to tutors and students. It gives tutors and students a general overview of active and passive groups in the learning community, as well as the structure of the community.

The statistical analysis shows that there might be correlation between the number of created fragments and the group performance. The validation procedure should be refined to confirm this assertion. We should also consider other variables that may affect the performance, such as group motivation, previous knowledge, and experience. The result from comparing the groups who preferred working in flexible modalities (high-flexibility groups) and those who worked mostly in f2f modalities (low-flexibility groups) supports the assumption that the Web-based learning environment is an *added value* for traditional engineering education (Gillet et al., 2005).

The evaluation loops also allow us to improve the user interface of the environment. This helps us know exactly what students really want in an online environment.

Conclusion

This chapter presents the iterative design and the evaluation of a Web-based experimentation environment deployed in engineering education, namely eMersion. The eMersion environment provides an excellent support for the deployment of a flexible learning paradigm in engineering curricula.

The chapter also presents the eJournal, an extended electronic laboratory journal, which is an implementation of what we called a mediation artifact or a collaboration artifact (Nguyen-Ngoc et al., 2004b, 2005b). The deployment and evaluation of the system over a long period of time have confirmed the adequacy of the chosen metaphor. It has also confirmed the important role of the laboratory journal in supporting collaborative learning activities in an online learning community.

This chapter proposes a model, namely the instrumentation feedback model for evaluation, for the assessment of online learning communities using Web-based experimentation environments. The model encourages an iterative evaluation process. The evaluation is carried out at different stages of the learning process through different evaluation loops. At each loop, different evaluation analysis methods—including qualitative and quantitative analysis, and Social Network Analysis—could be combined to provide evaluators with a maximum of data representing the different aspects of the online community. These analysis methods are fed with data coming from different sources, meeting the need for capturing different forms of interaction in the usage of a Web-based experimentation environment. The model opens up a new set of ways for evaluating online learning communities in engineering education. This model has been generalized from and validated by the experience gained from successive semesters. Although so far the model is only used for evaluating the automatic control laboratory courses and the eMersion environment at the EPFL, it is generic enough to apply to other pedagogical scenarios and other learning environments.

This chapter describes the results and analyses of the evaluation process carried out in the automatic control laboratory courses from the 2002 winter to the 2005 summer semesters at the EPFL.

Acknowledgments

The elements presented in this chapter result from various e-learning projects and activities carried out with the support of the Board of the Swiss Federal Institutes of Technology and of the European Union in the fifth and sixth framework programs (http://www.prolearn-project.org).

References

Atkan, B., Bohus, C. A., Crowl, L. A., & Shor, M. H. (1996). Distance learning applied to control engineering laboratories. *IEEE Transactions on Education, 39*(3), 320-326.

Card, S. K., Moran, T. P., & Newell, A. (1983). *The psychology of human-computer interaction.* London: Erlbaum.

Dourish, P., & Bellotti, V. (1992). Awareness and coordination in shared workspaces. *Proceedings of the Conference on Computer-Supported Cooperative Work* (CSCW), Toronto, Canada.

Fakas, G., Nguyen-Ngoc, A. V., & Gillet, D. (2005). The electronic journal: A collaborative and cooperative learning environment for Web-based experimentation. *Computer-Supported Cooperative Work (CSCW): The Journal of Collaborative Computing, 14*(3), 189-216.

Faltin, N., Böhne, A., Tuttas, J., & Wagner, B. (2002, August 18-21). Distributed team learning in an Internet-assisted laboratory. *Proceedings of the International Conference on Engineering Education,* Manchester, UK.

Frechtling, J., & Sharp, L. (1997). *User-friendly handbook for mixed method evaluation.* Directorate for Education and Human Resources, Division of Research, Evaluation, and Communication.

Gillet, D. (2003). Towards flexible learning in engineering education. In W. Aung, R. Altenkirch, T. Cermak, R. W. King, & L. M. S. Ruiz (Eds.), *Innovations 2004: World innovations in engineering education and research* (pp. 95-102). New York: iNEER in cooperation with Begell House Publishing.

Gillet, D., Geoffroy, F., Zeramdini, K., Nguyen-Ngoc, A. V., Rekik, Y., & Piguet, Y. (2003). The cockpit, an effective metaphor for Web-based experimentation in engineering education. *International Journal of Engineering Education, 19*(3), 389-397.

Gillet, D., Nguyen-Ngoc, A. V., & Rekik, Y. (2005). Collaborative Web-based experimentation in flexible engineering education. *IEEE Transactions on Education, Special Issue on Web-based Instruction, 48*(4), 696-704.

Grudin, J. (1988). Why CSCW applications fail. *Proceedings of the Conference on Computer-Supported Cooperative Work* (CSCW), Portland, OR.

Kazmer, M. M., & Haythornthwaite, C. (2005). Multiple perspectives on online learning. *ACM SIGGROUP Bulletin, 25.*

Leifer, L. (1997). Design team performance: Metrics and the impact of technology. In *Evaluating corporate training: Models and issues.* Kluwer.

Lewis, J. R. (1993). *IBM computer usability satisfaction questionnaires: Psychometric evaluation and instructions for use.* Boca Raton, FL: IBM Corporation.

McCormack, J. B., Morrow, R. K., Bar, H. F., Burns, R. J., & Rasmussen, J. L. (1991). The complementary roles of the laboratory notebooks and laboratory reports. *IEEE Transactions on Education, 34*(1).

Mosterman, P., Dorlandt, A. M., Olin Campbell, J., Burow, C., Bouw, R., Brodersen, A. J. et al. (1994). Virtual engineering laboratories: Design and experiments. *Journal of Engineering Education, 83*(3), 279-285.

Myers, J. D., Fox-Dobbs, C., Laird, J., Le, D., Reich, D., & Curtz, T. (1991). Electronic laboratory notebooks for collaborative research. *Proceedings of the Proceedings of the 5th Workshop on Enabling Technologies: Infrastructure for Collaborative Enterprises,* Stanford.

Nguyen-Ngoc, A. V., Gillet, D., & Sire, S. (2004a). Evaluation of a Web-based learning environment for hands-on experimentation. In W. Aung, R. Altenkirch, T. Cermak, R. W. King, & L. M. S. Ruiz (Eds.), *Innovations 2004: World innovations in engineering education and research* (pp. 303-315). New York: iNEER in cooperation with Begell House Publishing.

Nguyen-Ngoc, A. V., Gillet, D., & Sire, S. (2004b, June 21-26). Sustaining collaboration within a learning community in flexible engineering education. *Proceedings of the World Conference on Educational Multimedia, Hypermedia & Telecommunications* (ED-MEDIA), Lugano, Switzerland.

Nguyen-Ngoc, A. V., Rekik, Y., & Gillet, D. (2005a). Formal assessment of a Web-based learning environment. *Proceedings of the iNEER Conference on Engineering Education and Research,* Tainan, Taiwan.

Nguyen-Ngoc, A. V., Rekik, Y., & Gillet, D. (2005b). A framework for sustaining the continuity of interaction in a Web-based learning environment for engineering education. *Proceedings of the World Conference on Educational Multimedia, Hypermedia & Telecommunications* (ED-MEDIA), Montreal, Canada.

Nielsen, J. (1993). *Usability engineering.* San Francisco: Morgan Kaufmann.

Ogot, M., Elliott, G., & Glumac, N. (2003). An assessment of in-person and remotely operated laboratory. *International Journal of Engineering Education,* 57-64.

Roppel, A. T., Hung, Y. J., Wentworth, W. S., & Hodel, A. S. (2000). An interdisciplinary laboratory sequence in electrical and computer engineering: Curriculum design and assessment results. *IEEE Transactions on Education, 43*(2), 143-152.

Rosson, M. B., & Carroll, J. M. (2002). *Usability engineering scenario-based development of human-computer interaction.* San Francisco: Morgan Kaufmann.

Schmid, C. (1998). The virtual control lab VCLAB for education on the Web. *Proceedings of the American Control Conference,* Philadelphia, PA.

Schmidt, K., & Bannon, L. (1992). Taking CSCW seriously: Supporting articulation work. *Computer-Supported Cooperative Work (CSCW): The Journal of Collaborative Computing, 1*(1), 7-40.

Sellen, A. G., & Harper, R. H. R. (2002). *The myth of the paperless office.* Boston: MIT Press.

Shneiderman, B. (1998). *Designing the user interface: Strategies for effective human-computer interaction.* Reading, MA: Addition-Wesley.

Tzafestas, C. S., Alifragis, M., & Palaiologou, N. (2005). Development and evaluation of a virtual and remote laboratory in robotics. In *Innovations 2005: World innovations in engineering education and research* (pp. 255-270). New York: iNEER in cooperation with Begell House Publishers.

Wasserman, S., & Faust, K. (1994). *Social Network Analysis: Methods and applications.* Cambridge: Cambridge University Press.

Woodreff, E. (1999). Concerning the cohesive nature of CSCL communities. *Proceedings of the Conference on Computer-Supported Collaborative Learning* (CSCL).