Comparing a fully online course to a blended one: the case of compilers

Salvador Sánchez Alonso [(1)], Daniel Rodriguez García [(1)], Robert Clarisó Viladrosa [(2)]

[(1)] Information Engineering research unit, Universidad de Alcala
Ctra. de Barcelona, Km. 33.6, 28817 Alcalá de Henares, Spain
{salvador.sanchez,daniel.rodriguezg}@uah.es

[(2)] IT, Multimedia and Telecommunications Department, Universitat Oberta de Catalunya
Rambla del Poblenou 156, 08018 Barcelona, Spain
rclariso@uoc.edu

Abstract

The compilers course is part of the degree in Computer Science that the Open University of Catalonia (UOC) –an online university located in Barcelona– offers in fully online mode. In the University of Alcala (UAH), located in Madrid, the same course is given in a more traditional, face-to-face format, blended with a virtual environment as a help for assignments. This paper reports on the study carried out by the teachers of these courses during the 2008 academic year. The study aimed at determining whether students at UOC, who take only online courses perform (and consequently learn) as much as students taking identical course in a blended environment (UAH). Preliminary results comparing their marks in the first assessment of the course indicate that the amount of learning is similar in both settings.

1. Introduction

The Open University of Catalonia (UOC) is a virtual university aimed at complementing the Catalan university system "by making university studies available to citizens who cannot enroll for a face-to-face university degree" (Borges, 1998). As teachers in a traditional (so to speak) university, the University of Alcala, we were curious about how different our UOC students would perform compared to our UAH students. After reading on a similar experience by Anstine & Skidmore (2005), we were interested in establishing whether our UOC students, taking a fully online course, were learning as much as our regular students at UAH, even if the latter had unlimited access to teachers, as they could arrange an appointment anytime to meet with a teacher. Our personal perception was that UOC students, more used to work autonomously, would perform better as the assessment task implied an important process of previous documentation that traditional students are not completely familiar with. So, we hypothesized that, contrary to what most people could expect, our UOC students would perform better and consequently would learn more, or at least as much as, those students attending to regular face-to-face classes.

This paper is structured as follows. Section 2 explains the irrespective settings of each of the universities participating in the study, namely UOC and UAH, as well as the motivation for this research. Section 3 details the design of the research and how the study was conducted, showing the most interesting results. Finally, some conclusions and outlook for more comprehensive studies are provided in section 4.

2. Background

2.1. The Open University of Catalonia

The Open University of Catalonia (UOC) is a virtual university with more than 54.000 students all around the world (UOC, 2008). UOC has offered university degrees in Computer Science since 1997 and currently more than 6.000 students are enrolled in Computer Science courses.

Teaching at the UOC is performed through a *virtual campus* (Borges, 1998), an online platform

which aggregates many services oriented towards students and academics, such as: virtual classrooms where they can access course self-study materials, submit assignments and communicate with their peers and the course instructors; virtual laboratories for exercising practical skills (Prieto, 2008); and a virtual library that provides access to recommended bibliography, databases and journals. Academics and students communicate asynchronously using public forums or personal mailboxes.

Three academic profiles participate in the UOC learning methodology: the *counselor*, the *subject tutor* and the *lecturer*. Counselors provide orientation throughout the degree, guiding the selection of courses and providing information regarding professional openings, etc. Subject tutors answer questions about a specific course, design the assignments used for the assessment of the course and grade the students. Courses with many students are split into several virtual classrooms, each with a different tutor. Finally, lecturers act as coordinators of the team of tutors for a given course and supervise the course design. While lecturers are full-time academics at UOC, subject tutors are professional or academics with full time commitment at another institution, selected according to their expertise in the area. This particularity, which allowed having the same teachers in courses given by different universities, was the germ of this study.

As part of its degree in Computer Science, the UOC offers two 6-month courses in compiler design and construction, which are compulsory to achieve the Spanish 5-year degree in Computer Science (*Ingeniería en Informática*). These subjects have many prerequisite courses (programming, data structure, automata theory and computer architecture) and therefore they are among the last courses taken by computer science students, typically in the last year of their degree. Table 1 describes the contents of these courses. The course design has not yet been adapted to the European Credit Transfer System, but the workload of each course is estimated at 6 ECTS credits.

Table 1. Compiler courses in the UOC Computer Science curriculum.

Compilers 1	**Compilers 2**
• Introduction to compilers: history, construction process, compiler architecture. . . • Front-end: lexical, syntactical and semantic analysis, symbol tables, error recovery • Back-end: code generation and optimization • Construction of a compiler front-end using JLex and CUP	• Construction of interpreters • Mark-up languages (XML) • Advanced code generation: procedure calls (pass-by-value vs pass-by-reference), object-oriented programming languages, . . . • Construction of abstract stack machine interpreters and use of XML technologies (DTDs, XSLT, Java APIs for XML)

From the point of view of the students, the workload and difficulty of compiler courses is considered “above average” within the Computer Science curriculum. Some points typically raised to argue for this difficulty are: the use of many concepts and techniques from several prerequisite subjects, highly theoretical contents (e.g. parsing theory), long programming assignments and the introduction of new software tools –Jlex for the construction of lexical analyzers and CUP for the construction of syntactical analyzers– and languages, among others.

2.2. The University of Alcala

On the other hand, the University of Alcala (UAH) is a public university located in Alcala de Henares, a Unesco world heritage town located 35 km from Madrid, the capital of Spain. The city,

founded in the early 16th century, was awarded the World Heritage status in 1988 because it was the world's first planned university city.

The UAH offers 3 different degrees in computer science, the most advanced of which has a duration of 4 years –equivalent to 240 ECTS– and allows the student to obtain the Spanish degree in Computer Science (*Ingeniería en Informática*). As part of its syllabus, this degree includes a compiler course as a mandatory subject whose main topics are detailed in table 2.

Table 2. Compiler course in the UAH Computer Science curriculum.

Compilers
• Introduction to compilers: history, construction process, compiler architecture. . . • Front-end: lexical, syntactical and semantic analysis, symbol tables, error recovery • Back-end: code generation and optimization • Construction of a compiler front-end using JLex and CUP

The course follows a blended learning methodology, where face-to-face lectures are combined to virtual learning and interaction through a Moodle parallel course. Lectures, which are approximately 50% of the total time of the course, address the theoretical foundations of the course, and are taught in a traditional classroom. The rest of the time is dedicated to problem solving cases and assignments. During the practical part of the course, lecturers provide some initial guidance on the tools that students must use as well as a few practical examples on simple applications during 3 to 4 face-to-face sessions. During these sessions, students raise questions about tools use, installation, technical problems in general and the like. Once they master this part, students face three individual assignments during the rest of the course. During their assignment elaboration, students have access to a virtual learning environment — specifically set up to enhance both student to student and student to lecturer communication. Students use the virtual learning environment to raise questions on the application of theory to their assignments, technical problems, documentation and other communication activities both with their peers and with lecturers.

2.3. Motivation

Two professors from UAH's Compilers course also work as subject tutors in UOC's Compilers 1 course. This fact provided the opportunity to compare two learning methodologies, fully online learning versus blended learning, in the context of a Compilers course. To this end, we studied the success rate of students in the assessment activities of both courses.

An issue which lends credibility to the results is the fact that students in both universities were evaluated using exactly the same assessment activities, and that the grading criteria were also equal (given that all assignments where graded by the same set of professors).

3. Design of the research

The comparison was performed during the first semester of the academic year 2008-2009. In the UOC course "Compilers 1", there were three virtual classrooms, two in the Catalan language with 53 and 57 students respectively, and another in Spanish with 27 students, but only the class in Spanish was used for this preliminary study. In the UAH course, there was only a virtual classroom for the 53 students, even though the students were split into two different groups for the lectures part only for their convenience in terms of space and teaching shifts.

There are important differences among the profiles of students at the UOC and the UAH. The UOC is oriented towards life-long learning and students which cannot attend a face-to-face university due to personal or work commitments. Therefore, the students from UOC tend to be older on average (UOC, 2008), have work experience and study part-time while working full-time and/or caring for their families. It is also common that many students from UOC either have a degree on another area, or started a degree which they interrupted for some reason. On the other hand, students at the UAH tend to be younger and studying their first degree on a full-time basis, possibly (but not very often) simultaneously with a part-time job. This student profile also has a greater urgency to complete their degree quickly in order to enter the full-time job market. These differences in student profiles are significant in the study of the success rate. For example, as time is a very important resource for UOC's students, students only invest time in solving an assignment if consider they have a reasonable chance of success (rather than failing, they simply will not submit the assignment). Meanwhile, UAH's students have incentives to work on assignments even if they know their solution will not be optimal (getting feedback from course instructors and avoid "wasting a semester").

3.1. Teaching and assessment mechanisms

In both universities, UOC and UAH, the assessment of the compilers course is carried out in a similar manner.

UOC offers theirs students two alternative paths to pass the compilers course:

1. Three continuous assessment activities, followed by a validation test.
2. A final exam at the end of the course.

Both paths assess the theoretical concepts of the course and the practical issues as it is shown in table 3 –see how practical aspects play a very important role in these activities and therefore have an important weight in the final mark–. According to experience, continuous assessment improves the assimilation and understanding of concepts, especially in a distance learning environment. Therefore, continuous assessment path is the one recommended by OUC course instructors and the one selected by most students.

Table 3. Organization of the assessment activities at UOC

Activity 1	**Theory (30%)**	• Structure of a compiler (front-end/back-end, phases, symbol table, . . .) • Compilation vs interpretation • Compiler construction (bootstrapping, cross-compilers, . . .) • Lexical analysis phase
	Practice (70%)	• Definition of regular expressions • Construction of lexical analyzers using JLex and Java
Activity 2	**Theory (30%)**	• Syntactic analysis phase • Bottom-up (LR, LALR) and top-down (LL) parsing • Syntax error recovery
	Practice (70%)	• Integration of JLex and CUP • Construction of a syntactic analyzer (parser) using CUP and Java
Activity 3	**Theory (30%)**	• Semantic analysis phase (attributed grammars, type checking, . . .)

		• Code generation phase (intermediate code, . . .) • Code optimization phase (goals, algorithms, . . .)
	Practice (70%)	• Construction of semantic analyzers using CUP and Java

The validation test consists of a set of questions that focus on specific aspects of covered by the continuous assessment activities performed throughout the course. This test has two goals: to certify the *identity* of the student (i.e. the student has solved the assessment activities herself) and to *validate* the learning process (i.e. the student has understood the concepts and techniques used in the assessment activities).

The final grade is computed as the average grade of all continuous assessment assignments, taking into account that in order to pass the student has to pass *all* continuous assessment assignments and the validation test.

Regarding UAH, the assessment is divided into 2 separate parts:

1. A final exam at the end of the course for assessing theoretical concepts and knowledge.
2. Three assessment activities, followed by a personal interview for validation purposes. These assessments address the practical part of the course only.

The final grade is computed here as the average grade of the final exam (60% weight in the final mark) and all continuous assessment assignments (40% weight in the final mark). As in the UOC course, the student must pass *all* continuous assignments to pass this part of the course, but no student assignment is considered valid if the face-to-face validation is failed, because this test has the same goals previously described for the UOC: to certify that the student has solved the assessment activities herself and to validate the learning process. Table 4 shows the organization of the three assessment activities at UAH.

Table 4. Organization of the assessment activities at UAH

Activity 1	• Definition of regular expressions • Construction of lexical analyzers using JLex and Java
Activity 2	• Integration of JLex and CUP • Construction of a syntactic analyzer (parser) using CUP and Java
Activity 3	• Construction of semantic analyzers using CUP and Java

3.2. How the assignments were set up and assessed

In this paper, we will analyze the results achieved in the first continuous assessment activity in the semester (from the UOC perspective) and the first assignment (from the UAH perspective) as they have exactly the same tutors/lecturers, the same problem statements and a similar period of time to submit their solutions.

Assignments consisted on a set of exercises –both theoretical and practical in the case of the UOC and only practical in the case of UAH– that students must solve on their own. Once the problem statements were published, students had a fixed time (which is usually two to three weeks) to submit their solutions. Each assignment typically requires 4-5 hours to be completed, in addition to the time devoted to revising the materials (course materials, recommended bibliography and collections of solved problems).

To have two comparable groups of students' results, only the practical part of the UOC assessment activity was used for the study, so the theory part (30%, see table 3) was removed from it. Thus, the UAH and the UOC assignment activities were exactly the same, which allowed using the same criteria to correct both the students' results of the UOC and the UAH groups.

In particular, the assessments were corrected according to a set of criteria classified into 3 groups: essential (E), positive (+) and negative (-). We considered essential those requirements that all the student programs must accomplish as a precondition to pass the assessment, while positive and negative criteria were lists of issues that could have an influence on the final mark. More detail on the criteria followed by the courses' tutors is given in table 5.

Table 5. Correction criteria

Type of criteria	Description	Comments
E	The archive.lex must be processed by JLex without errors	
E	The resulting Java file must compile without errors	Some minor errors were not considered, e.g. using a class name for the scanner other than the one requested
E	The scanner created by JLex cannot end unexpectedly (e.g. due to a not handled exception or the like)	
+	The specification file is legible and include comments	
+	Lexical patterns are simplified through the use of macros	Usually linked to the previous criteria
+	Java code is well structured	The specification file has a method for each functionality instead of having the java code repeated several times
-	The specification file does not include the directive %class to change the default scanner class name	One mark was cut out of 10
-	The specification file includes too many states, most of them unnecessary	One mark was cut out of 10
-	Bad Java code in the specification file	0.5 mark was cut out of 10
-	The student makes use of inefficient or inadequate data structures	0.5 mark was cut out of 10

During all the period of assessment, and up to the deadline for submission, the course instructors were available to answer questions regarding the materials or the comprehension of the assessment wording. In the case of UOC, this help was of course available through the virtual environment, whereas in the UAH the lecturers allocated 2 hours per week to face-to-face help sessions in the lab, in parallel to the help they provided through the virtual environment.

3.3. Analysis of the results

Our initial thought was that students in blended learning would achieve better results than students in an exclusively on-line environment, as they have more direct access to lecturers both through the online environment and in the face-to-face sessions. However this was discarded by just comparing basic descriptive statistics as shown in table 6. Also, and even though the marks do generally follow a normal distribution, this is also confirmed as both standardized skewness and kurtosis are inside the range of ±2.

Table 6. Descriptive Statistics

	UAH	UOC
Count	53	27
Average	6.24717	6.95185
Variance	3.57831	6.03413
Standard deviation	1.89164	2.45645
Minimum	3.0	1.4
Maximum	9.5	10.0
Range	6.5	8.6
Standard skewness	-0.48213	1.48045
Standard kurtosis	1.29985	0.837939

This is also corroborated by the box-and-whisker plot of Figure 1. As it can be observed, although there is more variance for the on-line learning groups, the mean of the marks is higher.

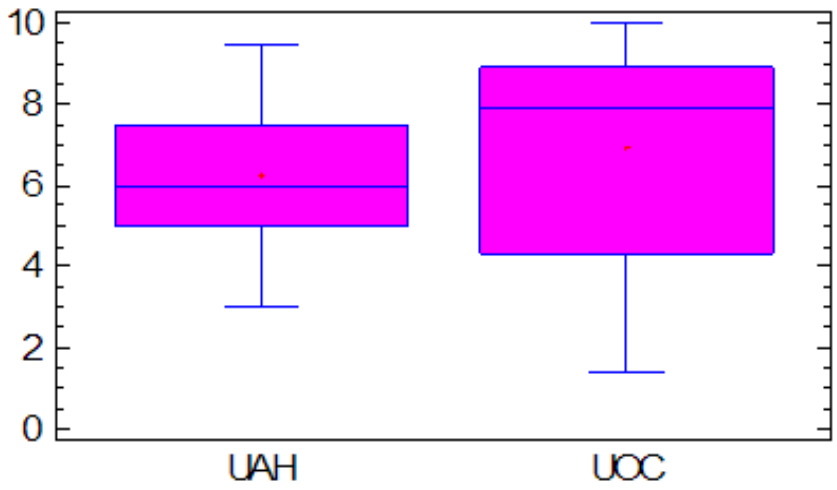

Figure 1. Box-and-whisker plots for marks

To check whether the difference between the means was statistically significant, we run a t-test with the following null hypothesis:

H_0: There is no difference between the means (i.e. the means are the same)

Therefore, the alternative hypothesis, H_a, is that there is a difference between the means.

Selecting *a*=0.05, the *f-test* for the equality of variances shows that there is no statistical difference between the variances of both groups and a t-test assuming equal variances could be applied (see table 7).

Table 7. F-Test Two-sample for variances

	UAH	*UOC*
Mean	6.24717	6.951852
Variance	3.578309	6.034131
Observations	53	27
df	52	26
F	0.593012	
P(F<=f) one-tail	0.054473	
F Critical one-tail	0.584925	

As it can be seen in table 8, the *t statistic* is lower than the *t critical* (.23727< 2.30601), and also the *p value* is greater than *a* (.0.079 > 0.05). According to this, we can accept the null hypothesis that there is no statistical difference between the means.

Table 8. t-Test: Two-sample assuming equal variances

	UAH	*UOC*
Mean	6.24717	6.951852
Variance	3.578309	6.034131
Observations	53	27
Pooled Variance	4.396916	
Hypothesized Mean Difference	0	
df	78	
t Stat	-1.42133	
P(T<=t) one-tail	0.079604	
t Critical one-tail	1.664625	
P(T<=t) two-tail	0.159208	
t Critical two-tail	1.990847	

Based on the marks obtained in our experiment, we can conclude that knowledge acquired in blended or exclusively online learning is similar.

3.4. Threats to validity

We faced 3 different kinds of threats to validity: those concerning the construct validity, those concerning the internal validity and those concerning the external validity of our experiment:

- **Construct validity** is the degree to which the variables used in the study accurately measure the concepts they purport to measure. In our case, marks were the only way to measure the student´s acquired knowledge.
- **Internal validity** is the degree to which conclusions can be drawn about the causal effect of the independent variable on the dependent variables. Potential threats include selection effects, non-random subject loss, instrumentation effect, and maturation effect. Although the students are different, it is noting that we only considered students handing in their assignments. Another possible threat is that both groups did not have the same time to work on the assignment (2 weeks for the UOC students vs. 3 weeks in the case of the UAH students).
- **External validity** is the degree to which the results of the research can be generalised to the population under study and other research settings. Although students in the online course are more mature than in the blended group, their initial knowledge was similar. The only possible difference is that the teaching materials were not identical in both cases.

4. Conclusions and future work

The research reported on the comparison of two courses on compiler basics, one fully online given at the Open University of Catalonia (UOC) and another blended one given at the University of Alcala (UAH). The study reported herein was restricted to the first assignment activity of the course as the other 2 were not finished by the UAH course by the deadline. Even though some factors change and the conditions and context cannot be demonstrated to be identical, most conditions were: the assessment to carry out, the teachers who acted as tutors and the evaluation criteria.

We first carefully studied possible threats to the study validity, considering all different aspects and concluding that those threats did not have a central influence on the results of our experiment. The tests run on the marks of both groups of students allowed us to conclude that the amount of knowledge acquired in blended learning (UAH) and in exclusively online learning (UOC) is similar.

Further work should progress on a more comprehensive study, covering all the stages of the compilers course, namely activities 1 to 3 and not only the activity 1. Also, studying new factors such as the lecturers influence (e.g. do students in other groups perform differently?), time given for assignment resolution, students' age, etc. would be helpful for getting more comprehensive conclusions. Replications over several years are also a target to address external validity.

Acknowledgements

The authors would like to thank Francesc Bagés and Alex Pajuelo for their collaboration in Compilers 1 at the UOC and discussions about the design of Compiler courses.

References

Anstine, J., & Skidmore, M. (2005). A small sample study of traditional and online courses with sample selection adjustment. *Journal of Economic Education*, 36(2), 107-127.

Borges, F. (1998). UOC and its 'Campus Virtual': A Model for Online University Education, In *Proceedings of the 2nd North American Web Developers Conference (NAWEB 98)*, New Brunswick, NJ, USA.

Open University of Catalonia (2008). Annual report 2007-2008. URL: http://www.uoc.edu/portal/_resources/EN/documents/memories/0708/memo_eng.pdf

Prieto, J. (2008). Ontology-based characterization and specification of virtual laboratories in Computer Science degrees. PhD thesis, Universitat Oberta de Catalunya.