Redefining dropping out in online higher education: a case study from the UOC

Josep Grau-Valldosera[1] and Julià Minguillón[2]

[1] Market Research Department
[2] Computer Science, Multimedia and Telecommunication Studies
Universitat Oberta de Catalunya, Barcelona, Spain
{jgrauv, jminguillona}@uoc.edu

Abstract. In recent years, studies into the reasons for dropping out of higher education (including online education) have been undertaken with greater regularity, parallel to the rise in the relative weight of this type of education, compared with brick-and-mortar education. However, the work invested in characterising the students who drop out of education, compared with those who do not, appears not to have had the same relevance as that invested in the analysis of the causes. The definition of dropping out is very sensitive to the context. In this article, we reach a purely empirical definition of student dropping out, based on the probability of not continuing a specific academic programme following several consecutive semesters of "theoretical break". Dropping out should be properly defined before analysing its causes, as well as comparing the drop-out rates between the different online programmes, or between online and on-campus ones. Our results show that there are significant differences among programmes, depending on their theoretical extension, but not their domain of knowledge.

Keywords: dropping out, higher education, online university, cohorts, distance education

1 Introduction

From an institutional perspective, university dropping out is very important, as dropping out needs to be seen as a failure of the university system to generate "product" (graduates) with an important quantity of public resources invested. In the analysis made in this paper – and based on the specific definition of dropping out presented in it – we see that during the first 26 semesters, the Universitat Oberta de Catalunya (Open University of Catalonia, UOC) received 62,450 new students enrolled in officially recognised degrees in Catalan; 13.3% of them finished a degree and 57.6% dropped out of their studies. These figures only include students who have been enrolled a certain number of semesters large enough to establish a criterion for dropping out, the main goal of this paper. In fact, establishing the exact moment when a student can be considered a drop-out is part of this goal. Dropping out is a highly relevant phenomenon that deserves to be analysed in detail. Furthermore, it is difficult

to establish comparisons with other centres given the specificities of the UOC, a purely online distance university.

One of the challenges posed to universities, be they brick-and-mortar or distance, is to define the concept of dropping out. The main difficulty lies in the fact that, faced with several successive semesters of non-enrolment by a student, it cannot be said with 100% assurance that this student has definitively dropped out of their studies, as it may happen that he or she is taking a longer or shorter break; this difficulty is even greater in distance studies, where the profile of the majority of students has more work and family commitments than that of brick-and-mortar type university students and where, therefore, the existence of breaks seems much more likely. Additionally, it should be considered that the official definition of dropping out does not reflect the particularities of online higher education.

The main aim of this paper is to define dropping out in online higher education following an inductive process based on an objective analysis of enrolment and non-enrolment (i.e. breaks) data of all students in officially recognised degrees in Catalan at the UOC between 1996 and 2008. The definition reached using this methodology will be valid for all teaching institutions that offer studies of a certain duration and with non-obligatory enrolment. In particular, it is highly adaptable, due to the possibility of the existence of breaks, to institutions offering distance university education. Dropping out analysis requires a certain historical database in order to be accurate. In this sense, the UOC, despite being a very young and pioneering university in the field of virtual education, has already been in existence for fifteen years and has, at present, almost 35,000 active students in officially recognised degrees in Catalan, which allows it to undertake a quantitative analysis on the basis of a strong statistical representation. To conclude, we should stress that this definition of dropping out will be established from an institutional perspective, i.e., without considering the perspective of the student; in this way of thinking, a student may "drop out" (not have achieved the aim of the qualification) from the point of view of the university, but he or she may be fully satisfied with the teaching experience, having achieved his or her personal learning objectives. Therefore, from an institutional point of view, the definition of dropping out will always be harder or more negative than reality.

The rest of this paper is organized as follows: Section 2 describes some approximations to the higher education dropping out phenomenon in brick-and-mortar universities and, especially, in distance learning institutions. Section 3 explains the data used and the methodology followed for analyzing it. Section 4 presents the main results and, finally, Section 5 deals with the conclusions and future research in this topic.

2 Prior analysis of university dropping out

Lassibille and Navarro [1] offer a recent view of the subject through the Spanish brick-and-mortar university system. They analysed 7,000 students enrolled at the University of Malaga starting in September 1996 and running to June 2004 (8 academic years). The authors ruled out analysis of students who took a break during a certain academic year, as they only accounted for 2% of the total (in the case of students analysed at the UOC – online students – this figure is around 14.6%). The definition of dropping out they give is very clear: a student is considered to have dropped out of a particular degree course even if they move to another one (we see later that this criterion is also used in the analysis of UOC students).

On the basis of this definition, overall dropping out by this cohort (i.e. the group of students who began at the start of 1996) during the period under consideration is a high 46.1%. The authors' work immediately concentrates on analysing the causes of dropping out. Other authors [2] put this figure at 40%-45% during the last 100 years. As we will see, the results obtained in this paper are coherent with these findings if we include the students for whom we do not yet have enough available data.

In 2004, Berge and Huang [3] conducted a synthesis of the bibliography available at the time on the problem of student drop-out rates in e-learning. They recorded the existence of a higher level of dropping out in virtual environments than in brick-and-mortar environments, as can also be seen in Frankola [4]; although they did point out that the problem of dropping out is complex and multidimensional. The authors established the definition of a holistic model of university dropping out, which concentrates its causes – and therefore the possible actions for reducing this dropping out – on institutional factors (including actions by lecturers and administrative personnel) and on the socio-demographic and academic characteristics of the students. This model provides a general framework that each institution should adapt to its specific characteristics.

In their definition of the complementary term to dropping out (which we could call "continuity"), Berge and Huang state that the ultimate aim of continuity needs not always to be seen as the qualification obtained through study (completion). For some students, successful learning experiences have more to do with "participating" and not with "completing" or "obtaining a degree", as stated in Kerka [5]. In this sense, Pappas and Loring [6] give the name "degree seekers" to students, who, from an institutional point of view, may be considered "drop-outs" when they stop their studies. However, we should remember that this paper adopts a totally institutional perspective on dropping out, which does not take into account, at least in this initial stage of analysis, the aims of students. It is supposed, therefore, that all students who enrol in a degree course intend to complete it and obtain the accreditations.

2.1 Dropping out in distance higher education

It is interesting to note how the UNED, the main Spanish distance university founded in 1972, analyzes the subject of dropping out by its students. In an extensive article, Callejo [7] takes dropping out analysis from a more long-term perspective than in a brick-and mortar university[1], in the same way as this article. However, in his definition of dropping out, he introduces a more qualitative element, namely the intention to continue or not to continue by students who are taking a break from their studies (which he asks them about in a survey). Callejo negates that the congestion of distance universities is one of the main causes of dropping out (it can be seen that degree courses with large numbers of students – e.g. law and psychology – have proportionally lower rates of dropping out than others with fewer students – such as engineering). The author cites the intrinsic difficulty of the contents of each programme as one of the main causes of dropping out.

Lastly, it is interesting to see how a benchmark university in the field of distance education on a world scale, the Open University (UK), has analysed its drop-out rates. Tresman [8] stresses the specificity of adult distance students and calls for an analysis that matches this profile to be undertaken. He states that, "the vast majority who withdraw – 94 per cent – still aspire to earn credit for the course/award upon which they embarked". He immediately collects their reasons for withdrawing from their studies during the various stages in their academic lives and, finally, proposes a strategy to improve student continuity. Tresman starts the work that Ashby [9] would undertake two years later, taking the Open University as a reference. The author calls the classic definition of continuity the "institutional dimension", understood to be "passing the course" and presents two new dimensions or definitions that relativise the importance of the final qualification: that of the student and that of the employer. The author gives the total drop-out rate (during the first academic period) as 40% and says that, if this figure needs comparing, it should be with other part-time educational institutions and not with other traditional university education (full-time brick-and-mortar) institutions.

3 Data and methodology

The data used in this paper are taken from UOC academic databases. For this initial study, only student enrolments are analyzed. The following variables are available: IDP (Person Identifier Code), unique to each student, it allows individual and at the same time anonymous monitoring; gender; student's date of birth; semester of the student's enrolment; codes of the subjects enrolled in by the student; final grades obtained in the subjects; number of credits that the subjects carry and, finally, the academic programme, e.g. Law or Computer Engineering. Specifically, there is a record for each subject enrolled in the officially recognised degrees in Catalan from the start of the university until the end of the 2008-2009 academic year (in all,

[1] Where taking a break during a single academic period is considered dropping out.

1,169,262 records). Observe that enrolment at UOC is opened each semester, which means bi-annually. A total of 19 degrees were offered during this period. Only valid enrolments have been included, i.e. ones that have been formalised and paid for, thus excluding enrolments that were subsequently cancelled. A total enrolment history was provided for 84,230 students, although only 62,450 (those of the 16 programmes with more available information) were analysed. This study ignores the pilot cohorts of the programmes that began at the start of the university, which limited student access during the first semester to a closed number and, for administration purposes; there was no access for new students during the second semester.

To analyse the data described in this article, only the "IDP", "semester of enrolment" and "academic programme" fields were considered. The information from these fields was used to generate 20 files, 19 for each programme and a general file for all programmes, each of which contains a record for each student. These students are those who enrolled in one or more semesters of the programme during the period in question. The records generated have the following coding:

10104;1;1;1;0;1;0;0;0;0;0;0;0;0;0;0;0;0;0;0;0;0

where the first number is the IDP and then a binary string for the semester record ("1" = student enrolled at least in one subject, "0" = student not enrolled in any subject). The specific nature of this string is that, for analysis purposes, all enrolment sequences have been put in the "same starting position", that is, the first semester when each IDP is enrolled in each degree is considered to be the same for all students. Obviously, the first element after IDP is always "1" (the first enrolment of each student). Finally, notice that the sequences "113000;1;0;0;0;0;0;0" and "10104;1;0;0" are different as more enrolment history about the first student is available for analysis (specifically, 7 semesters as opposed to 3).

Once the enrolment sequences file of each programme is generated, the frequency of break sequences (that is, of sequences of one or more "0") can then be analysed. This is performed via a process that detects the longest break sequence (with "1;0;...;0;1" format) within each enrolment sequence of each individual, with the particularity that if, for example, a student has taken a break once over 5 semesters and over 2 during another semester, he or she will only be calculated as having taken a break over 5 semesters. Notice that this process does not take graduates into consideration, as they may be considered as taking a break or abandoning their studies, when they have in fact obtained their degree. For exemplification purposes, Table 1 shows the probability of having a break of N semesters for the Law degree (with 7,938 students and a history of 24 semesters) and the Market Research and Techniques (MR&T) degree (with 1,718 students and a history of 14 semesters).

Table 1. Analysis of the break sequences from Law (left) and MR&T Studies (right)

	Law degree			**MR&T degree**		
Number of semesters	**Number of students**	**Percentage of students**	**Accumulated percentage**	**Number of students**	**Percentage of students**	**Accumulated percentage**
19	2	0.03	0.03	---	---	---
18	1	0.01	0.04	---	---	---
17	0	0	0.04	---	---	---
16	9	0.11	0.15	---	---	---
15	9	0.11	0.26	---	---	---
14	8	0.11	0.37	---	---	---
13	18	0.23	0.60	---	---	---
12	14	0.18	0.78	---	---	---
11	12	0.15	0.93	---	---	---
10	15	0.19	1.12	---	---	---
9	27	0.34	1.46	---	---	---
8	37	0.47	1.80	5	0.29	0.29
7	29	0.37	2.27	3	0.17	0.46
6	50	0.63	2.90	6	0.35	0.81
5	**69**	**0.87**	**3.77**	7	0.41	1.22
4	107	1.35	5.12	3	0.17	1.39
3	173	2.18	7.30	**30**	**1.75**	**3.14**
2	304	3.83	11.13	40	2.33	5.47
1	815	10.27	21.40	141	8.21	13.68
0	6239	78.60	100	1483	86.32	100

The column description of Table 1 is as follows: the first column is the number of semesters of consecutive breaks (namely N); the second is the number of students enrolled in the Law degree that take a break of length N; the third and fourth columns are the percentage of such students with respect to the total of students in the degree and the accumulated percentage, respectively. Columns 5-7 reproduce the same for the MR&T degree. In the light of these results, there are two students in the Law degree that take a break of 19 consecutive semesters, which may be surprising but shows the wide diversity of online students' interests and behaviour. Nevertheless, in order to define dropping out, we are interested in establishing a threshold for what we consider a reasonable period of break time. As shown in bold in Table 1, only 3.77% of Law students take a break of 5 or more semesters. In the case of MR&T students, a similar percentage (3.14%) is found but only with 3 semesters or more, showing a relevant difference among academic programmes. In short, if we define dropping out as taking a break of 5 or more semesters for the Law degree, we are assuming an error smaller than 5%, which can be considered reasonable. On the contrary, dropping out is defined in the MR&T degree as having a break of only 3 semesters with the same

error assumption. Notice that the fact that a Law student has the "1;0;0;0;0" string in his or her enrolment sequence is not sufficient information to see whether he or she will drop out, as we need an additional semester as mentioned above. Following this criterion we are able to label each student with a sequence of N or more "0" as dropping out.

Therefore, a definition of the drop-out rate for a specific programme would be reached inductively as being the proportion of students who have taken a break for N or more semesters out of the total number of students enrolled in the programme during the period in question. N is determined using the maximum probability of the 5% error rate in classifying the student as a drop-out once they have taken a break of N or more semesters in that specific programme; in other words, a maximum of 5% of students on that specific programme return to their studies after taking a break of N or more consecutive semesters, so the maximum error is upper-bounded.

4 Results

On the basis of the work set out in the section above, the definition of dropping out for each programme of officially recognised degrees in Catalan is reached. The specificity of the programme in question is highly relevant. Although, logically, the definition of dropping out in qualitative terms is the same for all courses; repetition of the probability analysis carried out for all programmes gives as the result different "quantitative definitions" depending on the values of the "parameter" of this definition, i.e. different N values for consecutive break semesters.

4.1 Differences among programmes

Table 2 provides a summary of the values associated with the graphs of the 16 programmes analysed (the three more recent programmes have not been included because more historical information was needed). The explanation of the fields of Table 2 is as follows. For each programme, the minimum number of consecutive break semesters needed to be considered dropping out is N, with the maximum error that is smaller than 5%, as well as the number of semesters defined in the curriculum of each programme[2], the number of semesters since the programme began and the number of students with at least N+1 semesters. Finally, the last two columns make reference to the percentage of students achieving the degree or dropping out, respectively.

2 Normally, the real duration is longer, as students enrol in fewer subjects than those theoretically defined in the curriculum, especially in distance universities.

Table 2: Results summary by programme and total of programmes

Programme	N	Error	Duration	Available data	Number of students	Accredited	Dropping out
Business Sci.	5	3,78%	6	26	16818	16.6%	54.3%
Tech. Eng. in CM	5	4,11%	6	22	5432	9.8%	66.8%
Tech. Eng, in CS	5	4,46%	6	22	7496	8.7%	65.6%
Tourism	3	3,38%	6	14	1889	9.6%	49.7%
Catalan	4	3,89%	8	22	1194	6.5%	58.9%
Law	5	3,78%	8	24	6149	10.2%	54.0%
Humanities	5	3,75%	8	24	5396	7.4%	64.3%
Psychology	3	4,58%	8	18	7674	3.8%	56.5%
Business	4	3,75%	4	22	3778	38.2%	40.9%
Labour Sci.	4	2,82%	4	16	3114	34.5%	44.8%
Political Sci.	3	4,27%	4	16	867	21.7%	49.5%
Audiovisual	3	2,67%	4	14	1070	21.9%	43.7%
Documentation	3	4,48%	4	20	2440	32.3%	50.3%
Market Res. & Tec.	3	3,14%	4	14	1374	32.4%	38.0%
Psychopedagogy	4	4,86%	4	26	4354	25.4%	54.2%
Computer Engineer.	4	3,36%	4	16	1541	30.1%	37.3%
TOTAL	4	4,35%	---	---	62450	13.3%	57.6%

As shown in Table 2, the number of semesters that define dropping out in each programme has particularly relevant variability: this value varies between 3 and 5 semesters. Notice that these figures are very conservative, since using an upper bound of 10% would have reduced the number of consecutive break semesters.

An initial analysis of these results shows that there appears to be no relationship between the type of programme content, i.e. technical or humanistic, and the number of semesters that determines dropping out. For example, in the case of Computer Engineering, the value is high (5 semesters), but it is the same in the case of Humanities. On the other hand, it does seem that in programmes where students have prior higher education experience related to the curriculum they are studying (in Spain they are known as “Second cycle” degrees, the ones with a duration of 4 semesters at the bottom of Table 2), dropping out is decided with fewer semesters than on programmes where this experience is not required (known as “First cycle” or “First and Second cycle”). This way, the average number of semesters that define dropping out of all the students who have dropped out from second cycle programmes is 3.69 semesters, while it is 4.61 in students of the other programmes. To validate this statement statistically, an independent-samples T-test[3] is carried out with SPSS package. Since the significance value of the test is less than 0.05, it can be safely concluded that the difference of 0.92 is not due to chance alone. It can also be seen

3 Equal variances not assumed.

that students in shorter duration programmes (second cycle) also take shorter breaks during their academic record, which seems logical.

4.2 Enrolment behaviour graphs

Armed with these parameters, graphs for each programme, as well as a general graph, can be drawn up to visually display the behaviour of enrolment, dropping out and accreditations of each programme. Figure 1 shows an example of the calculated graphs, in particular for the Law programme. For simplification purposes, only arcs with a minimum proportion of students (1%) are shown. The graph shows in some way the "survival history" of students of each programme. For example, in the first node of the graph shown, 6,149 students of Law (who enrolled at least 5+1 semesters ago) appear in the starting position (at the first semester). This number is greatly reduced in the second semester, in which only 4,048 students (65.83%) continue from the original 6,149. Noticeably, the main cause of this loss of 2,101 students is dropping out after the first semester (1,643 students). That is, one out of four students does not continue after the first semester. On the other hand, except one student obtaining the degree after his or her first semester, the rest (7.4%) take a break, which is also remarkable.

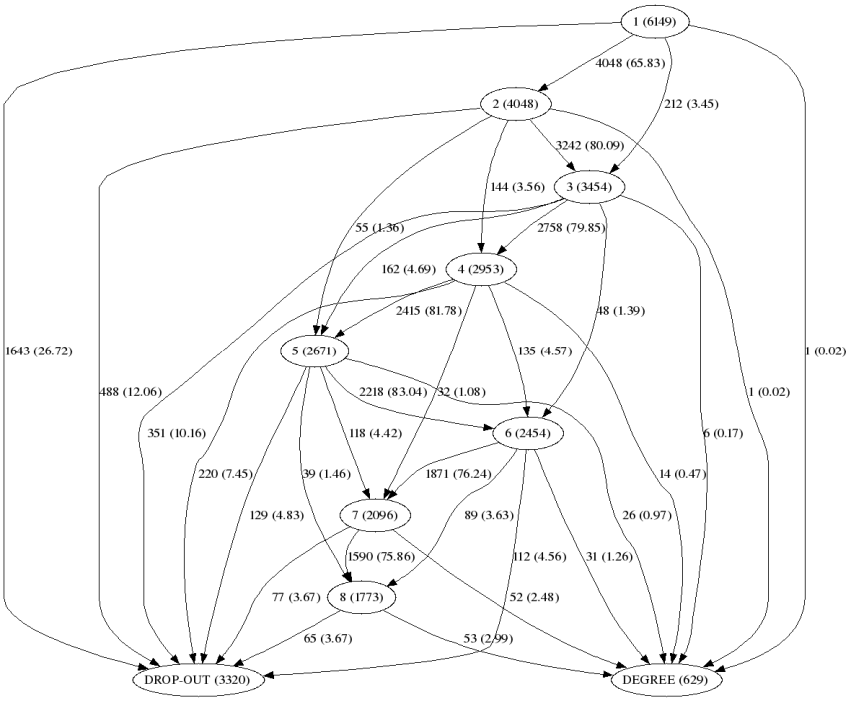

Fig. 1. Enrolment behaviour for the first 8 semesters of Law students

5 Conclusions

This paper deals with the formulation of a definition of dropping out that is suitable for the students of a distance higher education institution such as is the UOC. This objective is quite important if it is considered that most UOC students are adult students with work and family commitments additional to that of continuing education and, therefore, with a natural tendency to take academic breaks. On the other hand, this objective is somewhat difficult to reach, due to the fact that these break periods could be interpreted as only that (periods of rest) or, at some point, as indicators of having definitely dropped out of their studies.

At the end of the analysis carried out in this work, it seems that the effort invested in reaching an inductive definition of dropping out, achieved using the enrolment and non-enrolment (break) behaviour of students in each programme or academic discipline, has led to particularly relevant results. This definition is highly sensitive to the reality of each programme. Consequently, it has enabled us to classify students as drop-outs when they take a break of just 3 consecutive semesters, for example in the case of Psychology or Tourism, and up to as long a period as 5 semesters, i.e. two years and a half, in the case, for instance, of Business Sciences or Humanities.

The definition was sustained in an analysis of the Universitat Oberta de Catalunya (UOC) enrolment data, running from 1996-1997 to the 2008-2009 academic years for official degrees in Catalan, with at least 14 semesters of record. This includes 16 different programmes. In particular, an analysis in terms of programmes was carried out, concerning the frequency of break sequences, that is, the consecutive periods of non-enrolment that culminate in a reincorporation of students into the same degree. The result of this analysis is the minimum break sequence (of N or more semesters) that has a very low associated probability of students returning (fewer than 5%, the assumed error). N, the number of consecutive semesters in blank, enables a student to be classified as a drop-out for a specific programme. We should stress, however, that this N value is different for each programme, and that herein lies the potential for the definition of dropping out that has been reached in this paper. Therefore, a single definition for dropping out cannot be established at the university level.

In an initial approach to the analysis of differences between programmes in terms of the number of semesters that trigger dropping out, it appears that in programmes where students have prior higher education experience, the decision to drop out is made more quickly (almost one semester before) than in programmes where such experience is not required. This may be due to students having clearer objectives in these types of programmes, which are based on completed studies (in a previous degree). It could also be related to the shorter theoretical duration of such degrees. We should also point out that total (accumulated) dropping out percentages for these kinds of programmes are significantly smaller than those for the other programmes.

The analysis conducted in this paper, which allows us to establish whether a student can be considered a drop-out or not, is the starting point towards undertaking a close study of the characteristics of students who drop out. Such a study will be based on data already collected, but not yet incorporated into the analysis, such as age and sex of the students or other variables related to number and kinds of subjects taken every semester. We should also consider the addition of new variables which are the result of a more qualitative analysis than the one undertaken in this paper.

Once the causes of dropping out have been detected, the establishing of corrective actions that have a positive effect on reducing dropping out should report benefits both at institutional and personal levels, especially for those students who have given up the fight with a sense of not having achieved their learning objectives.

References

1. Lassibille, G. and Navarro Gómez, L. Why do higher education students drop out? Evidence from Spain. Education Economics, 2008, 16(1), p. 89-105.
2. Tinto, V. Dropouts from higher education: A theoretical synthesis of the recent literature. A Review of Educational Research, 1975, 45(1), p. 89-125.
3. Berge, Z. and Huang, Y. A Model for Sustainable Student Retention: A Holistic Perspective on the Student Dropout Problem with Special Attention to e-Learning. DEOSNEWS, 2004, 13(5).
4. Frankola, K. Why online learners drop out. Workforce, 2001, 80, p. 53-58.
5. Kerka, S. Strategies for retaining adult students: The educationally disadvantaged. ERIC Digest, 1988, 76.
6. Pappas, J. P. and Loring, R. K. Returning Learners. Increasing student retention. In L. Noel, R. Levitz, D. Saluri, and. Associates (eds.), 1985.
7. Callejo, J. Estudio de Cohorte de Estudiantes de la UNED: una Aproximación al Análisis del Abandono. Revista Iberoamericana de Educación a Distancia, 2001, 4(2), p. 33-69.
8. Tresman, S. Towards a strategy for improved student retention in programmes of open and distance education: a case study from the Open University, International Review of Research in Open and Distance Learning, 2002, 3(3), p. 1-11.
9. Ashby, A. Monitoring student retention in the Open University: definition, measurement, interpretation and action. Open Learning, 2004, 19(1), p. 65-77.