FROM PARTICIPATION TO DROPOUT:

QUANTITATIVE PARTICIPATION PATTERNS IN ONLINE UNIVERSITY COURSES

doi:{ 10.1016/j.compedu.2010.02.026 }

Please refer this manuscript as:

Nistor, N., & Neubauer, K. (2010). From participation to dropout: Quantitative participation patterns in online university courses. *Computers & Education*, *55*(2), 663-672.

From participation to dropout:

Quantitative participation patterns in online university courses

Nicolae Nistor & Katrin Neubauer

Ludwig-Maximilians-Universität München
Faculty of Psychology and Educational Sciences
Leopoldstr. 13, D-80802 München, Germany

Corresponding author: Nicolae Nistor
E-mail: nic.nistor@uni-muenchen.de
Phone: +49 89 2180 5284, Fax: +49 89 2180 99 5284

1 Introduction

The academic e-learning practice has to deal with various participation patterns and types of online learners with different support needs. Among these, special attention is requested by dropouts, a ubiquitous phenomenon in academic online courses. Online students are more likely to dropout than campus based students (Patterson & McFadden, 2009). Many researchers report dropout quotes up to 50% and more in online courses (e.g. Aragon & Johnson, 2008; Hesse & Giovis, 1997; Levy, 2007; Morris, Wu & Finnegan, 2005; Nichols & Levy, 2009; Patterson & McFadden, 2009). The consequences of dropouts for the online courses are critical especially in collaborative scenarios. Small learning groups can shrink to single learners, thus disturbing and demotivating the remaining students who were initially committed to participation, and maybe leading them to dropout, too. To avoid this, it is important to understand participation, to identify the different participation patterns and learner types, and to offer them appropriate support.

The factors that influence participation and dropout have already been discussed in previous literature. The most prominent theoretical models of academic attrition are synthesized by Rovai (2003) and include student characteristics and skills, external and internal factors, all these having an influence on the students' persistence decision, i.e. completing the course vs. dropping out. Among the internal factors, pedagogy is represented as learning and teaching styles, expected to fit to each other in order to support online learning. Successful participation in such a learning environment comprises learner accomplishing all the activities required by the underlying didactical concept. Unlike variables that may be difficult to measure (e.g. study habits) or restrictedly available (e.g. prior academic performance, for reasons of data privacy), participation is directly and easily observable. Also, Rovai's model implies that the observation of participation at early moments in the process of learning is likely to give insight about the learners' later persistence decision. Such insight was also intuitively experienced by the first author in the practice of online teaching and this motivated him to search for corresponding empirical evidence.

The present study aims at identifying quantitative participation patterns and exploring the interrelation between participation and later persistence in online academic courses. In the long run, we aim at giving online instructors a reliable, non-invasive, easy-to-handle method to predict dropouts and thus to avoid perturbances of online collaboration. Authors' intuition should thus be confirmed through empirical evidence.

2 Theoretical background

2.1 Problem-based online courses

Academic courses include frequently the use of electronic resources and online discussions, sometimes also online collaborative learning. They are not mere online images of the traditional courses, but rather expected to have specific didactics built up on the advantages and limitations of the used learning technologies. [anonymized] (…; also Scripture, 2008) proposes a problem-based didactical concept for online courses based on the principles of (a) problem-based learning with authentic problems, (b) participants' self-regulated learning including decisions like where and at what time of the day they learn, which learning materials and resources they use, how they share this amount of work with their co-learners etc., (c) collaborative learning in small online groups, and (d) instructional support especially in the form of instructors' recommending steps and strategies to successfully carry out the process of learning. From these principles, (a) addresses the didactic design, (b) and (c) the individual and collaborative process of learning and (d) the direct instruction. Especially through (b) and (c), the problem-based design is aimed at stimulating the learners to individual and cooperative learning activities, building thus the basis of participation.

From the perspective of organizing online learning, we can distinguish (A) an initial phase including registration, e-mail contact with the instructors, a face-to-face meeting, self introduction in the yellow pages of the electronic learning platform, finding a name for the virtual work group, and committing to a certain amount of work for the seminar, (B) carrying out the actual process of learning, and (C) an end phase consisting of the reflection and summarizing of the learning process, and preparing the final documentation requested to receive the credit points. From the qualitative point of view, the participants have to climb up from the level of (1) simple access and motivation, (2) online socialization, and (3) information exchange, up to the levels of (4) knowledge construction and (5) personal development, as described by Salmon (2004, p. 28; also Hrastinski, 2008).

From the perspective of online collaboration, the interactions between learners in the online environment are aimed at producing positive social interdependence (Johnson & Johnson, 2009) through designing the process of learning so that all the members of a small group work to reach the same learning goals and receive all the same grade for the group performance (goal, outcome and reward interdependence). The learning tasks comprise handling with large amounts of information, partially from the recommended research literature, partially from the Internet (resource interdependence). In every small group, the participants rotate in taking on the moderator role (role interdependence). The combination of collaboration and problem-based learning is expected to shape the interaction within the online course in the form of learners' dividing the original problem into smaller problems that can be more easily solved, which finally leads to the collaborative knowledge construction (Cognition and Technology Group at Vanderbilt, 1997). Participation becomes thus visible as written communication needed to coordinate the collaboration.

2.2 Participation in online university courses

Participation in an online course consists of learners' completing the activities specified in the seminar's didactical concept. Thus, according to the didactical design of the learning environment in question, participation takes complex and diverse forms. Among them, communication, i.e. writing, sending and reading messages, appears to be the central activity of the online learners.

To draw an overview of the previous research literature (Table 1), we differentiate first quantitative from qualitative participation. The former relies on aspects such as numbers of actions performed, frequency and length of messages exchanged etc., the latter on the quality of actions and contents of communication. Although we find qualitative participation to be highly relevant for learning, we concentrate in this study on quantitative aspects, which we consider to be a prerequisite of qualitative participation.

Within quantitative participation, passive participation is designated in the Internet jargon as “lurking”, an online activity in which a person only consumes without producing information. Lurkers hardly leave traces of their activity in online environments, this is why the passive participation is methodically difficult to trace, and from the available research few studies (e.g. Hesse & Giovis, 1997) address passive participation. In contrast, there is a wealth of studies of active participation that rely upon the various traces left by the learners in their environments (see overview by Hrastinski, 2008). Most of them (e.g. Carell, 2006; Caspi et al., 2006; Caspi, Chajut & Saporta, 2008; Davies & Graf, 2005; Gao & Wong, 2008; Hesse & Giovis, 1997; Joyce & Kraut, 2006; de Laat et al., 2007; Wise et al., 2006) analyze communication by criteria such as number, length and regularity of messages, or media choice. More recent studies apply the social network analysis to online learning environments (de Laat et al., 2007). Further studies observe the learners’ activities by means of log file analysis (e.g. Caspi et al., 2008; Davies & Graf, 2005; Hesse & Giovis, 1997); others rely on the learners’ subjective rating (e.g. Bürg, 2005; Chen et al., 2008; NSSE, 2008).

Table 1: Types of participation and data collection methods

	Types of participation			Data collection methods	Authors
Participation	Quantitative	Active (“participation”)	Online communication	Message analysis	Carell (2006) Caspi et al. (2006) Caspi et al. (2008) Davies & Graf (2005) Gao & Wong (2008) Hesse & Giovis (1997) Joyce & Kraut (2006) de Laat et al. (2007) Wise et al. (2006)
			Online activity	Log-file analysis	Davies & Graf (2005) Hesse & Giovis (1997)
				Learners’ subjective estimation	Bürg (2005) Chen et al. (2008) NSSE (2008)
		Passive (“engagement”)	Reading	Log-file analysis	Caspi et al. (2008) Davies & Graf (2005) Hesse & Giovis (1997)
	Qualitative	(not discussed in this study)			

Most of the studies concentrate on sustained and successful participation up to the end of the online courses. Few researchers (e.g. Hesse & Giovis, 1997; de Laat et al., 2007) mention the evolution of participation within the process of learning, and these go no further than counting the dropout participants and relating them to the number of learners who completed the course (Hesse & Giovis, 1997). This research deficit can be easily explained by the fact that dropout students can either be hardly reached for questioning, or they give inconclusive answers such as “The course was really great, unfortunately I had no time to complete it” (Aragon & Johnson, 2008).

2.3 Typologies of online learners

Tondeur, van Braak and Valcke (2007) suggest that the educational use of computers, i.e. learners’ participation in online environments cannot be studied as an isolated variable. Therefore they call for defining typologies of learners in order to identify frequent combinations of learner characteristics and predict learning behavior. This may be further helpful for designing instructional support or, from a wider perspective, educational policies. Del Valle and Duffy (2009) used cluster analysis to identify types of approaches to learning in online courses. Based on students’ learning strategies, three clusters could be identified, i.e. “mastery oriented”, “task focused” and “minimalist in effort”. The difference between the clusters was mainly related to participation variables such as course duration, number of sessions used to complete the course, time on learning resources, time in the mail system etc. It is remarkable that these variables can be collected using non-invasive methods such as log file analyses. On the other hand, the analysis was made on a global level, the course didactics were not taken into consideration. The clusters displayed differences in satisfaction with the course and self reported learning; dropouts were not considered.

Attrition in online learning environments is a significant phenomenon that takes place more frequently than in face-to-face courses (Patterson & McFadden, 2009). Given a typology of online learners, some learner types will be probably more likely to dropout. Therefore, we would expect a participation based typology of online learners to predict learners’ persistence in the learning environment and eventually their successful course completion. The persistence vs. dropout phenomenon in academic environments, either online or campus based, is extensively represented in the research literature. A prominent model was first formulated by Tinto (1975) and later extended by Kember (1989) and Rovai (2003). These comprise numerous variables such as individual attributes, commitment to goals and institution, performance, social interactions, academic and social integration, and complex relationships. Only the later form of the model (Rovai, 2003) mentions the course pedagogy, i.e. learning and teaching styles, however without further details. Dropouts were predicted in previous research with accuracy up to 75% (Morris, Wu & Finnegan, 2005) or even 85% (Parker, 1999). From the methodological point of view, the available studies (Aragon & Johnson, 2008; Caspi et al., 2008; Levy, 2007; Morris et al., 2005; Nichols & Levy, 2009; Patterson & McFadden, 2009) compare the completers vs. dropout subgroups with respect to various variables depending on the research model used in each case, and conclude that significant differences indicate the dropout predictors. Again, none of these studies takes the course didactics into account.

3 Research questions

In this study we adopt a quantitative view of participation that takes the didactics of the online course into consideration, relies on non-invasive measures (i.e.

observation and log file analysis), and includes dropouts into the analysis. On this background we propose the following research questions:

1. How high is the attrition in the online courses and when does it occur?
2. Which are the significant differences between the completion and the dropout subgroups with respect to the quantitative participation indices?
3. Which types of participants can be identified with respect to learners' quantitative participation patterns?
4. How accurately can learner's persistence vs. dropout be predicted on the ground of the quantitative participation patterns?

4 Method

Design. To answer the research questions, similarly to the cited previous studies we collected longitudinal data over the entire academic term (14 weeks) and compared the completion and the dropout subsamples. Since the courses belonged to the regular academic activity, our study can be regarded as field research.

Sample. The study was conducted at the Ludwig-Maximilians-Universität (LMU) of Munich, Germany, which has a total of over 40,000 students (aprox. 25,000 female and 15,000 male), registered to 18 faculties. Among these, the Faculty of Psychology and Educational Sciences has approx. 6,300 active students (approx. 800 male and 5,200 female). From these, 1,600 (1,300 female, 300 male) have a major in Educational Sciences (excluding teacher education) or Psychology. The Chair of Education and Educational Psychology is one of the 25 research units at the Faculty, and has a main research focus in technology based learning. From a total of 20-25 courses offered each term by the chair, most of them use computer support, but no more than five are entirely online; three of these, “Introduction to Knowledge Management”, “Online Courses Evaluation” and “Online Courses Development and Implementation” were subject of the present study. These were offered regularly each term as optional courses mainly for undergraduate students of Educational Sciences. The participants were besides local students also remote students from other Bavarian universities. All of the involved universities are campus-based, have similar structures, and offer usually a small number of online courses, which are generally coordinated by the Virtual University of Bavaria, at the moment of the study with a total of 172 courses and over 27,000 course registrations yearly.

During two terms (each with a regular duration of 14 weeks), a total of N = 209 students took part in the studied courses, from which 144 were female and 65 male, aged between 17 and 50 (M = 25.36, SD = 4.64). 75 studied Educational Sciences, 105 had other fields of study such as Psychology, Computer Science or Business Administration; 29 did not specified their field of study. 90 of the participants were local students of the Munich University, and 115 were remote students form the univerisities Bamberg, Regensburg, Bayreuth, Erlangen-Nürnberg, Weingarten and Würzburg.

Setting. Independently of the contents, the online courses consisted of six learning units, each over a period of two or three weeks. The first was dedicated to the learning organization (registration, contact between students and instructors, face-to-face kick-off meeting, self-introduction in the Yellow Pages of the learning platform). The following were conceived as problem-based learning and started from authentic cases. The last unit comprised summarizing of the course. A more detailed description of the course concept can be found in [anonymized].

All learning tasks were to be solved in small collaborative groups; the first task however was based on plenum discussions. This design particularity was due to two observations experienced by the instructors. First, dropouts would occur most frequently at the course beginning. Second, the most active, as well as the most inactive learners would preserve their activity level for the entire course period. Even though these observations were not yet empirically sustained before this study, the most and the least active participants were uniformly distributed over the groups. Thus, a minimum of three and a maximum of five learners were assured for each group, so that none of the groups would be at risk to disintegrate because of dropout. The learning groups were kept for the entire course period; inactive participants were excluded after a warning message by e-mail; inactive groups were dissolved and the remaining active members were redistributed to active groups. Finally, only the group performance was considered for the final degree.

Instruments. Measuring participation was based on observation during the entire course duration. The operationalization of the quantitative participation construct was built on the online course didactics and resided into a set of ten variables deduced as learners' responses to the various assignments: registration, organization tasks, e-mail contact with the instructors, active participation in the online discussions and in the course evaluation. The registration time was divided into three subintervals (early, middle and late), in which the occurring sing-ups were counted. The fulfilled punctual, organization tasks (e-mail contact with the instructor after registration, participation in the presence session, personal introduction in the "yellow pages" of the learning platform) were proven and counted. The e-mails addressed to the instructors in the two weeks of the course were also counted, as well as the answered questionnaires of the course evaluation and the messages from the discussion forums. Whenever noticed that a participant does not take part actively in a discussion block, the instructor sent him or her an e-mail message to ask if he or she is still participating in the course. If the answer was negative or missing, a dropout was registered. Dropouts were also registered when students explicitly withdrew from the course. The participation at the course evaluation was measured as the number of questionnaires responded and the number of demand notes received by the participants to remind them of the questionnaires. (Longitudinal questionnaire data were also collected along the courses, but proved inconclusive for predicting dropouts; therefore this part of the study will not be discussed here.)

Course delivery and data collection. The students registered to the online course and carryed out the assigned organization tasks. After each of the six learning units we counted the active participants. Due to the slightly different timetables of the courses, the time intervals between the measure points varied between two and three weeks; the first measure point was set in all three courses 3.5 weeks after the course begin. At the end of the course, we analysed the written messages, and processed statistically the collected data.

5 Results

5.1 Attrition during the online course

From the initial number of 209 participants (144 female, 65 male), 159 (76.1%) completed the course and the rest of 50 (23.9%) droped out. The dropouts of the first two weeks were 16 in total (7.7% from all or 10.1% of the dropouts), from which 14 female (87.5% from the dropout sub-sample, 9.7% from all female) and 2 male (12.5% from the dropout sub-sample, 3.1% from all male) participants.

The moment of dropout was distributed all along the course period, however more frequent at the beginning. The evolution of total, female and male dropouts along the courses is displayed in absolute values in fig. 1a and in procentual values in fig. 1b. (The first measure point took place 3.5 weeks after the term start, i.e.1.5 weeks after the registration of the participant numbers given above.)

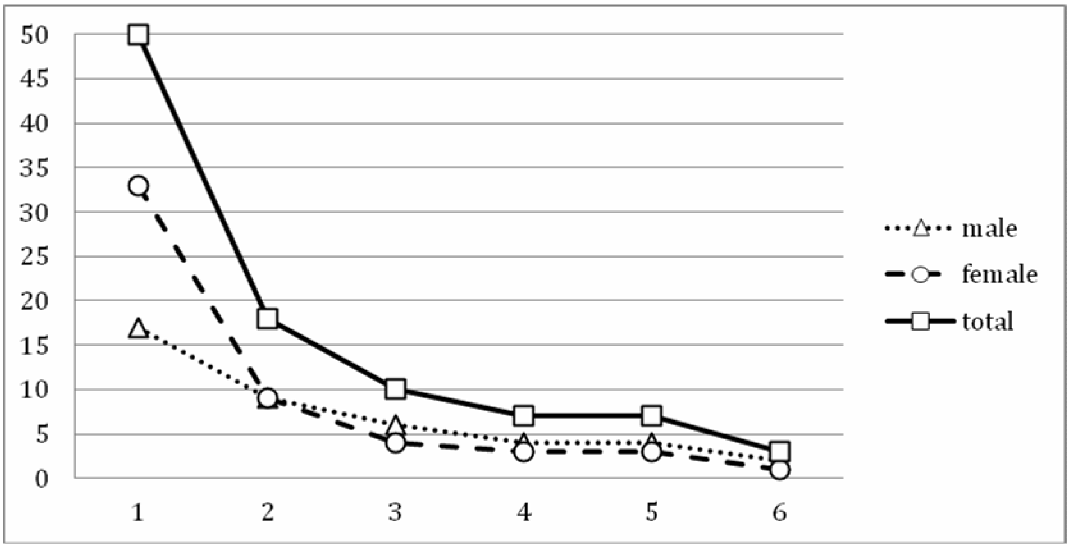

Fig. 1a. The evolution of the participants' numbers in the dropout subgroup during the online courses (absolute values in the six measure points)

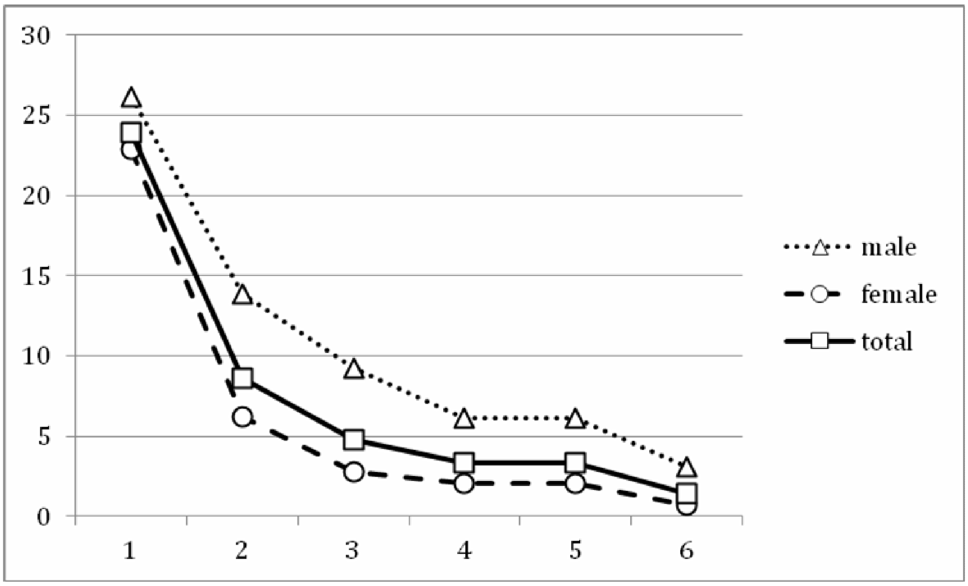

Fig. 1b. The evolution of the participants' numbers in the dropout subgroup during the online courses (relative values pro hundred dropouts in the six measure points)

5.2 Differences between completion and dropout subgroups

Learners' participation in the online course was operationalized as a set of ten variables deduced as learners' responses to the various assignments of the course. We tested which of these components made a significant difference between the completion and the dropout subgroup. A one-tailed ANOVA variance analysis (table 2) showed that only the participation in the presence session, personal introduction in the "yellow pages" of the learning platform, e-mails to the instructor in the first two weeks, total number of the sent messages, total length of the sent messages, and the participation at the course evaluation displayed significant differences ($p < .01$) between the completion and the dropout subgroups. These variables were chosen for the further description of the quantitative participation.

Responses to assignments and communication with the instructor. The completion and the dropout subgroup participated differently to several activities of the course, i.e. the dropout subgroup displayed significantly less intensive participation than the completion subgroup, as found out through variance analysis (table 2). Only 6.8% of the dropout subgroup (vs. 26.3% of the completion group) came to the presence session, 44.7% (vs. 97.4%) wrote a personal introduction in the "Yellow Pages" of the learning platform, and 36.4% (vs. 71.4%) sent e-mail to the instructor as requested after the registration (figure 2).

Moment of registration. The period of one month, in which the students could enroll to the course, was divided into three equal parts with a length of ten days. Observing the distribution of the enrolled students over the three thirds, dropout students show a strong tendency to register early. Thus, in the early third, 31% of the enrolled students (28 persons, 19 female and 9 male) proved to be dropouts. In the next two thirds, there were only 21% (5 persons, 3 male and 2 female) and respectively 18% (16 persons, 11 male and 5 male) dropouts. In comparison, from the completion subgroup 61 participants (42 female and 19 male) registrated in the early third, 19 participants (15 female, 4 male) in the middle third and 75 participants (51 female, 24 male) in the late third. Thus, the moment of registration tended to a significant difference (*p* = .097) between the dropout and the completion subgroups, i.e. the dropout subgroup tended to “last minute registration”. Therefore, the moment of registration was added to the list of relevant participation indicators.

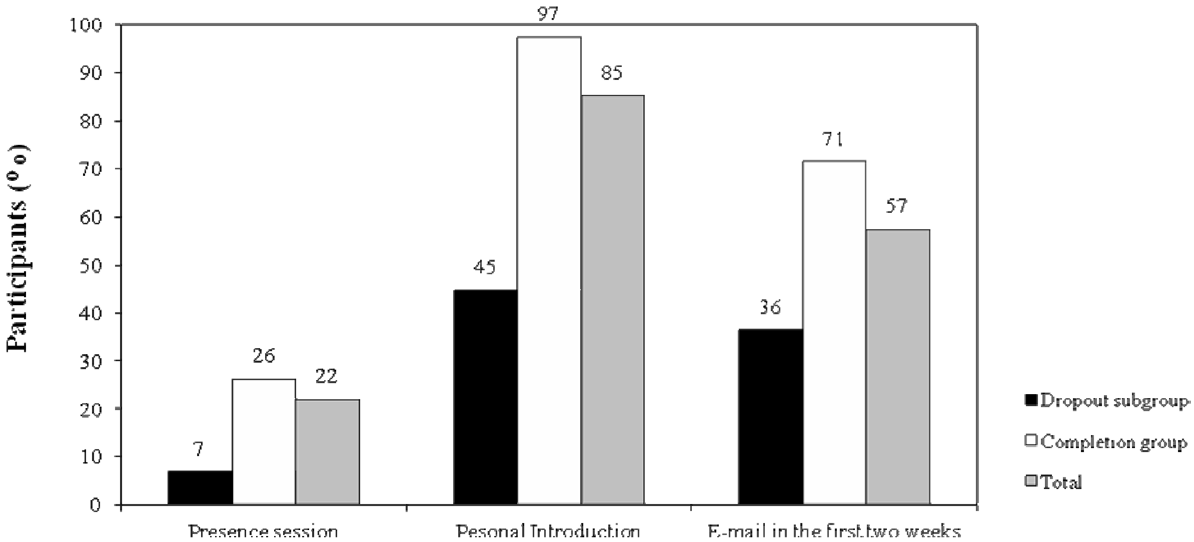

Figure 2. Responses to assignments and communication with the instructor at the beginning of the online course

Table 2: Indicators of the quantitative participation in the dropout vs. completion sub-groups (one-tailed ANOVA variance analysis); for indicators 1 to 5, 8 and 9, in brackets the absolute values and percentual values (from all male and respectively female students of the dropout, respectively completion subgroup); ** $p < .01$

		Dropout subgroup			Completion subgroup			df_{treat}	df_{error}	F	*p*
		male N = 17 (%)	female N = 33 (%)	total N = 50	male N = 48 (%)	female N = 111 (%)	total N = 159				
1.	E-mail contact with the instructor after registration	3 (17.6)	4 (12.1)	7	4 (8.3)	11 (9.9)	15	1	85	1.372	*.245*
2.	Further e-mail to the instructor before the begin of the course	0 (0.0)	4 (12.1)	4	5 (10.4)	16 (14.4)	21	1	85	.455	*.502*
3.	Participation in the presence session	1 (5.9)	2 (6.1)	3	12 (25.0)	29 (26.1)	41	1	85	15.571	*.000***
4.	Personal introduction in the “yellow pages”	6 (35.3)	17 (51.5)	23	48 (100.0)	104 (93.7)	152	1	85	40.763	*.000***
5.	E-mails to the instructor in the first two weeks	3 (17.6)	9 (27.3)	12	9 (18.8)	26 (23.4)	35	1	85	13.709	*.000***
6.	Total number of messages	140	177	317	3362	10875	14237	1	85	155.353	*.000***
7.	Total length of the messages in characters (nubmer of senders)	120972 (7)	202023 (12)	322995 (19)	4134718 (48)	12408553 (108)	16533271 (156)	1	85	43.548	*.000***
8.	Participation at the course evaluation as number of questionnaires completed	5 (29.4)	12 (36.4)	17	48 (100.0)	105 (94.6)	153	1	85	178.252	*.000***
9.	Participation at the course evaluation (number of demand notes)	7 (41.2) (10)	9 (27.3) (17)	16 (27)	36 (75.0) (60)	76 (68.5) (152)	112 (212)	1	85	24.710	*.000***

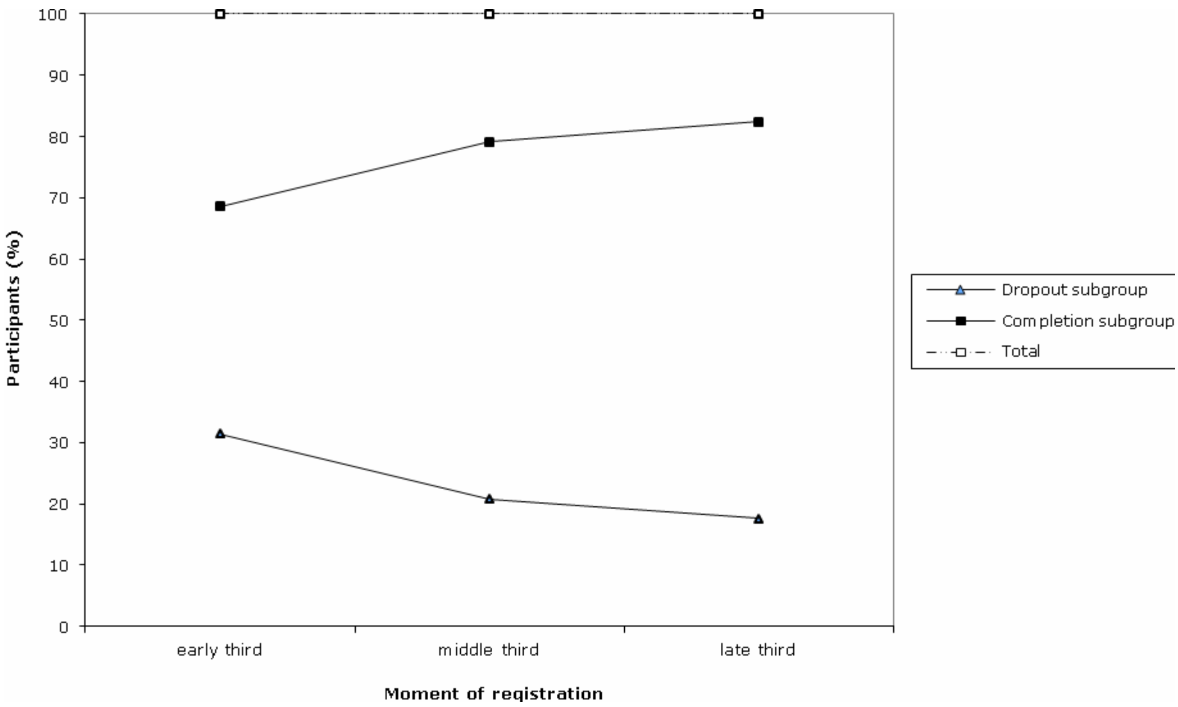

Figure 3. Participants' moments of registration to the online course

5.3 Participant types

To identify the learner types with respect to participation, a two-step cluster analysis was performed. From the entire data sample (N = 209), 143 cases included missing values and hat to be rejected, thus the cluster analysis could be performed with 66 cases. First, a hierarchical cluster analysis indicated the number of 4 optimal clusters. However, the hierarchical cluster analysis is not suitable for variables with different scale levels, so the analysis was done with a predetermined clusters number of 4. The learners were classified as follows (see also the data in table 3 and the cluster profiles in figure 4).

Cluster 1: Highly committed students. The first cluster is the largest one with 25 learners, 6 male and 19 female students from the own university (13 students of Educational Science), representing 37.9% of the analyzed sample. This subgroup consists of active learners, most of which participated in the face-to-face session, introduced themselves in the "yellow pages", wrote at least one e-mail message to the instructor, contributed more to the discussions than the other subgroups (M = 83.40, SD = 28.69) and with the larges messages (M = 180,238.68 characters, SD = 144,580.89). Also, they participated to the evaluation study and responded to 5.76 from 6 questionnaires, being reminded to do this 1.84 times.

Cluster 2: Minimalist remote students. The second and smallest cluster consists of 9 remote participants, 7 male and 2 female from other, cooperating universities (2 students of Educational Science), representing 13.6% of the sample. None of them took part in the presence session but all of them introduced themselves in the "yellow pages". Only 4 of them wrote messages to the instructor. Their communication behavior was below average, their contributions to the group discussions were fewer (M = 57.78, SD = 22.60) and shorter (M = 70,825.56 characters, SD = 42,871.93) than the messages from clusters 1 and 3. They participated to the evaluation study responding to 5.33 from 6 questionnaires, being reminded 2.22 times.

Table 3. Cluster variables (z-standardized mean values)

Variable	Cluster 1 Highly committed students (N = 25)	Cluster 2 Minimalist remote students (N = 9)	Cluster 3 Average local Ed Sci students (N = 16)	Cluster 4 Dropouts (N = 16)
University (own vs. other)	-.866	1.443	-.289	-.289
Study	.124	1.050	-1.359	.185
Sex	.183	-1.461	.730	.548
Participation in the presence session	1.472	-.760	-.357	-.357
Personal introduction in the "yellow pages"	.654	.654	.148	-1.457
E-mails to the instructor in the first two weeks	1.470	-.490	-.245	-.735
Total number of messages	.874	.085	.461	-1.420
Total length of the messages (characters)	1.098	-.365	.468	-1.201
Participation at the course evaluation (number of questionnaires responded)	.629	.669	.629	-1.492
Participation at the course evaluation (number of demand notes)	.161	.669	.629	-1.459

Cluster 3: Average local Educational Science students. Cluster 3 is formed by 16 learners, 15 female and 1 male students of Educational Science, building 24.2% of the sample. Few of them (n = 2, i.e. 12.5%) participated in the face-to-face session, however most of them (n = 15, i.e. 93.8%) introduced themselves in the "yellow pages". Their communication behavior was average, half of them wrote e-mail messages to the instructor and their contributions were between clusters 1 and 2 with M = 70.00 (SD = 28.50) messages and M = 133,145.38 characters (SD = 107,193.93). They participated mostly to the evaluation study responding to 5.5 from

6 questionnaires, nevertheless having to be reminded about it 2.19 times.

Cluster 4: Dropouts. Finally, cluster 4 consists of 16 learners, 14 female and 2 male students, most of them (12, i.e. 75%) from the own university, representing 24.2% of the sample. Half of them studied Educational Science. Only two of them participated in the presence session, and only 6 of them (37.6%) wrote at least one e-mail message to the instructor; 12 (75%) introduced themselves in the "yellow pages". The participants from this subgroup communicated less than others, they wrote a mean number of M = 8.94 messages (SD = 19.29) with a mean length of 8,230.56 characters (SD = 19,982.49). Their participation in the evaluation study was weak, having responded to 2.06 from 6 questionnaires while being reminded 0.6 times. Unlike all the other participants, the learners from cluster 4 droped out during the online course.

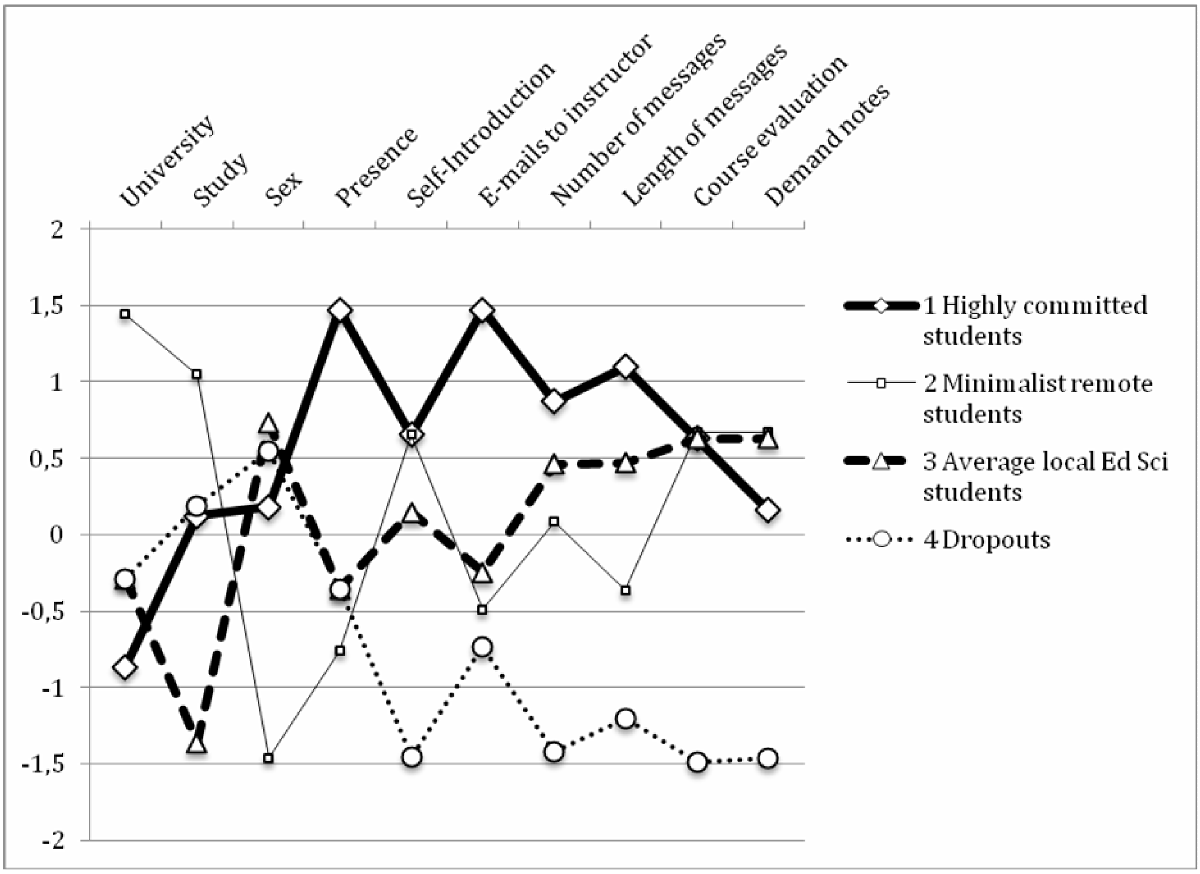

Figure 4. Results of the two-step cluster analysis (z-standardized mean values)

5.4 Dropout prediction

The existence of a statistically identified cluster consisting of all dropout students suggests that learners could be predicted relying on the same data, i.e. collected in the first two weeks (from a total of 13) of the course. Therefore, we conducted a predictive discriminant analysis using as variables the relevant participation components identified as response to out second research question. Only three of these, i.e. presence in the face-to-face session, self-introduction in the "yellow pages" and e-mails to the instructor in the first two weeks proved to be predictive (table 4).

In the next step, we verified the accuracy of the predicted classification. From the total sample (N = 209), 122 cases included missing values and hat to be rejected, thus the discriminant analysis could be performed with 87 cases. From these, based on the three discriminant variables (table 5) 48 cases (34 female, 14 male) were classified correctly to the completion and 20 (13 female, 7 male) to the dropout subgroup. Hence, the classification accuracy reached 78.2%.

Table 4. ANOVA with the three variables used in the discriminant analysis (** p < .01)

	df_{treat}	df_{error}	F	Sig.
Presence in the face-to-face session	1	85	15.571	.000**
Self-introduction in the “yellow pages”	1	85	40.763	.000**
E-mails to the instructor in the first two weeks	1	85	13.769	.000**

Table 5. Measured vs. predicted membership in the completion vs. dropout subgroups

Measured appartenence to subgroups	Predicted subgroups*					
	Dropout			Completion		
	Female (N = 62)	Male (N = 25)	Total (N = 87)	Female (N = 62)	Male (N = 25)	Total (N = 87)
	n (%)	n (%)	n (%)	n (%)	n (%)	n (%)
Dropout	13 (20.1)	7 (28.0)	20 (54.1)	14 (22.6)	3 (12.0)	17 (45.9)
Completion	1 (1.6)	1 (4.0)	2 (4.0)	34 (54.8)	14 (56.0)	48 (96.0)

* 78.2% of the initial classifications (75.8% of the female, 84.0% of the male participants) were correct

6 Discussion

In summary, we studied several online university courses with a total number of participants N = 204 and we found attrition to affect less than a quarter of the participants, which is not unusual for facultative courses, no matter if online or presence (Levy, 2007; Patterson & McFadden, 2009). Problem-based leaning had induced intensive cooperative learning activity in the online courses and thus offered rich material for the analysis of participation.

Which indices of the quantitative participation are relevant for the course completion? Unlike previous persistence studies, where mainly student characteristics are focused, we operationalized participation as a sequence of compulsory or at least highly recommendable steps responding to the didactical course model (Hrastinski, 2008; Salmon, 2004). Our method was nevertheless similar to previous studies (Aragon & Johnson, 2008; Caspi et al., 2008; Levy, 2007; Morris et al., 2005; Nichols & Levy, 2009; Patterson & McFadden, 2009), i.e. we compared the completion and the dropout subgroups. After the variance analysis, only some of these activities proved relevant. Indicators like e-mail contact with the instructor were irrelevant, which is consistent with the self-regulated learning didactics. The moment of registration only tended to be statistically significant, i.e. dropout participants tended either to register very early (and then probably find other, more important courses for their study plan) or very late (and then probably struggle with an overbooked study plan).

The relevant indicators of the successful participation were the learning activities addressing the collaborative process of learning, i.e. participation in the presence session, personal introduction in the “yellow pages”, e-mails to the instructor in the first two week; total number of messages, total length of the messages (characters); participation at the course evaluation (number of questionnaires responded); participation at the course evaluation (number of demand notes). A similar difference was found between dropouts and completion participants’ communication.

Which types of participants could we identify? The cluster analysis of the participation data measured in the first two weeks of the course (from a total of 13 weeks) resulted into four learner subgroups: Highly committed students with extensive contributions to the group learning (similar to the mastery oriented cluster found out by del Valle and Duffy, 2009); minimalist remote students from other, cooperating universities, who accomplish a minimum of learning activity (probably similar to the minimalist cluster found out by del Valle and Duffy, 2009); average local Educational Science students; and dropouts (not taken into consideration by del Valle and Duffy, 2009). Two of these clusters were remarkable. First, from the instructional point of view, the remote students subgroup appear to be isolated from the local students, even if, according to the collaborative scenario, they were spread over the small learning groups. The lack of communication seems to affect remote students’ performance and it calls for special attention from the instructor. Second, from the empirical point of view, the dropout subgroup appears to be determined very early in the process of learning, which opens instructors the possibility of interventions that minimize the perturbations of the collaboration, if not even prevent the dropouts. Also, the existence of a clearly separated dropout cluster suggests the possibility of predicting dropouts using early participation data.

How accurately could we predict learners’ persistence in the online course? The participation data gathered in the first two weeks allowed a prediction accuracy of nearly 80%, similar to those reached by Morris et al. (2005) or Parker (1999). Unlike these, our prediction relayed on participation data related to the course didactics and collected by obervation, not by questionning. The “intuitive prediction” based on authors’ experience as online instructors and on the observation of learners and learning activities was confirmed. The prediction was useful for 90% of the dropouts; the rest of 10% had already left the course at the moment when the prediction was done.

What impact has learners’ gender on the participation patterns? Our investigation was situated in a social environment dominated by female students. The basic proportion was 85% female to 15% male among the students with a major in Educational Sciences or Psychology, and 70% female to 30% male in our online courses. The dropout rate was above the baseline for the male students (65% female, 35% male), however female students droped out earlier (88% female, 12% male dropouts in the first two weeks). During the courses, female students displayed stronger social presence (i.e. had more contact with the instructor, introduced themselves more frequently in the “yellow pages”, wrote more messages, participated more intensively to the course evaluation) and were higher in performance (i.e. more frequent in the participants clusters 1 and 2).

Apparently, the online courses seemed to attract more male students than the traditional teaching; males also droped out more frequently, however not as fast as the female students. Females appeared nevertheless to comply more with the requirements of the didactical concept. As for the learning performance, two clusters of higher performance were dominated by females, i.e. 76% female vs. 24% male among the “highly committed students” and respectively 94% female vs. 6% male among the “average local Educational Science students”. On the other hand, males dominated the cluster “minimalist remote students” (78% males vs. 22% females). Both the female compliance with the course requests (the learning activity of female students reflects less the individual attitude and intentions, which may lead to dropout) and the dropout speed (the prediction is based on the assumption of a constant participation behavior) may explain why the dropout prediction was somewhat better for males (85% accurate clasification) than for females (75%).

The internal and external validity of this study may be however restricted on the conceptional level by having focused on the participation indices dictated by the online course didactics. Further factors such as financial aid (Morris et al., 2005) or study plans (often mentioned by the students as a reason for dropout; see also Aragon & Johnson, 2008) were not taken into consideration. As for the methodic aspects, even if a number of over 200 participants represents a relatively solid empirical basis, incomplete data were inevitable in a field study. These may also restrict the external validity of our conclusions.

7 Conclusions

As a practical consequence, the operationalization we chose for measuring the quantitative participation proved successful i.e. lead to significant results. Therefore, for the academic e-learning practice it is recommendable to measure participation at the beginning of the course as a form of checking the learning prerequisites and make an appropriate intervention if needed. This operation may be easily implemented in learning management systems.

As a theoretical consequence, we showed that online course didactics play a crucial role for student participation patterns, and attrition as well. Previous attrition models (e.g. Kember, 1989; Rovai, 2003; Tinto, 1975) should be refined including beside individual, institutional and social variables, also the course didactics.

Future work continuing this study should offer, on the one hand, a more in-depth view of quantitative participation in correlation with the attrition variables formulated by Rovai (2003). On the other hand, the qualitative participation and its learning patterns should be explored. Both should provide a better understanding of the impact induced by dropout in the online courses, and of possible remedies through which the instruction could be improved.

References

Aragon, S. R. & Johnson, E. S. (2008). Factors influencing completion and noncompletion of community college online courses. *The American Journal of Distance Education, 22* (3), 146-158.

Bürg, O. (2005). *Akzeptanz von E-Learning in Unternehmen. Die Bedeutung von institutionellen Rahmenbedingungen, Merkmalen des Individuums und Merkmalen der Lernumgebung für die Akzeptanz von E-Learning in Unternehmen*. Berlin: Logos.

Carell, A. (2006). *Selbststeuerung und Partizipation beim computerunterstützten kollaborativen Lernen*. Münster: Waxmann.

Caspi, A., Chajut, E., Saporta, K. & Beyth-Marom, R. (2006). The influence of personality on social participation in learning environments. *Learning and Individual Differences, 16*, 129-144.

Caspi, A., Chajut, E. & Saporta, K. (2008). Participation in class and in online discussions: Gender differences. *Computers & Education, 50* (3), 718-724.

Chen, P.S.D., Gonyea, R. & Kuh, G. (2008). Learning at a distance: Engaged or not? *Innovate – Journal of Online Education, 4* (3). Retrieved on December 9, 2009 from http://www.innovateonline.info/pdf/vol4_issue3/Learning_at_a_Distance-__Engaged_or_Not_.pdf

Cognition and Technology Group at Vanderbilt (1997). *The Jasper Project: Lessons in curriculum, assessment, and professional development*. Mahwah, NJ: Lawrence Erlbaum.

Davies, J. & Graff, M. (2005). Performance in e-learning: Online participation and student grades. *British Journal of Educational Technology, 36* (4), 657–663.

Gao, F. & Wong, D. (2008). Student engagement in distance learning environments: A comparison of threaded discussion forums and text-focused Wikis. *First Monday, 13* (1). Retrieved on December 9, 2009 from http://firstmonday.org/htbin/cgiwrap/bin/ojs/index.php/fm/article/view/2018/1921

Hesse, F. W. & Giovis, C. (1997). Struktur und Verlauf aktiver und passiver Partizipation beim netzbasierten Lernen in virtuellen Seminaren. *Unterrichtswissenschaft, 25* (1), 34-55.

Hrastinski, S. (2008). What is online learner participation? A literature review. *Computers & Education, 51* (4), 1755-1765.

Johnson, D. W. & Johnson, R. T. (2009). An educational psychology success story: Social interdependence theory and cooperative learning. *Educational Researcher, 38* (5), 365-379.

Joyce, E. & Kraut, R. E. (2006). Predicting continued participation in newsgroups. *Journal of Computer-Mediated Communication, 11* (3). Retrieved on December 9, 2009 from http://jcmc.indiana.edu/vol11/issue3/joyce.html

Kember, D. (1989). A longitudinal-process model of drop-out from distance education. *Journal of Higher Education, 60* (3), 278-301.

de Laat, M., Lally, V., Lipponen, L. & Simons, R.-J. (2007). Investigating patterns of interaction in networked learning and computer-supported collaborative learning: A role for Social Network Analysis. *International Journal of Computer-Supported Collaborative Learning, 2* (1), 87-103.

Levy, Y. (2007). Comparing dropouts and persistence in e-learning courses. *Computers & Education, 48* (2), 185-204.

Morris, L. V., Wu, S. S. & Finnegan, C.L. (2005). Predicting retention in online general education courses. *American Journal of Distance Education, 19* (1), 23-36.

Nichols, A. J. & Levy, Y. (2009). Empirical assessment of college student-athletes' persistance in e-learning courses: A case study of a U.S. National Association of Intercollegiate Athletics (NAIA) institution. *Internet and Higher Education, 12* (1), 14-25.

National Survey of Student Engagement (NSSE)(2008). *Promoting engagement for all students: The imperative to look within*. Bloomington, IN: Indiana University Center for Postsecondary Research. Retrieved on December 9, 2009 from http://nsse.iub.edu/NSSE%5F2008%5FResults/

Parker, A. (1999). A study of variables that predict dropout from distance education. *International Journal of Educational Technology, 1* (2). Retrieved on December 9, 2009 from http://www.ascilite.org.au/ajet/ijet/v1n2/parker/

Patterson, B. & McFadden, C. (2009). Attrition in online and campus degree programs. *Online Journal of Distance Learning Administration, 12* (2). Retrieved on December 9, 2009 from http://www.westga.edu/~distance/ojdla/summer122/patterson112.html

Rovai, A.P. (2003). In search of higher persistence rates in distance education online programs. *The Internet and Higher Education, 6* (1), 1-16.

Salmon, G. (2004). *E-moderating: The key to teaching and learning online*. London: RoutledgeFalmer.

Scripture, J. D. (2008). Recommendations for designing and implementing distributed problem-based learning. *American Journal of Distance Education, 22* (4), 207-221.

Tinto, V. (1975). Dropout from higher education: A theoretical synthesis of recent research. *Review of Educational Research, 45* (1), 89-125.

Tondeur, J., van Braak, J. & Valcke, M. (2007). Towards a typology of computer use in primary education. *Journal of Computer Assisted Learning, 23* (3), 197-206.

del Valle, R. & Duffy, T.M. (2009). Online learning: Learner characteristics and their approaches to managing learning. *Instructional Science, 37* (2), 129-149.

Wise, K., Hamman, B. & Thorson, K. (2006). Moderation, response rate, and massage interactivity: Features of online communities and their effects on intent to participate. *Journal of Computer-Mediated Communication, 12* (1).