

# Learning Analytics to Explore Dropout in Online Entrepreneurship Education

Ignacio Toledo<sup>1</sup>, Carlos Albornoz<sup>1</sup>, Kaethe Schneider<sup>2\*</sup> 

<sup>1</sup>University of Desarrollo, Santiago de Chile, Chile

<sup>2</sup>Friedrich-Schiller University of Jena, Jena, Germany

Email: \*k.schneider@uni-jena.de

**How to cite this paper:** Toledo, I., Albornoz, C., & Schneider, K. (2020). Learning Analytics to Explore Dropout in Online Entrepreneurship Education. *Psychology*, 11, 268-284.  
<https://doi.org/10.4236/psych.2020.112017>

**Received:** January 9, 2020

**Accepted:** February 8, 2020

**Published:** February 11, 2020

Copyright © 2020 by author(s) and Scientific Research Publishing Inc.  
This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).  
<http://creativecommons.org/licenses/by/4.0/>



Open Access

## Abstract

The technological disruption of E-Learning offers Entrepreneurship Education (EE) an unprecedented opportunity to leverage the affordances of web 2.0 and to widen the scope of entrepreneurship educational programs. However, in practice, high dropout rates in online courses (near 90%) pose major challenges to EE researchers. In this paper, we use learning analytics to explore the case of dropout during a female-oriented online entrepreneurship educational program. We observed that the evolution of dropouts and learning behavior of the participants in this program is not linear in time. Persistence decays in a two-step process: The first dropout phase reaches approximately 30% before the middle of the program and then stabilizes around that number. A second dropout phase is triggered in the last quarter of the program and it continues declining until the end of the program. This dropout pattern corresponds to the interaction indicators at aggregated level: Before the middle of the program the level of interaction drops together with the number of active users. However, the reduction in the interaction frequency is disproportionate in relation to the fall of number of the active users. After dividing by the percentage of active users the number of interactions still drops before the middle of the program. Using social network analysis (SNA) we show that initial dropouts have a considerable effect on the connectivity of the communication network, this is consistent with the observed decrease in social interactions. We also found significant correlations between entrepreneurial competencies and indicators of learning behavior and persistence, at the individual level. In line with the relevant literature, the most significant feature associated with persistent behavior was found to be the risk-taking orientation. Our findings suppose the first step towards an empirical model of persistence in Entrepreneurship Education online programs providing valuable insights for future research and for developing retention methods in online courses.

---

## Keywords

Entrepreneurship Education, E-Learning, Dropout, Learning Analytics, Social Network Analysis, Women Entrepreneurs, Learning Behavior, Persistence

---

## 1. Introduction

*Women Entrepreneurs* has been a four-year RISE (Research, Innovation and Staff Exchange) project in the Horizon 2020 framework of the European Union. Motivated by the need to improve indicators of success among female-owned businesses in the European Union, *Women Entrepreneurs* aims to design, implement and evaluate an online educational and training program that enables early-stage women entrepreneurs to increase their entrepreneurial success.

In contrast to traditional face-to-face entrepreneurship educational programs the *Women Entrepreneurs* project set the goal of turning towards effective interdisciplinary research through the integration of recent research trends in Entrepreneurship Education (EE) and in E-Learning. The output of this project is an educational and training program focused on the development of fundamental entrepreneurial competencies, which is embedded in a Learning Management System (LMS) allowing women entrepreneurs to self-direct their learning process, to seek tutoring and networking with peers.

Alongside the affordances of Web 2.0 and the tools available in Learning Management Systems, online education also poses major challenges in practice, particularly regarding attrition, persistence and student retention. Most open online courses suffer from high dropout rates, typically around 90% (Jordan, 2014).

One of the most widespread and rapidly growing delivery models of e-learning is the Massive Open Online Course (MOOC). Typically, MOOC participants are able to engage in social networking, freely explore course content and create and share content in forums (McAuley, Stewart, Siemens, & Cormier, 2010), exploiting at least three of the key features of Web 2.0 (Blaschke, 2012). Probably the most important aspect of MOOCs relies on massive enrollments of “students who self-organize their participation according to learning goals, prior knowledge and skills, and common interests” (McAuley, Stewart, Siemens, & Cormier, 2010). This kind of complex dynamics naturally results in social networks with core-periphery and community structures (Agudo-Peregrina, Hernández-García, & Iglesias-Pradas, 2012).

In spite of the advantages of online education, it faces challenges regarding attrition, persistence and student retention. Online educators and MOOC providers need to design and implement retention strategies (Lee & Choi, 2011; Yang, Sinha, Adamson, & Rose, 2013; Jordan, 2014; McKinney, Novak, Hagedorn, & Luna-Torres, 2018). Understanding the evolution of dropouts is, therefore, critical to online education.

The *Women Entrepreneurs* (WE) is an online training program funded by the European Union and executed by an alliance of universities and consulting companies' experts on education. It aimed to design, implement and evaluate an online program that enables early-stage women entrepreneurs to increase their entrepreneurial success. The WE project was based on the developmental model of entrepreneurial competencies proposed by [Schneider and Albornoz \(2018\)](#). The program was implemented in Germany and Ireland with 25 or more women entrepreneurs from each country. The duration of the program was four months, and the estimated work-load was five to eight hours per week, with a maximum of 128 hours. Participants who completed the four months of instruction received a certificate of participation.

The structure of the program consisted of four modules, one face-to-face opening session, and three online modules delivered via a digital platform. The digital platform was implemented in Moodle. Moodle is designed to host online courses based on constructivist and self-directed learning approaches ([Dougiamas & Taylor, 2003](#)).

In terms of dropout rates, The *Women Entrepreneurship* program was no exception: a high proportion of the participants of the first round did not complete the program. In this work, we analyze the collected data during the first round of the program and explore the relationships between the entrepreneurial competencies, the learning behavior, and dropouts.

## 2. Literature Review and Theory

We address knowledge transferability from the perspective of Polanyi's theory of knowledge ([Polanyi, 1962](#)), which states that each act of knowing is personal and social. It is personal, because it occurs within individuals, and it is social, because it can be shared among individuals. The personal component of knowledge is called tacit, as is difficult to put into words and refers to the product of personal meaning-making related to experience. The social component of knowledge is called explicit, as it can be stated in formal propositions and transferred within a group.

The practice of autonomous learning requires an enabling environment where learners can give free rein to their self-determined learning potential. This is what Web 2.0 provides, a ubiquitous learner-centered environment with a broad range of affordances. The key technological features of Web 2.0 that support the self-determined approach can be reduced to the following four ([Blaschke, 2012](#)): 1) Connectivity and social rapport/communication, 2) Content discovery/sharing (individually and collaboratively), 3) Content creation (individually and collaboratively) and 4) Knowledge and information aggregation and content modification.

New delivery models of e-learning have been developed to deal with high dropout rates. According to [Lee and Choi \(2011\)](#), the main sources of attrition in online courses can be classified into three categories, namely: student factors (academic background, relevant experience, skills, psychological attributes), course/program factors (course design, institutional support), and environmen-

tal factors). We added a fourth category, namely behavioral factors, which group online behavior registered in log databases during the online sessions including interactions and time allocation. In the proposed classification, interaction factors are switched to the behavior category since it is a better match than course/programs category.

Regarding student psychological attributes, recent studies in higher education suggest that self-determined motivation positively affects a student's persistence (Ryan & Deci, 2017; Jenó, Danielsen, & Raaheim, 2018). Self-Determination Theory (SDT) (Ryan & Deci, 2000) distinguishes two sources of motivation: autonomous (intrinsic) motivation and controlled (extrinsic) motivation. Findings in higher education STD-based models of persistence show that controlled motivation is positively correlated to dropout, while autonomous motivation is negatively correlated to dropout (Jeno, Danielsen, & Raaheim, 2018).

A recent study on course design factors finds that the length of the program in weeks is also positively correlated with attrition rates (Jordan, 2014).

Environmental factors are related to work commitment and to a learning supportive environment.

Most studies on interaction factors within E-Learning programs draw on Moore's (1989) "three types of interaction" approach to persistence in distance education, which made a distinction between student-student, student-teacher and student-content interactions. This approach has gained recent popularity in students' performance assessment in online courses, since log data allow directly measuring the frequency and duration of different types of interaction (Lee & Choi, 2011; Buckingham Shum & Ferguson, 2012). This has led to a proliferation of new types of interaction, namely, student-self interactions, student-system interactions, teacher-content interactions, among others (Agudo-Peregrina, Hernández-García, & Iglesias-Pradas, 2012). Yang, Sinha, Adamson, and Rose (2013) found that three behavioral factors positively correlate with persistence: being enrolled in the first cohort, extent of written content production in forums, and authority centrality in social networks inferred from posting behavior in forums. Time allocation behavior, such as uninterrupted time-on-task, also correlates with longer persistence (Lee, 2018).

Studies of e-learning behavior show that course completion in online environments is quite rare, and persistence is affected by multiple entry factors and online behaviors. Although great advances have been made in identifying these factors, the interactions between them still constitute a gap to be bridged. For example, how do students' psychological attributes affect student online behavior? Answering this question may contribute to a better understanding of persistence behavior in online course and therefore to the development of new theoretical approaches.

### 3. Methodology

Participants of the program were early-stage (first three years of founding phase)

micro and small women entrepreneurs from the service sector in Germany and Ireland. We did not establish any other eligibility criteria. Women entrepreneurs who planned to start a business but had not yet done so were also included. A population of 72 participants were initially enrolled, 60 from Germany, 12 from Ireland, but only 39 subjects from Germany and 8 from Ireland successfully filled out the entry questionnaire (T0). In the end, the study sample narrowed to only 23 participants from Germany who answered T0 and gave special permission for tracking their online behavior. The age of the participants ranges from 33 to 55. A disproportional high share of women with university and/or high school degree characterizes the educational background of the participants; the branch experience ranges from 0 to 8 and the managerial experience from 1 to 25.

Two sources of data were collected to assess the WE program: the T0 and the log data of the digital platform. The entry questionnaire consisted of several scales used to measure entrepreneurial competencies (Schneider, 2017). Among the scales included are subscales such as Risk-taking (RK), Innovation (IN), Aggressiveness (AG), Entrepreneurial Self-Efficacy, and Entrepreneurial Identity among others (Schneider, 2017). The log data correspond to the participants' interactions with the digital platform. Each recorded interaction consists of a time mark, the user name, the event name and, if any, the affected user name. The structure of the data allows us to process the information and to construct relevant indicators that reflect participants' e-learning behavior. Log data allowed building five indicators to describe online interaction within the platform (Table 1).

1) Student-Student Interactions (SSI). These refer to the social and academic interaction among students enrolled in a course, included communication and content exchange, which may lead to cooperation and construction of a social space that supports learning (Agudo-Peregrina, Hernández-García, & Iglesias-Pradas, 2012). SSI can be assessed applying Social Network Analysis' centrality measures. In this study weighted out-degree will be used.

**Table 1.** Categories of student interactions.

Interactions	Event Name	Theoretical Framework
Student-Student	Blog entries viewed, Comment created, Message contact added, Message sent, Message viewed, User profile viewed	Moore (1989); Dougiamas & Taylor (2003); Lee & Choi (2011); Agudo-Peregrina, Hernández-García, & Iglesias-Pradas (2012); Agudo-Peregrina, Hernández-García, & Pascual-Miguel (2014); Proskunina, Möhring, Schneider, & Storch (2019)
Student-Tutor	Message contact added, Message sent Message viewed	
Student-Content	Book printed, Chapter printed, Chapter viewed, Course module viewed, User enrollment in course, Zip archive of folder downloaded	
Student-Forum	Blog entry added, Blog entry updated, Discussion created, Discussion subscription created, Discussion viewed, Post created, Post updated, Some content has been posted, Subscription created	Romero, Ventura, & García (2008); Fai, Mak, Williams, & Mackness (2010); Yang, Sinha, Adamson, & Rose (2013)
Student-Program	A file has been uploaded, A submission has been submitted, Course module completion updated, Response submitted, Submission created.	Agudo-Peregrina, Hernández-García, & Iglesias-Pradas (2012)

2) Student-Tutor Interactions (STI). The tutor's role in self-directed learning programs consists mainly of guiding and facilitating the learning process if a student asks for help, and this can only be achieved through one-to-one communication. These communicative interactions, STI, are measured as the total number of actions related to message exchanges between students and tutors.

3) Student-Content Interactions (SCI). Interactions with content occur when students explore and make use of available learning material in the program's digital platform. This material may include texts, videos, among others. SCI will be measured for each student as the total number of actions involving content resources.

4) Student-Forum Interactions (SFI). In E-learning, platform forums are spaces designed for creation, use, and sharing of content. Participation in forums represents the students' willingness to explore and engage with their learning spaces (Kolb & Kolb, 2005). SFI is measured individually as the total number of actions related to participation in forums.

5) Student-Program Interactions (SPI). Meeting program academic requirements such as uploading homework or completing a module is a direct measure of progress in means of exposure to content and knowledge assimilation. This measure should correlate with persistence and with effective learning. SPI is measured by the total number of activities associated with program requirements.

The log data also included information related to time allocation of students to the training program. *Sessions Mean Duration* correspond to uninterrupted time-on-task is closely related to student self-regulated learning ability and is a strong predictor of performance in online courses (Lee, 2018). In this study, we estimate the average duration of sessions on the digital platform for each participant. The measure consists of the difference in time marks between the start of a sessions (Event Name = "Log in") and the last consecutive activity (before the next "Log in"). *Total Invested Time*. The total invested time is estimated as the sum of the duration of all sessions for each participant. The quality of this measure as proxy of persistence is assumed.

## 4. Results

All statistical analysis was performed using the software RStudio. The descriptive statistics for the variables of T0 are presented in **Table 2**. In the same table the results of the validation study reported by Schneider (2017) are included for comparison.

The test results for Risk-Taking and Entrepreneurial Self-Efficacy were found not to be significant at  $p < 0.05$ . This may suggest the existence of a self-selection effect on some of the attributes of participants. Time allocation indicators from Log Database are presented in **Table 3**.

When contrasting the results of the T0 to the results of the validation study the descriptive statistics showed that the possibility of a self-selection bias among

**Table 2.** Descriptive statistics of T0.

Variable name	Code	Present Study (N = 23)			Schneider (2017) (N = 305)			t-test
		Value Range	Mean	SD	Value Range	Mean	SD	p-value
Identity Complexity	T0_IC	1 - 8	2.11	1.02	0 - 4	2.74	1.03	*2e-16
Managerial Skills	T0_MS	1 - 5	3.33	0.62	1 - 5	3.71	0.70	*0.017
Aggressiveness	T0_AG	1 - 5	2.61	0.92	1 - 5	3.31	0.97	*7e-4
Risk-Taking	T0_RK	1 - 5	3.04	0.78	1 - 5	2.78	1.02	0.088
Innovation	T0_IN	1 - 5	1.50	0.88	1 - 5	4.33	0.74	*2e-16
Entrepreneurial Self-Efficacy	T0_ESE	1 - 7	6.08	0.79	1 - 7	5.79	1.05	0.053
Self-Determined Index	T0_SDI	(-12) - (+12)	8.41	1.93				
Intention to Learn	T0_IL	1 - 5	4.43	0.84				

**Table 3.** Log database time allocation indicators.

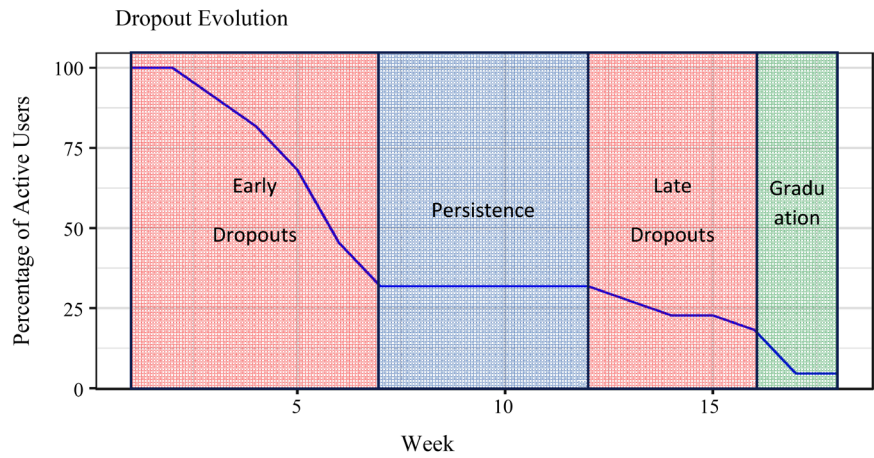
Variable name	Code	Mean	SD	Max
Sessions Mean Duration (min)	LD_SMD	30.17	18.02	72
Total Invested Time (hrs)	LD_Time	5.91	7.51	29.8

the participants could not be discarded for four indicators. This could mean that participants who enrolled in the WE program are more likely to perceive a need for improvement in those four entrepreneurial competencies, namely: Identity complexity, managerial skills, aggressiveness and innovation. Risk-taking orientation and Entrepreneurial Self-Efficacy did not show statistical differences with the results of the validation study, which may be interpreted as the participants' self-confidence in their learning skills and willingness to take risks to develop their entrepreneurial competencies. This insight may be useful in future research for early assessing the educational needs of each participant in order to personalize curricula and provide timely and more effective tutoring.

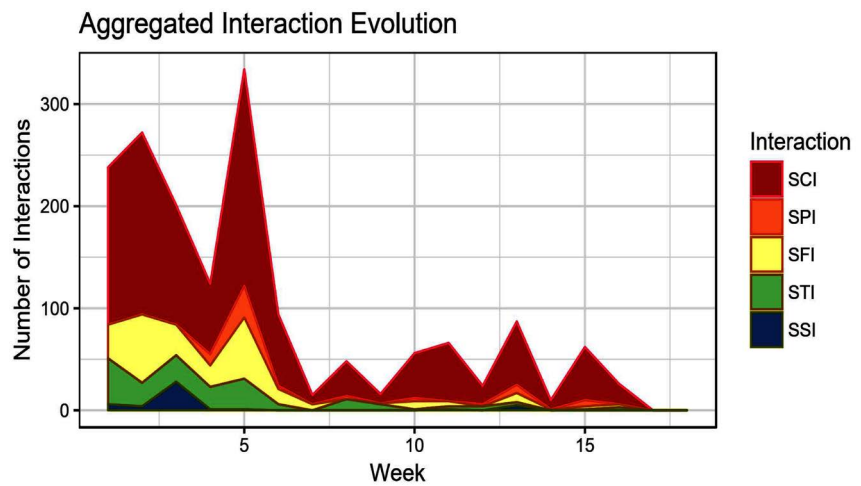
**Figure 1** depicts the evolution of dropouts; the trend shows a two-stage pattern of dropout. Since the registration date, the number of active users rapidly decays until it reaches approximately 30% at the 7th week and then it stabilizes around that number. A second dropout phase is triggered after the 13th week and it continues declining until graduation. The total number of graduated users of the program is four which is approximately 10% of the total enrolled users, and 17% of the sample of users whom gave consent to participate of the study using Log Data.

This dropout pattern is also reflected in the interaction indicators at aggregated level as is shown in **Figure 2**. After the 7<sup>th</sup> week the level of interaction drops together with the number of active users. **Kloft, Stiehler, Zheng, and Pinkwart (2014)** report a similar but smoother trend.

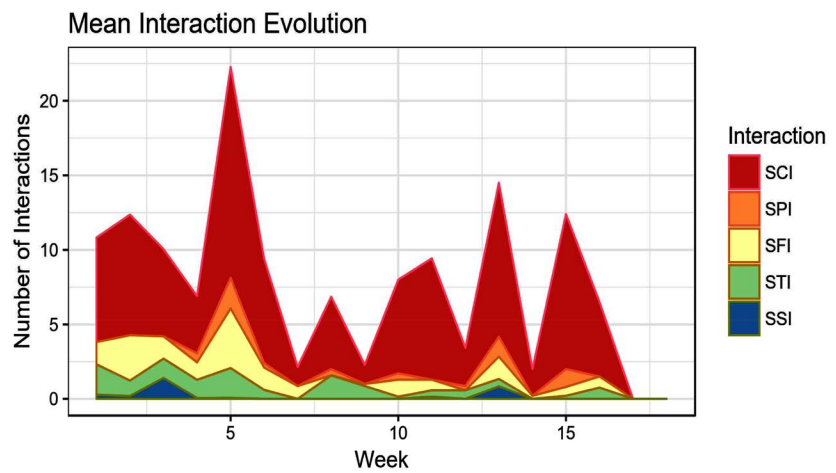
However, the reduction in the interaction frequency is disproportionate in relation to the fall of number of the active users. **Figure 3** shows how after dividing



**Figure 1.** Evolution of aggregated interaction indicators from log data.



**Figure 2.** Evolution of aggregated interaction indicators from log data. (Colored areas represent the aggregated values of each type of interaction on a weekly basis).



**Figure 3.** Evolution of normalized interaction indicators from log data. (Colored areas represent the aggregated values of each type of interaction on a weekly basis divided by the percentage of active users).



by the percentage of active users the number of interactions still drops after the 7<sup>th</sup> week. This disproportion is particularly true in the case of the indicator of student-student interactions, which drops to zero.

In order to explore differences in e-learning behavior, users are classified into three categories according to their individual level of persistence, namely: Early dropout, Late dropout, Graduated (**Table 4**).

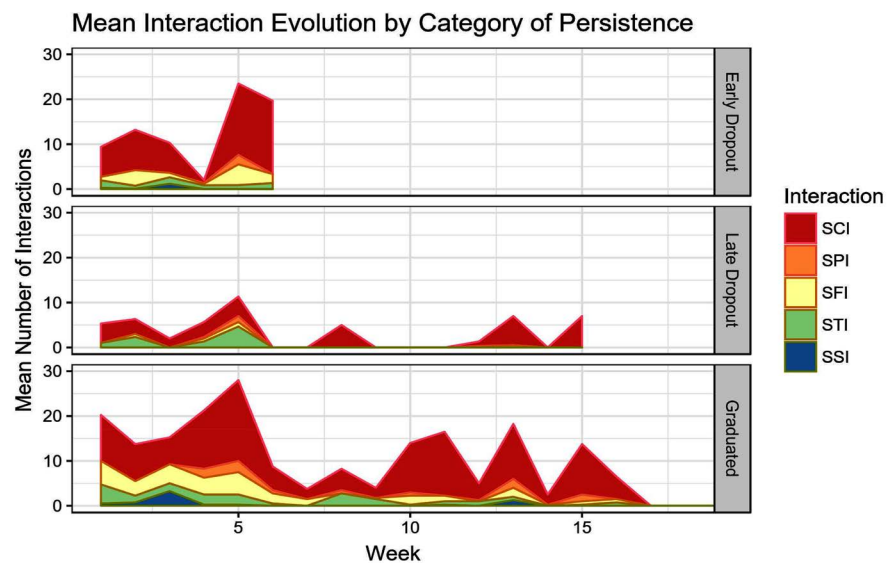
As might be expected, the results of the test show significant differences between categories for two interaction indicators: student-content (SCI) and student-program (SPI), which are explained by the cumulative effect of persistence. The more time invested in the platform, the higher the probability of interaction. However, the other three interaction indicators (SFI, STI and SSI) seem to violate this trend.

The comparison of the evolution of the indicators of interaction showed in **Figure 4** reveals some unexpected behavior. Not only the early dropouts but all three groups show higher levels of activity at the beginning of the course which tend to diminish in time. In the case of student-student interactions the indicators drop virtually to zero when it is expected to increase in time. The early dropout group shows a sudden spike in students' interactions with content (SCI) occurred in the midst of the exodus at the fifth week. At this point around 30% of the users had already left. This pattern may be a sign that students are downloading material just before departure.

Users in the late dropout group actively interact with tutors during the first weeks. However, during the rest of the course they exhibit long periods of

**Table 4.** Descriptive statistics of answers to entry questionnaire T0 by categories of persistence.

Variable name	Code	Early Dropouts (N = 16)		Late Dropouts (N = 3)		Graduated (N = 4)		K-W test
		Mean	SD	Mean	SD	Mean	SD	p-value
Identity Complexity	T0_IC	1.97	1.04	2.33	1.15	2.50	1.00	0.46
Managerial Skills	T0_MS	3.33	0.60	3.33	0.61	3.35	0.85	0.98
Aggressiveness	T0_AG	2.62	0.80	1.67	0.76	3.25	1.04	0.09
Risk-Taking	T0_RK	2.91	0.61	3.00	0.50	3.63	1.38	0.58
Innovation	T0_IN	1.56	1.01	1.17	0.29	1.50	0.58	0.72
Entrepreneurial Self-Efficacy	T0_ESE	6.17	0.69	5.89	0.51	6.08	1.42	0.61
Self-Determined Index	T0_SDI	8.13	1.86	10.2	1.46	8.13	1.77	0.20
Intention to Learn	T0_IL	4.25	0.93	4.66	0.58	5	0	0.19
Student-Content Interaction	LD_SCI	32.31	30.67	32.00	12.12	131.25	97.96	*0.03
Student-Program Interaction	LD_SPI	1.38	1.89	2.33	2.51	10.5	7.42	*0.02
Student-Forum Interaction	LD_SFI	7.75	9.21	2.26	2.08	32.5	33.72	0.13
Student-Tutor Interaction	LD_STI	4.56	4.67	9.67	2.08	18.5	21.99	0.08
Student-Student Interaction	LD_SSI	1.25	2.46	0	0	6.50	7.05	0.07



**Figure 4.** Evolution of mean interaction indicators from log data by persistence categories. (Colored areas represent the mean values of each type of interaction on a weekly basis grouped by category of persistence).

inactivity and sporadic interactions with content. In spite of having persisted for a longer time than the early dropouts, users in the late dropout group did not directly engage in communicative interactions with peers nor participated in the forum. The graduated group is the largest contributor to student-student interaction during the first weeks. After the 6th week interaction with content become increasingly predominant meanwhile communicative interaction with peers (SSI) drops. For the rest of the course graduated users remain active but with a lower rate of interaction.

A communication network is elicited from the indicators of student-student interactions and student-tutor interactions. The resulting communication network can be visualized in **Figure 5(a)**, in which red nodes represent the tutors; brown nodes represent the students who departed during the early dropout period; olive green nodes represent late dropout students; and green nodes represent the graduated students who completed the program and received their certificates. Size of nodes is proportional to the number of neighbors and the edges width is proportional to the logarithm of the total number of messages exchanged between two nodes.

The colored polygons represent the underpinning communities identified by using the community detection algorithm proposed by **Clauset, Newman and Moore (2004)**. **Figure 5(b)** shows that each community is composed of a heterogeneous set of nodes, among which tutors stand out by their centrality. By the nature of their role, tutors tend to concentrate a high ratio of the incoming messages from many students. However, this kind of communication also shapes the network structure, since a link between two students connected to a tutor is approximately four times more likely to occur than between two random students.

$$P_{\text{link between students connected to tutors}} = \frac{N^{\circ} \text{ triangles closed by a tutor}}{N^{\circ} \text{ all possible triangles closed by a tutor}} = \frac{8}{47} = 0.17$$

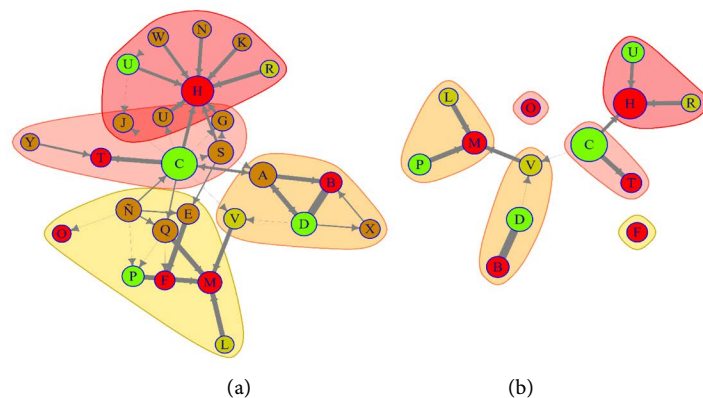
$$P_{\text{link between random students}} = \frac{N^{\circ} \text{ of edges between students} - N^{\circ} \text{ triangles closed by a tutor}}{N^{\circ} \text{ of all possible edges between students} - N^{\circ} \text{ all possible triangles closed by a tutor}}$$

$$= \frac{22 - 8}{380 - 47} = 0.042$$

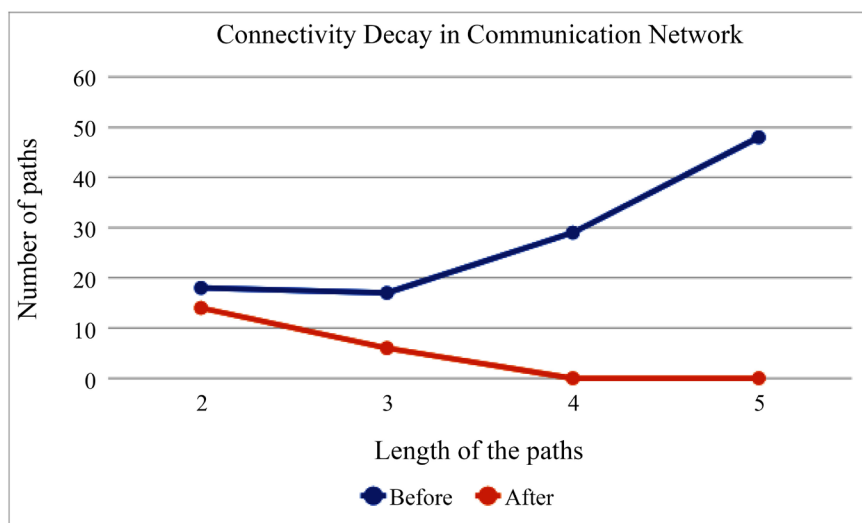
A high ratio of the early dropout nodes are located in the overlapping zones between communities, structurally these nodes play an important intermediating role. Right after the departure of early dropout users, the connection between communities resulted greatly compromised, impeding information exchange and the social communication between students, view **Figure 5(b)**.

Departure of broker nodes during the initial stage may be a plausible explanation of the drop of student-student interactions in the later periods, by causing a reduction of the social neighborhood of the remaining users and by increasing the disconnectedness between them, as shown in **Figure 6**. While the number of paths of length two does not show a serious difference between before and after the early dropouts, for longer paths the number is seriously affected.

Despite the fact that the communication network was active only during a short period of time the effect of the social interactions that took place may have affected the overall persistence of users. **Table 5** shows the first ten participants with higher levels of interaction; graduated participants are highlighted in bold. It can be noticed that graduated students who show high levels of interaction with both peers and tutors (D & C), achieve better performance in both academic progress and total invested time, than those who tend to interact only with tutors rather than with peers (U & P).



**Figure 5.** Student-Student & Student-Tutor Communication Network. ((a) Full communication network. (b) Network after removing the early dropout users. Edge's arrow represents the flow direction of messages and Edge thickness is proportional to the logarithm of the number of messages sent. Red, brown, olive and green nodes represent tutors, early dropout students, late dropout students and students graduated from the program, in that same order. Size of nodes is proportional to the number of neighbors. Colored areas represent communication clusters identified using a community detection algorithm).



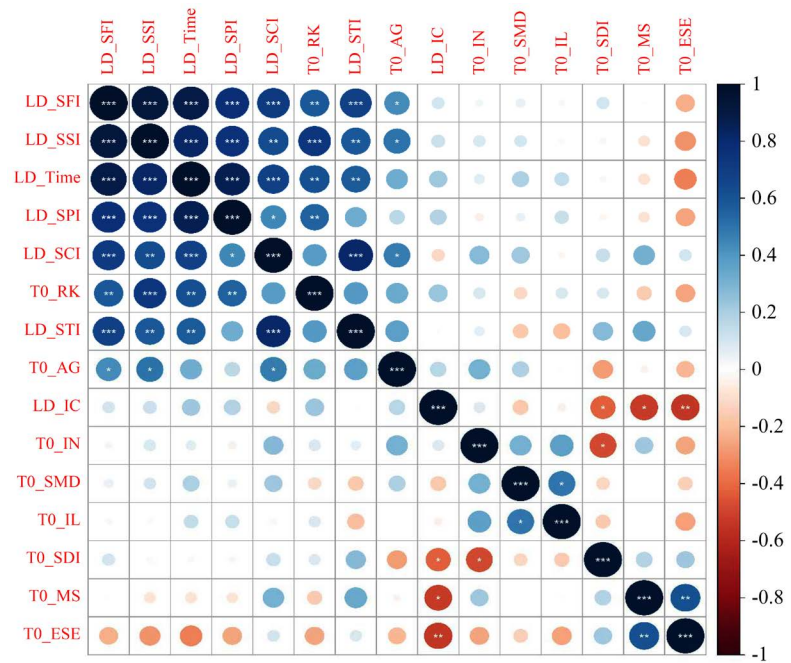
**Figure 6.** Connectivity decay in communication net.

**Table 5.** Interactions and Performance. First ten participants with higher SSI and STI values.

Participant ID	Peer Group Interaction (LD_SSI)	Tutor Interaction (LD_STI)	Academic Progress (LD_SPI)	Total Invested Time (LD_Time)	Completion (Certificate)
D	14	51	11	1567	Yes
C	11	5	20	1788	Yes
A	7	10	6	272	No
N	7	1	4	542	No
S	4	4	1	135	No
Q	1	13	1	252	No
U	1	5	9	408	Yes
W	1	5	0	157	No
P	0	13	2	537	Yes
E	0	13	1	409	No

The visualization of the correlation matrix is shown in **Figure 7**. Correlations were calculated using Pearson's correlation factor. Results show a positively correlated cluster between some variables from the T0 Questionnaire and most of Log Database Indicators.

In respect of indicators of students' interactions, a positively and significantly correlated cluster was found. The correlated indicators cluster was composed of student-student, student-tutor, student-content, student-forum and student-program interactions. Only the correlation between student-tutor and student-program interactions was found non-significant. Regarding time allocations indicators, persistence measured by the indicator of total invested time was found significantly and positively correlated with the interactions cluster. On the other hand, Intention to Learn (T0\_IL) significantly correlates with the sessions' mean duration (SMD) indicator but did not with persistence. Entrepreneurial Orientation



**Figure 7.** Correlation matrix of questionnaire indicators and log data indicators. (Color and size represent magnitude of Pearson’s Correlation Coefficient  $\rho$ . Significance levels are represented by: \*\*\*  $p \leq 0.001$  \*\*  $p \leq 0.01$  \*  $p \leq 0.05$ ).

dimensions, in particular Risk-Taking and Competitive Aggressiveness did significantly and positively correlate with some of the variables of the interactions cluster. This finding suggests an interaction between entry psychological characteristics of participants related to entrepreneurial competencies, and their E-Learning behavior during the Women Entrepreneurs program.

### 5. Discussion

In the process to further understanding of the learning behavior and the dropout in online entrepreneurship educational programs for women entrepreneurs, we have been able to collect scarce but nonetheless valuable data. In the light of the diverse theoretical streams, the results of the quantitative and computational analysis of the available data can be summarized as follows:

*Intention to Learn and the Boost at the beginning.* The intention to learn construct measured the participants’ intention to commit to a plan of time allocation which included hours-per-week and total time in months. The intention to learn is significantly correlated with longer sessions but not with longer persistence, this suggests that the construct does not have a significant effect on the equal distribution of participation over time. Analyzing the dynamics of the interaction indicators, we find that general activity spiked right after the beginning of the program. This pattern persisted even after normalizing by the number of active members in each week and separating the groups by their persistence profiles. This intertemporal inconsistency of the planned participation may be explained by the Theory of Planned Behavior as a lack of behavioral control, which

could be affected by many factors beyond the scope of this research.

*Self-Determined Motivation and Dropout Behavior.* The absence of a correlation between Self-Determination as the level of autonomous (intrinsic) versus controlled (extrinsic) motivation and persistence, measured as the total invested time, suggests that the effect of autonomous motivation has a different effect on nascent women entrepreneurs compared to higher education students. The contextual differences between higher education and entrepreneurship education may play a role for the motivational to environmental factors. In entrepreneurship education, the idea of success of nascent entrepreneurs is to keep their business growing. Consequently, the pursuit of knowledge acquisition must serve this purpose. On the other hand, higher education motivation for academic achievement is framed in a long-term incentive scheme where the idea of success is directly associated with completion and obtainment of a degree. Future research could benefit from including external sources of motivation into the environmental factors list.

*Dropout and Loss of Communication.* After the boost stage general activity drops to a lower level, this affects a higher proportion of student-student interactions. Social Network Analysis shows that the departure of early dropout participants modifies the structure of the communication network in such way that remaining participants are more disconnected than before, which could be a plausible cause to the drastic decline in student-student interactions. This result suggests that massive dropouts may have a network level effect in the communication networks.

*Risk-Taking Orientation, Social Interaction and Persistence.* According to [Nicholson, Soane, Fenton-O'Creevy and Willman \(2005\)](#), people who have set up their own business score higher on the career risk-taking, as well as financial, social and overall risk-taking. In the same study, social risk-taking is found to be highly correlated with the personality traits of extraversion and openness. This suggests that participants with higher social risk-taking levels should be more willing to explore their learning space ([Kolb & Kolb, 2005](#)) and to build up communities of practice ([Wenger, 1998](#)), which gives a straightforward explanation for the found high correlation between the risk-taking orientation and the social-communicative interaction with the peer group, through both messages and forum posts. Additionally, both risk-taking orientation (T0\_RK) and the indicator of social interaction (LD\_SSI) were found to be significantly and positively correlated with the total invested time (LD\_Time) which is assumed to be a proxy of persistence. This finding is in line with the theory since social interaction with peers is expected to be positively associated with persistence ([Lee & Choi, 2011](#)), suggesting that risk-taking orientation have an indirect effect over persistence. A second path between risk-taking orientation and persistence is related to the uncertainty condition in which entrepreneurs decide to enroll and stay in an educational and training program. Nascent entrepreneurs' time, attention and material resources are limited assets and sometimes scarce, and there-

fore engaging in entrepreneurship education should be viewed as an investment decision with high potential but uncertain return. Risk-taking orientation should increase the tendency to persist when there is high subjective perception of usefulness and return expectation for acquired knowledge. An extensive body of evidence supports the fact of gender differences in risk-taking behavior, particularly the understanding that women are more risk-avoiding than men. Our results show that risk-taking orientation is a key factor predicting persistence and performance in a female oriented online course. Therefore, retention methods in online courses for nascent women entrepreneurs should include measures for a reduction of participants' risk perception, further lowering barriers to social interaction and emphasizing the future benefits of entrepreneurship education.

Understanding attrition in online courses is a compelling task for researchers and practitioners in Entrepreneurship Education who aim to take advantage of new technologies and use them to benefit future generations of entrepreneurs. In this regard, we have centered our efforts to further the understanding of the learning behavior and dropout in online entrepreneurship educational programs for women entrepreneurs. The dropout pattern informs the practitioners in Entrepreneurship Education about the importance of establishing retention strategies before the middle and the end of the program that also affects the interaction patterns.

## 6. Limitations

The results found were obtained from a small number of observations, and this is the most serious limitation of the study and restricts the number of feasible analyses as well. Low completion rates of questionnaires T1 and T2 resulted in the impracticality of pre-test/post-test assessment. Future research should aim at implementing timely retention methods to increase completion rates.

The inclusion criteria for participants of the study were women entrepreneurs from Ireland and Germany, but only Germany data were used for analysis. Therefore, it is not possible to generalize the results obtained from this study to other countries or continents, since perceptions of the measurement instruments may vary among different cultural environment.

Gender differences in the use of online educational resources such as MOOCs have been studied with the Technology Acceptance Model (TAM) approach. This framework was not included in the present study's literature review, because it does not directly address persistence phenomena, but it may provide insights for future research. TAM's theoretical constructions such as "perceived usefulness" and "perceived ease of use" seem particularly attractive as a way to describe the persistence decision-making process. Integration of the TAM approach into persistence models may help achieve a more comprehensive understanding of enrollment and persistence in online courses.

Although drawing policy implications from the model is still not recommended, this is the first step towards an empirical model of persistence and E-learning

behavior in online entrepreneurship educational programs and can still provide valuable insights for future research and for developing retention methods in online courses.

### Acknowledgements

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 645441. The authors would like to thank the European Commission for funding this research and innovation project.

### Conflicts of Interest

The authors declare no conflicts of interest regarding the publication of this paper.

### References

- Agudo-Peregrina, Á. F., Hernández-García, Á., & Iglesias-Pradas, S. (2012). Predicting Academic Performance with Learning Analytics in Virtual Learning Environments: A Comparative Study of Three Interaction Classifications. In *Proceedings of the 2012 International Symposium on Computers in Education (SIIE)* (pp. 1-6). IEEE.  
<https://ieeexplore.ieee.org/abstract/document/6403184>
- Agudo-Peregrina, Á. F., Hernández-García, Á., & Pascual-Miguel, F. J. (2014). Behavioral Intention, Use Behavior and the Acceptance of Electronic Learning Systems: Differences between Higher Education and Lifelong Learning. *Computers in Human Behavior*, *34*, 301-314. <https://doi.org/10.1016/j.chb.2013.10.035>
- Blaschke, L. M. (2012). Heutagogy and Lifelong Learning: A Review of Heutagogical Practice and Self-Determined Learning. *The International Review of Research in Open and Distance Learning*, *13*, 56-71. <https://doi.org/10.19173/irrodl.v13i1.1076>
- Buckingham Shum, S., & Ferguson, R. (2012). Social Learning Analytics. *Journal of Educational Technology & Society*, *15*, 3-26. <https://doi.org/10.1145/2330601.2330616>
- Clauset, A., Newman, M. E., & Moore, C. (2004). Finding Community Structure in Very Large Networks. *Physical Review E*, *70*, Article ID: 066111.  
<https://doi.org/10.1103/PhysRevE.70.066111>
- Dougiamas, M., & Taylor, P. C. (2003). Moodle: Using Learning Communities to Create an Open Source Course Management System. In *Proceedings of EdMedia 2003 World Conference on Educational Multimedia, Hypermedia & Telecommunications* (pp. 171-178). Waynesville, NC: Association for the Advancement of Computing in Education (AACE).
- Fai, S., Mak, J., Williams, R., & Mackness, J. (2010). Blogs and Forums as Communication and Learning Tools in a MOOC. In *Proceedings of the 7th International Conference on Networked Learning* (pp. 275-284). Lancaster: Lancaster University.
- Jeno, L. M., Danielsen, A. G., & Raaheim, A. (2018). A Prospective Investigation of Students' Academic Achievement and Dropout in Higher Education: A Self-Determination Theory Approach. *Educational Psychology*, *38*, 1163-1184.  
<https://doi.org/10.1080/01443410.2018.1502412>
- Jordan, K. (2014). Initial Trends in Enrolment and Completion of Massive Open Online Courses. *The International Review of Research in Open and Distance Learning*, *15*, 133-160. <https://doi.org/10.19173/irrodl.v15i1.1651>



- Kolb, A. Y., & Kolb, D. A. (2005). Learning Styles and Learning Spaces: Enhancing Experiential Learning in Higher Education. *Academy of Management Learning & Education, 4*, 193-212. <https://doi.org/10.5465/amle.2005.17268566>
- Kloft, M., Stiehler, F., Zheng, Z., & Pinkwart, N. (2014). Predicting MOOC Dropout over Weeks Using Machine Learning Methods. In *Proceedings of the EMNLP 2014 Workshop on Analysis of Large Scale Social Interaction in MOOCs* (pp. 60-65). Stroudsburg, PA: Association for Computational Linguistics. <https://doi.org/10.3115/v1/W14-4111>
- Lee, Y. (2018). Effect of Uninterrupted Time-on-Task on Students' Success in Massive Open Online Courses (MOOCs). *Computers in Human Behavior, 86*, 174-180. <https://doi.org/10.1016/j.chb.2018.04.043>
- Lee, Y., & Choi, J. (2011). A Review of Online Course Dropout Research: Implications for Practice and Future Research. *Educational Technology Research Development, 59*, 593-618. <https://doi.org/10.1007/s11423-010-9177-y>
- McAuley, A., Stewart, B., Siemens, G., & Cormier, D. (2010). *The MOOC Model for Digital Practice* (pp. 1-63).
- McKinney, L., Novak, H., Hagedorn, L. S., & Luna-Torres, M. (2018). Giving Up on a Course: An Analysis of Course Dropping Behaviors among Community College Students. *Research in Higher Education, 60*, 184-202. <https://doi.org/10.1007/s11162-018-9509-z>
- Moore, M. G. (1989). Editorial: Three Types of Interaction. *American Journal of Distances Education, 3*, 1-7. <https://doi.org/10.1080/08923648909526659>
- Nicholson, N., Soane, E., Fenton-O'Creevy, M., & Willman, P. (2005). Personality and Domain-Specific Risk Taking. *Journal of Risk Research, 8*, 157-176. <https://doi.org/10.1080/1366987032000123856>
- Polanyi, M. (1962). *Personal Knowledge. Towards a Post-Critical Philosophy*. Chicago, IL: University of Chicago Press.
- Proskunina, U., Möhring, J., Schneider, K., & Storch, M. (2019). Training Female Entrepreneurs: A Self-Governed or Tutor Driven Learning Process. *International Journal of Teaching and Education, 7*, 73-91. <https://doi.org/10.20472/TE.2019.7.1.006>
- Romero, C., Ventura, S., & García, E. (2008). Data Mining in Course Management Systems: Moodle Case Study and Tutorial. *Computers & Education, 51*, 368-384. <https://doi.org/10.1016/j.compedu.2007.05.016>
- Ryan, R. M., & Deci, E. L. (2000). Self-Determination Theory and the Facilitation of Intrinsic Motivation, Social Development, and Well-Being. *American Psychologist, 55*, 68-78. <https://doi.org/10.1037/0003-066X.55.1.68>
- Ryan, R. M., & Deci, E. L. (2017). *Self-Determination Theory: Basic Psychological Needs in Motivation, Development and Wellness*. New York: The Guilford Press.
- Schneider, K. (2017). Entrepreneurial Competencies of Women Entrepreneurs of Micro and Small Enterprises. *Science Journal of Education, 5*, 252-261. <https://doi.org/10.11648/j.sjedu.20170506.14>
- Schneider, K., & Albornoz, C. (2018). Theoretical Model of Fundamental Entrepreneurial Competencies. *Science Journal of Education, 6*, 8-16. <https://doi.org/10.11648/j.sjedu.20180601.12>
- Wenger, E. (1998). Communities of Practice: Learning as a Social System. *Systems Thinker, 9*, 2-3.
- Yang, D., Sinha, T., Adamson, D., & Rose, C. P. (2013). "Turn on, Tune in, Drop out": Anticipating Student Dropouts in Massive Open Online Courses. In *Proceedings of the 2013 NIPS Data-Driven Education Workshop* (pp. 1-8). Lake Tahoe, NV.