Learning Analytics in a Virtual Learning Environment: the challenge of mapping socio-affective scenarios

Jacqueline Mayumi Akazaki, Leticia Rocha Machado, Ketia Kellen Araújo da Silva, Patricia Alejandra Behar

Federal University of Rio Grande do Sul

Brazil

Abstract

Virtual courses are increasingly being offered in Brazil, making it imperative to develop technological resources and research to help in the teaching and learning processes in this modality. One approach is to analyze student's socio-affective profile in Virtual Learning Environments (VLE). The co-operative learning network (ROODA) VLE has two features called the Social Map (SM) and Affective Map (AM), which can both contribute to the visualization of data regarding social interaction indicators and students' moods in the environment. The SM presents the social relations formed through indicators, which are the absence; collaboration; the distance from the class; evasion; informal groups and popularity, enabling the identification of the participating subjects in the form of sociograms. The AM identifies students' moods graphically through indicators, which are excitement, discouragement, satisfaction, and dissatisfaction. Thus, this article aims to map the possible recurrent socio-affective scenarios in a VLE using Learning Analytics (LA). LA is defined as measurement, collection, analysis, and reporting of data about students and their contexts to understand as well as optimize learning and the environments in which it occurs. It can also contribute to the understanding of student's learning profile, based on social and affective aspects, thus allowing the teacher to develop pedagogical strategies consistent with the needs of each subject. The importance of integrating the possible social and affective scenarios was verified using LA, making it possible to deepen the comprehension of the subjective and qualitative questions regarding the students' interactions in the VLE. In this study, the scenarios are understood as the intersection between the Affective Map and Social Map indicators identified in a VLE. It has both a qualitative and quantitative approach. The choice is qualitatively justified because the research object involves social and affective phenomena that were subjectively expressed in texts and social interactions manifested in the ROODA VLE. It is quantitatively justified by the need to measure the mapping of socio-affective indicators through social parameters and moods applying LA. The subjects were undergraduate students who participated in distance learning courses at a Brazilian public university that used the ROODA VLE in the second semester of 2019. Data were collected from social and affective maps to identify if there was a relationship between them. As a result, based on the existing indicators of social interactions and moods, the socio-affective indicators were created using LA in order to analyze the students' behavior in relation to the forms of interaction and communication that occur in the ROODA VLE.

Keywords: Socio-affective; Virtual Learning Environment; Learning Analytics.

1. Introduction

There has been significant growth in Distance Learning (DL) over the last ten years in Brazil, particularly in the consolidation of practices aimed at bringing education to all areas of the country [1]. This modality is flexible both in terms of the time and place of study, but it hinders the physical and simultaneous interaction between the parties involved in the teaching and learning process. Considerable challenges exist in terms of accompanying each student’s distance learning trajectory. Therefore, it is necessary to develop and employ different strategies to increase the success rates of DL students [2].

The lack of more autonomous and expressive performance, as well as competences such as planning and organization, make students feel unmotivated and lacking confidence to continue with their studies. The lack of interaction and close proximity to colleagues in the virtual space, often make it difficult to cope with these difficulties [3].

The field of education itself has been contemplating social and affective dimension in hopes of valuing a more integrative education. It is noteworthy that this research considers the construction of knowledge according to Piaget’s work [4, 5] in which social interactions and affectivity play specific and vital roles. From the Piagetian perspective, it is understood that social interactions form a link between the subject-environment and thus foster discussions about the learning object, enabling new cognitive structures to be built. Affect, on the other hand, is linked to the motivation to discover, creating interest in investigation, which serves as the driving force of the individual's actions [5].

According to Dolle [6], the individual learns not only through internal affective and cognitive processes, but especially from the demands caused by their social relations. In fact, the relationships established between the teacher, student, object of knowledge, and environment represent essential aspects in teaching and learning. These exchanges are intended to prepare the subject beyond the construction of knowledge, such as how to live in society. Immersed in them are affective attributes and social processes, which deeply condition the cognitive processes [5; 7].

From this perspective, it is understood that although the tools for providing DL are constantly being improved and perfected, there are still issues that require further attention. Primarily the recognition of the students’ affective manifestations and the interactions possible in these spaces. This information can provide teachers with essential elements to meet the demands of their students by offering them adequate help [8].

The need to analyze the interactions performed in these spaces is necessary in order to enhance the relationships that occur in the teaching and learning processes, understand the interests and particularities of the students, and bring the actors involved closer together. The area of Learning Analytics (LA) research emerged in 2010 as a possible solution. Originating in web analytics, it initially was intended only to serve students who presented difficulties. It is currently used to monitor and accompany all students, enabling individualized analysis [9; 10].

However, studies applying LA in the Brazilian context of VLEs are still quite recent and scarce. In fact, only 20 Brazilian educational institutions using LA were found [11]. Although LA can be applied in any modality, most of the work covers distance learning courses. For example, in the Moodle VLE, the student's navigation route is analyzed using two parameters, which allow the sections the student views and the use of available resources to be tracked. The indicators allow for the monitoring of performance, interaction, and individual student’s trajectory within the environment [12].

Hence, this research starts with a reflection on the social and affective aspects found in VLEs. In this context, the objective of the research is to map the possible recurrent socio-affective scenarios in the ROODA VLE using LA. In this study the scenarios are understood as the intersection between the Affective Map and Social Map indicators identified in a VLE. With this information, the teacher can then apply strategies and make decisions based on the student's particular socio-affective profile.

This article is organized into six sections. The following section discusses the ROODA environment, as well as Affective and Social Maps. The third section develops an understanding of Learning Analytics. The fourth describes the research methodology, and the fifth section presents the results. Lastly the final considerations are elaborated.

2. ROODA Virtual Learning Environment – Cooperative Learning Network: A focus on Affective and Social Maps

The ROODA (in English the Cooperative Learning Network) VLE began to be developed in 2000 by a research group at a Brazilian public university. ROODA is user-centered. It allows students access to materials and tools and also provides spaces for exchanging and sending activities enabling interaction between participants. Because it is an institutional environment, it is constantly updated in order to keep up with the emerging changes in the academic community [13].

The research for this article was developed in the ROODA VLE because it is the VLE used for the courses studied here. The ROODA virtual learning environment has a total of 26 synchronous and asynchronous communication features, including the Social Map (SM) and the Affective Map (AM). The SM and AM are used to identify a student's social and affective aspects. Maps are features used exclusively by professors to graphically visualize the aspects manifested by students participating in the ROODA VLE. The data are obtained from communication resources such as the journal, forum, contacts (similar to e-mail) and chat, as well as comments inserted in the Webfolio and Library [13; 14].

The AM was designed to infer the student's mood (excited, discouraged, satisfied, or dissatisfied) by analyzing their interactions in the environment and present them graphically [14]. Inference is conducted through probabilistic reasoning, based on the data collected which are subjectivity in text, actions carried out in the virtual environment, and personality traits. The teacher can therefore choose to view the student by week or month. Clockwise, the first quadrant indicates a satisfied mood; the second, excited; the third, discouraged; and the fourth, dissatisfied, as can be seen in Figure 1.

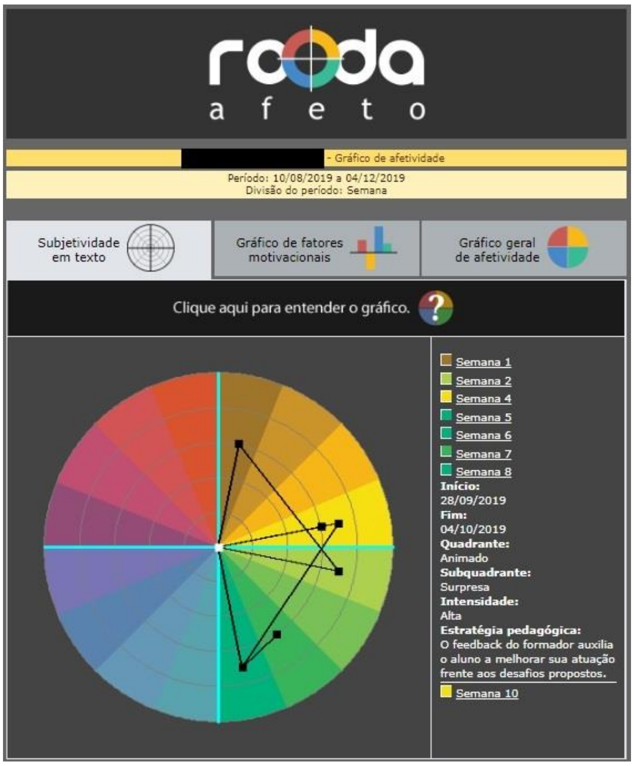

Figure 1. Affective Map. Source: https://ead.ufrgs.br/rooda/.

The first quadrant, satisfied, indicates that the student reveals satisfaction, joy, enthusiasm, and pride in the task accomplished. The second, excited, shows that the student demonstrates surprise, interest, hope and serenity in some way to face the learning challenges. The third quadrant, discouraged, suggests that the student in some way expresses sadness, fear, shame, and/or guilt for not being able to keep up with the course content. And finally, dissatisfied students express irritation, contempt, aversion, and/or envy [14].
The SM is a tool that aims to present the social relations formed in the environment, enabling the identification of the participating subjects as sociograms. This makes it possible to visually demonstrate the position occupied by the individual in the group and the nucleus of relationships that form simultaneously around them. Through a sociogram, one can therefore perceive the social position of each participant and their relationship with the rest of the group [15]. Indicators of social interaction make it possible to visualize the bonds, influences, and preferences that exist in a certain course or in a group [13]. Based on this SM, the social category level is calculated, namely: absence, collaboration, distance from the class, evasion, informal groups and popularity [16], as shown in Figure 2.

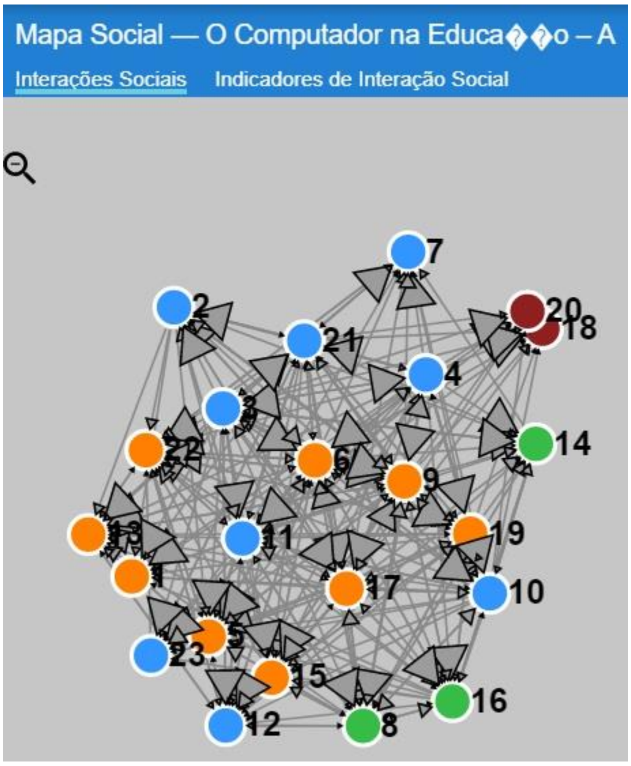

Figure 2. Social Map. Source: https://ead.ufrgs.br/rooda/.

The following section develops an understanding of Learning Analytics, the reference model, and the overall process on which this work is based.

3. Learning Analytics (LA)

LA emerged as a solution to address the need to enhance the relationships that occur through technology in the teaching and learning processes. It also helps to understand the interests and needs of students, especially to meet the needs of students that face difficulties in building knowledge [9; 17; 18]. Thus, LA is defined as the measurement, collection, analysis, and reporting of data about students and their learning contexts. It aims to bring the principal actors involved closer and analyze the students' interactions in virtual spaces [10; 19; 20].

In this study the work of Moissa et al., [19] and Dyckhoff et al., [21] are used as a reference model. This model considers four dimensions, which are: what, why, how and who, as can be seen in Figure 3.

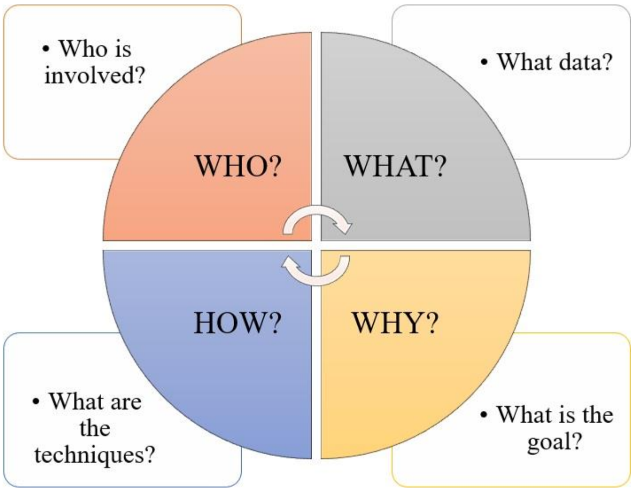

Figure 3. Reference Model of LA. Source: Based on Anna Lea Dyckhoff et al., [21] and Barbara Moissa et al., [19].

The four dimensions of this model are:

(i) **What?** Refers to the types of data collected. These can come from virtual learning environments, instructional sources, social networks, among others;

(ii) **Why?** Is related to the objectives and results of the analysis, which may be: monitoring and analysis, prediction and intervention, tutoring and monitoring, evaluation and feedback, adaptation, reflection, personalization, and recommendation.

(iii) **How?** Is linked to the different techniques that can be employed to detect patterns contained in the data.

(iv) **Who?** Is aimed at the subjects involved, which may be students, teachers, educational institutions, researchers, system designers, among others.

The types of data collected in this research come from the ROODA VLE, the objectives and results are monitoring and analysis; prediction and intervention; and customization and recommendation. The technique employed was manual mapping of the indicators and the subjects were teachers and students.
The importance of LA can be further emphasized in terms of understanding and optimizing learning, accompanying the students' learning trajectory, enabling specific and individualized analysis. Thus, the overall LA process is an interactive cycle performed through three main steps, as illustrated in Figure 4: (1) data collection and preprocessing, (2) analysis and action, and (3) postprocessing [22].

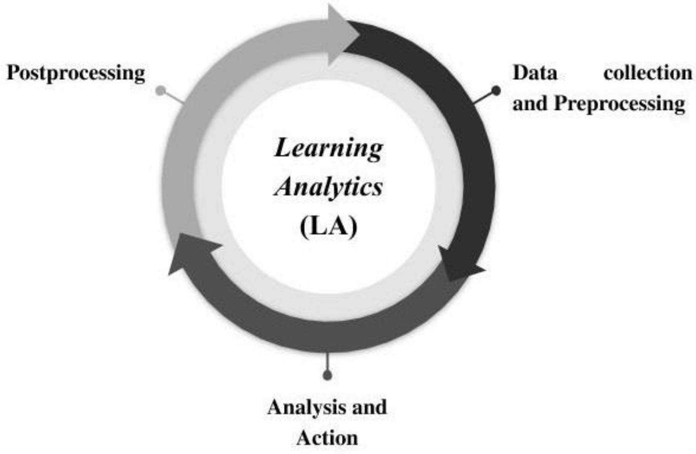

Figure 4. General LA process. Source: Based on Mohamed Amine Chatti [22].

Data collection and preprocessing: The first step involves collecting data in educational environments and systems. The second, *preprocessing*, aims to eliminate irrelevant attributes. At the end of this step the data are made available in a format that can be used as input for LA.

Once the data have been *preprocessed*, different techniques can be applied in order to discover useful hidden patterns in the data. This is the purpose of the *Analysis and Action* phase, which includes analysis, action, and visualization of the information. Taking action is the main objective of every analytical process, which incorporates adaptation, analysis, evaluation, *feedback,* intervention, mentoring, monitoring, personalization, prediction, recommendation, reflection, and mentoring.

Next is *postprocessing,* where the data are compiled and refined from additional sources, attributes for iterations are established, indicators/metrics are identified, and the analyzed variables are modified.

Thus, considering the phases of the LA, as well as the indicators mapped in the SM and AM, the methodology used to conduct the research is presented below.

4. Methodology

This research aims to map the possible recurrent socio-affective scenarios in a Virtual Learning Environment using Learning Analytics. A qualitative and quantitative case study approach was therefore developed with two groups in order to map the indicators. The case study format, based on Yin [23], was chosen because it allows for comparisons of contemporary phenomenon within context, enabling the researcher to broaden the perception of circumstances that may not be clearly evident, exploring various aspects and points of view.

The case consisted of two undergraduate classes at a Brazilian public university, totaling 33 students, over a period of fifteen weeks, corresponding to one semester. To participate in the research, subjects had to meet the following criteria:

(i) Be a student registered in ROODA;
(ii) Have basic computer knowledge;
(iii) Have access to a computer with Internet;
(iv) Sign the Informed Consent Form (ICF).

To address the ethical questions of the research, all participants were informed about the intended objectives and the methodology that would be used. The Informed Consent Form (ICF) was shared with the participants in order to formalize the research and clarify possible doubts, giving them the choice of whether or not to participate in the research. Privacy was also highlighted, emphasizing that all information including their identities would be kept confidential.

Based on the data collected in the case study, we used Learning Analytics. The following three steps were performed:

1. *Data collection and preprocessing*: Collection was performed through ROODA in two undergraduate classes. Of the initial total of 33 students, three dropped out and one canceled their enrollment, leaving 29 students. Of these, 17 were women and 12 men. Firstly, the results collected from the four mood indicators in the AM (excited, discouraged, satisfied, and dissatisfied) were organized in a table. Then, the data from each quadrant were compared with the social indicators available in the social map (absence, collaboration, distance from class, evasion, informal groups and popularity) in the same table. Lastly, the scenarios were created based on the data collected.
2. *Analysis and action*: Analysis was performed by comparing the mood and social indicators. Similarities were found in the 15-week period and all of the 29 students' data were compiled, totaling 435 socio-affective scenarios. Of this amount, a total of 95 had no student representation that week and were thus indicated as "undefined" on the Affective Map. Therefore, these did not appear that week on the Subjectivity Chart, in addition to not being linked to a social indicator. Thus, when an indicator was missing or "undefined", it was accounted for but not analyzed because the focus was on mapping interactions and communications. After removing the aforementioned cases, a total of 340 scenarios remained.
3. *Post-processing*: Lastly, the final table was refined by accounting for recurring scenarios in order to identify possible socio-affective indicators. Thus, of the 340 total scenarios, 234 did not have one of the social or affective indicators. Thus, for the purposes of this study, these were not included, since the objective was the intersection of both indicators. Therefore, a total of 106 scenarios were considered. The following section presents the results obtained from this work.

5. Results

This research aimed to present the mapping of socio-affective indicators that help Learning Analytics for use in the ROODA VLE and its Social Map and Affective Map features. Thus, for this study we considered the Subjectivity Graph based on the AM and the social indicators from the SM. The Subjectivity Graph does not show the weeks in which the subject had an "undefined" state, that is, weeks where the student did not communicate and/or did not enter the environment. Thus, it was not possible to visualize the student's quadrant during that period. The SM examined students in terms of each of the indicators every week. It should be noted that a student may have had more than one social indicator in a given week, but only one affective indicator. In this study, all possible intersections were analyzed, totaling 106 resulting scenarios.

Figure 5 was organized to present the final result of the intersections between affective and social indicators in the scenarios found. The first column presents the four indicators of students' moods, followed by the social indicators that corresponded to each week. The recurrence of combinations, i.e. the number of times each scenario appeared forming relationships between the indicators, was calculated and presented. Thus, the last column has the number of times the combination occurred in the student group. Then, the results were analyzed for recurring scenario indicators based on the combinations.

This study determined an average to determine whether the resulting crossover was standard. To do so, the number of scenarios found (106 scenarios) was divided by the number of students (29 in all). The result of this calculation in this study was 3.65, and was thus established as the cut-off value to identify the most commonly recurring scenarios that were flagged with the color blue, as shown in Figure 5. The scenarios that appear the most frequently per quadrant are presented.

Affective Indicator	Social Indicator		Number of times
excited	absence	-	**15**
	absence	collaboration	9
	collaboration	-	9
	collaboration	distance from the class	1
	collaboration	popularity	6
	distance from the class	-	1
	popularity	-	1
satisfied	absence	-	8
	absence	collaboration	4
	collaboration	-	**18**
	collaboration	popularity	6
	distance from the class	-	4
	popularity	-	1
discouraged	absence	-	2
	collaboration	-	5
	collaboration	distance from the class	1
	collaboration	popularity	8
	popularity	-	2
dissatisfied	absence	-	2
	collaboration	distance from the class	1
	collaboration	popularity	1
	popularity	-	1
	Total		**106**

Figure 5. Intersections of the affective and social indicators based on the number of times they were identified in the analysis. Source: The authors (2019).

Therefore, it is possible to analyze that over the 15 weeks, the majority of students were “satisfied” and “excited”, and the social indicator “collaboration” was identified the greatest number of times. Another analysis was the comparison of recurrence of the scenario of the mood "Satisfied" with "Collaboration", since it was repeated during the 15 weeks of class.

The moods "discouraged" and "dissatisfied" had few repetitions, almost always less than 4 and are linked to the social indicator of "absence" and "distance from the class". This indicates that these students most likely had difficulties at some point during the course, but did not drop the class.

The data also points out that many students who were “excited” were also “absent”. The hypothesis is that these occurrences were due to the pedagogical strategies adopted by the teachers during these weeks analyzed, since they were in the initial classes (2 scenarios), or in shorter weeks with holidays (6 scenarios). In other scenarios it was not possible to find a constant.

For the affective indicator “satisfied”, it can be concluded that in most scenarios (18) when the student is in this quadrant, they are also “collaborative”. On the other hand, in most scenarios (8) where the student is in the “discouraged” quadrant, they were also in the “collaboration” and “popularity” indicator. Therefore, the student interacted and even posted materials in ROODA even though they were discouraged. Finally, the student in the "dissatisfied" quadrant did not engage in social interactions with other users. Figures 6, 7 and 8 show the main intersections of the scenarios that were mapped.

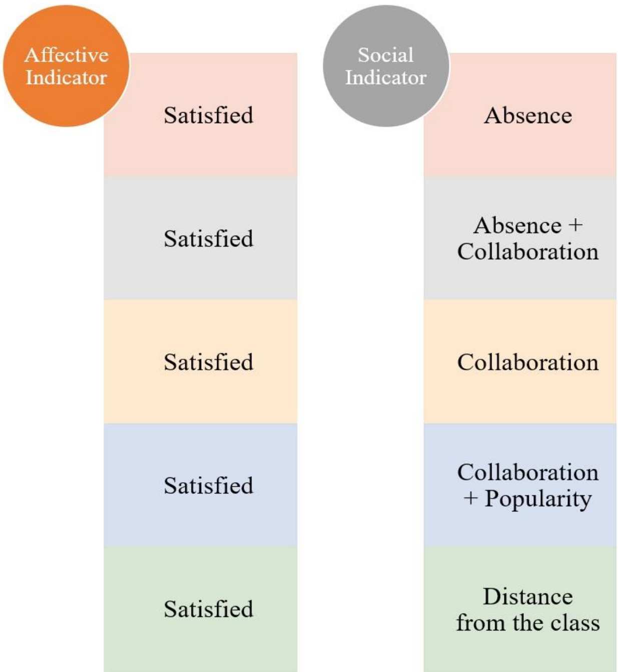

Figure 6. Main intersections of the scenarios mapped from the affective quadrant "Satisfied".
Source: The authors (2019).

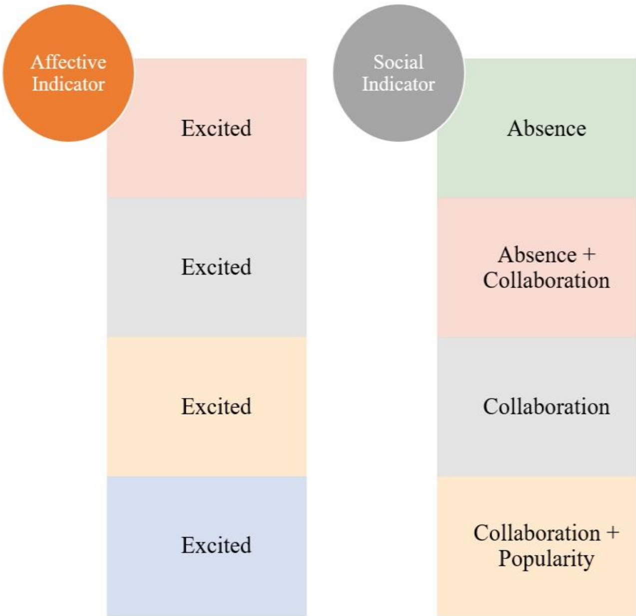

Figure 7. Main intersections of the scenarios mapped from the affective quadrant "Excited".
Source: The authors (2019).

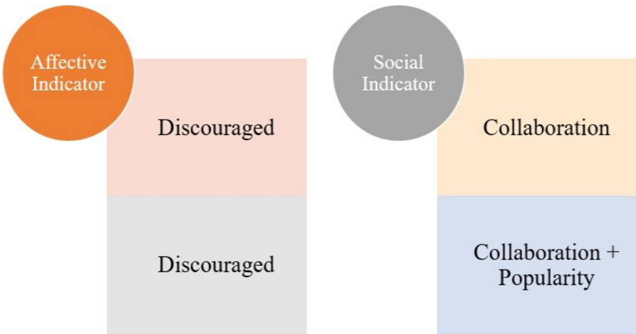

Figure 8. Main intersections of the scenarios mapped from the affective quadrant "Discouraged".
Source: The authors (2019).

Thus, the likelihood that the socio-affective scenarios found may be repeated in other situations depends on the pedagogical practices, communication tools adopted in each situation, and the student profile. Therefore, it is pertinent to perform new mappings in order to find other possible socio-affective patterns of students within a virtual learning environment for these scenarios. Hence, the main contribution of this work is to point out the existence of certain socio-affective scenarios that can be inferred and teachers can then use these findings to know the students' needs and apply pedagogical strategies based on each student's particular profile.

6. Conclusions

In DL the physical distance between the actors in the educational process makes their relationships unique. Given this, the ways of knowing the other, communicating and acting in a VLE are elements that must be analyzed for the continuous qualification of this teaching modality. Given these particularities, we must also consider the relevance that socio-affective processes play in learning. Therefore, it is necessary to develop analytical tools to provide new VLE features able to promote pedagogical practices that are more appropriate for the new educational paradigm. Indeed, it is assumed that distance education should not diminish or disregard the influence of social and affective relationships on learning. Thus, it is advantageous for the teacher to master these new tools and understand the student's socio-affective profile. Aiming to contribute to the process of knowledge construction through the socio-affective profile, this research aims to meet students' interests and individual needs as well as help teachers to analyze and extract this information. This profile will help reveal the interests and preferences for student's learning, promoting teacher reflection on the teaching process and providing resources for creating strategies and making decisions. This may help with strategies to prevent students from dropping out as well as optimize the process of analysis and definition of actions that can motivate, guide, and evaluate students.

7. References

[1] ABED. 2018. Associação Brasileira de Educação a Distância. Informações do Anuário Brasileiro Estatístico de Educação Aberta e a Distância (Abraead/2017). Recuperado em 05 dezembro, 2018, de <http://www.abed.org.br/site/pt/midiateca/noticias_ead/1526/2018/04qualidade_na_educacao_superior_a_distancia_no_brasil_onde_estamos,_para_onde_vamos>.

[2] Kalyana Monalyza Fernandes Câmara. 2016. Perfil dos alunos do curso de Pedagogia EaD do polo de Macau da UFRN: desafios, vantagens e sucessos da aprendizagem (Bachelor's thesis, Universidade Federal do Rio Grande do Norte).

[3] José Manuel Moran and José Armando Valente. 2015. *Educação a distância*. Summus Editorial, São Paulo, BR.

[4] Jean Piaget. 1973. *As operações lógicas e a vida social*. Forense, Rio de Janeiro, BR.

[5] Jean Piaget. 2014. *Relações entre a afetividade e a inteligência no desenvolvimento mental da criança*. Wak, Rio de Janeiro, BR.

[6] Jean-Marie Dolle. 1993. *Para além de Freud e Piaget: referenciais para novas perspectivas em psicologia*. Vozes, Petrópolis, BR.

[7] António Damásio. 2012. *O erro de Descartes: emoção, razão e o cérebro humano*. Editora Companhia das Letras, São Paulo, BR.

[8] Carla Barvinski, Gislaine Ferreira R. Madureira, Leticia R. Machado, Magalí Teresinha Longhi and Patricia Alejandra Behar. 2019. Construction of a Socio-affective Profile Model of Students in a Virtual Learning Environment. In *Smart Education and e-Learning,* Springer, Singapore, pp. 159-168.

[9] Ryan Shaun Baker and Paul Salvador Inventado. 2014. *Educational data mining and learning analytics. Learning analytics.* Springer, New York, NY, pp. 61-75.

[10] SoLAR. 2019. Society for Learning Analytics Research (SoLAR). Recuperado em 15 abril, 2019, https://solaresearch.org/about/.

[11] Pedrina Célia Brasil, Tainá Jesus Medeiros and Isabel Dillmann Nunes. 2018. Uma Revisão Sistemática Sobre o Uso de Learning Analytics em Ambientes Virtuais de Aprendizagem Brasileiros. In *III Congresso sobre Tecnologias na Educação (Ctrl + E)*, Fortaleza, pp. 371-380.

[12] Luciano Henrique Gomes de Almeida and Edna Gusmão de Góes Brennand. 2016. Learning Analytics em ambiente virtual de aprendizagem Moodle: um estudo de caso em componentes curriculares para cursos semipresenciais. In *Gestão & Aprendizagem*. Paraíba, 4, 2, pp. 76-93.

[13] Patricia Alejandra Behar. 2019. *Recomendação Pedagógica em Educação a Distância* (1 ed.). Penso Editora, Porto Alegre, BR.

[14] Magalí Teresinha Longhi. 2011. Mapeamento de aspectos afetivos em um ambiente virtual de aprendizagem. 2011. 273f. Tese de Doutorado (Doutorado em Informática na Educação), Centro Interdisciplinar de Informática na Educação. Universidade Federal do Rio Grande do Sul, Porto Alegre.

[15] Jacob Levy Moreno, García Bouza and Saúl Karsz. 1972. *Fundamentos de la sociometría*. Paidós, Buenos Aires, ARG.

[16] Carla Barvinski, Ana Carolina Ribeiro Ribeiro, Magalí Longhi and Patricia Alejandra Behar. 2017. Proposta de Modelo Socioafetivo de Aluno para a Recomendação de Estratégias Pedagógicas. In *Brazilian Symposium on Computers in Education,* Recife, pp. 1637-1646. DOI: 10.5753/cbie.sbie.2017.1637.

[17] George Siemens and Ryan S. de Baker. 2012. Learning analytics and educational data mining: towards communication and collaboration. In *Proceedings of the 2nd international conference on learning analytics and knowledge* (LAK' 12). ACM Proceedings, pp. 252-254. DOI: https://doi.org/10.1145/2330601.2330661.

[18] Barbara Moissa, Isabela Gasparini and Avanilde Kemczinski. 2015. Educational Data Mining versus Learning Analytics: estamos reinventando a roda? Um mapeamento sistemático. In *Brazilian Symposium on Computers in Education,* pp. 1167-1176. DOI: ttp://dx.doi.org/10.5753/cbie.sbie.2015.1167.

[19] Barbara Moissa, Isabela Gasparini and Avanilde Kemczinski. 2014. Learning Analytics: um mapeamento sistemático. In *Nuevas Ideas en Informática Educativa TISE*, pp. 283-290.

[20] Julian Moreno Cadavid and Andrés Pineda Corcho. 2018. A systematic literature review in Learning Analytics. In *Workshops do Congresso Brasileiro de Informática na Educação*, Rio de Janeiro, BR, pp. 429-438. DOI: http://dx.doi.org/10.5753/cbie.wcbie.2018.429.

[21] Anna Lea Dyckhoff, Dennis Zielke, Mareike Bültmann, Mohamed Amine Chatti and Ulrik Schroeder. 2012. Design and implementation of a learning analytics toolkit for teachers. In *Jornal de Tecnologia e Sociedade Educacional*, pp. 58-76.

[22] Mohamed Amine Chatti, Anna Lea Dyckhoff, Ulrik Schroeder and Hendrik Thüs. 2013. A reference model for learning analytics. In *International Journal of Technology Enhanced Learning*, pp. 318-331.

[23] Robert K. Yin. 2015. *Estudo de Caso: Planejamento e métodos*. Bookman editora, Rio Grande do Sul, BR.