

# Qualitative Evaluation of Student Participation in Distributed Learning Communities

Arnaldo Santos

PT Inovação SA

Aveiro, Portugal

arnaldo@ptinovacao.pt

Paulo Gomes

*AILab*, Centro de Informática e Sistemas da

Universidade de Coimbra

DEI – Pólo II

Universidade de Coimbra, Portugal

pgomes@dei.uc.pt

## ABSTRACT

Student participation in forums and other collaborative learning environments are effective ways for students to acquire new knowledge. But teachers have difficulty to evaluate, in an objective way, the student participation in this kind of environments. With this motivation, we have developed an approach for the evaluation of student participation in collaborative learning environments. This paper presents a set of measures that quantify the quality of student participation in a discussion forum. The developed evaluation model is based on the analysis of statistical data that measures the average quality of student messages. This model is integrated in a Learning Management System providing useful information about student participation for the teacher. We also present an application example and an overview about related work.

## Keywords

bLearning, eLearning, Qualitative Evaluation, Quantitative Evaluation, LMS, Distributed Learning Communities.

## 1. INTRODUCTION

The massive use of the world wide web introduced a new learning model: distance learning. The educational processes began to be seen from a different perspective, giving rise to new learning methodologies as eLearning and bLearning. The effective use of these new methodologies, and associated technologies, in educational and formation environments requires constructivist support frameworks. These frameworks must provide to the students a friendly platform for thinking, cooperation, exchanging of ideas, and learning in a virtual class environment. But teachers also have to make a similar adaptation, regarding teaching, pedagogical strategies, evaluation methodologies, and technological abilities [1].

The use of technological environments based on Learning Management Systems (LMS), can make the analysis and qualitative evaluation of asynchronous student participation in Distributed Learning Communities (DLC) complex. This problem discourages teachers to evaluate student participation in collaborative environments. With the appropriate evaluation tools, the teacher can give a higher grade importance to the student participation, providing an objective evaluation and stimulating students to participate in forum discussions.

This paper presents a set of measures that quantify the quality of student participation in a discussion forum, enabling the teacher to associate a higher part of the student grade to the forum participation. This evaluation model is based on the analysis of

statistical measures that measure the average quality of student messages, and is integrated in a Learning Management System providing useful information about student participation to the teacher.

Our approach is based on research works by Rob Phillips [2] and Andrew Stapleton [3], and was developed in the PERSONA project, a collaboration between PT Inovação and the Artificial Intelligence Lab of CISUC (Center for Informatics and Systems of the University of Coimbra). The developed model uses also statistical information about the quality of student messages.

The next section describes the role of DLC in eLearning environments. Section 3 presents research work related with our evaluation model. Section 4 describes the evaluation criteria usually used in DLC environments, and section 5 presents the developed evaluation model. Section 6 shows an application example, and then this paper concludes with some final remarks and future work.

## 2. DLC IN THE ELEARNING CONTEXT

According to Rheingold [4], a DLC represents “a group of individuals, separated geographically, that works sometimes in small nucleus, or individually, constituting communities that work in behalf of common interests”. Rheingold classifies a group of participants in a virtual community, as social aggregations and distinguishes four essential factors: the amount of persons involved that are capable of obtain diversity and communication capacity to preserve the community; the time that allows the establishment and the community's development; the human feelings, support of relationship; and the personal relationships.

Other definitions of DLC help us to understand this communication phenomenon between people, through a certain objective. As Wilson and Ryder [5] declares, DLC are “groups of people that support themselves towards their needs of learning and constitute communities of distributed learning”.

For Preece [6], an online community is a group of people that interact socially, while sharing purposes for fulfillment of their interests, needs, swap of information or services, obeying a code, protocol, ritual or norms of conduct that are accept by all and having an information system as a support infrastructure, making easier the interaction and upholding the sense of unity.

Paulo Dias [7] considered that the Web is “an excellent mean to assemble interactions in the learning communities, with their own sociability in the virtual environment, through which is possible to develop the involvement processes, the sharing and collaborative construction of knowledge”. According to the same

author, DLC represents an “multidimensional and flexible environment that allows the communication through a net where the emerging community is characterized by its dynamics in the share of interests and ideas, for the display and confrontation of the individual understandings with others community's members, transforming their practices of social interaction in a process of collaborative learning and distributed representation, creating, this way, a knowledge community”.

The eLearning and, more recently the bLearning (blended Learning), introduced in distance teaching methodologies, are presented as formative strategies and innovative methodologies that enhance the DLC’s development.

According to Santos [8], it’s possible to identify 5 appropriate components for each formation model in the implementation of a DLC for eLearning or bLearning context:

- The Materials or Contents, with scientific quality and prepared for self-learning or collaborative learning, in several formats.
- The e-teachers or Teachers, with scientific, pedagogic and technological competences.
- The Interaction Systems, appropriate for the population and the objectives of the learning that can be merely informative, advisement or of remote help.
- The Technology that should be used as a mean for learning or for teaching (platform, communications and necessary equipments).
- The Evaluation methodologies, rigorous and transparent, in a way to evaluate the students, the formation, the administration and communication systems.

This work refers, particularly, to this last one component: the pedagogic evaluation of student participation in asynchronous environments of learning.

### 3. RELATED WORK

Benign and Trentin [9] present a reflection about the online evaluation of courses. In the case of the discussion online forums, Benign and Trentin suggest a quantitative and qualitative analysis of the exchange messages.

The quantitative analysis intends to quantify the number of exchange messages, while the qualitative analysis intends to quantify the quality of the content of the exchange messages.

The authors emphasize some elements of the evaluation’s process of the courses as:

- the participation dimension
- the participants' individual characteristics
- the interpersonal communication
- the support given by the teacher
- the content of the messages and of the collaborative works
- the participants' reactions
- the social environment

For the referenced study, the most important aspect is the message content and the participation dimension of pupils.

Rob Phillips' paper [2] describes two case-studies that discuss issues involving the efficient use of electronic forums, in an

educational scope. The author starts by describing the two case-studies referring several issues involving the use of forums in an eLearning course. Later it is established a group of guidance principles used for pedagogical, planning and logistic purposes.

The element that is more related with the project PERSONA is the evaluation of the quality of the messages exchanged by students in the forum. In this field, Rob Phillips explains that a grade is attributed to the students' participation in the forum, with the objective to stimulate the use of the forum for exchanging ideas and as a learning tool.

Each one of the forum messages are evaluated and assigned a grade by the teacher. He used a scale with five categories, presented in the Table 1. After defining this scale, the author defines an evaluation table that establishes the correspondence between the student's grades and the amount and quality of that same student's messages (see Table 2).

**Table 1 – Scale of categories used for the evaluation of the messages exchanged by students in the forum.**

Category	Description
A	Excellent message, demonstrates a great understanding of the themes in discussion, making possible the orientation of the debate for new reflection areas.
B	Message of good quality, demonstrates understanding of the themes and it makes possible the progression of the debate.
C	Involvement attempt, however it doesn't demonstrate an understanding of the themes in debate, does not contributes for the discussion evolution.
D	Shows a following of the forum’s discussion.
E	Irrelevant.

**Table 2 – Grade assignment criteria for student participation in the forum.**

Grade	Number and type of messages
6	2*A or (1*A and 4*B)
5	1*A and 3*B
4	1*B and at least 3*C
3	At least 4*C
2	Any number of Ds
1	Any number of Es

The recommendations that the author makes in relation to pedagogical issues associated with the use of forums for educational purposes, are:

- The forum use should happen in a discussion context and of debate of ideas, and not of mere demonstration of knowledge. It is necessary to know how to use the acquired knowledge and to apply it in the form of arguments in a debate of ideas.

But that means that the teacher has to prepare this type of discussions based in problems or in tangible subjects.

- It is necessary to establish, in accurately and precision way, which are the objectives of these discussions, assigning them an objective in the course scope. This way, students can be aware that the debate has an objective and that their results/conclusions will be useful for a certain course activity. Grades should be used to encourage student participation.
- A model centralized in the student, where the teacher is not more than a communication facilitator serving as catalyst when the forum's participation is stagnated.

In logistic terms there are some points that should be distinguished:

- The choice of an appropriate number of students to participate in the forum is very important because it can imply the impossibility of the evaluation of all of the messages of the forum.
- The separation of the discussion by topics can be a solution for avoiding student confusion, and consequently maintaining order and organization in the forum.
- A problem that can happen is the initialization of a discussion. In this point, the author advises the existence of a first presentational meeting with all the students, so they can meet personally. Other suggestions are also presented by the author.

Related with this approach, we highlight the establishment of a scale and the respective way of evaluation of the messages. Despite this, we think that the evaluation scale is suitable, though it does not considers several other factors that we find to be important.

In the work presented by Andrew Stapleton [3], we highlight the way that messages of an online forum are evaluated. This evaluation is made in three different levels: individual, for discussion topic, and for group. The evaluation of the individual level has the objective of evaluating the interaction of the individual's discussion. At the level of the topic, the evolution of the discussion of the respective topic is evaluated. And at the level of the group, the evolution of the content's discussion of a social group is analyzed.

In the case of the individual level, it's focused two subjects: the interaction and content resulting from the interaction. Relating to specific topic analysis, the authors detach the context evaluation or the theme in debate. In terms of group analysis, the duration and intensity of the interaction of the group are points that we highlight.

#### **4. EVALUATION CRITERIA IN CLD PARTICIPATION**

The evaluation of a student participation in an educational forum involves several issues. We focus the following issues: quality of the messages sent, quantity of messages, temporal message distribution, topic message distribution, minimal participation, type of communication of each message and number of users of the forum.

#### **4.1 Quality of Messages**

One of the most important parameters to be evaluated in messages is its quality. Because this parameter is evaluated by the responsible teacher (it only makes sense this way), it's assumed that this evaluation is correct. Any way, we could analyze the quality of the messages as a whole. This way we could analyze the global consistency of the messages. To do this analysis we can use statistics measures of sample centrality (average and median), of the messages sent by the student. To complete the quality measure we also use the standard deviation as a variability measure.

#### **4.2 Quantity of messages**

The quantity of messages is easier to analyze than the quality. It's possible to perform two types of analysis: normalization and relativization of the absolute value of messages sent by the student compared with the rest of the users (this way it's possible to get an independence relative to the absolute criteria of the Rod Phillips style, that misses when the conditions changes significantly); or with the crossing of quantity with quality of messages doing a sum of the various messages classifications.

#### **4.3 Temporal and topic distribution of messages**

A better way to show data to the teacher is to present a graphical analysis of temporal and topic distribution of student messages. In this analysis, the aim is to evaluate the student's behavior in a temporal axis. This way we can locate sporadic participations and regular participations. The graphic of the messages temporal distribution can be a useful tool for the teacher justify the student's behavior.

The analysis of the topic messages distribution could indicate unbalance (or not) of participations on the various forum threads. With this option the teacher can chose between: assigning more importance to a deeper participation or the other way around. The graphic of the distribution could be a important tool for evaluation.

#### **4.4 Minimum participation**

In the topic of messages quantity, a problem that can emerge: the quality of global participation in the forum can be too low, or too high. If there is a high participation, it can be assumed that there is no problem. However, if there is a very low participation, note that the definition of low participation depends from teacher to teacher, we believe that can become a problem for the evaluation of the various participations, especially if the classification is normalized. In these situations we think that the definition of a minimal participation threshold is a possible and effective solution. The teacher can define what s/he believes by minimal participation, and can specify the minimum limit for the number messages that student send to the forum.

#### **4.5 Type of communication message associated**

The analysis of the type of message communication can be an important criterion. A message represents an act of communication between to or more forum participants. In this perspective, a message can be classified based on three elements: sender, communicative act, and destination. A simpler approach based on this classification can be the following:

- Student, answer, student
- Student, answer, teacher
- Student, asks, student
- Student, asks, teacher
- Student, asks, forum

These classifications can be used in various ways, from which we highlight two: to define a different importance (ex. answering a question can be more important than making a question, however we point out the relativity of this assumption); or the analysis of the message distribution by each classification can be used as a tool for the teacher to get more information about the student.

#### 4.6 Number of forum participants

The numbers of forum participants has a direct influence in classifications. Be it because there are too many students and the teacher has logistical problems evaluating all the messages sent. Or because there are few students and it becomes difficult to initiate forum threads. Any way, we think that the number of the forum participants is something to be considered when grading students. This factor can be taken into account using classification normalization, or with some correction factor.

### 5. PROPOSED EVALUATION MODEL

The proposed model is based on five different aspects: quality of messages, quantity of messages, minimum level of participation, time distribution of messages and topic distribution of messages. During the integration of these factors, we took into account the participation of the group, using the total number of messages as reference.

#### 5.1 Message Quality

Message quality is based on the classification given by the teacher to each message sent by the student. This classification is based on a five value scale, similar to the one proposed by Rob Phillips. This scale is defined as:

- 4 – The message denotes comprehension of the related topics beyond what was expected from the student, giving space for progress and discussion in new and relevant dimensions.
- 3 – The message denotes comprehension of the related topics, giving space for progress and discussion.
- 2 – Attempt of enrolment, however does not demonstrate an understanding of the related topics, not contributing for the evolution of the discussion.
- 1 – The message only denotes an accompaniment of the discussion, not introducing something new or relevant.
- 0 – The message is irrelevant and or of context.

Two measures are used to evaluate the quality of messages: average and standard deviation of the individual evaluation of each message. Considering that each message has a given classification ( $MsgClassification[i]$ ), average and standard deviation are defined by the formulas (where  $NMsg$  is the number of messages sent by the student):

$$AverageQuality = \frac{\sum_{i=1}^{NMsg} MsgClassification[i]}{NMsg}$$

$$QualitySD = \sqrt{\frac{\sum_{i=1}^{NMsg} (MsgClassification[i] - AverageQuality)^2}{NMsg - 1}}$$

The average is used as a measure of central tendency of the message's classification, while the standard deviation is used as a measure of variability, giving information about the consistency of the student participation.

#### 5.2 Quantity of Messages

The evaluation of the quantity of messages, sent by the student, can be done in a simple way, using the percentage of messages sent by the student weighted by the total number of messages of the forum ( $TNMsg$ ):

$$ParticipationPercentage = \frac{NMsg}{TNMsg}$$

However, this measure has some limitations, for instance, a student can achieve a good participation percentage by sending a lot of messages to the forum, even if all of those messages have a bad classification. Another measure of the participation, that also takes into account the quality of messages, is given by the formula:

$$ParticipationPercentage = \frac{\sum_{i=1}^{NMsg} MsgClassification[i]}{\sum_{j=1}^{TNMsg} MsgClassification[j]}$$

In this formula we divide the sum of classifications of the messages, sent by the student, by the sum of the classifications of all the messages in the forum. This way, we assure that a student with a bad participation, independently of the number of messages s/he sent to the forum, will have a bad evaluation in this parameter.

However, this formula still has some limitations because it evaluates the participation as a fraction of the total participation in the forum instead of evaluating it in a percentage scale.

So, we propose two alternative formulas, one that evaluates the participation of the student relatively to the best participation, and another that evaluates the participation of the student relatively to a minimum level of participation defined by the teacher.

$$Participation(StudentX) = \frac{ParticipationPercentage(StudentX)}{\text{Max}[\forall Y; ParticipationPercentage(Y)]}$$

The calculation of the participation of a student is based on his participation percentage, normalized by the best participation percentage of the group. This way, the spectrum of possible classifications goes from 0 to 100. However, this formula has the characteristic of giving always the maximum classification to the best participation of the group.

$$Participation(StudentX) = \begin{cases} \frac{ParticipationPercentage(StudentX)}{ParticipationLimit} < \\ \frac{ParticipationPercentage(StudentX)}{ParticipationLimit} < 1 \\ 1 < \frac{ParticipationPercentage(StudentX)}{ParticipationLimit} \geq 1 \end{cases}$$

The second formula uses the minimum level of participation defined by the teacher to establish the participation percentage of the student between 0 and 1.

### 5.3 Minimum Level of Participation

To deal with situations where the number of messages sent by the student is clearly unsatisfying, we find important that the teacher initially defines a level of minimum satisfactory participation. This level should be defined as a percentage value between 0 and 100, which represents the percentage of participation of the student relatively to the total participation of the group.

### 5.4 Temporal Distribution of Messages

The analysis of temporal distribution of messages uses a graphical representation to assist the teacher in the evaluation. This graphic is used to justify the grade recommendation made by our evaluation model. This graphic is relative to each forum.

The information used in the analysis of time distribution is:

- Time distribution of the messages sent by the student.
- Time distribution of the messages read by the student.

The evaluation of time distribution of messages of each forum is based on the formula:

$$\text{AverageQualityOfTimeDistribution} = \frac{\sum_{i=1}^{TNMsg} \text{MsgClassification}[i]}{N\text{Days}}$$

### 5.5 Messages Read

The number of messages read by the student is another factor to take into account. This factor qualifies the capacity of the student to follow the discussions in the group, quantified by the number of messages read by the student. There are two possibilities for the calculation of this parameter. Once a message is read, it is no longer taken into account. Or every time a message is accessed, it counts as a different read.

If a message is read several times but it is considered as being consulted only one time, we use the formula:

$$\text{ReadMsgPercentage}(\text{StudentX}) = \begin{cases} \frac{\text{ReadMsg}(\text{StudentX})}{TNMsg - NMsg} < \frac{\text{ReadMsg}(\text{StudentX})}{TNMsg - NMsg} < 1 \\ 1 < \frac{\text{ReadMsgs}(\text{StudentX})}{TNMsg - NMsg} \geq 1 \end{cases}$$

Basically, the formula considers the number of messages read by the student and normalizes the result by the total number of messages.

In the case that a message is read several times and is considered as being consulted the number of times it has been read, we use a formula based in a limit for the maximum number of read messages defined by the teacher:

$$\text{ReadMsgPercentage}(\text{StudentX}) = \begin{cases} \frac{\text{ReadMsg}(\text{StudentX})}{\text{ReadMsgLimit}} < \frac{\text{ReadMsg}(\text{StudentX})}{\text{ReadMsgLimit}} < 1 \\ 1 < \frac{\text{ReadMsg}(\text{StudentX})}{\text{ReadMsgLimit}} \geq 1 \end{cases}$$

## 5.6 Final Grade

To recommend a student grade to the teacher, all the previous factors are taken into account. The final classification of each student in each forum is given by:

$$\text{Classification} = \omega_1 \times \text{AverageQuality} + \omega_2 \times (4 - \text{StandardDeviation}) + \omega_3 \times 4 \times \text{Participation} + \omega_4 \times \text{TimeDistribution} + \omega_5 \times 4 \times \text{Presentation} + \omega_6 \times \text{ReadMessages}$$

The parameters used in the formula are explained below:

- **$\omega_1, \omega_2, \omega_3, \omega_4, \omega_5$  and  $\omega_6$ :** Are real values between 0 and 1, representing the weight of each parameter. The sum of these six weights must be equal to 1, so that classification will be between 0 and 4.
- **Average Quality:** Is the average of the classification given to the messages sent by the student. The messages which have been given a classification of 0 aren't considered.
- **Standard Deviation:** Is the standard deviation of the classification given to the messages sent by the student. This parameter is used to evaluate the regularity in the classification obtained by the student.
- **Participation:** Is the student level of participation relatively to the group. If a limit of participation is defined, every student that reaches that limit will obtain the maximum classification in this parameter. If there is not a limit of participation, the best classification of participation in the group is used as reference. The computation of the participation level is based in the quality of messages, and the quantity of messages.
- **Time Distribution:** Is the evaluation of the regularity in the participation of the student along the time.
- **Presentation:** Evaluates if the student presented himself to the group or not. In the case he presented himself to the group, he will have the maximum classification in this parameter, otherwise he will have a classification of 0 in this parameter.
- **Read Messages:** Evaluates the participation of the student relatively to the read messages.

## 6. APPLICATION IN A SIMULATED ENVIRONMENT

This research work is implemented as a .NET DLL and is integrated in the LMS FORMARE (version 4.0), providing a new feature: quantitative and qualitative evaluation of student participation in forum discussions. FORMARE allows the teacher to configure all the parameters of the message evaluation (see Figure 1), providing the teacher a flexible evaluation environment.

Parâmetros a avaliar:	Peso
A. Qualidade Média:	0,3
B. Desvio Padrão:	0,2
C. Participação:	0,1
D. Distribuição Temporal:	0,2
E. Apresentação:	0,1
F. Mensagens Lidas:	0,1

  

Parâmetros Opcionais:	
Limite Participação [parâmetro opcional]:	20

Figure 1 – Parameter selection for the evaluation formula.

After parameter definition, which is made at the beginning of the course, the teacher must evaluate all the messages sent to the forum (see Figure 2). The evaluation parameters can be changed by the teacher any time during the course.

Figure 2 – An example of message evaluation

## 7. CONCLUSIONS

This paper describes a model for evaluation of student participation in forums. Based on a evaluation scale of five values, the model measures the quality and quantity of messages that students sent and reads from the forum. The model implementation includes also the visualization of this data in a user friendly way, as graphs of temporal and thematic distribution.

The main advantage of our approach is providing an evaluation model that works as a tool for helping the teacher measuring the

student participation. These measures can be used by the teacher as arguments for student grading, enabling the assignment of a higher evaluation component to student participation. Another positive aspect of our approach is the automation of the evaluation process, relieving the teacher of complex computations.

One limitation of this approach is the selection of the weights to be used in the evaluation model. Each teacher has her/his own specific evaluation criteria, which makes weight selection a difficult task. In future work, we plan to incorporate a feedback mechanism in the model with the capability to automatically adapt the weights to the teacher preferences. This can be achieved using the teacher feedback on the suggested grades, and applying machine learning algorithms.

## 8. REFERENCES

- Mendes, I. and A. Dias. *Uma experiência de b-Learning no âmbito de uma disciplina de licenciatura da Universidade do Minho*. in *Conferência eLES '04/eLearning no Ensino Superior*. 2004. Aveiro: Universidade de Aveiro.
- Phillips, R. *Facilitating online discussion for interactive multimedia project management*. OTIS e-Workshop 2000 [cited 2005 15/8/2005].
- Stapleton, A., et al. *Analysing hypertextual discussion for connected knowing: Units of analysis*. in *8th Annual Teaching Learning Forum*. 1999. Perth, Australia: The University of Western Australia.
- Rheingold, H., *The Virtual Community: Homesteading on the Electronic Frontier*. 1993, New York: Addison-Wesley.
- Wilson, B.G. and M. Ryder, *Distributed Learning Communities: An Alternative to Designed Instructional Systems*. Educational Technology Research and Development, 1998(September).
- Preece, J., *On-line Communities: Designing Usability, Supporting Sociability*. 2000, Chichester: John Wiley & Sons.
- Dias, P., *Comunidades de aprendizagem e formação online*. Nov@Formação, 2004. 3(3): p. 14-17.
- Santos, A., *Ensino a distância e Tecnologias de informação e-learning*. 2000, Lisboa: FCA - Editora de Informática.
- Benigno, V. and G. Trentin, *The evaluation of on-line courses*. Journal of Computer Assisted Learning (JCAL), 2000. 16: p. 259-270.