Monitoring and Analyzing collaboration in e-learning environment: two case studies applied to vocational training

Felice Carugati, Patrizia Selleri, Maria Cristina Matteucci
Carlo Tomasetto, Elvis Mazzoni, Pietro Gaffuri

SEFoRA Lab – Faculty of Psychology – ALMA MATER STUDIORUM (University of Bologna)

1. Theoretical Background

Nowadays, the three key concepts that could best depict e-learning are the following: *learning objects*, *web groups* (or *web communities*) and *social networks*. We can conceive learning objects and web groups/communities as elements characterizing e-learning 1.0, while the social networks are more informal environments typical of e-learning 2.0. These three different key aspects are also representative of different ways of conceiving knowledge transmission and construction in e-learning environments. In everyday discussions, and often improperly, the concept of e-learning (*electronic-learning*) involves multiple aspects of distance education, which range from content selection to the organization and coordination of specific on-line courses. On the one hand, e-learning may be identified principally with forms of learning and training which are essentially based on interactions between group or community members. Learning processes that lie behind this mode of conceiving e-learning found their theoretical references on socioconstructivism (Doise & Mugny, 1997) and sociocultural approach to human cognitive development inspired by Vygotskij. From this point of view, individual cognitive development is conceived as a result of social interaction in which:

- the support and the sustain of either adult or expert peer/partner is determinant;
- there is the possibility of a simultaneous presence of different points of view, and the consequent necessity of a negotiation of common meanings or objects (i.e., the notion of *sociocognitive conflict*; Doise, Mugny, & Perret-Clermont, 1974; Carugati & Gilly, 1993).

In this paradigm we can insert also the social networking, one of the emergent phenomena of the Web 2.0. The essential characteristic of a social networking system is the definition of the personal relational sphere. User has to act for selecting personal contacts, for constructing the personal area and for defining which persons can access in it (usually a *blog*). Further, user can create his personal interest discussion groups and decide which persons can participate in it. Finally, user can also choose the level of the interaction: private one-to-one interaction (only the receiver reads messages), personal interaction (the message is sent to the receiver but it is public), interaction with the groups of personal contacts or interaction with the entire network.

On the other hand, e-learning is also conceived as a mere online transposition of typical educational models of face-to-face classes. According to this approach, learning is conceived as a simple content supply. Therefore, the "e" component (electronic) refers only to the content in terms of design, supply and fruition. This is the case of learning objects, by which one tends "to break educational content down into small chunks that can be reused in various learning environments, in the spirit of object-oriented programming" (Wiley, 2000, p. 7). Thus, content selection, construction and organization by educators, and content supply by web artefacts, become the very critical phases for learning processes.

2. From Individual to Collective Actions

In this contribution, the interest for the two conceptions of e-learning is connected to the monitoring of on-line actions performed by people involved in on-line courses. When we refer to actions, we consider the perspective of Leont'ev (1978) about human activity, in which activity is seen as always collective and sustained by some social motive or necessity. Each human activity is constituted by individual actions, achieved by individual or groups and directed to specific goals (fig. 1). Each action consists of operations, i.e. automatic acts without a voluntary control performed by the individual in the execution of some action (e.g., the mouse control performed by an expert user). Since actions could be performed by a single person (e.g., the student's utilization of the resources proposed by the teacher in web platform), but also by a group (e.g., the discussions in a web forum), we can consider actions as individual (a student interacts with web artefacts for downloading/reading/learning contents, e.g., a web platform,) or as collective (a student interacts with other students through web artefacts, e.g., a web forum).

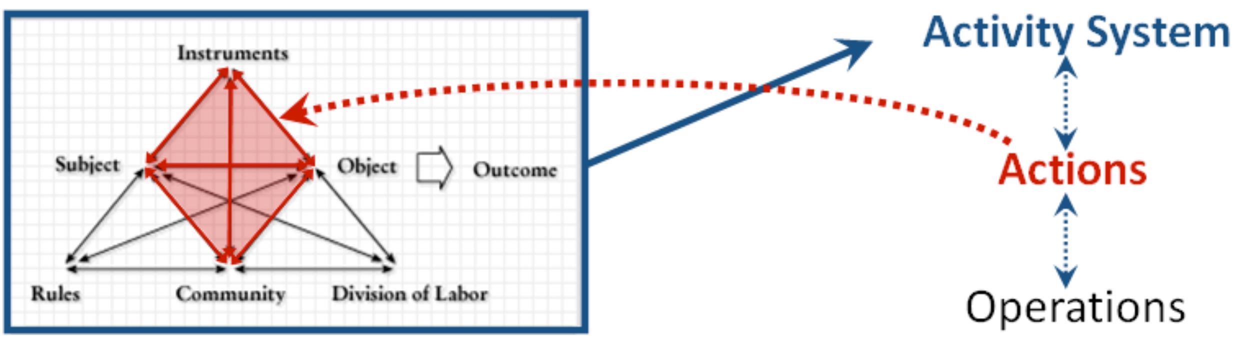


Figure 1: Activity system representation from the web site of the Centre for Activity Theory and Developmental Work Research (http://www.edu.helsinki.fi/activity/)

In all of these cases, actions related to the student's activity may be considered in terms of competence acquisition; in fact, actions involve the use of web artefacts for knowledge acquisition (individual actions), and the manage of on-line interactions with others for collective knowledge, sharing and construction (collective actions). If we consider the importance of competences and learning outcomes both at university level (e.g., Dublin Descriptors), and at vocational level (e.g., the Lifelong Learning Programme 2007-2013 launched by the European Union, in which web technologies are seen as one of the key tools for achieving the objectives of the programme - Pépin, 2007), we may easily realize that this issue is crucial not only in the field of academic research, but also in the field of European policies.

Starting from these considerations, how can we monitor students' on-line action in e-learning environments?

3. From Web Tracking to Social Network Analysis

A quantitative technique for data collection about "what user do" in an on-line environment is the *web tracking* (Calvani, Fini, Bonaiuti, & Mazzoni, 2005; Mazzoni, 2006; Proctor & Vu, 2006). Through web tracking it is possible to collect a number of data about the frequency of visits and time spent on web pages during the surfing on web artefacts (e.g., web site or web platforms). This data collection technique is a feature that characterizes almost all of the existing web platforms, and it is also provided by the Italian legislative decree concerning Distance University as means for monitoring and evaluating students' on-line activities. If, on the one hand, we can consider web tracking as a good technique for collecting data about individual actions, i.e., about the frequentation and the usage of web contents by students (e.g., learning objects, documents, hypertexts, etc.), we cannot affirm the same as far as the application of this technique to web communities or social networks is concerned. Of course, web tracking allows us to collect data on interactions between students, which may consist, e.g., of sent or received messages or replies. However, these data refer to individual level (how many messages a student has sent, received, etc.) and do not provide any indication about addressees. Relational aspects, therefore, are not taken into consideration by web tracking. Nevertheless, this information is available. In other words, web tracking may be employed also in order to collect data about to whom a message/reply is sent, and about the identity of the receiver of a given message/reply (the so called relational data), but these data are normally used only for summing and displaying the quantity of messages sent and received by single students.

Now, if we consider web groups or web communities in e-learning environment, we have to consider that the final outcome of a collective activity does not derive from simple individual actions, but principally from collective actions performed by the group/community. In this case we consider individual actions as separated from collective actions, and we have to take into account that group performance does not derive from a sum of individual actions, but rather from indicators that allow us to map the collective actions of a group/community.

As previously outlined, relational data of web group/community could be collected by web tracking; this possibility, besides facilitating the application of quantitative analysis, allows to construct the adjacency matrix of relational data for applying the *Social Network Analysis* (SNA) to group exchanges. Starting from the transposition of relational data in a matrix, SNA allows, on the one hand, to graphically represent the network of relations by sociograms and, on the other hand, to transform this network in concepts for describing the communicative structure of the network. Now, a very interesting aspect is that we can develop an analysis considering two levels, i.e., by focusing on the single members and their relations in the network (*ego-centred analysis*) or by focusing on the network and its structural characteristics (*whole network* or *full network analysis*). Obviously, these two aspects are related. This means that for each whole network structural indexes we have also specific individual measures. E.g., the density of a network, i.e., "the proportion of possible lines that are actually present in the graph" (Wasserman & Faust, 1994, p. 101) or more simply the percentage of aggregation of its members, derives from the degree of each member, i.e., the totality of direct contacts he/she has activated or received by others. Considering the centralization, i.e., the dependence of a network from its "most important" actors, we have, together with this whole index, also the centrality index of each member, i.e., his/her importance/prominence for the communicative structure. Thus, these related networks and individual measures allow us to perform map description of collective actions of a community. On the one hand, we can monitor and depict the role and function of each member in the community knowledge exchange (e.g., wideness and aggregation of his/her neighbourhood or direct contacts, central or peripheral role in information exchanges/transmission, participation in subgroups, etc.); on the other hand, we can monitor the group/community while considering the aggregation of the communicative structure, the reciprocity in discussions, the number and density of possible subgroups, etc..

4. Web Tracking, SNA and Web artefacts for online communication: some critical issues

Compared with classical analysis models based on the characteristics of single subjects, like those coming from web tracking collection, the SNA focuses attention on relational data such as links, contacts or ties between a group of people or organizations such as family, associations, societies, nations etc.. Application of the SNA starts with the creation of an *adjacency matrix* (e.g., with contacts, connections or exchanges within a particular social network or a web forum). The relational interaction network data is presented in the resulting adjacency matrix where each line corresponds to the sender of a message and each column to the recipient. The interaction box between sender and receiver shows the number of messages exchanged, or the existence (1) or absence (0) of a contact (fig. 2).

		Receivers					
		Stud1	*Stud2*	*Stud3*	*Stud4*	*Stud5*	*Stud6*
Senders	Stud1	0.0	0.0	3.0	4.0	0.0	0.0
	Stud2	0.0	0.0	2.0	0.0	0.0	0.0
	Stud3	3.0	4.0	0.0	0.0	0.0	0.0
	Stud4	2.0	0.0	3.0	0.0	0.0	0.0
	Stud5	0.0	0.0	0.0	0.0	0.0	0.0
	Stud6	1.0	3.0	2.0	2.0	0.0	0.0

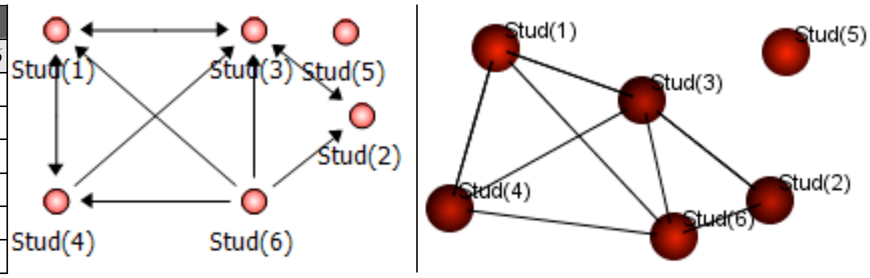


Figure 2: Adjacency matrix of students' exchanges between a web forum and bidimensional and tridimensional sociogram representation elaborated with Cyram NetMiner 3 (2007).

Despite the apparent simplicity, creating an adjacency matrix is strongly influenced by the type of communication found in the web artefact used. The first important aspect to consider in creating an adjacency matrix is a correct understanding of the terms *sender>recipient*. In the case of an e-mail, it is not difficult to identify the sender/receiver, as well as the contents of the message, for whom it is explicitly intended and, therefore, the connection between a reply and the message it refers to. It's the same with social networks personal communications and contacts, in which there is an explicit definition by the sender of the receivers of messages or contacts. This is more complicated when considering the adjacency matrix referring to web forum exchanges which have the essential characteristic of sharing information, contents and resources between participants, with the aim of reaching a common objective. Many important consequences arise regarding the definition of sender/receiver. Above all, gathering relational data to create an adjacency matrix in relation to web forum exchanges can be carried out in two principal ways:

- tracking the web data in its specific data base, e.g. *a log file*;
- based on a content analysis, which allows recipients of web forum messages to be identified.

As far as web tracking is concerned, the correct identification of senders/receivers must take into consideration two important factors in the creation of an adjacency matrix: 1) the opening messages in a discussion are not tracked as they do not have a specific receiver, but are posted to a common area; 2) the display interface of messages posted on a web forum influences the way data is tracked. As web forums are based on the idea of sharing and collaboration within the group or community, therefore, unlike e-mails or personal messages and contacts on social networks, these messages do not explicitly identify a receiver but are posted to the entire group or community.

Although it is possible to track the data regarding the display of a message on the screen, a correct understanding of it depends on the type of display interface used by the web forum. Two different interfaces are considered:

- SI (*Sequential Interface* of messages),
- DDI (*Deep Discussion* Interface).

The SI web forum makes it impossible to identify which messages a user has read as he/she clicks only on the discussion for opening all the messages and not on the single messages for seeing it. In the DDI web forum, on the other hand, the user has to click on the message to display it, thus facilitating the creation of an adjacency matrix on the basis of messages read. The user is also well aware that the reply command refers only to that message and the answer is linked to it. As can be seen, a rather critical picture emerges regarding the use of relational data from web tracking to create an adjacency matrix and apply it to the Social Network Analysis. Despite the criticisms outlined, the adjacency matrix built on the DDI basis of reading messages is a reasonably efficient way to represent the network of information and contents created within a web group/community.

Alternatively, the researcher can always base his/her research on an analysis of the message contents. Compared with tracking, the analysis of message contents allows a much more precise definition of sender, and above all receiver. As outlined above, the SI web forum often has single messages or replies which in reality have more receivers than the one who is automatically linked to the message through tracking. The content analysis allows creating an adjacency matrix starting with the sender and therefore re-establishes the correct direction of the interaction, i.e. sender>receiver. A content analysis also makes it possible to differentiate between the types of message content and therefore to create adjacency matrixes with multiple factors regarding the information and content networks activated (e.g. social support, help requests, information, awareness building etc.).

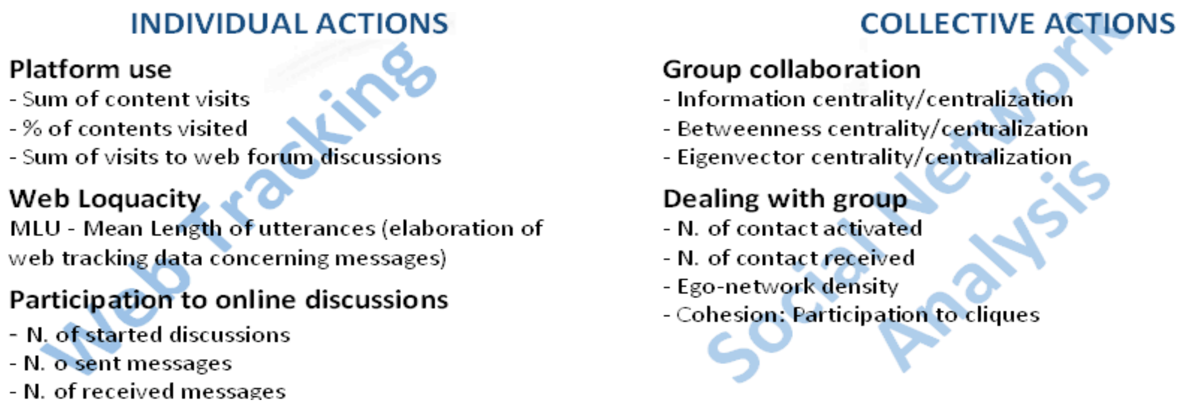
Despite these positive aspects of content analyses, it should be borne in mind that this type of analysis cannot give information about who has opened and read a message since it is taken almost for granted that the message has influenced only those who have answered it, while at the core of web forums are the presuppositions of total sharing and collaboration. Therefore, even the simple reading of a message can have a profound influence on the reader, albeit in a latent manner, and on any of his/her subsequent postings. Therefore the reading of messages appears to have a bearing on understanding how information and content is propagated within the group/community/network considered.

5. A model for monitoring on-line activities

In order to illustrate how web tracking data and SNA indexes may be utilized, we proposed a model for monitoring and analyzing on-line activities used during three blended-learning experiences (two groups of teachers in vocational training and one group of university students). This model permits to represent individual and groups profiles based on both individual (coming from web tracking indicators) and collective (SNA indexes) actions.

The first step of our model is the data collection by web tracking. These data have been elaborated for having indicators that are not simple frequencies, but measures of individual's actions in web environment. The second step is the application of Social Network Analysis (SNA) to web interactions carried out on collective discussions in web environments. The model we propose is constructed considering both types of indicators. Web tracking data are considered indicators of individual actions, while the SNA indices concern two levels: collective indices of the activities carried out by groups but also individual indices about the role members play in collective e-Learning activities.

The model consists of 5 areas of actions: 3 areas of individual actions, collected by web tracking (platform use; loquacity; participation to discussions) and 2 areas of collective actions collected by SNA (role in group collaboration; dealing with group).



All web tracking indicators and SNA indexes have been elaborated so that we could obtain a graph for each participant, which describes his/her actual performance levels in each area in relation to the maximum performance level attained by his/her group. The same may be done for the entire group, in order to obtain a graph displaying the average performance of participants in each area in relation to the maximum performance level attainable by the group (fig. 3).

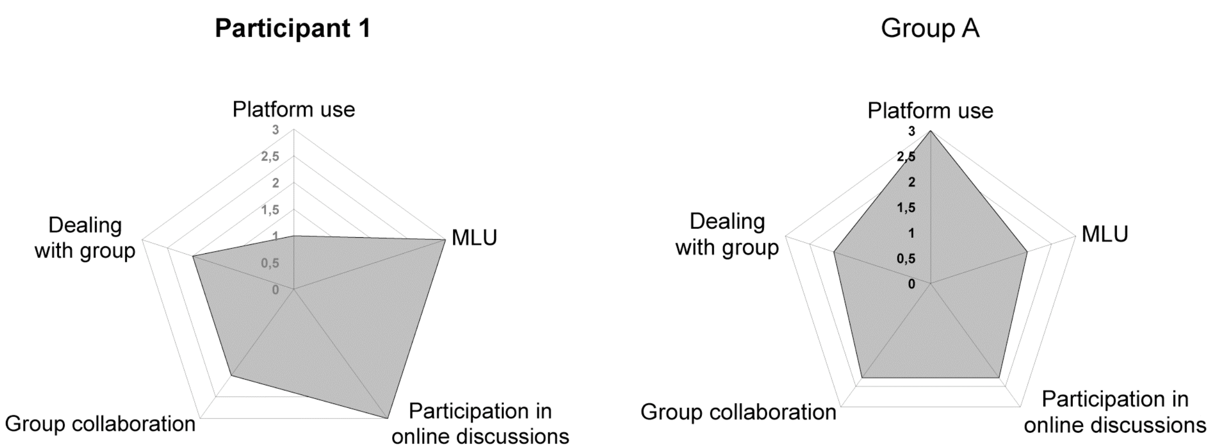


Figure 3: An example of performance attained by a participant and by his/her group.

By collecting and analyzing data longitudinally, e.g., every month for the during of a web courses, this model allow us to monitor and describe the evolution of the student's/group's performance in a determined period within his/her Zone of Proximal Development (Vygotskij, 1978).

In summary, this model allows us to take into consideration and represent not only the individual actions a student performs within an e-learning environment, in order to interact with contents, but also the collective actions he/she accomplishes for interacting with his/her colleagues during on-line group collaboration. Further, as we show in figure 3, we can use this model for representing group performances, and thus for comparing different groups involved in virtual learning environment characterized by collaborative activities. Finally, with this model it's possible to monitor and analyze the evolution of student's/group's performance during the period of a specific e-learning courses.

Bibliography

Carugati, F., & Gilly. M. (1993). The multiple sides of the same tool: cognitive development as a matter of social construction of meaning. *European Journal of Psychology of Education*, *8*(4), 345 - 354.

Doise, W., Mugny, G., & Perret-Clermont, A. N. (1974). Ricerche preliminari sulla sociogenesi delle strutture cognitive. *Lavoro Educativo, 1,* 33-50.

Doise, W., & Mugny, G. (1997). *Psychologie sociale et développement cognitif*. Paris : Armand Colin.

Vygotsky, L.S. (1978). *Mind in Society: The development of higher psychological processes*. Cambridge, MA: Harvard University Press.

Wasserman, S., Faust, K. (1994). *Social Network Analysis. Methods and Applications*. Cambridge: University Press.

Wiley, D. A. (2000). Connecting learning objects to instructional design theory: a definition, a metaphor, and a taxonomy. In D. A. Wiley (Ed.). *The Instructional Use of Learning Object* (pp. 1-35). Bloomington, Indiana: Agency for Instructional Technology And Association For Educational Communications Of Technology. Retrieved March 19, 2008, from http://reusability.org/read/