Detecting Self-Regulated Learning in Online Communities by Means of Interaction Analysis

Giuliana Dettori and Donatella Persico

Abstract—Interaction analysis is increasingly used to study learning dynamics within online communities. This paper aims to investigate whether Interaction Analysis can help understand the practice and development of Self-Regulated Learning (SRL) in Virtual Learning Communities (VLCs). To this end, a set of SRL indicators is proposed to spot clues of self-regulated events within students' messages. Such clues have been identified and classified according to Zimmerman's SRL model and some subsequent studies concerning SRL in Technology Enhanced Learning Environments (TELEs). They have been tested on the online component of a blended course for trainee teachers, by analyzing the messages exchanged by a group of learners in two modules of the course. The results of this analysis have been compared with those of a previous study carried out, with more traditional methods, on the same course. The similarity of the results obtained by the two approaches suggests that Interaction Analysis is an effective, though rather labor-intensive, methodology to study SRL in online learning communities.

Index Terms—Collaborative learning, computers and education, distance education, education.

1 Introduction

This paper proposes the use of Interaction Analysis (IA) for investigating the practice of Self-Regulated Learning (SRL) in Virtual Learning Communities (VLCs). This technique allows one to gather data of a different nature than those obtained with traditional methods, such as questionnaires and interviews. Hence, it offers the possibility of complementing and reciprocally validating the outcomes of different studies.

SRL is based on a set of relevant cross-curricular skills able to facilitate learning at all ages and in different learning situations. Its potential, which is illustrated in Section 2, makes it a central topic of interest for the improvement of education.

VLCs, and in general Computer-Supported Collaborative Learning (CSCL), are a way of learning that has been increasingly gaining attention and diffusion in the past decades. Their main features and relationships with SRL are described in Section 3. This way of learning is likely to further grow and expand in the near future, due to the continuous improvement of Web technology and the increased attention to social practices induced by the diffusion of Web 2.0 applications. Analyzing learning in such environments is therefore a major issue of educational research in the current technological, cultural, and social contexts.

IA is a research method that can be successfully employed to investigate the dynamics indicated by written interactions between subjects, for example, in collaborative activities in online learning environments. Therefore, this method is increasingly applied to the analysis of learning dynamics in CSCL, as explained in Section 4. It can be applied to a wide variety of learning-related aspects, provided one has at their disposal a set of indicators related to the aspect of interest.

In this paper, we propose a set of indicators of SRL that allowed us to analyze students' interactions in order to investigate the self-regulation of online collaborative learners. This is described in Section 5.

We also report, in Section 6, on the application of these indicators in an exploratory study on an online teacher training course in Educational Technology. The outcomes of this study are then discussed and compared with those of a previous study carried out with more traditional means (questionnaires).

Finally, in Section 7, the feasibility, reliability, and cost-effectiveness of the IA approach are evaluated, with the aim to encourage its diffusion and application on larger and diverse sets of data.

2 What Is Self-Regulated Learning

The term SRL identifies a process based on a set of competencies allowing learners to improve their learning efficacy, as well as to apply and adapt their knowledge and strategies across different subjects. Research in this field investigates the behavioral, emotional, motivational, cognitive, and metacognitive aspects involved when students try to control their own learning processes [1], [2], as well as the pedagogical approaches that can help learners gain and improve self-regulation competence.

SRL is not a mental ability nor an operative skill but rather a student-directed process that transforms mental abilities into operative skills in relation to a specific task [1] and in a given context [3]. Self-regulated learners master and deliberately control their own learning by setting their own learning goals, choosing and applying different learning strategies according to such goals and reflecting on their own learning, as well as evaluating their progress and consequently adapting their plans, in a cyclical process.

• *The authors are with the Institute for Educational Technology (ITD), Italian National Research Council (CNR), Via De Marini 6, 16149 Genoa, Italy. E-mail: {dettori, persico}@itd.cnr.it.*

Manuscript received 21 Mar. 2008; accepted 19 June 2008; published online 25 July 2008.
For information on obtaining reprints of this article, please send e-mail to: tlt@computer.org, and reference IEEECS Log Number TLTSI-2008-03-0026. Digital Object Identifier no. 10.1109/TLT.2008.7.

They are often intrinsically motivated, have a good degree of self-efficacy, and see learning as a proactive activity; in other words, they actively control rather than passively endure the learning process. It is not surprising, therefore, that SRL has rapidly gained attention in the educational field over the past couple of decades, because it appears a fundamental component of both academic success and the ability to effectively cope with lifelong learning needs.

Such a wide range of competences obviously requires time and care to develop. The literature indicates that some aspects of SRL, such as metacognitive knowledge and skills, generally improve as students get older. It also points out, however, that the acquisition of general SRL competence is not automatic nor spontaneous [4] but, rather, requires suitable teaching and practice. Several authors suggest that it should be explicitly fostered, by including it in classroom instruction [5]. This can be done by setting up flexible, student-centered environments promoting active learning [6], providing students with suitable feedback, and encouraging them to evaluate their outcomes and revise them consequently. In order to become self-regulated learners, both individual and social learning experiences appear to be necessary [7], [1].

Moreover, the literature reports that SRL competence is, to some extent, context dependent: it certainly includes cross-curricular components that may be applied in all contexts, such as metacognition, self-efficacy, and awareness of the importance of using effective learning strategies, but part of these skills and abilities depend on the learning context [3]. For instance, people who are very effective in individual, traditional learning may not be as good at learning collaboratively, let alone learning collaboratively online, because this approach entails negotiating objectives, strategies, and concepts, which is rarely practiced in individual learning.

Research into SRL is currently carried out by analyzing students' observed actions, that is, by trying to understand to what extent they set their goals, plan their learning, evaluate their progress, and practice metacognition and self-reflection. Such investigations mostly rely on interviews where learners are requested to describe, ex post, the strategies and methods they used during the learning process, or on questionnaires aimed at eliciting information from the learners' about their strategic planning and the other choices made during the leaning process. A checklist to analyze the features of technology-enhanced learning environments (TELEs) was also proposed [8], to evaluate, possibly a priori, whether a TELE potentially supports the practice of SRL.

It should be noted that none of such methods are able to directly evaluate the practice of SRL, but rather they try to deduce its presence from students' opinions. A research method allowing a direct analysis of the learning process, based on the interactions taking place throughout it, would therefore yield data that could usefully complement those data that are mediated by the subjects' post hoc reflections.

3 Learning in Online Environments

3.1 Characterizing Online Collaborative Learning

VLCs and CSCL deal with the implementation of collaborative learning in online environments. Both rely on Computer Mediated Communication (CMC) to support group interaction at a distance among trainees, with the guidance of facilitators and tutors.

In such environments, communication takes place mostly in a textual and asynchronous way. This has important consequences on how learning is stimulated and takes place. In written communication, interaction times are dilated hence participants have the possibility to reflect, before sharing their ideas with their peers, for longer than occurs in oral interactions. Moreover, all contributions, which are posted in forums or blogs, remain at the disposal of all participants, hence facilitating precise reference during discussion as well as further revision and reflection [9]. Finally, the possibility to carry out more than one discussion stream at a time gives space to everybody to actively take part in the opinion exchange. These three features facilitate the implementation of socio-constructivist learning activities much more than can be done in face-to-face courses with a high number of students.

3.2 CSCL and SRL

The relationship between CSCL and SRL is quite complex because effective use of CSCL environments appears both to require and to improve the ability of learners to self-regulate their own activity [10].

There are many reasons why CSCL is believed to foster certain SRL skills. First, SRL competence, and in particular metacognitive skills, is often among the explicit or implicit objectives of CSCL learning activities. This is primarily due to the fact that learners who are new to this training method usually lack some of the metacognitive and self-direction skills needed to take full advantage of this learning approach. Well-designed courses, therefore, try to stimulate learners in this respect. Moreover, learning with CMC is heavily based on textual interaction, and this supports reflection not only on content but also on the learning process itself. As a consequence, such learning environments foster the practice of SRL by putting into play several SRL-related skills, to the point that they are regarded as promising for its development [11], [12], [13]. At the same time, some initial SRL competence is necessary in order to make good use of learning experiences within VLCs not only because students need to control time and pace of their learning process but also because collaborative activities entail negotiating objectives, strategies, and concepts with peers.

4 Investigating Learning Dynamics

CSCL environments lend themselves very well to investigate learning dynamics because interaction is in written form. Moreover, a variety of information is available to researchers due to the fact that communication platforms usually record meaningful events, such as logins and logouts, access to folders and opening messages, downloads and uploads, and so forth. Several research studies, therefore, use IA to investigate learning dynamics in CSCL. In particular, a research methodology that has been increasingly used for this purpose is Content Analysis [13], [14]. It consists in detecting phrases and expressions that reveal aspects of interest in the written messages exchanged by the learners. This allows one to analyze and elaborate data about the frequency and nature of the detected expressions, therefore

combining qualitative analysis of individual messages with quantitative elaboration of results. This method, taking advantage of the nonintrusive capability of CMC to track events during the learning process, can potentially replace or at least complement other, more traditional ways for gathering data on learning, such as questionnaires and interviews. For this reason, IA is considered a powerful source of information and is increasingly applied in research on Web-based learning, even though it takes a large amount of time to extract data from the messages, as well as to analyze and interpret them [15]. In some cases, parsing techniques can be of help, but only if some specific expressions can be identified that consistently and exhaustively characterize the clues searched for.

Content analysis may be used to investigate different aspects of learning, of both cognitive and affective kinds, therefore looking at content of various nature [13]. The variables investigated may be manifest, that is, visible and objectively recognizable, or latent, i.e., implicit in message content.

Manifest variables are related to explicit communication features, and therefore, they are easier to detect. An example of manifest content is the number of times students address each other by name. In general, manifest content can be investigated with a good degree of objectivity by seeking specific expressions; the coding process, therefore, is relatively easy to automate.

In other cases, however, the aspects under study cannot be directly connected with well-defined expressions or syntactical constructs in the analyzed texts, but rather they need to be inferred on a semantic basis. In these cases, content analysis is said to rely on the detection of "latent variables" [16]. Detection of latent content is rather complex and subjective, in that it requires interpretation and application of some heuristics in the analysis of the messages. Nevertheless, latent content is worth the attention because it is often related to very interesting research questions.

Investigating SRL in online environments involves the detection of latent content, in that self-regulation cannot be associated with the use of particular expressions or constructs. Rather, it is revealed by the fact that learners carry out certain kinds of actions, therefore entailing an analysis on the semantic level.

5 Working Out SRL Indicators

The study of SRL by means of IA is complicated by the fact that, despite the variety of approaches that have been applied to investigate the nature and extent of SRL [17], this competence is usually characterized in terms of general, rather than specific, skills and actions. It is therefore necessary to start by defining SRL indicators that can guide the search for latent content items.

We based our analysis on the characterization of SRL proposed by Zimmermann [1], [2], which is rather detailed and widely adopted. We also took into consideration some subsequent elaborations of these studies on the potential support to SRL granted by Technology Enhanced Learning Environments (TELEs) [8], [18], [19].

Based on the work of all these authors, SRL appears to be characterized by two orthogonal sets of aspects, which we will call, respectively, "process" model and "component" model of SRL. The process model views SRL as consisting of three phases that are cyclically repeated during learning activities of self-regulated learners and influence each other: planning, monitoring, and evaluation. The component model, on the other hand, distinguishes among the cognitive (behavioral), metacognitive, motivational, and emotional aspects of SRL. The two models can meaningfully be considered both at the individual and social levels. This characterizes SRL as a kind of 3D process, in which three independent sets of features can be observed.

TABLE 1
Groups of SRL-Related Aspects Giving Rise to the Classification Indicators Used in Our Approach

		Planning	Monitoring	Evaluation
cognitive and meta-cognitive	individual	PCI	MCI	ECI
	social	PCS	MCS	ECS
motivational and emotional	individual	PMI	MMI	EMI
	social	PMS	MMS	EMS

Based on this characterization, and taking into consideration the fact that individual activity and social construction of knowledge are both very important in VLCs and strictly intertwined, we devised the following orthogonal features. Their combination allowed us to classify and determine SRL indicators to guide IA in online learning activities. Here, they are

- the learners' abilities to plan, monitor, and evaluate their own learning process; these can be investigated by spotting the learners' active contribution to: choosing learning objectives and contents; working out or adapting learning strategies; suitably configuring the learning environment; evaluating learning results by comparing one's outcomes with the outcomes of peers and with models possibly provided;
- the learners' abilities to cope with cognitive, metacognitive, emotional, and motivational challenges imposed by the learning process, throughout the above-mentioned phases; these can be captured by identifying clues that show deliberate application of strategies to solve complex problems, to cope with stress and anxiety, to keep up motivation, and to relate with peers in a smooth and profitable way;
- the learners' abilities to practice all the above actions both in individual study and in a collaborative learning context.

The indicators of SRL abilities proposed in this paper derive from this theoretical framework, by combining these three kinds of features. Table 1 shows the 12 groups of aspects raising from such combination. Following Garrison et al. [14], we grouped cognitive with metacognitive aspects since it is often difficult to clearly mark the separation between them, especially in a context, like VLCs, that usually fosters metacognitive activities along with cognitive ones. Similarly, we grouped motivational

TABLE 2
Description of Planning Indicators

Code: PCI - Making personal plans on how to proceed in the learning process: breaking up tasks in sub-tasks, establishing deadlines, detecting priorities, etc. - Detecting plan changes necessary to overcome failures.
Code: PCS - Making proposals on how to proceed in the learning process. - Discussing and negotiating planning aspects. - Working out joint plan changes necessary to overcome failures.
Code: PMI - Exploring one's expectations about the current learning activity. - Anticipating possible emotional aspects.
Code: PMS - Discussing expectations and motivations about the current learning activity and learning in general. - Sharing motivations for own commitment - Getting peers involved in planning.

The codes refer to Table 1.

and emotional aspects since the border between them is quite blurred.

Tables 2, 3, and 4 show a description of the 12 groups of aspects mentioned in Table 1. These tables illustrate what should be observed in students' messages in order to support the claim that their activity in an observed learning experience was self-regulated.

The underlying assumption of this study is that, when a message contains one of the above indicators, that is, a clue that the sender has carried out a self-regulated action, then we can think that she/he, taking such action, has practiced

TABLE 3
Description of Monitoring Indicators

Code: MCI - Working consistently on the assigned task. - Monitoring plan fulfilment. - Making syntheses of individual work done and objectives reached.
Code: MCS - Quoting peers' contributions, asking questions, reacting to peers messages - Mediating among peers. - Checking understanding. - Summarising the ideas suggested by all group members. - Encouraging peers to act.
Code: MMI - Expressing one's emotions and motivations. - Looking for appropriate support when needed. - Disclosing oneself to peers.
Code: MMS - Encouraging peers to express their emotions and motivations. - Encouraging peers and providing them emotional support. - Taking care of group functioning by informing peers of one's intentions.

The codes refer to Table 1.

TABLE 4
Description of Evaluation Indicators

Code: ECI - Assessing own learning. - Analysing results, spotting difficulties and causes of failures. - Reflecting on individual learning achieved. - Comparing one's work with that of peers. - Expressing awareness of individual time management.
Code: ECS - Assessing group learning. - Commenting on group achievements. - Reflecting on group learning. - Encouraging peers to express their opinions on the work done.
Code: EMI - Comparing one's current motivation and emotions with the original ones. - Understanding the reasons for possible changes to plans. - Commenting on emotional aspects developed during the learning process.
Code: EMS - Expressing appreciation for peers' efforts, contributions and results. - Spotting group's malfunctioning and analysing its causes.

The codes refer to Table 1.

self-regulation to some extent. For example, let us suppose that a student sends a message commenting on strengths and weaknesses of the outcomes of the group's work on some task and another student answers by proposing a plan to go on with the next assignment. In our approach, we assume that the first student has carried out some kind of evaluation of the work done, while the second student has engaged in a form of planning (see Table 5 for examples of possible quotes for each indicator). The opposite, however, cannot be claimed, because if a student does not express in her/his messages something that allows us to infer a self-regulated activity, this does not mean that self-regulation did not take place, but simply that the student did not feel the need, or simply did not happen, to express it. This holds, in general, independently of the chosen set of indicators and entails that IA, as a method to investigate SRL, can possibly underestimate its presence but is unlikely to overestimate it.

6 A Case Study

6.1 Background

We used the SRL indicators described above to analyze the learning dynamics that took place in the online component of a blended teacher training course in educational technology. This course was run in 2005 by ITD-CNR for the Specialization School for Secondary Education of the Italian region Liguria [20]. The course lasted 12 weeks (see course structure in Fig. 1) and involved 95 students and eight tutors, who exchanged, in total, 7,605 messages. Among these, the student messages were approximately 77 percent of the total.

We selected for this study the activities of Modules 3 and 4, to which we will refer in the following as Activity 1 and Activity 2. Because of the exploratory nature of the current

TABLE 5
Examples of Message/Indicator Associations

Code: PCI - "I'd better analyse one of the assigned web sites each day, if I want to meet the deadline".
Code: PCS - "Let us devote a couple of days to the readings, and then try to summarise them".
Code: PMI - "I must try to spot the interesting facets of this topic, that instinctively does not attract me!"
Code: PMS - "OK, all of us have been taking it easy to see the others' reactions. What about really tackling the task, now?"
Code: MCI - "I'm working on the first part. I'll post it by tonight, as planned".
Code: MCS - " I agree with what you wrote because...".
Code: MMI - "I'm really thrilled by this new activity…". - "I must confess that I don't understand why…".
Code: MMS - "Don't you think we should all try to cooperate better?" - "I won't be able to take part in the discussion tomorrow".
Code: ECI "I'm quite happy about my work, although I know I took longer than I should have".
Code: ECS - "We have done a very careful report, don't you think so?"
Code: EMI - "At the beginning of this course I was a bit worried about not being able to handle the technology, but in the end I found it easier than I thought".
Code: EMS - "Thank you Sara for working so hard! You did a great job! "

study, we did not analyze the whole mass of exchanged messages but focused on one subgroup of eight students with one tutor. The two activities lasted three weeks each and included a total of 249 messages, 218 of which posted by the students. All students involved contributed to these posts, in slightly different measure, as shown in Table 6.

The group of students whose interactions were analyzed is a good representative of the whole cohort of course participants, in that it has similar characteristics: similar ratio between males and females, similar mix of backgrounds, and average grade in the final assessment very close to the average grade of all the students (27.5/30 versus 27.9/30).

Both the considered modules were based on collaborative learning but involved different ways to organize the group activity. The first was a role play, where students were required to take the role of strongly characterized teachers (e.g., the technology enthusiast, the technology detractor, the bureaucrat, the pragmatist, and so forth) and to discuss from these different points of view strengths and weaknesses of a WebQuest. The second was a case study on

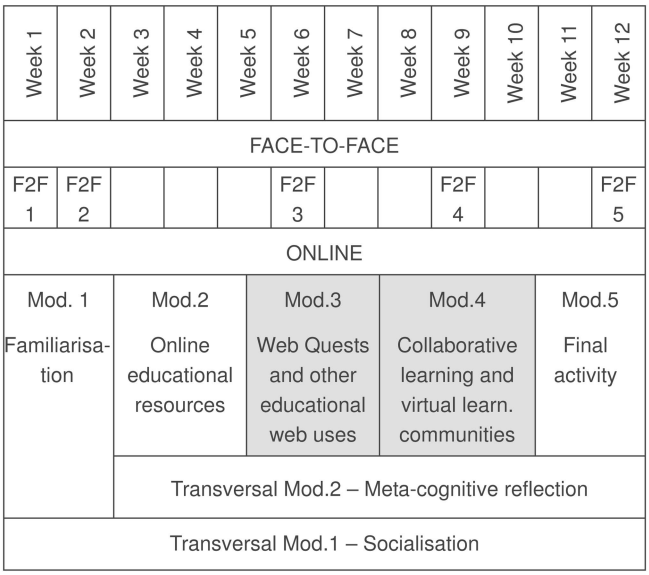

Fig. 1. Structure of the considered course. Interactions were analyzed for Modules 3 and 4.

TABLE 6
Number of Messages Posted in the Two Activities of the Analyzed Sample

	Student msgs			Tutor msgs		Total
	N	mean	SD	N	%	N
Activity1	90	11,3	5,4	14	13,5	104
Activity2	128	11,1	17,1	17	11,7	145

school-based learning communities. Trainees were supposed to discuss pros and cons of a school project recently carried out by a few teachers with their classes. The features of the proposed project were explained to the student teachers by its designers and the related documentation (instructional design, students' products, and assessment results) was made available to them.

6.2 Coding Procedure

Two coders examined all the messages of the selected sample and classified the SRL-related expressions detected according to the codes presented in Table 1. One of the coders had been involved in designing and running the course; the other had moderate experience with CSCL activities and a good level of expertise on SRL. In order to get trained for this analysis, the two coders separately searched in messages examples of the various indicators and then compared and discussed their selections.

After coding, the interrater reliability was calculated (Holsti's method) and resulted above 80 percent globally. After the computation of the interrater reliability, the coders discussed the controversial cases until they reached 100 percent agreement. The data reported in the following refer to the agreed coding.

The fact that these values are quite acceptable is a point in favor of the replicability of this approach. Table 7 shows that the percentage of significant messages was not very

TABLE 7
Basic Data about the Two Activities

	Activity 1	Activity 2
Total number of students' messages	90	128
Messages containing SRL indicators (N and %)	32 (35.6%)	49(38.2%)
Total number of SRL indicators	39	70
Average number of indicators per SRL-related message	1.2	1.4
Inter-rater reliability	0.89	0.80

high, which might mean that SRL did not take place extensively or that students did not often express the self-regulated actions they were carrying out.

Table 8 shows a comparison of the SRL-related expressions detected by the two coders. Coder 1 ratings are always slightly higher than those produced by Coder 2, which suggests a more open attitude of Coder 1 rather than a disagreement on the way to interpret students' messages. This was confirmed by the comparison and discussion of the selected expressions and explains why it was easy to reach a complete agreement after comparing the differences.

The high rate of agreement also suggests that it was not difficult to classify the considered messages against the grid given in Table 1. This is important from the methodological point of view, in relation to the feasibility of the suggested method, in spite of the difficulties inherent to the use of latent variables.

More accurate measures of the interrater reliability were not deemed necessary, given the exploratory nature of this study and its aims, focusing on the feasibility of the method and on the formulation of hypotheses to be investigated with subsequent studies. In case of similar studies on much bigger samples of messages, it would be advisable to adopt more advanced measures of reliability, which take into consideration chance agreement [21], such as Kohen K [22], along with accurate statistical analysis.

TABLE 8
A Comparison of the SRL Indicators Detected by the Coders for Activity 1

	Coder 1	Coder 2	Agreement
PCI	4	3	3
PCS	9	4	4
PMI	2	2	2
PMS	0	0	0
MCI	0	0	0
MCS	12	10	10
MMI	2	1	1
MMS	6	6	6
ECI	7	6	6
ECS	1	1	1
EMI	0	0	0
EMS	6	6	6
Total	49	39	39

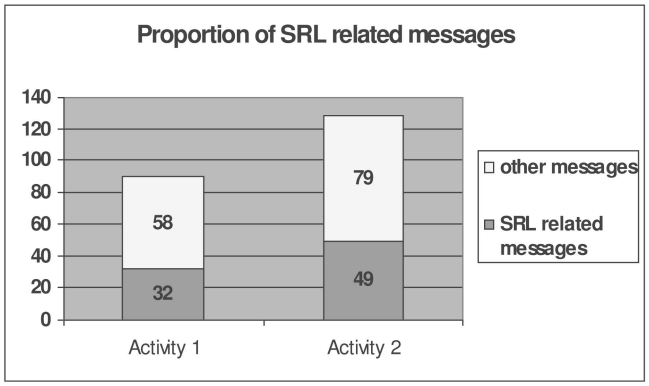

Fig. 2. Number of total messages posted by the students in the two activities and number of messages containing SRL indicators.

The chosen unit of analysis was the message. This choice appeared advantageous in that messages are objectively identifiable, their extent is determined by the message authors, and they consist of a possibly large but still manageable set of cases. The analyzed messages turned out to contain almost all the indicators proposed in Table 1. On the other hand, several messages contained more than one occurrence of the same indicator or of different ones. This made the analysis of the data slightly more difficult to interpret, since, for instance, the percentage of messages containing SRL-related expressions does not give an exact idea of the concentration of indicators detected.

Some quantitative data about the two activities were also considered, such as the number of messages exchanged per day and the contribution of individual students to the discussion. These data helped us gain a global picture of the learning dynamics in the two activities but did not provide much information on the development of self-regulation and, therefore, will not be reported in this study.

6.3 Outcomes of the Study

The main results of the content analysis are reported in Table 7 and Figs. 2, 3, 4, and 5. These figures show the raw data, without statistical elaborations on them, because the limited size of the sample analyzed makes them easier to read than complex elaborations. In most cases, we will refer

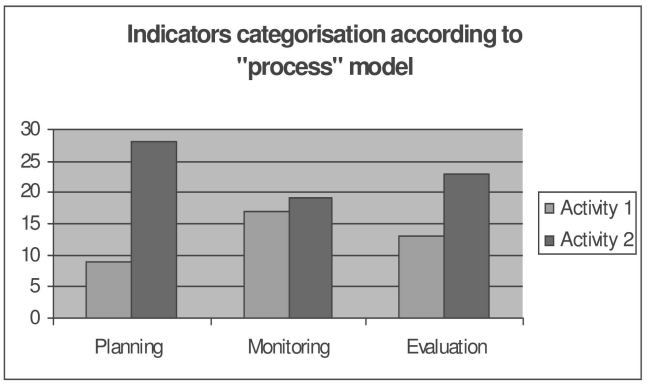

Fig. 3. Coding results along the categories of the process model, that is, highlighting the planning, monitoring, and evaluation phases of SRL.

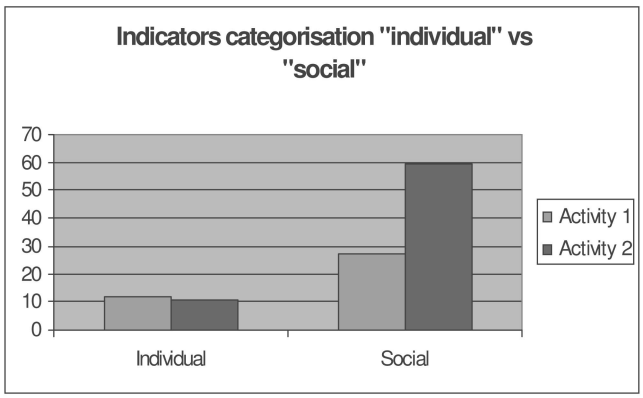

Fig. 4. Coding results along the individual versus social categories.

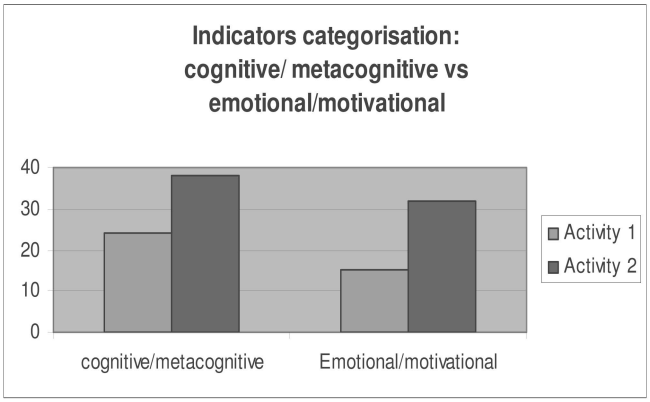

Fig. 5. Coding results along the categories cognitive and metacognitive versus emotional and motivational.

to the actual number of indicators found rather than percentages of messages, because, as pointed out above, several messages contained more than one SRL indicator, so that it does not make much sense to reason in terms of percentages of SRL-related messages. It is useful to remind that the two activities had the same duration, which makes the comparison of the raw data meaningful.

The data in Fig. 2 show that trainees participated more in Activity 2 (the case study) than in Activity 1 (the role play). This is true not only in terms of number of messages but also as concerns "SRL density." This clearly appears from Table 7, which shows that the percentage of SRL-related messages and the average number of indicators per SRL-related message were higher in Activity 2.

These data may be due to the different natures of the tasks to be carried out in the two activities, but they can also support the hypothesis that the students, over the course, were learning to participate and to self-regulate themselves. Most likely, both explanations contributed to determine this distribution of SRL occurrences, together with other possible causes that do not appear from the data used in this study.

The difference between the tasks carried out in the two activities can also explain the data in Fig. 3, which show that indicators of planning-related events in Activity 1 are significantly less than those in Activity 2. One reason for this may be that Activity 1, being a role play, had an inherent plan: once taken a role, the participants were requested to adapt their behavior to the role constraints and this partially limited their freedom of planning. However, Activity 2 shows a higher concentration of SRL-related events than Activity 1 also as concerns monitoring and evaluation tasks, which again supports the idea that students generally self-regulated their learning more in this module.

Fig. 4 shows that SRL-related indicators at a social level were definitely more frequent than indicators showing SRL at individual level. Once again, there are two possible reasons behind these data and it is likely that both have contributed to determine the situation. One reason is that VLCs tend to favor the social aspects of SRL more than its individual aspects (for example, students feel encouraged to plan, monitor, and evaluate the group work, more than they do with their own individual work). The second explanation is that in online collaborative environments students feel the need to express, when writing messages, the social aspects of their learning activity more than they do with the individual aspects. In other words, they might be planning, monitoring, and evaluating their own individual work as well, but they do not always communicate it in their messages.

6.4 Comparison with Data of Different Origin

The considerations raising from this analysis are much in line with the outcomes of a previous study where a different method was used to investigate SRL development in the same course [10]. That study presented the results of a survey carried out with two questionnaires, one filled in by SRL experts and another by 72 of the 95 trainees taking part in this course. Both concerned the interviewees' opinions about the support granted in the course to practice SRL. The survey showed that the potential of the environment used was deemed valuable especially as concerns the social aspects of SRL: students claimed that they felt a strong social support to their own SRL development from tutors and, even more, from peers.

Fig. 5 shows the message categorization according to the component model. From these data, the cognitive/metacognitive aspects appear to have been supported more than the emotional/motivational ones. This is not surprising, since the considered modules were devoted to cognitive activity on the course content knowledge. A module exclusively devoted to socialization was run throughout the course in parallel to all other modules, as shown in Fig. 1, and students had been explicitly invited not to invade the content-related conferences with out-of-topic conversations. To confirm this, most of the expressions related to motivation and emotion detected in the analyzed modules were always somehow related to the learning procedure and outcomes, such as appreciation for the good work carried out, expressions of one's feelings and expectations in relation with the learning activity, or encouragements not to give up.

In the study by Dettori et al. mentioned above [10], the comparison of these two categories was the only point of disagreement between the data related to experts' and students' opinions. As shown in Fig. 6, according to SRL experts, the emotional and motivational components of such support were stronger than the cognitive/metacognitive ones. According to the trainees, the former was weaker than the latter. IA and, in particular, the data

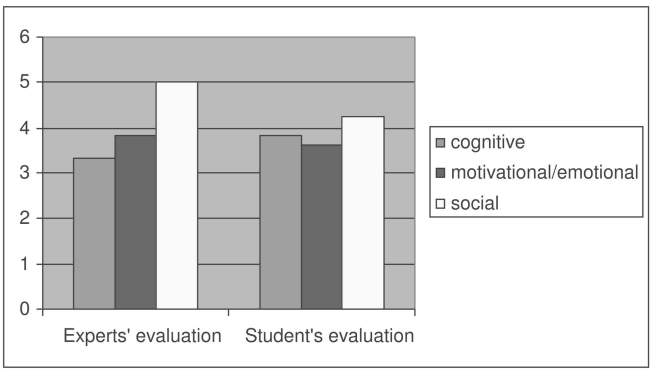

Fig. 6. Comparison between the average values obtained from the experts' evaluation and students' evaluation of the same course (from [9]).

shown in Fig. 4 seem to confirm the outcomes of the students' questionnaires.

7 Discussion

The exploratory nature of this study determined the choice to work on a small sample, with a manual method and with limited statistical tools. Its aims were

- to find out whether content analysis and our taxonomy of indicators are suitable to investigate SRL in VCLs, that is, whether they provide data consistent with those obtained by means of other research methods;
- to understand if the indicators are sufficiently well defined to grant an acceptable reliability and to refine them as necessary;
- to verify whether there are ways to partially automate the textual analysis process;
- to understand if the method is cost-effective and worth applying.

As for the first point, the consistence of the collected data with the outcomes of a previous study carried out with different means is encouraging and suggests that our approach can be adopted to investigate the presence of SRL in bigger sets of messages and in different contexts.

It is worth reminding that evidence of the presence of one SRL indicator does not—per se—prove the *development* of SRL. It only supports the claim that a particular aspect of SRL was *practiced*. Zimmerman's [1] studies on SRL, however, suggested that these abilities develop through social support and practice. In addition, increased frequency of the indicators during the learning process can be regarded as a clue of SRL development. The opposite, however, is not necessarily true. The fact that SRL indicators are not found in students' messages does not necessarily mean that the students did not control their learning: they might simply have not made the process explicit in their messages. Researchers who intend to use this method, therefore, should be aware that what can be found in messages is likely to be correct but it may not provide a complete picture.

Also, on the second point, we can make positive considerations. The indicators' list used appeared to be quite complete and apt to classify all the SRL-related situations encountered. The fact that the interrater reliability turned out to be high suggests that the indicators are not difficult to interpret and suitable to guide the detection of the latent variables of interest. Globally, the structure and most of the original indicators, which had been derived from the literature on SRL, were fit to the purpose. Some refinements were made to the indicators' list while rating the messages, since reading students' messages allowed the coders to identify learning actions, which were clearly self-regulated but were not identifiable as such according to our indicators. Table 1 reports the final version of the taxonomy of indicators.

As for the third point, we realized that there is no easy way to automate the analysis process. As a matter of fact, while in studies focused on manifest content the analysis can be carried out by means of software tools that look for expressions related to the searched clues, in the case of SRL there does not seem to be any typical expression to spot the clues we are looking for. For instance, planning actions can be introduced in many different ways, such as "I propose...," "Why don't we do...," "We could do ...," and many others (or their equivalent in other languages). The same holds true for monitoring and evaluation sentence patterns: there are so many ways to introduce a sentence where monitoring or evaluation considerations are brought forward, that it appears hardly possible to employ typical text analysis software tools to find them. This means that the search must necessarily be done on a semantic level and this makes content analysis for SRL an inherently subjective and interpretative process.

These considerations, together with the fact that SRL-related messages are not a high percentage of the examined ones, suggest that the rating work is not very cost-effective. This is not surprising, since it is widely acknowledged that content analysis on any aspect is usually a quite labor-intensive research method, especially if the search is made on latent variables and therefore cannot be automated. In order to try to overcome this problem, a very interesting applied research direction would be to develop CMC tools that expressly support content analysis, for example, by allowing one to associate raters' annotations to each message and to compute statistics about them. Such tools would be very useful for content analysts regardless of the aims of the research study they are carrying out.

To conclude, one might wonder why one bothers to apply such a labor-intensive method as content analysis to investigate SRL in VCLs. In general, information about SRL abilities is sought after through interviews with the subjects involved into the learning process, questionnaires, or observation. Questionnaires and interviews collect *opinions* and other information that are reported by the learners or their teachers. On the other hand, observation and content analysis of exchanged messages allow us to analyze directly what students actually did. Messages do not give us access to all that has been taking place during the learning process and, certainly, do not reflect the totality of students' thoughts and actions, but they allow us to work on data that are not affected by learners' opinions, nor biased by observers' point of view.

Moreover, observation and messages are distributed along the whole duration of a course. This means that we can analyze the evolution of self-regulation over time,

which is not possible if such study is made by means of end-of-course questionnaires, since these elicit students' opinion when the questionnaire is administered. For all these reasons, we believe that IA can provide a valid tool to study SRL in VLCs, especially when complemented by other methods of analysis.

It is true that the outcomes of IA are affected by coders' discretion, but the related risks can be reasonably reduced by establishing a valid coding procedure, including well-defined indicators and a way to keep coding differences under control.

REFERENCES

[1] B.J. Zimmerman, "Developing Self-Fulfilling Cycles of Academic Regulation: An Analysis of Exemplary Instructional Models," *Self-Regulated Learning. From Teaching to Self-Reflective Practice,* Shunk and Zimmermann, eds., pp. 1-19, The Guildford Press, 1998.

[2] B.J. Zimmerman, "Attaining Self-Regulation: A Social Cognitive Perspective," *Handbook of Self-Regulation,* M. Boekaerts, P. Pintrich, and M. Zeidner, eds., pp. 13-39, Academic Press, 2000.

[3] M. Boekaerts, "Self-Regulated Learning: Where We Are Today," *Int'l J. Educational Research,* vol. 31, pp. 445-457, 1999.

[4] M. Boekaerts, "Self-Regulated Learning: A New Concept Embraced by Researchers, Policy Makers, Educators, Teachers and Students," *Learning and Instruction,* vol. 7, no. 2, pp. 161-186, 1997.

[5] D.H. Schunk and B.J. Zimmerman, *Self-Regulated Learning, from Teaching to Self-Reflective Practice.* The Guilford Press, 1998.

[6] G. Corrigan and N. Taylor, "An Explanatory Study of the Effect a Self-Regulated Learning Environment Has on Pre-Service Primary Teachers' Perceptions of Teaching Science and Technology," *Int'l J. Science and Math. Education,* vol. 2, pp. 45-62, 2004.

[7] S. Bolhuis, "Towards Process-Oriented Teaching for Self-Directed Lifelong Learning: A Multidimensional Perspective," *Learning and Instruction,* vol. 13, pp. 327-347, 2003.

[8] K. Steffens, "Self-Regulated Learning in Technology-Enhanced Learning Environments: Lessons of a European Peer Review," *European J. Education,* vol. 41, nos. 3/4, pp. 353-380, 2006.

[9] G. Van den Boom, F. Paas, J.J.G. Van Merrienboer, and T. Van Gog, "Reflection Prompts and Tutor Feedback in a Web-Based Learning Environments: Effects on Students' Self-Regulated Learning Competence," *Computer in Human Behavior,* vol. 20, pp. 551-567, 2004.

[10] G. Dettori, T. Giannetti, and D. Persico, "SRL in Online Cooperative Learning: Implications for Pre-Service Teacher Training," *European J. Education,* vol. 41, nos. 3/4, pp. 397-414, 2006.

[11] J.B. Arbaugh, "Learning to Learn Online: A Study of Perceptual Changes between Multiple Online Course Experiences," *Internet and Higher Education,* vol. 7, pp. 169-182, 2004.

[12] R. Lynch and M. Dembo, "The Relationship between Self-Regulation and On-Line Learning in a Blended Learning Context," *Int'l Rev. Research in Open and Distance Learning,* vol. 5, no. 2, http://www.irrodl.org/content/v5.2/lynch-dembo.html, 2004.

[13] L. Rourke, T. Anderson, D.R. Garrison, and W. Archer, "Methodological Issues in the Content Analysis of Computer Conference Transcripts," *Int'l J. Artificial Intelligence in Education,* vol. 12, pp. 8-22, 2001.

[14] D.R. Garrison, T. Anderson, and W. Archer, "Critical Inquiry in a Text-Based Environment: Computer Conferencing in Higher Education," *The Internet and Higher Education,* vol. 2, nos. 2/3, pp. 87-105, 1999.

[15] F. Pozzi, S. Manca, D. Persico, and L. Sarti, "A General Framework for Tracking and Analysing Learning Processes in CSCL Environments," *Innovations in Education and Teaching Int'l,* vol. 44, no. 2, pp. 169-179, 2007.

[16] W. Potter and Levine-Donnersteirn, "Rething Validity and Reliability in Content Analysis," *J. Applied Comm. Research,* vol. 27, pp. 258-284, 1999.

[17] B.J. Zimmerman, "Theories of Self-Regulated Learning and Academic Achievement: An Overview and Analysis," *Self-Regulated Learning and Academic Achievement,* B.J. Zimmerman and D.H. Schunk, eds., pp. 1-37, Lawrence Erlbaum Assoc., 2001.

[18] P. Banyard, J. Underwood, and A. Twiner, "Do Enhanced Communication Technologies Inhibit or Facilitate Self-Regulated Learning?" *European J. Education,* vol. 41, nos. 3/4, 2006.

[19] "Self-Regulated Learning in Technology Enhanced Learning Environments," *Proc. Targeted Cooperative Network Conf. (TACONET '05),* R. Carneiro, K. Steffens, and J. Underwood, eds., 2005.

[20] M. Delfino, S. Manca, and D. Persico, "Harmonizing the Online and Face-to-Face Components in a Blended Course on Educational Technology," *Proc. Int'l Educational Technology Conf. (IETC),* May 2007.

[21] B. De Wever, T. Schellens, M. Valke, and H. Van Keer, "Content Analysis Schemes to Analyze Transcripts of Online Asynchronous Discussion Groups: A Review," *Computers and Education,* vol. 46, pp. 6-28, 2005.

[22] M. Capozzoli, L. McSweeney, and D. Sinha, "Beyond Kappa: A Review of Interrater Agreement Measures," *The Canadian J. Statistics,* vol. 27, no. 1, pp. 3-23, 1999.

Giuliana Dettori received a degree in mathematics from the University of Genoa, Italy, in 1978. Since then, she has been a researcher at the Italian National Research Council, first in the Institute for Applied Mathematics, then, from 2002, in the Institute for Educational Technology, both in Genoa. During her professional career, she spent two years in the USA, where she was a visiting scientist at Syracuse University, New York, and five months in Ecuador at Chimborazo Polytechnic School. Her research interests have covered different topics in educational technology and applied mathematics. She is an author of more than a hundred scientific papers, mostly at the international level. She is a member of the International Group for the Psychology of Mathematics Education (IG-PME), an associate editor of the *International Journal of Distance Education Technologies*, and a member of the Book Reviewers Panel of the *British Journal for Educational Technology*. She has been in charge, for her Institute, of the European projects Waterman, Telepeers, and Efelcren, as well as of a project on representations within the European Network of Excellence Kaleidoscope, where she also coordinated the Special Interest Group "Narrative and Learning Environments." For the dissemination activity setup for this group, in 2007, she won the Kaleidoscope's "Press and Communication Award."

Donatella Persico received a degree in mathematics from the University of Genoa, Italy, in 1981. Since then, she has been a research scientist in the Institute for Educational Technology, Italian National Research Council, Genoa. In 2003, she was appointed as a senior researcher. During her professional career, she was a visiting scientist at Bradford University, United Kingdom, in 1985 and Exeter University, United Kingdom, in 1986. She was a lecturer in educational technology in the Post-Graduate School for Secondary Teaching, University of Genoa, from 2000 to 2006. She has been in charge, for her Institute, of several national and international projects. Her major research interests include instructional design, e-learning, and teacher training. She is the author of educational material and scientific publications of various kinds, including books, educational software, multimedia material, and several research papers concerning aspects of educational technology. She sits on the editorial board of the international journal *Innovations in Education and Teaching International* and of the national magazine *Tecnologie Didattiche*, as well as on the advisory boards of several other journals. She was awarded the Best Presentation Award at the Computer Assisted Learning Conference (CAL 1997) in Exeter.