Original article

The effects of animations on verbal interaction in computer supported collaborative learning

M. Sangin,* P. Dillenbourg,* Cyril Rebetez,† Mireille Bétrancourt† and Gaëlle Molinari*

*CRAFT, School of Computer and Communication Sciences, Ecole Polytechnique Fédérale de Lausanne, Switzerland
†TECFA, School of Psychology and Education, University of Geneva, Switzerland

Abstract

This paper focuses on the interaction patterns of learners studying in pairs who were provided with multimedia learning material. In a previous article, we reported that learning scores were higher for dyads of an 'animations' condition than for dyads of a 'static pictures' condition. Results also showed that offering a persistent display of one snapshot of each animated sequence hindered collaborative learning. In the present paper, further analyses of verbal interactions within learning dyads were performed in order to have a better understanding of both the beneficial effect of animations and the detrimental effect of the presence of persistent snapshots of critical steps on collaborative learning. Results did not show any differences in terms of verbal categories between the two versions of the instructional material, that is, static versus animated pictures. Pairs who were provided with persistent snapshots of the multimedia sequences produced fewer utterances compared to participants without the snapshots. In addition, the persistent snapshots were detrimental both in terms of providing information about the learning content and in terms of producing utterances solely for the purpose of managing the interaction. In this study, evidence also showed that these two verbal categories were positively related to learning performances. Finally, mediation analyses revealed that the negative effect of persistent snapshots was mediated by the fact that peers of the snapshots condition produced less *information providing* and *interaction management* utterances. Results are interpreted using a psycholinguistic framework applied to computer-supported collaborative learning (CSCL) literature and general guidelines are derived for the use of dynamic material and persistency tools in the design of CSCL environments.

Introduction

Should the design of educational software be different when one knows there will be two users in front of a single machine? This question probably locates the emergence of computer-supported collaborative learning (CSCL) as a specific research field. The work of Roschelle (1992), based on a physics micro-world, was

an early contribution to the field. He observed that task representations used by the micro-world shape social interactions among students, a principle that he summarized as 'designing for conversations'. This influence of communication tools on cognition, which has been stressed for many years by socio-cultural theories, became a principle for designing CSCL environments. A second example of the mediation of interactions is the synchronous argumentation tool, Belvedere (Suthers *et al.* 1995). Two students, working side by side or online, construct a graphical representation of a scientific debate. They share a graphical editor in which they

connect text boxes. The palette of boxes ('hypothesis', 'data') defines the argumentation grammar and the links between boxes (e.g. 'X supports Y', 'X contradicts Y') define the dialogue structure. The way in which a task representation shapes social interaction is referred to as 'representational guidance' by Suthers and Hundhausen (2003). Both examples illustrate a CSCL approach in which social interactions are influenced in a subtle and indirect way through task representations. In parallel, CSCL has also developed more coercive methods referred to as 'scripts' (Weinberger *et al*. 2002; Dillenbourg & Jermann 2007). This paper, however, contributes to the first approach, the understanding of how interface features indirectly shape social interactions.

Why do task representations influence social interactions? As pointed out in socio-cultural theories, they shape the language students use for discussing and reasoning about a task. Peers may refer to external representations in order to co-construct a shared understanding of the task at hand; it is not the intrinsic epistemic value of task representations that matters but the extent to which they can be instrumented by learners for grounding their understanding (Roschelle 1992). Grounding is the means through which a speaker checks if an addressee has understood what he meant and repairs possible misunderstandings (Clark & Brennan 1991). Indeed, two interlocutors initiate in a collaborative effort in order to improve the effectiveness of communication (Clark & Brennan 1991). In other words, verbal interaction triggers 'grounding' needs that lead to more elaborated talks. For van Boxtel *et al*. (2000), some of these research results are interpreted within the Piagetian framework and focus on the role of conflict in social interaction (Doise & Mugny 1984; Brown & Palincsar 1989); other results are explained by the Vygotskian notion of scaffolding and show relationships between elaborated information seeking (King 1990) and elaborated information providing activities (e.g. in order to be helpful, explanations have to be given at an appropriate level of detail (Webb 1989), and collaborative learning outcomes. In this sense, we understand 'designing for conversation' as providing peers with representational support for the fruitful elaboration of a shared understanding.

In many CSCL environments, students jointly construct or manipulate a task representation and thus, the study of grounding mechanisms became a central element in CSCL. However, even if these task representations are manipulated by students, they nonetheless are rather static. We found that persistency is critical for supporting group processes (Dillenbourg & Traum 2006), and therefore questioned how students can activate grounding mechanisms when the task representation is an animated picture. We present an empirical comparison of collaborative processes in dyads learning with static versus animated pictures.

Related work

The benefits of animated pictures over static ones received a lot of attention within instructional psychology and educational multimedia. Animated pictures are considered particularly suitable for representing dynamic systems such as lightning storms, human respiratory system, bicycle pumps, etc. (Mayer & Moreno 2002). However, empirical research failed to establish systematic benefits of using dynamic graphics instead of static ones in learning (Bétrancourt & Tversky 2000; Bétrancourt *et al*. 2001). Several studies compared the effects of static graphics and animated ones, but the results diverge. Some of them found an added value in using animations instead of static pictures and texts (Palmiter & Elkerton 1993), while others did not (Rieber 1990; Pane *et al*. 1996). The main drawback of animated pictures is that they induce a high perceptual load (Bétrancourt & Tversky 2000). The flow of information sometimes generates an extraneous cognitive load that prevents students from devoting their resources to processing the information and leads to what Lowe (2004) called an 'overwhelming' effect. Most studies explain these results in terms of cognitive load (Sweller 2003). This theory has been developed around individual learning, but what happens when animations are shown to a group instead of an individual?

Schnotz *et al*. (1999) carried out an empirical study in which they measured learning from animated versus static pictures in a collaborative situation. The learning material used was a hypertext explaining time zones on earth with either animated (simulations) or static graphics. They found that pairs assigned to the simulation condition performed less well than pairs assigned to the static graphics condition. In a previous study conducted with the same material but with individual learners, they obtained the opposite results. That is, learners in individual settings took advantage of animations
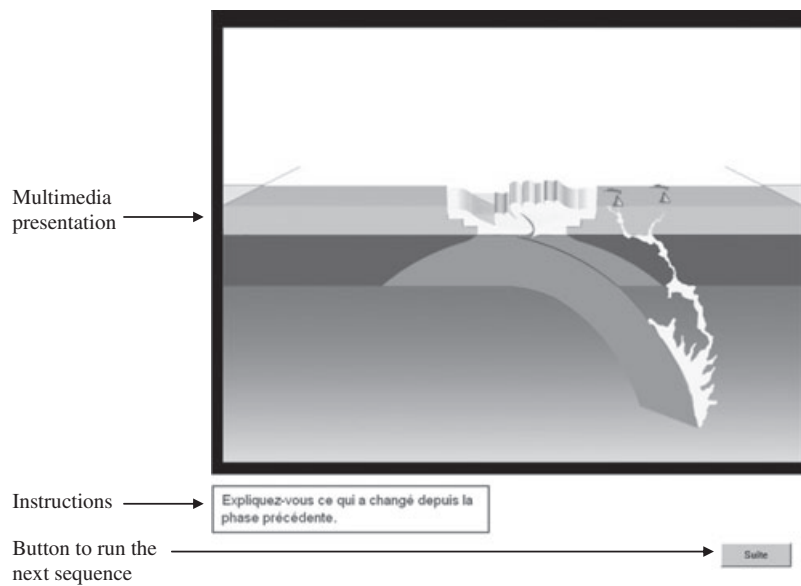
Expliquez-vous ce qui a changé depuis la phase précédente.

Suite

Fig 1 Screen capture of the display ('geology' material; *dynamic/no-persistency* condition) taken during the pause after the 11th sequence. The presentation could neither be controlled nor viewed a second time. An instruction appeared during pauses to stimulate participants to explain themselves the changes. A button was provided to run the next sequence.

whereas animations were detrimental to learning in collaborative settings. The authors interpreted these results in terms of cognitive load: the animations imposed a heavy cognitive load on learners in terms of information processing, which was added to the load of interacting with peers and managing the collaboration. Therefore, the cognitive overload impaired the construction of an appropriate mental model.

In line with his idea of 'designing for conversation', Roschelle (1990) proposed the following design criteria: manipulation for communication, minimalism, authentic activity and persistence. This last principle is of especially great importance for animated external representations designed for a collaborative setting. Indeed, talking about transient features of a display that are prone to appear and disappear in the course of the interaction implies difficulties that may hinder the co-construction of shared knowledge.

We investigated (Rebetez *et al*. 2005; Rebetez *et al*. 2008) the cognitive load hypothesis by adding a persistency tool to standard dynamic multimedia materials. We hypothesized that cognitive load might be decreased if learners were provided with a snapshot describing the previous episode of the animation: learners would not need to memorize it, they could simply click on a thumbnail to see it again (see Fig 1). We also expected the snapshots to act as artefacts that would support the co-construction of shared knowledge for learners in pairs. We will refer to the interfaces that include this persistency tool as 'persistent snapshots'. An empirical

study conducted with 160 university students was then carried out with a $2 \times 2 \times 2$ factorial design that combines the factors mentioned so far: (1) learning from static pictures versus with animated pictures; (2) individual learning versus learning in pairs; and (3) with snapshots versus without snapshots. Participants were provided with two consecutive multimedia presentations on the astronomical phenomenon of Venus transit and the geological phenomenon of tectonic plates (please refer to the method section for more details). The learning outcomes were measured with a learning test in which we differentiated between the students' scores at simple retention questions versus their scores at inference questions (questions that required elaboration from the learning material). Here, we summarize the main results.

The students using movies obtained significantly higher scores than those learning with static pictures. The two other factors, individual/collaborative and snapshots/no-snapshots, had no significant effects on post-test scores. More interestingly, two interaction effects were found:

1 Participants working in pairs obtained significantly higher inference scores with movies than with static pictures, while there was no difference between static and dynamic pictures for participants working individually. These results question the hypothesis of a cognitive overload resulting from the sum of the cognitive load generated by movies and the load
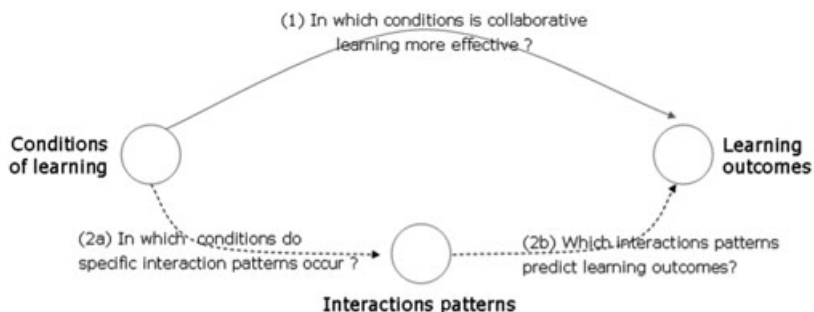
Fig 2 Research questions in computer-supported collaborative learning.

induced by the management of social interactions. To our surprise, in a NASA-TLX questionnaire (a subjective task-load index; see Hart and Staveland (1988) on cognitive load, students working in pairs reported a significantly lower load than those working individually.

2 Even more surprising was the effect of the persistent snapshots on the scores of inference questions; the presence of snapshots led to higher scores for individuals, but it led to lower scores for dyads.

The complexity of these results calls for further analysis of what actually happens in different collaborative learning situations. Research on collaborative learning has evolved from a traditional experimental paradigm that controls a few conditions to an 'interactions' paradigm; searching for interaction patterns (see Fig 2) that mediate the effect of independent variables on dependent variables (Dillenbourg *et al*. 1996, p. 201). Different patterns of verbal interactions have been investigated such as explanations, argumentation, negotiation and conflict resolution.

The answers to Fig 2, question (1) have been summarized above and are detailed in (Rebetez *et al*. 2005; Rebetez *et al*. 2008). The present contribution investigates Fig 2, questions (2a) and (2b), and addresses research questions regarding the socio-cognitive processes that may explain the results of the collaborative setting. We therefore explore the verbal interactions among peers – as it is most often the case in CSCL research – as process variables potentially mediating the main results.

Research questions

While the main goal of the previously described study (Rebetez *et al*. 2005; Rebetez *et al*. 2008) was to investigate the effects of animated pictures on learning outcomes in individual and collaborative settings (Fig 2, question 1), the questions we address here are aimed at further examining the results obtained from the collaborative setting: How do multimedia variants (static/animated; persistent/not) influence collaborative processes (Fig 2, question 2)? We therefore conducted several in-depth analyses of the group processes in terms of verbal interactions. These analyses aim to answer two questions:

1 Do different interaction patterns emerge in different conditions?
2 Are differences in interaction related to learning outcomes?

Concerning the first research question, our hypothesis is that different learning contexts or conditions imply different types of interaction that in turn mediate the effects. Concerning the second question, we expect that meaningful verbalizations such as elaborated explanation providing is related to better learning performance (Webb 1991; Teasley 1995). Consequently, further analyses of verbal interactions may lead to a better understanding of the beneficial effect of dynamic pictures and the detrimental effect of the presence of persistent snapshots of critical steps on collaborative learning performances.

Method

Participants

Participants in the collaborative setting were 80 French-speaking first and second semester university students (33 female, 47 male; age mean = 20 years). Their prior knowledge of the topics described in the learning material was tested through a multiple-choice questionnaire. All participants were complete novices in the proposed subject areas. Twenty out of the 80 were randomly assigned to each of the four experimental conditions derived from a 2 (animation factor: *dynamic* versus *static* pictures) × 2 (*persistent snapshots* factor: *snapshots* versus *no-snapshots*) intersubject design. The four conditions are *dynamic/snapshots*, *static/snapshots*, *dynamic/no-snapshots* and *static/no-snapshots*. Experimental dyads were created: 26 same-gender dyads (16 male and 10 female dyads) and 14 mixed dyads. The effect of the same-gender versus mixed dyads was tested and was not significant. Half of the dyads were involved in engineering curricula and half were involved in social science curricula (psychology and educational science). The same number of dyads from each curriculum was associated to each of the four experimental conditions. The effect of the curriculum was tested and was not significant. Participants were remunerated 20 Swiss francs for their participation.

Because of technical problems in videotapes, interactions from two dyads (one *static/no-snapshots* dyad and one *dynamic/no-snapshots* dyad) were not available and removed from the analyses; 38 dyads were finally included in the analyses (ten in the *dynamic/snapshots*, ten in the *static/snapshots*, nine in the *dynamic/no-snapshots* and nine in the *static/no-snapshots*).

Instructional multimedia material

The instructional material used for the experiment included two different scientific subjects presented consecutively and always in the same order. The first topic was an explanation of the astronomical phenomenon of the Venus transit. The second was an explanation of the geological phenomenon of tectonic plates. The material was designed under advice from experts in both fields. These scientific phenomena were chosen for two reasons. First, due to their intrinsically dynamic nature, their comprehension would require learners to form a dynamic mental model (Mayer 1989). Second, they include changes (translations, transformations and transitions) suitable for animations Lowe (2003). The two subjects were totally independent and no knowledge transfer possibilities could be expected.

Four different versions of the learning materials were developed for each of the two instructional subjects (i.e. 'astronomy' and 'geology'), according to the four experimental conditions. Each version consisted of 12 consecutive media sequences. Each of these consisted of an animated picture for the dynamic condition or a static picture for the static condition. Each sequence was selected in order to introduce a state or a change in the dynamic process. The static and dynamic pictures had the same size (800 × 600 pixels) and were laid out on the top-right part of the interface (see Fig 2). They were accompanied by the same audio commentary (mean duration of a sequence: 17 s). The audio commentary was synchronized with the dynamic pictures. To create the static pictures of the *static* condition, we used the most informative and complete frame of each of the corresponding animated pictures of the *dynamic* condition. After each multimedia sequence, the presentation was paused and participants had to click on a button to run the next sequence (see Mayer & Chandler 2001). Participants could see every episode and hear the associated audio commentary only once.

Under the 'persistent' condition, a set of 12 greyed labels (320 × 200 pixels) were provided from the beginning of the learning phase on the left side of the interface (see Fig 1). After each presentation of a sequence (static or dynamic), a new label providing access to a key-snapshot of the previously displayed sequence became coloured and usable. Participants were thus able to see a bigger version (640 × 480 pixels) of the snapshot by moving the mouse over the corresponding label only during the pauses between sequences. Only labels of the previous sequences were activated (see Fig. 1). It is important to note that the activated labels did not give access to the whole sequence (media plus audio commentary) but only to a relevant static snapshot of it. These snapshots were smaller versions of the static pictures of the static condition, namely the most representative snapshot of the corresponding dynamic sequence.

Procedure

Our material was completely automated using AUTHOR-WARE (Adobe, San Jose, CA, USA) and the procedure was exactly the same for each of the different conditions.

At the beginning of the experiment, participants sat at two different computers in the same room and started by individually performing a short test concerning their previous knowledge about the 'Venus transit'. The test was aimed at selecting only novices in this topic. After a short introduction to the instructional material, one of the participants was asked to join his/her peer in front of the same computer for the collaborative learning phase.

The multimedia material was presented to the dyads according to the experimental condition (static pictures versus dynamic pictures, with or without snapshots depending on the condition). Between each sequence, participants were prompted to 'explain what has changed since the previous sequence'. The goal of this instruction was to stimulate interaction between the peers. They had to press a button to move to the next episode. Participants of the *dynamic/snapshots* group and the *static/snapshots* group were allowed to access the key frames of previous sequences only during pauses. The elapsed time between episodes ('elaboration time') was measured. Participants could not stop or act on what they saw in the movie while it was running. Participants in all conditions did not have the opportunity to see a sequence's content or to hear the commentary a second time.

After the collaborative learning phase, participants were separated again onto two different computers and had to perform a learning test individually, in order to assess their learning about the 'Venus transit'. It included 16 multiple-choice questions: nine retention questions (questions used to test memorization of elements of the material) and seven inference questions (questions used to test the general comprehension of the material). The answers to the retention questions were explicitly described in the audio comments or depicted in the images. The right answer for inference questions was not explicitly presented in the material but had to be inferred by participants. Each question comprised four to seven possible answers with one correct answer.

After the completion of the astronomy learning test, exactly the same procedure was repeated for the second instructional subject (i.e. tectonic plates). Prior knowledge of the participants about tectonic plates was tested individually. After a short introduction, the instructional material corresponding to a particular experimental condition was provided to collaborators at the same computer.

After the collaborative learning phase, participants were separated to answer the 'geology learning test' individually. This test was also composed of 16 questions (nine retention questions and seven inference questions).

Coding of verbal interactions

Dyads' verbalizations were recorded on videotapes and transcribed using a video-annotation software (TRANSANA, Madison, WI, USA). Turn-taking was used as a criterion for the corpus segmentation. The coding scheme (see Fig 3) was inspired by the communicative act approach (Bunt 1989; Baker *et al*. 2005) and is illustrated in Fig 4. Categories are treated as mutually exclusive. Information-seeking utterances are questions. Degree of elaboration in information providing was considered 'low' when students repeated information contained in the last multimedia episode, as 'medium' when the utterance integrated elements provided in different past episodes and as 'high' when the utterance integrated elements not mentioned in the multimedia material but are added by the students on the basis of their prior knowledge or their inferences. The 'task management' category included interactions on how to use the interface. The 'interaction management' category contained utterances about managing coordination and collaboration. The entire corpus comprised 5350 utterances and was coded by two independent coders with an intercoder reliability of Cohen's Kappa = 0.81, $P < 0.0001$. Table 1 provides some examples of each category.

Results

To further analyse our previous results (Rebetez *et al*. 2005; Rebetez *et al*. 2008) on the collaborative process level, we addressed two sub-questions: (1) Do different interaction patterns emerge in the different conditions? (2) Are differences in interaction related to learning outcomes? Accordingly, the first section explains the effects of the animation and snapshots factors on the interaction variables. The second section describes correlational data about the relation between the interaction variables and the learning outcomes. The final
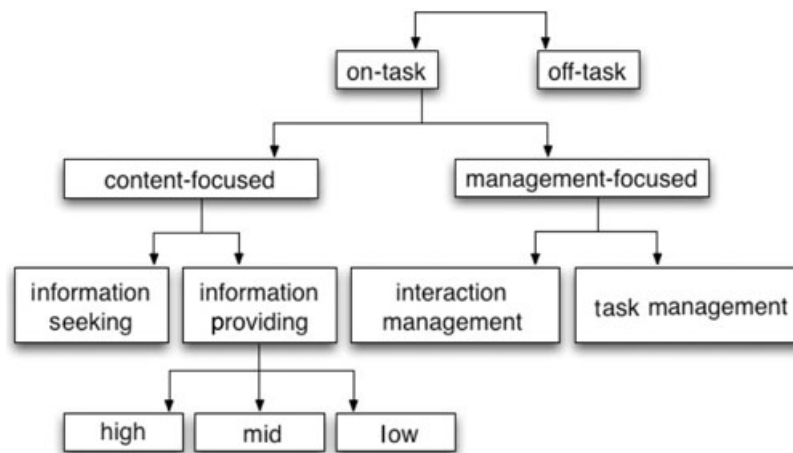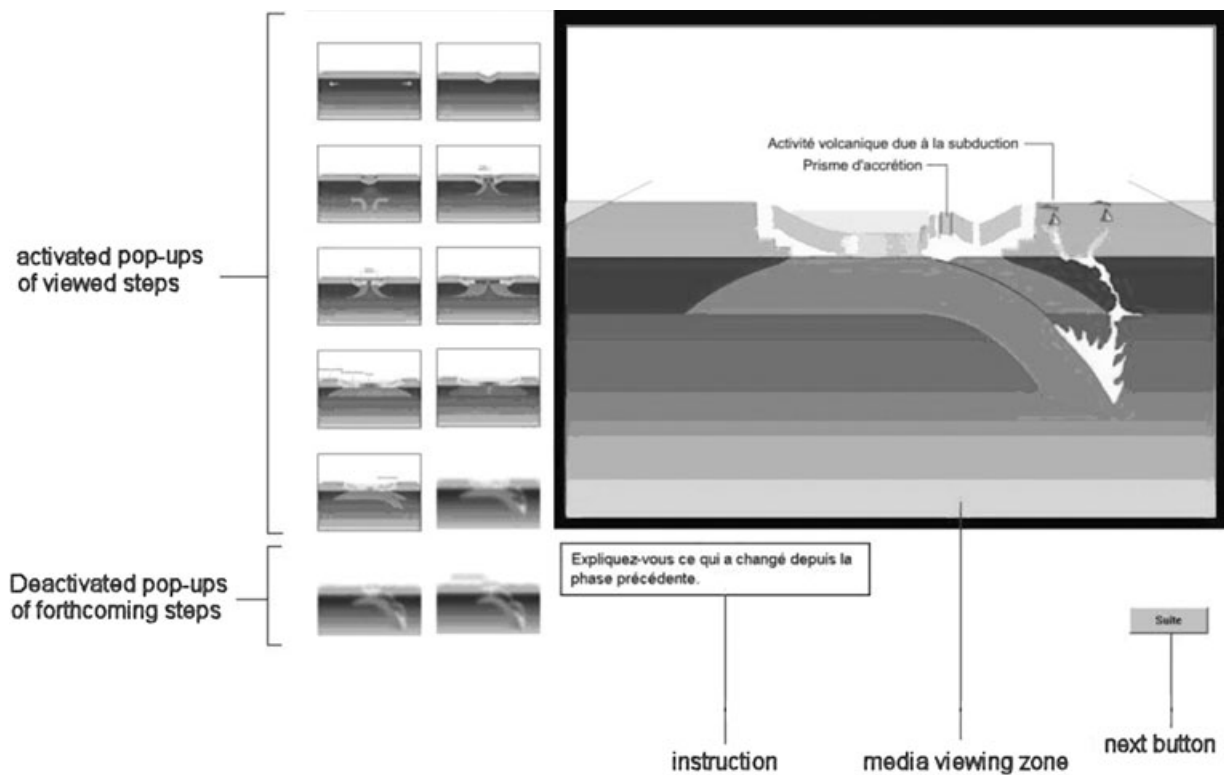
Fig 3 Coding scheme for verbal interactions.

Fig 4 Screen capture of the display ('geology' material, *dynamic/persistency* condition) taken during the pause after the ninth sequence. The nine coloured labels on the left side are active labels providing access to snapshots of the past sequences. The grey labels are the forthcoming sequences not yet active.

section focuses on the mediation effects of interaction variables on the relation between the main factors and the learning outcomes.

The time between two episodes was dedicated to reflection and interaction. On average, this 'elaboration' time was 26.7 s. *Dynamic* pictures did not induce different elaboration times compared with *static* pictures ($M = 27.3$ s for movies and $M = 26$ s for pictures; $F = 0.15$, $P = ns$). *Persistent snapshots* led to longer elaboration phases but the difference is also not significant

Table 1. Examples of message for each category of interaction.

Categories	Examples
Information seeking	'Do you know what the asthenosphere and the lithosphere are?'
Information providing: Low-elaboration	'Well! It says that Venus is faster than Earth.'
Mid-elaboration	'This one is called the partial transit. The last one was the peripheral transit and the first one was the central transit.. [1.4].. OK?'
High-elaboration	'I think this is melting because it's hot there in the asthenosphere. It's probably why we have volcano growing on the surface there'
Interaction management	'Hold on! Let's summarize it again'
Task management	'Open this last snapshot. I would like to verify something'
Off-task	'I like this screen. I would like one'

Table 2. Means and standard deviations for the second and third-level coding categories.

	Content-focused				Management-focused	
	HI	MID	LO	IS	IM	TM
				Means (standard deviations)		
Static	0.98 (0.8)	0.33 (0.24)	1.40 (0.84)	0.41 (0.44)	0.20 (0.19)	0.22 (0.19)
Dynamic	1.10 (0.95)	0.28 (0.23)	1.35 (0.74)	0.37 (0.24)	0.24 (0.24)	0.32 (0.34)
Persistency	0.92 (0.84)	0.25 (0.21)	1.16 (0.73)	0.34 (0.29)	0.16 (0.14)	0.33 (0.30)
No-persistency	1.15 (0.91)	0.35 (0.25)	1.62 (0.81)	0.44 (0.39)	0.29 (0.16)	0.21 (0.24)

HE, high-elaborated; MID, mid-elaborated; LO, low-elaborated; IS, information-seeking utterances; IM, interaction management utterances; TM, task management utterances.

($M = 29$ s for *snapshots* group; $M = 24.1$ s for *no-snapshots* group; $F = 2.95$, $P = ns$). Within this elaboration time, pairs produce on average 39.7 utterances (as segmented by our coders), a number that does not vary between conditions ($M = 38.7$ utterances in snapshots condition, $M = 40.7$ in no-snapshots condition; $F = 1.97$, $P = ns$). Therefore, in the analyses reported in the next section, we used utterance-per-minute frequencies. In addition, given the very low number of off-task utterances (less than 0.03% of the corpus), the associated results are not reported here.

How do the main factors affect the process variables?

Univariate factorial analyses of variance were performed to assess the way the learning conditions affected the socio-cognitive process variables. An α level of 0.05 was used for all statistical analyses.

Animation factor
Concerning the total amount of utterances produced during collaboration, the results showed no significant difference between the *static* ($M = 3.14$, SD = 1.24) and *dynamic* ($M = 3.24$, SD = 1.46) conditions [$F(1,75) = 0.09$, $P = ns$].

Concerning the first level of coding categories (i.e. content-focused and management-focused categories), the animation factor did not affect the 'content-focused' utterances [$F(1,75) = 0.02$, $P = ns$, *static* condition: $M = 2.72$; *dynamic* condition: $M = 2.70$]. Participants of the *dynamic* condition produced more 'management-focused' utterances ($M = 0.54$, SD = 0.40) compared to participants of the *static* condition ($M = 0.41$, SD = 0.30). However this result was not significant [$F(1,75) = 2.41$, $P = 0.12$].

In the second and third coding levels, Table 2 indicated that the animation factor did not affect the production of 'information seeking' and 'interaction management' utterances. Nonetheless, the animation factor marginally affected the production of 'task management' utterances. Participants of the *dynamic* group tend to produce more 'task management' utterances than the participants of the *static* group [$F(1,75) = 3.01$, $P = 0.08$].

	HI	MID	LO	IS	IM	TM
Participants (*N* = 76)						
Retention	0.367**	0.195	0.090	−0.057	0.346**	0.002
Inference	0.404**	0.165	−0.065	−0.070	0.298**	0.082

Table 3. Table of correlations between learning scores and interaction categories.

**P < 0.01.

HE, high-elaborated; MID, mid-elaborated; LO, low-elaborated; IS, information-seeking utterances; IM, interaction management utterances; TM, task management utterances.

Persistent snapshots

The participants of the *no-snapshots* condition produced more utterances ($M = 3.62$, SD = 1.46) than the participants of the *snapshots* condition [$M = 2.8$, SD = 1.14, $F(1,75) = 7.41$, $P = 0.008$]. Concerning the second level utterances, the *no-snapshots* group's participants produced significantly more 'content-focused' utterances ($M = 3.13$, $SD = 1.27$) than the *snapshots* group participants [$M = 2.34$, SD = 1.06, $F(1,75) = 8.57$, $P = 0.005$]. Nonetheless, the results did not reveal significant differences between the *no-snapshots* ($M = 0.496$) and *snapshots* ($M = 0.468$) groups about the 'management-focused' utterances [$F(1,75) = 0.11$, $P = ns$].

The *persistent snapshots* affected four second and third-level utterance categories of the corpus. Participants of the *no-snapshots* group produced significantly more utterances of the 'mid-elaborated' type [$F(1,75) = 4.44$, $P = 0.038$], 'low-elaborated' type of information providing [$F(1,75) = 6.71$, $P = 0.012$] and 'interaction management' type [$F(1,75) = 6.71$, $P = 0.012$] than the participants of the *snapshots* group (see Table 2 for mean and standard deviations). Conversely, they produced significantly less 'task management' utterances than the *snapshots* group [$F(1,75) = 3.84$, $P = 0.05$]. To sum up, the *persistent snapshots* group negatively affected the amount of mid- and low-elaborated information providing and 'interaction management' utterances and positively affected the amount of 'task management' utterances produced per minute.

How are the process variables related to the learning outcomes?

In order to answer our second research question, that is, how are peers' interactions related to the learning outcomes, we performed Pearson's correlations between process variables and learning performances.

Results confirmed the importance of verbalizations during collaborative learning by reporting a positive and significant correlation between the total amount of utterances produced per minute and both retention [$r(76) = 0.38$, $P = 0.001$] and inference scores [$r(76) = 0.32$, $P = 0.005$]. More precisely, concerning the first level categories, Pearson's correlations were significant concerning the relation between the 'content-focused' utterances and both the *retention* [$r(76) = 0.36$, $P = 0.002$] and *inference* scores [$r(76) = 0.28$, $P = 0.015$]. Concerning the 'management-focused' utterances, the results showed a significant correlation only with the *inference* score [$r(76) = 0.24$, $P = 0.035$]. The results confirm that more than managing the task and the interaction, it is the interaction about the learning content that is positively related to learning outcomes.

Table 3 presents the Pearson's correlations between the two learning outcomes (i.e. retention and inference scores) and the six verbal categories of the second and third levels (i.e. 'high-elaborated', 'mid-elaborated', 'low-elaborated', 'information seeking', 'interaction management' and 'task management'). Two out of the six categories were significantly and positively related to learning performances. As shown on Table 3, the production of 'high-elaborated' utterances was significantly correlated to the retention [$r(76) = 0.367$, $P < 0.001$] and inference scores [$r(76) = 0.404$, $P < 0.001$]. Therefore, the production of high-elaborated utterances is positively related to the information retention and conceptual comprehension. Surprisingly, there was also a strong relation between the amount of 'interaction management' utterances and learning performance. The results reported a positive and significant correlation between the production of 'interaction management' utterances and the retention [$r(76) = 0.345$, $P = 0.002$] and inference scores [$r(76) = 0.297$, $P < 0.009$].

To sum up, in general, the production of 'content-focused' utterances seems to have a strong relation with learning performance. Further analyses of the content-focused utterances revealed that only the production of high-elaborated utterances are significantly and positively related to learning performance. Concerning the 'management-focused' utterances, a significant relation was found only with the inference scores. In-depth analyses revealed that this relation is mainly explained by the positive correlation between the 'interaction management' utterances and learning performance.

Mediation analyses

We addressed the question of whether some of our interaction categories mediated the effects of our main factors on the learning outcomes. Following Baron and Kenny's (1986) recommendations, we used simple linear regressions and multiple regressions. The authors recommend testing three main steps. The first step is reached if the variations on the independent variables significantly account for the variation on the presumed mediating variable. The second step is reached when the variation of presumed mediating variable significantly accounts for the variation in the dependent variable. The third step is reached when in a multiple regression equation, a previously significant relation between the independent and dependent variable is no longer significant, when controlling for the presumed mediation variable.

Given the fact that the *animation* factor did not affect our process variables, the first step is not met. Mediation did apparently not occur in the effect of animation on the learning performances. However, some of the process variables appeared to be good enough candidates to mediate some of the negative effect of the persistent snapshots: 'information providing' and 'interaction management' utterances. The respective results are reported in the next paragraph.

Concerning the first step of the mediation model, simple linear regressions were used. The results showed a negative relation between the *persistent snapshots* and frequencies of respectively 'information providing' ($\beta = -0.323$, $P = 0.004$, $r^2 = 0.11$) and 'interaction management' ($\beta = -0.294$, $P = 0.01$, $r^2 = 0.087$) utterances. The second step was also verified for these two types of utterances; *inference* scores were positively related to frequencies of 'information providing' ($\beta = 0.278$,

$P = 0.015$, $r^2 = 0.08$) and 'interaction management' utterances ($\beta = 0.297$, $P = 0.009$, $r^2 = 0.09$).

Concerning the last step, following Baron and Kenny (1986), we tested the relation between the persistent snapshots and the *inference* test by controlling for the presumed mediating variables. Concerning the 'information providing' variable, the multiple regression showed that the persistent snapshots factor was no longer a significant predictor of inference test ($\beta = -0.138$; $P = 0.26$) whereas the 'information providing' was still a significant predictor ($\beta = -0.235$, $P = 0.05$). Concerning the frequency of 'interaction management' utterances, this last step showed that the *persistent snapshots* factor was also no longer a significant predictor ($\beta = -0.134$; $P = 0.252$) whereas the 'interaction management' category was still a significant predictor ($\beta = 0.257$, $P = 0.029$, $r^2 = 0.11$). Sobel tests were performed to measure the significance of the indirect effect for the 'information providing' ($z = -1.90$, $P = 0.05$) and 'interaction management' utterances ($z = -1.95$, $P = 0.05$), respectively.

We can conclude that the frequencies of 'information providing' verbalizations and the 'interaction management' verbalizations both mediated some of the *persistent snapshots* factor's effect on the participants' inference scores.

Discussion

For the sake of clarity, we will first discuss the results related to our second research question, namely the relation between the process variables and the learning outcomes. The second part will be devoted to the results related to our first research question: the effect of the *animation* factor and the *persistent snapshots* factor on the interaction variables. In the final section, we will discuss the way some of the verbal interaction variables mediated the detrimental effect of the *persistent snapshots* on participants' inference scores.

Pearson's correlation provided evidence of a relationship between the high-elaborated utterances and the learning performances. These results are consistent with previous findings (Webb 1991; Teasley 1995; Van Boxtel *et al*. 2000). Producing elaborated verbalization and explanation triggers activities such as analogies and reformulations that result in more elaborated knowledge building. As described by Webb (1991), elaborated information providing allows the user to become aware

of inconsistent reasoning and knowledge gaps. 'interaction management' utterances also proved to contribute to learning. This result is more difficult to interpret. Following a strictly cognitive load model, interaction management processes such as conflict management should be expected to imply an extraneous load on learners. Thus, the more management is needed, the more collaboration 'issues' are supposed to have been occurred and the more it should have impaired learning. Another more 'CSCL-flavoured' interpretation is that the frequency of these utterances is an indicator of an explicitly managed collaborative process. As stipulated by Dillenbourg (1999; Dillenbourg & Bétrancourt 2006), a certain collaborative effort is needed to achieve shared understanding. Interaction regulation activities such as management of turn-taking and misunderstandings trigger pragmatic efforts (in terms of grounding, Clark & Brennan 1991) which benefit collaboration. In other words, fruitfully managing interaction guarantees symmetry in participation and a better mutual regulation that has been proven to benefit collaboration (Dillenbourg 1999).

Previous results reported that pairs who studied the dynamic pictures learned more than pairs studying static pictures. However, further analyses on the process variables did not provide evidence of a reliable effect of the *animation* factor on the participants' verbal interactions. One possible explanation for this is that our coding might have missed some important aspects of interaction related to the positive effect of dynamic pictures on collaborative learning performances. Furthermore, this main effect may also be explained by process variables other than social interaction, (i.e. cognitive processes). However, this result remains intriguing since, as reported in the main effects (Rebetez *et al.* 2005), the animation factor had no effect on individual learning on the inference score. Thus, the enhancing effect of animated pictures cannot be attributed only to cognitive processes. The fact that this effect appeared only in collaborative settings implies that the interaction among peers triggered socio-cognitive processes that could also not only be explained on the pragmatic level (i.e. degree of elaboration, management of interaction etc.). In any case, this result contradicted past research on animated pictures in collaborative settings (Schnotz *et al.* 1999). One plausible explanation is that the animations were fragmented into meaningful sequences. Contrary to the study reported by Schnotz *et al.* (1999),

in our study, the collaboration was made of alternating multimedia phases and social interaction phases. Since pairs were almost never studying the material and discussing it at the same time, there was no reason to expect the accumulation of cognitive loads (i.e. load due to the multimedia content coupled to the load due to the social interaction) as stipulated by Schnotz *et al.* Anyway, this hypothesis explains only the fact that the effect of the dynamic pictures was not detrimental to dyads.

The analysis of the second factor – the presence of persistent snapshots of critical steps – leads to more consistent results. Persistent snapshots negatively affected the frequencies of utterances produced. More precisely, three categories of utterances were reduced: low- and mid-elaborated information providing and 'interaction management'. Exchanging information is the basis of collaborative learning. By reducing these types of utterances, persistent snapshots seem to have impoverished collaboration. This assertion is supported by the mediation analyses. The persistent snapshots reduce the frequency of verbal interactions that are related to learning outcomes, namely the 'information providing' utterances and the 'interaction management' utterances. The detrimental effect of *persistent snapshots* on the learning outcomes appeared to be mostly due to the fact that offering a mnemonic and social 'prosthesis' impaired a crucial aspect of the collaboration: the exchange of information and its coordination. On the other hand, snapshots appeared to improve significantly the production of utterances related to the management of the interface. However, these utterances are of less relevance to the learning purpose and may distract the learners from more critical aspects of learning like content-related information exchanging.

Conclusions

This article presents analyses on the process level that aim at furthering results of a previous study in which dyads learning with animations performed better than those learning from static pictures. These results are in contradiction with those found in the literature reporting detrimental effects of animations in collaborative settings (Schnotz *et al.* 1999). We believe that animated pictures serving as external representation provide good opportunities for collaborative learning of dynamic material but under some restricted conditions. Segmenting the animated material provides opportunities for col-

laborants to discuss the material between episodes by distributing over time the moment dedicated to viewing the material and the moment dedicated to discussing it.

Another counter-intuitive result we wanted to explore further was the fact that providing persistent information seemed to strain collaborative learning performances. The process analyses on the interaction level showed that persistent display of critical information reduces the amount of verbalization related to the learning material and enhances verbalization about interface handling. What we can take away from this is that instructional designers should think carefully before adding (socio-)cognitive prostheses that aim to facilitate cognitive and interaction processes. The side effect of these prostheses can be detrimental to social processes. By adding interaction with the interface, these tools can hinder social interaction among learners through what we may call a 'split-interaction effect'.

Acknowledgements

References

Baker M., Andriessen J., Lund K., van Amelsvoort M. & Quignard M. (2007) Rainbow: a framework for analyzing computer-mediated pedagogical debates. *International Journal of Computer-Supported Collaborative Learning* **2**, 247–272.

Baron R.M. & Kenny D.A. (1986) The moderator-mediator variable distinction in social psychological research: conceptual, strategic, and statistical considerations. *Journal of Personality and Social Psychology* **51**, 1173–1182.

Brown A.L. & Palincsar A.S. (1989) Guided cooperative learning and individual knowledge acquisition. In *Knowing, Learning and Instruction: Essays in Honor of Robert Glaser* (ed. L.B. Resnick), pp. 395–451. Lawrence Erlbaum, Hillsdale, NJ.

Bétrancourt M. & Tversky B. (2000) Effect of computer animation on users' performances: a review. *Le Travail Humain* **63**, 311–330.

Bétrancourt M., Bauer-Morrison J. & Tversky B. (2001) Les animations sont-elles vraiment plus efficaces? *Revue d'Intelligence Artificielle* **14**, 149–166.

Bunt H.C. (1989) Towards a dynamic interopretation theory of utterances in dialogue. In *Working Models of Human Perception* (eds B.A.G. Elsendoorn & H. Bouma), pp. 419–455. Academic Press, London.

Clark H.H. & Brennan S.E. (1991) Grounding in communication. In *Perspectives on Socially Shared Cognition* (eds L.B. Resnick, J. Levine & S.D. Teasley), pp. 127–149. APA, Washington, DC.

Dillenbourg P. (1999) What do you mean by collaborative learning? In *Collaborative learning: Cognitive and Computational Approaches* (ed. P. Dillenbourg), pp. 1–19. Elsevier, Oxford.

Dillenbourg P. & Bétrancourt M. (2006) Collaboration load. In *Handling Complexity in Learning Environments: Research and Theory* (eds J. Elen & R.E. Clark), pp. 141–165, *Advances in Learning and Instruction Series*. Elsevier, Amsterdam.

Dillenbourg P. & Jermann P. (2007) Designing integrative scripts. In *Scripting Computer-Supported Collaborative Learning – Cognitive, Computational, and Educational Perspectives* (eds F. Fischer, H. Mandl, J. Haake & I. Kollar), pp. 275–301, Computer-Supported Collaborative Learning Series. Springer, New York.

Dillenbourg P. & Traum D. (2006) Sharing solutions: persistence and grounding in multimodal collaborative problem solving. *The Journal of the Learning Sciences* **15**, 121–151.

Dillenbourg P., Baker M., Blaye A. & O'Malley C. (1996) The evolution of research on collaborative learning. In *Learning in Humans and Machine: Towards an interdisciplinary learning science* (eds E. Spada & P. Reiman), pp. 189–211. Elsevier, Oxford.

Doise W. & Mugny G. (1984) *The Social Development of the Intellect*. Pergamon Press, Oxford.

Hart S.G. & Staveland L.E. (1988) Development of NASA-TLX (Task Load Index): results of empirical and theoretical research. In *Human Mental Workload* (North Holland edition, Vol. 52) (eds P.A. Hancock & N. Meshkati), pp. 139–184. Elsevier Science Publishers B. V, Amsterdam.

King A. (1990) Enhancing peer interaction and learning in the classroom through reciprocal questioning. *American Educational Research Journal* **27**, 664–687.

Lowe R.K. (2003) Animation and learning: selective processing of information in dynamic graphics. *Learning and Instruction* **13**, 157–176.

Mayer E.R. (1989) Models for understanding. *Review of Educational Research* **59**, 43–64.

Mayer R.E. & Chandler P. (2001) When learning is just a click away: does simple interaction foster deeper understanding of multimedia messages? *Journal of Educational Psychology* **93**, 390–397.

Mayer R.E. & Moreno R. (2002) Animation as an aid to multimedia learning. *Educational Psychology Review* **14**, 87–99.

Palmiter S. & Elkerton J. (1993) Animated demonstrations for learning procedural computer-based tasks. *Human Computer Interaction* **8**, 193–216.

Pane J.F., Corbett A.T. & John B.E. (1996) Assessing dynamics in computer-based instructions. Paper presented at the *Conference on Human Factors in Computing Systems*, 13–18 April, 1996, Vancouver, Canada.

Rebetez C., Bétrancourt M., Sangin M. & Dillenbourg P. (2008) Learning from animation enabled by collaboration. *Instructional Sciences* (in press).

Rebetez C., Bétrancourt M., Sangin M. & Dillenbourg P. (2005) Collaborer pour mieux apprendre d'une animation. *Proceedings of the 17th International Conference on Francophone sur l'Interaction Homme–Machine*; 27–30 Sep, Toulouse, France (pp. 171–176), International conference proceedings series. ACM Press, Toulouse.

Rieber L.P. (1990) Animation in computer-based instruction. *Educational Technology Research and Development* **38**, 77–86.

Roschelle J. (1990) Designing for conversations. Paper presented at the *AAAI Symposium on Computer Based Environments for Learning and Teaching*; March 1990, Stanford, CA.

Roschelle J. (1992) Learning by collaborating: convergent conceptual change. *The Journal of the Learning Sciences* **2**, 235–276.

Schnotz W., Böckheler J. & Grzondziel H. (1999) Individual and co-operative learning with animated pictures. *European Journal of Psychology of Education* **14**, 245–265.

Suthers D. & Hundhausen C. (2003) An empirical study of the effects of representational guidance on collaborative learning. *Journal of the Learning Sciences* **12**, 183–219.

Suthers D., Weiner, A., Connelly J. & Paolucci, M. (1995) Belvedere: engaging students in critical discussion of science and public policy issues. Paper presented at *AI-Ed '95, the 7th World Conference on Artificial Intelligence in Education*; 16–19 Aug, Washington DC.

Sweller J. (2003) Evolution of human cognitive architecture. In *The Psychology of Learning and Motivation*, Vol. 43 (ed. B.H. Ross), pp. 215–266. Academic Press, New York.

Teasley S. (1995) The role of talk in children's peer collaboration. *Developmental Psychology* **3**, 207–220.

van Boxtel C., van der Linden J. & Kanselaar G. (2000) Collaborative learning tasks and the elaboration of conceptual knowledge. *Learning & Instruction* **10**, 311–330.

Webb N.M. (1989) Peer interaction and learning in small groups. *International Journal of Educational Research* **13**, 21–39.

Webb N.M. (1991) Task related verbal interaction and mathematics learning in small groups. *Journal for Research in Mathematics Education* **22**, 366–389.

Weinberger A., Fischer F. & Mandl H. (2002) Fostering computer supported collaborative learning with cooperation scripts and scaffolds. In *Proceedings of the Conference on Computer Support for Collaborative Learning (CSCL 2002)* (ed. G. Stahl), pp. 573–574. Lawrence Erlbaum, Hillsdale, NJ.