Collaborative learning with animated pictures: The role of verbalizations

Mirweis Sangin, Gaëlle Molinari, Pierre Dillenbourg,
Ecole Polytechnique Fédérale de Lausanne (EPFL), School of Computer and Communication Sciences CRAFT,
EPFL-CRAFT CE 1 631 Station 1 CH-1015 Lausanne Switzerland,
{mirweis.sangin, gaelle.molinari, pierre.dillenbourg}@epfl.ch

Cyril Rebetez, Mireille Bétrancourt,
University of Geneva TECFA, Acacias 54 CH-1227 Carouge Switzerland,
{cyril.rebetez, mireille.betrancourt}@tecfa.unig.ch

Abstract: The present study examines the benefits of instructional animations in a collaborative learning situation. We investigated the interactions that occur during collaborative learning from different materials: animated versus static pictures, both with or without persistency of information. The animation factor had no effect on the types of interactions but the persistency of information led to fewer content-focused interactions. We also analyzed the impact of these interactions on learning performances: Two types of interactions (content-focused and social relations and coordination management) mediate some of the effects of the permanence factor on learning performances.

Introduction

The present study is aimed at deepening some findings from a larger investigation of the benefits of multimedia instructions in two different settings: individual and collaborative learning. We used two complementary theoretical frameworks, cognitive research on multimedia learning (e.g., Mayer, 2001; Schnotz & Lowe, 2003) and sociocultural studies of collaborative learning (e.g., Baker, 2002; Dillenbourg, 1999). The latter were further explored to highlight the mediating functions of the peers' interaction on learning effects.

Learning with animation

Animated pictures are used in multimedia environments to represent the dynamic aspects of complex subject matters in an explicit way (Lowe, 2004). According to Schnotz and Rasch (in press), animations have two different positive functions in learning. First, they enable learners to perform more cognitive processing (*enabling function*) by providing them with additional information that cannot be displayed by static pictures (e.g., the temporal relationships of entities within a system). Second, they help learners to build a dynamic mental representation (*facilitating function*) by giving them an external support for simulating the behavior of the system depicted.

Although there is some evidence that animations have positive effects on the understanding of dynamic situations (e.g., Mayer & Anderson, 1992), research failed to establish systematic benefits of using animated graphics instead versus static ones (e.g., Bétrancourt & Tversky, 2000). In some cases, animations may even prejudice learning (Lowe, 2004). Lowe (2004) suggested two different types of problems to explain these negative results, i.e., *overwhelming* and *underwhelming*. First, given the limited capacity of working memory, learners may not be able to meet the additional processing demands associated with animations. When learning from dynamic visualization, they have to select relevant elements (for the purpose of the task) from a larger amount of information provided by multiple frames (rather than by the single frame of a static picture) in a very limited time. They have also to keep in memory and integrate information distributed either spatially across the display area or temporally through the different frames of the animation. In other words, the overwhelming effect appears when some aspects of animation (e.g., the overflow of information, their transient aspect) impose a cognitive load that leads learners to reduce their cognitive resources available for learning (see also Sweller, 1999). In contrast, underwhelming may arise when dynamic presentations induce in learners either an illusory feeling of understanding or an investment withdrawal (usually caused by the complexity of the animations). In consequence, these learners do not process information at a deep enough level to ensure learning. This type of problem may occur when learners do not have adequate prior knowledge on the domain to be learned: To avoid cognitive overload, they may focus their attention on some partial aspects of the animation. Moreover, Schnotz and Rasch (in press) showed that the facilitating function of animations can be harmful for experts who do not need external support for cognitive manipulations.

Collaborative learning with animation

Over the past decades, research on collaborative learning has been predominantly carried out within what has been termed the "interactions paradigm" (Dillenbourg, Baker, Blaye, & O'Malley, 1996). They therefore mean that "*The question 'under which conditions is collaborative learning efficient?' is splited into two (hopefully simpler) sub-questions: Which interactions occur under which conditions and what effects do these interactions have*" (Dillenbourg et al., 1996). Researchers focused on "constructive" interactions such as conflict resolution, explanation or mutual regulation, and identified them as potential vehicles of conceptual learning. Indeed, they have in common that they lead students to verbalize knowledge (that would remain otherwise tacit) necessary for building a shared representation of the task.

According to Roschelle (1992) or Crook (1998), collaborating students may benefit from external representations such as pictures or animations, because they can play the role of referential anchors that may resource the construction of a shared understanding. Schnotz, Böckheler and Grzondziel (1999) also argued that computer-based learning environments with animations would be well appropriate for collaborative learning because they could enhance the possibilities for interaction and exploration. In order to test this hypothesis, they carried out an empirical study in which they compared learning from animated versus static pictures in a collaborative situation. Contrary to their previous study on individual learning with the same materials, learning with animations resulted in lower performances with two different tasks. The authors interpreted these negative results in terms of cognitive load (Sweller, 1999). The animations would impose a heavy extraneous cognitive load that would be increased with the social interaction load needed to manage interaction with the peer.

Goal and research questions

This paper reports a study conducted to investigate the effectiveness of animated and static pictures under the conditions of individual and collaborative learning (Rebetez, Sangin, Bétrancourt, & Dillenbourg, 2005; Sangin, Rebetez, Dillenbourg, & Bétrancourt, 2005). Our first results provided evidence that animated pictures result in higher retention and inference performances compared to static pictures. Moreover, contrary to Schnotz et al. (1999), the positive effect of animations on inference performance was higher in collaborative situations than with individual situations. Our results also indicated that, in the individual condition, learners benefit from snapshots of previous steps of animations (permanently presented on the screen). In contrast, this "persistency of information" (or permanence tool) led to lower scores in the collaborative condition. According to Sangin et al. (2005), two possible explanations for this negative (and also counter-intuitive) result could be assessed. First, a *split-interaction effect* could occur due to the addition of cognitive load associated with the interaction with the computer interface on one hand, and with the interaction between peers on the other hand. The second explanation could be some sort of *socio-cognitive underwhelming* effect: The permanence of information could lead the learners being insufficiently engaged in the construction of a shared understanding.

The analysis presented here focuses on the collaborative learning settings. It is conducted according to the "interactions paradigm", and dissociates 3 types of questions (see Figure 1):

A) What is the effect of learning conditions ("animation" and "permanence" factors) on intermediate variables (types of interaction)?

B) What is the relationship between categories of interaction and learning outcomes, i.e., the effects these interactions may have on retention and inference performances?

C) How the intermediate variables (types of interactions) mediate the relationship between learning conditions and learning effects?

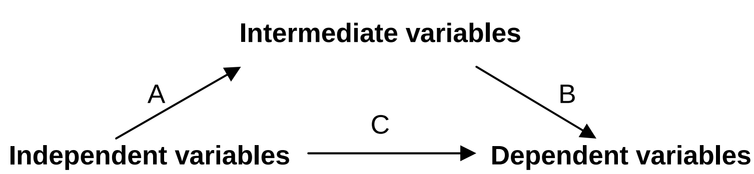

Figure 1. The interactions paradigm (Dillenbourg et al., 1996)

.

Method

Participants and design

Participants were 80 French-speaking students from the Swiss Federal Institute of Technology (Lausanne) and the University of Geneva. We used a factorial 2 x 2 between-subjects design with the first factor being the mode of presentation (static slides vs. animations) and the second factor being the permanence of information ("with-permanence" vs. "without-permanence") according to whether permanent snapshots of the previous sequences were presented or not. In each group, there were 20 participants who worked in dyads. Moreover, all participants reported low prior knowledge in the 2 learning topics (Venus transit and tectonic plates) through an individual pre-test questionnaire (5 multiple-choice questions per topic).

Material and instruments:

Instructional material. The material consisted of 2 successive multimedia instructions (images accompanied by audio narration), the first one designed to explain the phenomenon of the Venus transit (astronomy), and the second one the phenomenon of the tectonic plates (geology). Two versions of each multimedia instruction were constructed. The first version was composed of 12 animated sequences ("animations" version), the second of 12 static pictures ("slides" version). In the static version, each slide corresponded to the most significant picture taken from the corresponding animation. In addition, the static and animated pictures had the same size (800 x 600 pixels), and were laid out on the top-right part of the interface. In the "with-permanence" conditions, 24 grayed labels (12 in the "Venus transit" phase, 12 in the "tectonic plates" phase) (320 x 200 pixels) were provided on the left side of the screen (see Figure 2). After either a slide (slides/with-permanence condition) or an animated sequence (animations/with-permanence condition) was shown, a new label became colored and gave access to a static snapshot. Participants could see a bigger version of the snapshot by moving the mouse over the corresponding label (640 x 480 pixels). The snapshots in the slides condition were the same as those in the animations condition, and thus represented the critical steps of each studied phenomenon. Finally, there were no snapshots in the "without-permanence" conditions (slides/without permanence condition; animations/without-permanence condition).

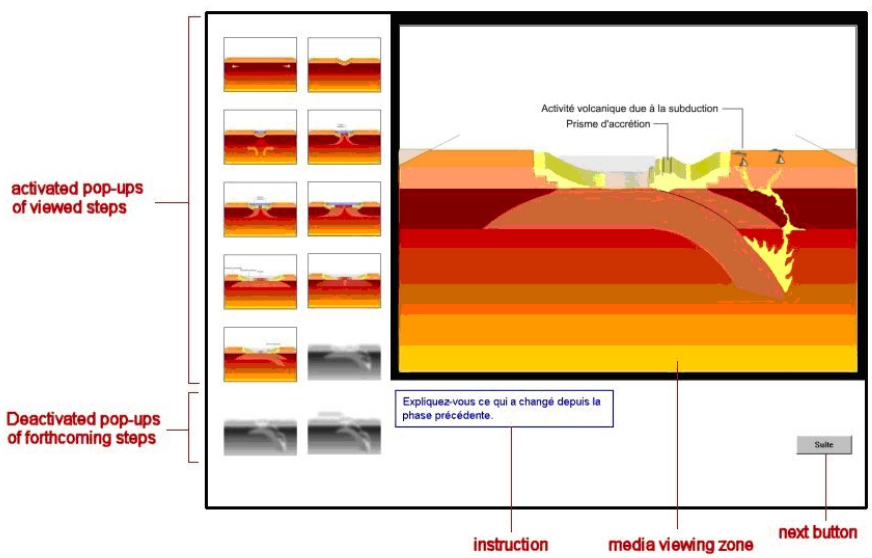

Figure 2. Interface screen-shot of the "with-permanence" condition after the ninth sequence

Learning tests: The testing procedure followed the conventional paradigm used to evaluate the mental model constructed during multimedia learning (Mayer et al., 1992). After the presentation of each multimedia instruction (Venus transit and tectonic plates), all participants were provided with an individual post-test consisting of 2 types of multiple-choice questions (presented in either a graphical or a textual format): 9 retention questions and 7 inference questions. The retention questions measured what participants can remember about each topic; their answers were explicitly quoted either in the images or in the audio narration. The inference questions tested a deeper understanding of each described phenomenon. Indeed, they involved inferring information from the learnt dynamic mental model. Finally, the retention and inference questions were presented in random order.

Apparatus: All instructional materials and (pre- and post-test) questionnaires were computer-based (Authorware software). Responses to multiple-choice questions were recorded by the computer. In addition, the verbal interactions of the 10 dyads were audio and video recorded.

Procedure and scoring

Procedure: The experimental session consisted of two phases – the "Venus transit" phase followed by the "tectonic plates" phase – of 3 parts each: (a) pre-test questionnaire (b) learning phase, (c) post-test questionnaire. Moreover, each learning phase was composed of 2 sub-parts. In the first sub-part, participants were asked to individually study a media-based introduction designed to provide them with prior knowledge necessary to understand the upcoming instructional material. In this introduction, 2 or 3 illustrated paragraphs were presented describing some general concepts on the solar system topic (e.g. movement of planets, transit phenomena) or on the tectonic plates topic (e.g. various layers, movements). In the second sub-part, one pair member had to leave his/her working station, and to join his/her peer to explore together the multimedia presentation. Between each slide or each animation, the instruction – "explain what has changed since the last part" – appeared on the screen; it aimed to encourage peers to interact with each other and to synthesize together the lastly presented information.

Interaction coding system: All videotaped interactions were transcribed verbatim. Turn-taking was used as criteria for the audio corpus segmentation. For purposes of statistical analysis, we decided to use a simple but relevant coding system to analyze the interactions. The coding system was inspired by the RAINBOW method (Baker, Quignard, Lund, Molinari, & Séjourné, 2005) and modified on the basis of our own research interests. The set of categories we used is based on three main distinctions: (a) between interaction that is "inside" the collaborative activity or "outside" (off-task activity) of it, (b) within "inside" activity, between interaction that is "content-focused" or "not content-focused" (social relation, interaction management, task management), and (c) within the "content-focused" interaction, between "information seeking" and "information providing". In addition, the last category was divided into 3 sub-categories corresponding to 3 levels of elaboration (see Table 1). A unique and mutually exclusive category code was associated to each corpus segment and was systematically counted for each participant (two researchers analyzed the interactions; inter-coder reliability was good: Kappa = .81, $p < .0001$).

Table 1. Mutually exclusive codes for the between peers' interaction corpus.

	category	code	description
Content-focus categories	elaborated	A	Elaborated segments based on the pedagogical content plus information not provided by the material (learners' prior knowledge, inferences, deductions etc.; high level elaboration)
		B	Summarizing segments based on all past multimedia sequences. (medium level elaboration)
		C	Summarizing segments containing information based only on the latest multimedia sequence (low level elaboration)
	Information seeking	D	Segments composed of any questioning and information requests
Non-content-focus categories	Interaction-management	E	Segments concerned with interaction management activities
	Task-management	F	Segments related to task management and interface manipulation activities

Results

According to the interactions paradigm, we present the results in 3 steps, respectively represented as A, B and C on Figure 1.

A) How factors determine group interactions?

To explore the influence of the main factors on the intermediate interaction variables, we performed a set of one-way ANOVA tests. The factors are “presentation mode” (slides/animations) and “permanence” (with-permanence/without-permanence). The verbal interactions are either 'content-focus' (categories A, B, C, D in Table 1) or non-content-focused (categories E & F). Since interaction time was not controlled, frequencies (number of interactions divided by the total time) are used for all the following analyses (instead of raw data). We found no significant effect of the “presentation mode” factor on verbal interactions, but an effect of the “permanence” factor: The “content-focused” interactions are more frequent in the “without-permanence” condition ($M = 1.43$) than in the “with-permanence” condition ($M = 1.07$), ($F(1, 75) = 10.95$, $p < .01$).This difference does not appear for “non content-focused” interactions ($F(1, 75) =.46$, $p > .1$).

B) Intermediate and performance variables’ relations

In order to test the relations between the interaction intermediate variables and the performance variables, regression analyses were conducted with the intermediate variables as predictors of the retention and inference scores. Concerning the relation between the retention scores and frequency of the “content-focused” interactions, the regression analysis showed a positive significant relation ($\beta = 0.331$, $p < .01$). A positive significant relation was also found between the “content-focused” intermediate variable and the inference scores ($\beta = 0.250$, $p < .03$). The more the peers provide content-focused utterances, the better their scores seem to be; both on retention and inference items. The regression analyses revealed no significant relation between the frequency of “non-content-focused” interactions and neither the retention scores ($\beta = 0.163$, $p > .1$) nor with the inference scores ($\beta = 0.078$, $p > .5$).

To deepen our understanding of the relations between the interaction variables and the learning scores, a regression analysis was conducted with each functional verbalization categories (namely A, B, C, D, E and F categories; see Table 1) as predictors and peers’ learning performance scores as dependent variables. The predictors explained more than 49% of the peers’ retention scores ($R2 = 0.49$, $p < .01$). Two of the 6 predictors reached the significant level: the A category ($\beta = 0.25$, $p < .02$) and the E category ($\beta = 0.27$, $p < .01$). When we used the inference scores as a dependant variable, the same model explained more than 51% of the peers’ inference scores ($R2=0.51$; $p < .01$). As for the retention scores, the same 2 categories reached the significant level, namely the A category ($\beta = 0.261$, $p < .02$) and the E category ($\beta = 0.260$, $p < .02$). Hence the 2 best predictors for the retention and inference scores are high elaborated information providing segments of the peers’ verbalizations (category A) and the segments concerned with interaction management (category E).

C) Mediation of the performance scores by the intermediate variables

Interaction variables were examined as variables potentially mediating the effect of the permanence factor on the inference scores, i.e., the measure on which the performance of the “without-permanence” conditions’ peers was higher than the “with-permanence” conditions’ peers (Sangin, Rebetez, Dillenbourg, & Bétrancourt, 2005). Four steps were taken to assess mediating effects using regression methods as suggested by Baron and Kenny (1986). The first step is met when variations in the level of the independent variables significantly account for the variation in the level of the presumed mediating variable (see Figure 1, path A.). The second step is verified when the variation of presumed mediating variable significantly account for the variation in the dependent variable (see Figure 1, path B.).

Regression analyses were performed to verify these 2 first steps by taking our interaction categories as potential mediators, the permanence factor (coded -1 for “without-permanence” and +1 for “with-permanence”) as an independent variable and the inference scores as a dependent variable. Two of our variables reached the significant level on the 2 steps. The regression test confirmed that the “content-focused” variable was significantly and negatively related to the "permanence" independent variable ($\beta = -0.357$, $p = .002$), and significantly and positively related to the "inference score" dependent variable ($\beta = 0.250$, $p = .03$). The dialogue category E (interaction management) was also significantly and negatively related to the permanence independent variable ($\beta = -0.293$, $p = .01$), and significantly and positively related to the inference dependent variable ($\beta = 0.298$, $p = .009$). The third step was to verify the relation between the independent and the dependent variables (see Figure 1, path C.) which was marginally significantly negative ($\beta = -0.210$, $p = .069$). This step is the same for our 2 mediating variables.

According to Baron and Kenny (1986), in the last step, we entered both the independent variable (permanence) and the mediator variable (“content focused” and Category E) at the same time as predictors for the dependant variable (inference scores). Concerning the "content-focused" mediating variable, this last regression showed that permanence was no longer a significant predictor (β = -0.138; p = 0.26) whereas the "content-focused" variable was still marginally a significant predictor (β = -0.210, p = .069). Thus, it can be concluded that the frequency of "content-focused" interactions mediates some of the permanence factor's effect on the peers' inference scores. Concerning the E category (interaction management), this last step showed that permanence was also no longer a significant predictor (β = -0.134; p = .252) whereas the E category was still a significant predictor (β = 0.259, p = .002). Thus, it can be suggested that E variable did also mediate some of the effects of the permanence factor on the learners' inference scores.

Discussion

Our previous research had suggested benefits of animated pictures for collaborative learning on both retention and inference performances (Rebetez et al., 2005; Sangin et al., 2005). A second factor was also used in this previous research which offered a permanence tool providing snapshots of the movie critical steps as a memory and grounding support. Surprisingly, the results indicate a negative effect of the permanence tool on collaborative learning. Indeed, peers who were not provided with this permanence tool performed better. Two possible explanations were provided to account for this negative effect, i.e., split-interaction effect and the socio-cognitive underwhelming effect. Do the new results enable us to choose between these explanations?

The interaction analyses suggest that only content-focused interactions are related with peers learning performances. Indeed, the results suggest positive and significant relation between the amount of "content-focused" verbalizations and the retention and inference scores whereas no significant relations occurred between "non-content-focused" verbalizations and the learners' performances. The results are consistent with those found by Webb (1991): Elaborated explanations are positively correlated with explainers' performances. Nonetheless, verbalizations about interaction management are also positively correlated to subjects' performances. A possible explanation of this second effect is that argumentation and negotiation of the collaboration process, even when students do not talk about the content to be learned, are still beneficial for learning (Dillenbourg, 1999).

Regarding our main factors, results suggest no effect of the animation factor on the social interactions. Peers do not verbalize more or less when provided with static pictures versus animated pictures. Regarding the fact that content-focused verbalizations are positively related to learners' performances, no evidence of a potential mediating effect of the verbal interactions on the benefits of animated pictures on learning performances could be provided. Conversely, concerning the permanence factor, the results showed that the permanence of information reduces learners' amount of "content-focused" verbalizations. Moreover, the results support a (partly-) mediating effect of the frequency of interactions on the detrimental effect of the permanence tool on inference scores. This decrease of interaction frequency induced by the permanence of information supports our socio-cognitive underwhelming hypothesis: Offering a permanence tool which primarily aimed to facilitate the grounding processes by stabilizing referential objects occurs to be detrimental to collaborative learning processes because it decreases the learners' verbalizations and interaction management processes.

In this paper, we focused on data concerning the relations between the subjects' verbal productions and their learning scores. Actually, a deeper analysis of our results requires to dissociate the role of the explainer and the explainee, and to treat separately the effects at the individual or the pair level through multilevel analyses.

References

Baker, M. J. (2002). Argumentative interactions, discursive operations and learning to model in science. In P. Brna, M. Baker, K. Stenning, & A. Tiberghien (Eds.), *The Role of Communication in Learning to Model* (pp. 303-324). Mahwah, NJ: Lawrence Erlbaum Associates.

Baker, M. J., Quignard, M., Lund, K., Molinari, G., & Séjourné, A. (2005). *The elaboration of argumentative knowledge in multi-representational computer-mediated debates*. Manuscript submitted for publication.

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.

Bétrancourt, M. & Tversky, B. (2000). Effect of computer animation on users' performances: A review. *Le travail Humain*, *63*(4), 311-330.

Crook, C. (1998). Children as Computer Users: The Case of Collaborative Learning. *Computers Education*, *30*(3/4), 237-247.

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

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

Lowe, R. K. (2004). Animation and learning: Value for money? In R. Atkinson, C. McBeath, D. Jonas-Dwyer, & R. Phillips (Eds.), *Beyond the comfort zone: Proceedings of the 21st ASCILITE Conference* (pp. 558-561). Perth, 5-8 December.

Mayer, R. E. (2001). *Multimedia Learning*. Cambridge: Cambridge University Press.

Mayer, R. E., & Anderson, R. B. (1992). The instructive animation: Helping students build connections between words and pictures in multimedia learning. *Journal of Educational Psychology*, *84*(4), 444-452.

Rebetez, C., Sangin, M., Bétrancourt, M., & Dillenbourg, P. (2005). *Learning from animations is enabled by collaboration*. Manuscript in preparation.

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

Sangin, M., Rebetez, C., Dillenbourg, P., & Bétrancourt, M. (2005). *Collaborative learning with animated pictures: From the persistency of information to the split-interaction effect*. Manuscript in preparation.

Schnotz, W., Böckheler, J., & Grzondziel, H. (1999). Individual and co-operative learning with animated pictures. *European journal of psychology of education*, *14*(2), 245-265.

Schnotz, W., & Lowe, R. K. (2003). External and internal representations in multimedia learning. *Learning and Instruction*, *13*(2), 117-254.

Schnotz, W, & Rasch, T. (in press). Enabling, facilitating, and inhibiting of animation in multimedia
learning: Why reduction of cognitive load can have negative results on learning. To appear in *Educational Technology Research and Development*.

Sweller, J. (1999). *Instructional design in technological areas*. Camberwell, Australia: ACER Press.

Webb, N. M. (1991) Task related verbal interaction and mathematics learning in small groups. *Journal for Research in Mathematics Education*, *22*(5), 366-389.

Acknowledgments

We gratefully acknowledge the contribution of Patrick Jermann, Florence Colomb, Shuja Parvez and Jean-Baptiste Hayé. This research was funded by the Swiss National Science Foundation (grant #11-68102.02).