collaborative learning from animated pictures

Control and collaboration in multimedia learning: Is there a split-interaction?

Cyril Rebetez

Work presented to obtain the

« Diplôme d'études approfondies en psychologie cognitive expérimentale »

Faculté de Psychologie et de Sciences de l'Education

Université de Genève

June 2006

Jury

Mireille Bétrancourt	Professeur, TECFA - Unige	Director
Olivier Renaud	MER, MAD - Unige	
Roland Maurer	MER, Psychologie Cognitive - Unige	

Abstract

This work presents an empirical research on the comprehension and learning from animated pictures explaining the functioning of dynamic systems (tectonic and astronomic phenomenon). In a previous study we found that learners studying in pairs benefited more of animated presentations than individual learners. But, participants had lower results when the interface contained interactive devices. We explained the results as a possible flawing of the learning process caused by the concurrent presence of two types of interaction (with the peer and with the screen). In order to further investigate this so called "split-interaction effect", we designed the present experiment. Two factors were used; the first one was the effect of user's control on the pace of the animation (low vs. high control). The second factor was the learning setting, depending if participants learned the experimental material individually or collaboratively. Results were inconclusive and not in favour of our split-interaction hypothesis. Participants in collaborative and low control condition obtained lower retention scores than other groups. This result goes against the split-interaction effect. Other analysis were led and showed that our factors, but not individual cognitive abilities, influenced the way participants explored the materials. Nevertheless the differences in the material exploration did not led to differences of understanding. These results question the existence of a split-interaction effect but also the importance of delivery features as opposed to instructional design questions.

1 Introduction

One of the most popular forms of visualization in multimedia instruction is animation. It is generally believed that computer animation should facilitate the comprehension of natural or human-made dynamic systems such as weather patterns, circuit diagrams, the circulatory system or the mechanics of a bicycle pump. Computer animation is a powerful medium for displaying how dynamic systems function in a space and time scale accessible to human perception. Some pieces of research supported the idea that computer animation is potentially beneficial to learning, particularly when it depicts the transition between steps that novices cannot infer from static pictures (Tversky, Morrison, & Betrancourt, 2002).

1.1 Learning from animated pictures

Research on animated pictures mostly emerged from the field of text and pictures integration (Mandl & Levin, 1989). The multimedia capabilities quickly made animated pictures an appealing and efficient way to present information or to explain a phenomenon. Nevertheless, numerous types of presentations can be called animation and the definition itself is very broad (Bétrancourt & Tversky, 2000).

Being more sophisticated than plain text or static images, animated graphics are often seen as evident improvement to explain or teach; especially dynamic concepts like mechanical systems or meteorological phenomena. Research by Thompson and Riding (1990) supports the hypothesis that animation facilitates learning since it presents the micro-steps of processes that are absent from static graphics. They used three materials to teach the Pythagorean Theorem to children through a spatial demonstration. The first group saw a static graphic, the second a discrete animation (constituted of two static graphics) and the third a continuous animation. The third group outperformed the two others in term of comprehension. The authors stated that the third material contained more information since it depicted the micro-steps of the process. In other studies, positive effects were found on student motivation and implicit learning (Kaiser, Proffitt, Whelan, & Hecht, 1992; Rieber, 1991). In a more recent study, Catrambone & Fleming Seay (2002) found that for complex inference problems (a sorting algorithm), the animation group performed slightly better than the static group, but not for easy problems. Schnotz, Böckheler and Grondziel (1999) worked with an interactive animation in order to teach time zones. In the first research presented, students using these animated materials obtained better detail encoding scores than the ones using static pictures. Nevertheless, no difference was found on mental simulation task.

1.2 Advantages of animated pictures

Animated pictures are studied extensively in the domain of comprehension and learning. In this perspective the learner need to select, organize and integrate information in order to build a mental model of the phenomenon (Johnson-Laird, 1983). By nature, the potential benefit of animated pictures is to demonstrate dynamic processes. Obviously, phenomenon of dynamic nature can see benefits of being depicted in a dynamic way. Nevertheless, different kinds of presentations are called animated pictures, as the broad definition of and the several different categories tend to show it. Lowe (2003) identified three components of animated graphics which are three kinds of change over time: Transformations are changes in size, shape and general appearance of objects; translations move objects from one place to another, and transitions make them appear or disappear. Lowe claims that these components are the main characteristics and advantages of animation. Learning materials depicting such changes could take benefit of animated pictures to improve their explanations.

To depict dynamic changes in a static graphic forces the designer to use symbols or conventions. Learners need to understand these symbols and to interpret them correctly. A learner studying static graphics needs to mentally infer the changes depicted. This supplementary cognitive processing can create erroneous representations since the learner could infer wrong changes from the picture. But above all, this supplementary cognitive processing requires supplementary cognitive resources. Maybe these resources could be allocated to other important learning processes if the changes were directly depicted in the graphical representation, like for animated pictures. Tversky (2002) underlined that animation adds the change over time as compared with static graphics. Thus, the critical gain of animation over static graphics is the visualization of dynamic processes micro-steps, difficult to mentally infer for novices (Tversky et al., 2002).

Schnotz and Rasch (2005) develop the idea of three possible effects of animated pictures on the learner. The first is a “facilitating” effect. Animation can facilitate the construction of a dynamic mental model, mainly by explicitly showing the micro-steps. The second, called “enabling” effect, goes further: Animated graphics could make possible the comprehension of specific dynamic processes, very hard to apprehend from static pictures for example. Very complex processes have to be seen in action to be understood. The third effect, negative to learning, is called “inhibiting effect”. Animated pictures could inhibit the learner to mentally animate the dynamic phenomenon. The result would be an illusion of comprehension and a poorer mental model. This last effect was comforted by several studies (Lowe, 2003, 2004).

1.3 Drawbacks of animated pictures

Research on animated pictures often failed to find strong benefits of this medium, even when the instructional animation was carefully designed (Bétrancourt & Tversky, 2000; Lowe, 1999). The most common explanation is based on the idea that learning from animated pictures may be cognitively too demanding for novices of a domain. Compared to static representations, animation increases processing complexity on the perceptive level as learners must be attentive to simultaneous changes in the display. The conceptual level is concerned as learners build a "runnable" mental model while they are studying the animation (Mayer, 1989). The working memory is also more involved as learners have to keep transient information in memory: the previous states and trajectory of each element of the system.

When learners are presented with static graphics, they have to mentally animate the presented information in order to understand the dynamism of the phenomenon. Dynamic graphics have the advantage to disambiguate the exact process of change by explicitly showing every transition point between the main steps. Though, presenting changes over time has the major inconvenient of changing over time! Indeed, the informational flow induced by the presentation makes all pictorial information transient. Information comes and goes with every frame. There is no way to get it back when it is gone. Static pictures do not have this additional time-related information but allow learners to quietly explore them and to integrate elements in their mental model.

1.4 Theories on multimedia learning

A large body of research demonstrated that graphics are beneficial to learning. One of the main explanations is provided by the mental model theoretical framework: Understanding a text requires the reader to form a mental representation, structurally analogical to the situation described. Providing an analogical visualization through the use of graphics would facilitate the construction of the mental model (Mayer, 1989). Other models that can be applied to the use of animation for learning purposes also claim that, at some point, the elements that interact with each other should be maintained simultaneously in working memory. For example, Mayer's multimedia learning theory (Mayer, 2001) explicitly refers to the working memory as a selection, organisation and integration point between perceived and encoded information. Models of dynamic system comprehension (Narayanan & Hegarty, 2002) and of text and picture integration (Schnotz & Lowe, 2003) are other inspiring explanations of what happens when one learns from animated pictures.

1.4.1 Cognitive theory of multimedia learning

One of the most widespread theories of multimedia learning is described by Mayer (2001). The model is based on well known cognitive sciences theories: The limited capacity of short term memory, described with the three stages of human memory (Atkinson & Shiffrin, 1968). Sensory memory, short-tem memory and long-term memory being the three stages information passes through in order to be remembered. Later, Baddeley & Hitch (1974) defined their multi-component model of working memory. In particular, they showed that phonological and visuo-spatial information are stored in short-term memory by different processes with different resources. The dual coding theory formulated by Pavio (1986) also supports the separated processing of verbal and non-verbal (or visual) information. Hence, a word encoded in a verbal way will be better recalled if also encoded in a visual form. Mayer also describes a principle of active information processing (Mayer, 1999), stating that learning is more efficient if reinforced by a real cognitive investment and work. A conscious activity from the learner, such as voluntary attention shifts to important elements or mental organization. In the end, Mayer's theory of multimedia learning is close from Attkinson & Shiffrin (1968) model, with three phases of information processing: selection, organisation and integration to a prior mental model. Mayer (2003) insists on the fact that these phases are not a fixed order, but more an iterative process (see figure 1 for a graphical representation).

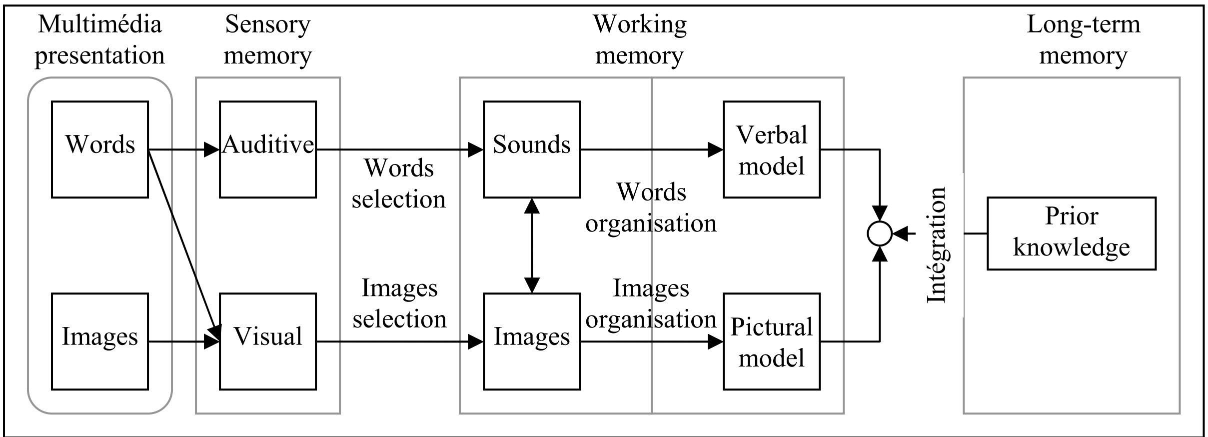

Figure 1 : Model of multimedia learning (Mayer, 2001)

Using this model, Mayer carried out several experiments and described seven principles for the elaboration of efficient multimedia messages (Mayer, 2001):

1. **Multimedia principle:** a message created with words and corresponding images is better retained than a message created with words alone.

2. **Spatial contiguity principle:** learning is improved when images and corresponding words are spatially integrated. For example, legends should be close to the corresponding picture elements.
3. **Temporal contiguity principle:** learning is improved when visual and verbal elements are presented together.
4. **Coherence principle:** learning is better when words, images and sounds not directly useful for comprehension are removed. Anecdotes, illustrations and ambient music are example of often unnecessary elements.
5. **Modality principle:** Animated pictures presented with an audio commentary are better understood than accompanied with on screen text.
6. **Redundancy principle:** learning is better when presenting an animation with an audio commentary than an animation, its commentary and the corresponding text.
7. **Individual differences principle:** these principles are stronger for learners that are novice in the domain; they are also stronger for learners with high visuo-spatial skills.

1.4.2 Text and graphic integration

Schnotz and Bannert's (2003) provided an elaborated model of how verbal-symbolic and depictive information are conjointly and interactively processed in order to form a mental model, which eventually may affect conceptual organization. They define the final organisation of knowledge in two parts. On one hand, a propositional representation gathers together semantic elements, in a symbolic structure. On the other hand, a mental model is formed from perceptive and visual organisation of the different elements in an analogical form, but also from semantic elements. Both representations are strongly related and have similar structures.

The selection of pertinent information uses top-down processing. Previous knowledge guides the gathering of information. In the absence of a pertinent mental model to guide visual exploration, other selection processes will be used. Lowe (2003; 2004) showed that novices learners were mostly relying on perceptive salience to extract information form a meteorological map.

Knowledge organisation is both based on bottom-up and top-down processing. Perceptive organisation of the elements as well as anterior knowledge are used in order to build a mental model linked with a propositional representation. Of course, these selective and organisational functions stand on working memory. Figure 2 gives a graphic representation of this model.

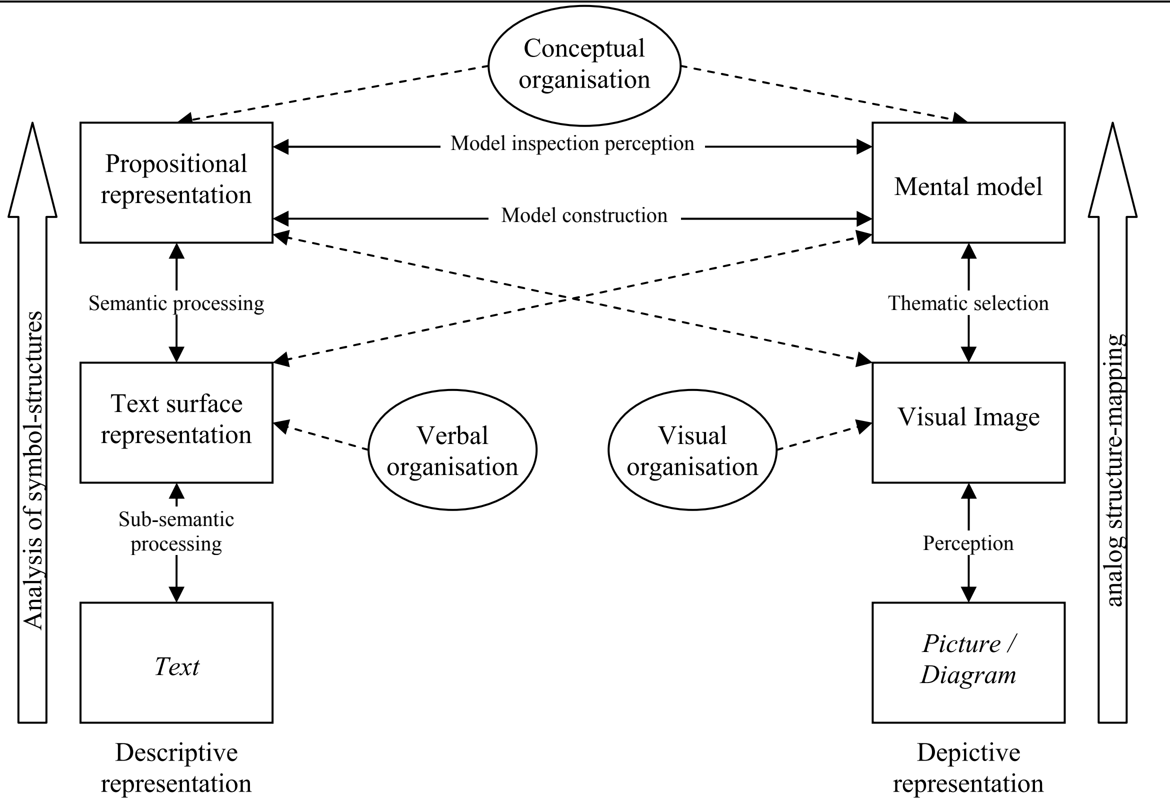

Figure 2 : model of text and graphic integration, from (Schnotz & Bannert, 2003)

1.4.3 Dynamic system comprehension

Narayanan & Hegarty (2002) presented a model of dynamic system comprehension including five steps for the construction of a dynamic representation. Steps involve 1) the decomposition and identification of all elements, depending on their visual or verbal nature. 2) A first organisation leading to static mental models (one verbal and one visual), followed by 3) the identification of referential links between modalities (which word correspond to which graphical element). The dynamic mental model is achieved after 4) the identification of cause-effect relationships and 5) a final integration. It is then possible for the learner to mentally simulate the system in motion, and therefore to infer its functioning under new conditions. On the quality of the learning material depends the quality of the mental model that one can construct. Again, the multiple integrations can only be done in working memory.

1.4.4 Cognitive load explanation

If animated pictures induce higher needs to process the information than static graphics and as human information processing resources are limited, it is plausible that learners studying animation could experience a cognitive overload.

The cognitive load model is often used to explain different experimental results when learning from animated pictures (Paas, Renkl, & Sweller, 2004; Sweller, 2003; Sweller, Chandler, Tierney, & Cooper, 1990). In this model, the authors distinguish three sources of cognitive load: The *intrinsic* cognitive load corresponds to the difficulty of a specific concept to be learned. It increases with the number and complexity of elements and relations. The *germane* cognitive load is generated by cognitive processes involved in the learning activity. Resources used for comprehension and organisation of every element in order to build the mental model are part of this source and necessary for the learning efficiency. The third source is called *extraneous* load. It gathers all additional cognitive processes not directly useful to learning. A deficient material, wrongly organised or incomplete, will firstly need to be reorganised by the learner, in order to be understood and learned. Such supplementary processing, called extraneous load, requires some of the already limited cognitive resources. If the remaining cognitive resources are insufficient for the schema to be elaborated in long term memory, the learning process can be strongly impaired. This theory is based on the assumption that working memory resources are limited to process new elements. Too much information to process at the same time means the loss of, at least, a part of it (Baddeley, 1986).

The poor results of animated pictures, compared to static graphics, can often be explained through this theory. The quantity and novelty of elements to process, mixed with additional selection of information or interface management can overload individual capacities. The extraneous load grows quickly and prevents the effective processing of the animated material. As a result the mental model would be less complete and less organized since fewer elements can be taken into account during elaboration.

1.4.5 Overwhelming and underwhelming effect

The high perceptual and cognitive computation necessary during learning from animations could induce what Lowe (2004) called an "overwhelming" effect. The cognitive cost is sometimes so high that learners decide not to use animations (Lowe, 2003; Pane, Corbett, & John, 1996). It happens when a too wide amount of information is provided to the learners to be processed.

Lowe (2003) also develops an opposite argument. An "underwhelming" effect is observed when learners are "guided" by the dynamic presentation and do not really get involved in its comprehension. Participants do not have to infer the temporal course and possibly do not do it at all. The result is an illusion of comprehension or an investment withdrawal caused by the complexity of the elements and interactions involved. When tested, learners would then report low results. This theory could find support in Mayer's focus on need of activity, also underlined by Ainsworth and Van Labeke (2004), and supported by Palmiter and Elkerton (1993). The underwhelming effect reminds the inhibiting function of animated graphics defined by Schnotz and Rasch (2005). Partisans of a cognitive load explanation would describe the overwhelming effect as the result of the sum of extrinsic, germane and intrinsic loads being higher than what the learner's cognitive abilities can take. On the other hand, an underwhelming effect would be the result of a lack of germane load involved in the learning process. Thus, the cognitive processing necessary for a correct integration are not deployed. However, if the cognitive load theory is often easy to apply to almost any learning situation (and result), it does not always explain the apparition condition of learning results from each other.

1.4.6 Lessons for instructional designers

These different models lead to ergonomic recommendations to build instructional multimedia documents that induce good comprehension. The effectiveness of a presentation is defined as the ability to induce a representation that enables the learner to correctly understand and predict the behaviour of the explained system. We already enumerated Mayer's (2001) seven principles to elaborate efficient multimedia messages. Narayanan and Hegarty's (2002) model also helps the construction of multimedia instructional materials in order to explain dynamic systems. Based on his cognitive load theory, Sweller (2003) describes a series of effects and guidelines to create learning materials:

1. **Goal free effect:** novice learners with a specific learning goal (like a precise question to answer) focus on the goal and pay no attention to other information. This is detrimental to learning.
2. **Worked examples effect:** using known and resolved examples diminish cognitive load and improves comprehension.
3. **Problem completion** effect: the worked out example should be followed by a similar but unresolved problem to maximise motivation.
4. **Modality effect:** two messages on similar elements should be provided through different sensory modalities.
5. **Split-attention effect:** occurs when learners have to process and integrate multiple and separated sources of information. For instance, a geometrical sketch is better understood when textual information is spatially integrated rather than separated (Sweller et al., 1990). This effect is very similar to Mayer (2001) spatial and temporal contiguity principles.

6. **Redundancy effect:** when the same information is presented more than once the multiple processing is negative for comprehension since it increases external cognitive load. If novices can benefit from partially redundant information (integrated text and picture for example), expert's performances can be impaired (Kalyuga, Chandler, & Sweller, 1999). These six first effects try to minimize extraneous cognitive load (to reduce the number of cognitive processes involved that are unnecessary for learning).
7. **Element interactivity effect:** interactivity with the material increases negative effects such as split-attention and redundancy effects.
8. **Isolated interacting elements effect:** with complex models containing multiple interacting elements it is advisable to begin with presenting every element separately.
9. **Imagination effect:** mentally simulating the functioning and interaction of elements allow experts to obtain better results.
10. **Expertise reversal effect:** with experts, several effects are inversed. In this case, classical design rules are advisable instead of those funded on cognitive load.
11. **Guidance fading effect:** as expertise is obtained, learners should be less guided in their exercises.

1.5 Control and interactivity

One simple way to enable learners to process all information and at the same time to keep them engaged with the presentation is to give them control over the pace of the animation. Mayer & Chandler (2001) investigated a very minimalist control device which consisted in breaking down the animation in short sequences. A pause was automatically included after each animated sequence and the learner had control to run the next sequence. The results showed that learners who studied the controlled animation had better transfer performance than learners who studied the continuous animation. The authors interpreted their results in terms of cognitive load, since the pauses allowed learners to integrate partial information before processing further, and thus save resources in working memory. Using a more advanced control panel, Schwan & Riempp (2004) showed that users who had control over the pace and direction of videos learned more rapidly how to tie nautical knots than learners who could not act on the video. The authors claimed that the learners with control could choose the parts on which they wish to allocate more attention and thus distributed their cognitive resources more effectively. Both studies show positive effects for learning when using a controlled animation. However, Bétrancourt & Realini (2005) used three different levels of control of an animation: without control, partial (only pause/play), or full control (also back and forward). Participants had 10 minutes, three times the total duration of the animation, to study it at their own pace, using the available controls. Results showed no difference of the level of control on retention questions. But participants using no control had higher inference scores than the other groups. This contradictory result was explained by the possible creation of a split-attention effect by the control panel (Sweller, van Merrienboer, & Paas, 1998). But this does not explain why other studies involving control did not have this problem. If benefits of control seem obvious concerning multimedia learning, few studies put it into practice and the conditions of application might not be as trivial as it first seems.

1.6 Collaborative learning setting

According to the distributed cognition theory (Hutchins, 1995), learning can be improved when the processing complexity is distributed over several cognitive systems, for example by using collaborative learning situations. The computer-supported collaborative learning (CSCL) field defines the effectiveness of learning through participants' efforts to integrate their different perspectives (Perret-Clermont, 1993; Roschelle & Teasley, 1995). Learners have to negotiate meanings, share and compare their points of view and construct common knowledge. Dillenbourg (1999) defined three main characteristics to characterize a collaborative situation:

1. Interaction between participants must contain a degree of symmetry; the knowledge levels should be relatively comparable, so what one brings can always be useful to the other.
2. Participants share common objectives and interest; this assures the willing of everyone.
3. Work distribution is finely tuned and very flexible. Participants work continuously together.

An important point for collaborative learning is the necessity for co-learners to build a shared representation of the task and of the different elements linked to it. This common ground has to be maintained all along the task, by a process called grounding (Clark & Brennan, 1991; Roschelle & Teasley, 1995). Dillenbourg (1999) claimed that the benefit of collaboration for learning is a 'side-effect' due to elaborating and maintaining a shared representation of the problem at hand (Dillenbourg & Bétrancourt, 2006; Jermann, 2004). Through this process, participants verify their mutual comprehension of the problem. The grounding can be processed through artefacts. These real-world objects are used as symbols or information source to create an array of accessible knowledge around an individual (Moore & Rocklin, 1998). These same artefacts can be used to mediate and communicate the same knowledge with others (Stahl, 2002). Their presence in CSCL environments are designed to support the elaboration and update of a common representation. But the type of grounding depends on interaction context. Setting up a richer environment for collaborating pairs facilitates grounding and support the elaboration of common representations. Sense and common references would be collaboratively built in these external devices, changing them from informative images to shared representation holders: artefacts.

According to collaborative learning theories, facilitating the grounding process would facilitate communication and thus grounding would induce better learning results. On the other side, as dynamic graphics are transient by nature, collaborative learners would have a harder time to refer to objects which are not on the screen anymore (Clark & Brennan, 1991). This argument could explain one result of Schnotz, Böckheler and Grondziel (1999). In their second experiment, participants in collaborative setting were asked to explore an interactive material in order to understand time zones. Participants in pairs showed lower learning results while learning from a dynamic material than from a static material (for both detail encoding and mental simulation tasks). These results were contrary to the ones of their first experiment in which individual learning setting were used. Participants in individual learning condition showed better comprehension using animated displays than static graphics (but only for detail encoding, not for mental simulation task). A lack of permanent references, or artefacts, for the pairs would force collaborative learners to constantly contextualise their discourse. Learners would have to do a lot of communication, not directly related to comprehension. This could be seen, as Schnotz, Böckheler and Grondziel (1999) reported, as an augmentation of extraneous cognitive load.

More recently, Dillenbourg & Bétrancourt (2006) defined the "collaboration load" as the supplementary cognitive processes involved with collaboration (division of labour, verbalization, grounding, mutual modelling). These processes need cognitive resources but also facilitate the learning process. However, like for Schnotz, Böckheler and Grondziel (1999), the collaboration load costs added to a dynamic presentation can induce too much supplementary processing to allow an efficient comprehension and the creation of a "runnable" mental model (Mayer, 1989). Collaborative setting can improve learning since each learner can act as an external memory and an assessment reference for the other learners. The contrary could also be claimed since learners working collaboratively should allocate cognitive resources to elaborate a shared understanding of the situation (Clark & Brennan, 1991; Roschelle & Teasley, 1995) and to maintain a shared representation of the problem (Dillenbourg, 1999).

1.7 The role of individual cognitive skills

All these studies, not always focused on animated pictures, give insights to build adequate multimedia presentations. However, other research fields can be helpful to understand and improve the use of animated pictures for learning. Instead of defining the best delivery features to present a multimedia material, one can ask what cognitive abilities are required to process multimedia and animated materials. Studies on individual differences when learning from animated pictures can give a spot to understand the cognitive mechanics implied.

Individual visuo-spatial abilities and working memory limitation are very important while learning from animation. Several studies from Hegarty and colleagues already showed the importance of visuo-spatial abilities while learning dynamic processes through static images (Hegarty, Kriz, & Cate, 2003; Hegarty & Sims, 1994; Hegarty & Steinhoff, 1997). Hegarty and Sims (1994) selected participants depending on their visuo-spatial abilities, using the paper-folding test (Ekstorm, French, Harman, & Dermen, 1976). They were asked to solve mechanical pulley tasks on the basis of static pictures. The authors showed that participants with a high spatial ability made fewer mistakes and were less impaired by the number of pulleys to mentally animate than low spatial ability subjects. More recently, Hegarty, Kriz and Cate (2003) presented static and animated graphics to explain the flushing cistern system. They showed no comprehension differences for static or dynamic presentations, but again, participants with better spatial abilities showed better performances than participants with low spatial abilities. Hegarty and Steinhoff (1997) allowed participants to take notes on the diagram to indicate the inferred motion of a pulley system. Consistently with their previous research, high-spatial participants solved the problems better than low-spatial participants. Nevertheless, low-spatial participants who could take notes on the diagram significantly improved their performance (but not high-spatial participants).

These individual abilities have an influence on processing instructional material involving visual information. If low-spatial participants can be helped in a problem solving task by a specifically designed material, can it also be the case for learning tasks? Participants with low spatial abilities would benefit from animations since they can only mentally simulate the system at a very high cognitive cost. This function is what Schnotz and Rasch (2005) called the “enabling” effect. Finding efficient learning tools to help learners with low spatial abilities regarding the learning task is definitely a goal for computer supported learning. In a way, the facilitating and enabling functions could need specific conditions to show up. On the other hand, bad design could lead a dynamic graphic to inhibit comprehension.

The individual visuo-spatial abilities create differences in the ability to process dynamic pictures, but what about other types of individual abilities? As the text and picture integration model suggests it, semantic and verbal processing is needed in order to build an efficient propositional representation. Verbal-processing abilities should be taken into account in further research involving individual differences for multimedia learning.

1.8 Collaborative learning from animated pictures

In a previous study (Rebetez, Sangin, Betrancourt, & Dillenbourg, submitted), we investigated the learning setting (pairs or individual); as well as the dynamism of the material (static or dynamic, whether the material was an animation or several static pictures); and permanence (with or without snapshots depicting previous steps of the material). The learning materials consisted in two animations (or series of pictures), explaining dynamic natural phenomenon. These domains were chosen for their intrinsic dynamism and the need to build a dynamic mental model in order to fully understand them. Thus, dynamism really could add something to the graphic representation (Bétrancourt & Tversky, 2000). Figure 3 presents a snapshot of the material in the permanence condition.

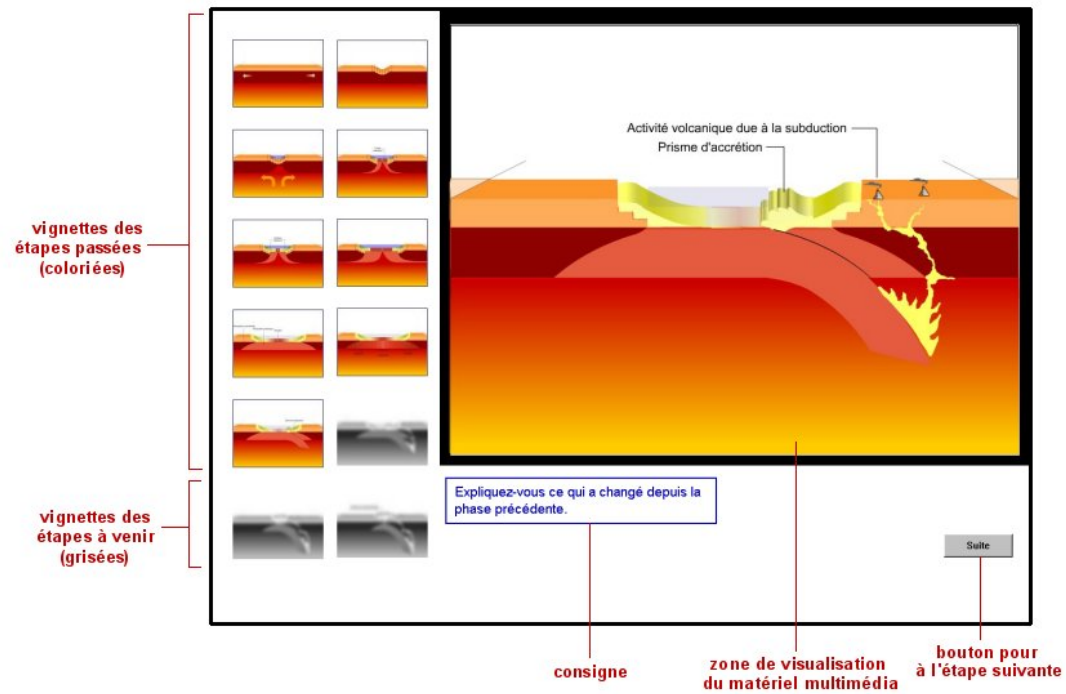

Figure 3 : Snaphot of the material in permanence condition (snapshots on the left).

The learning performance was measured through retention and inference questionnaires for each of the materials. Visuo-spatial skills were also measured for individual participants with the paper folding test (Ekstorm et al., 1976). Our hypothesis implied improved retention and comprehension (inference) when learning from animations. Collaborative learning setting would also improve these variables, as well as the presence of snapshots, especially for individual learners. Principal results were:

- A main effect of dynamism on retention: participants learning with animated pictures remembered more elements than participants learning with static graphics. A positive effect of dynamic presentation appeared. No other main effects were observed on learning performance (neither collaboration nor permanence).
- An interaction between dynamism and learning setting on inference: participants in collaborative condition benefited from animated pictures and attained higher inference scores than collaborative participants with static pictures. In parallel, participants in individual setting obtained comparable results using dynamic or static pictures, but still lower than in collaborative-dynamic condition. Only participants in pairs completely benefited from the potential of animated pictures for learning.
- The spatial skills were strongly correlated with the retention and comprehension. In all conditions, participants with high spatial skills obtained better scores than participants with low spatial skills.
- An interaction between permanence and collaboration showed that participants in individual condition obtained higher inference scores with permanence than without. However participants in collaborative condition had higher scores without permanence than with permanence. If one factor could help understanding, both together were negative.

More detailed results and discussions are available in my Tecfa master thesis (Rebetez, 2004). If the first three main results presented here corresponded to our hypothesis, the fourth one was rather unexpected. The explanation we obtained for this result was called “split-interaction”.

1.8.1 Split interaction

This hypothesis is based mostly on Sweller's cognitive load theory (2003). The split-attention effect, redundancy effect and the fact that participants can process a limited amount of information at one time are premise to this hypothesis.

During the collaborative learning phase, a certain amount of cognitive resources are devoted to maintain the interaction with the peer. Knowledge must be shared and mutual understanding maintained. Dillenbourg & Bétrancourt (2006) call this the collaboration load, but Sweller (2003) would call this an addition of germane and extraneous cognitive load. Anyway, this allows a benefit for the learner in terms of permanence of information and can be used to unload the limited cognitive processors. This has a payoff since in our previous experiment; participants in collaborative conditions better understood the dynamic materials than participants in individual learning conditions.

When using permanent snapshots, participants have to manage their interaction with the computer interface. The additional load of the interface with thumbnails and the handling decisions by the participants add extraneous cognitive load. Some of these negative effects were described by Sweller (2003), such as the element interactivity effect, the split attention effect and the redundancy effect.

If those two aspects have payoffs on the learning side when separated, our previous study shows that a conjunction of them is negative. Both effects are inhibited as if none was present. We named this conjunction a split-interaction effect, since it is created by the simultaneous interaction with the peer and with the computer interface. This hypothesis assumes that both types of interactions at the same time are generating too much extraneous cognitive load to allow the learner to correctly process the information. In other words, learners studying in pairs could handle the complexity of animated information more easily than individual learners, provided that the instructional material does not require a high level of interaction.

Similarly, this hypothesis could explain the results of Schnotz, Böckheler & Grzondziel (1999) described previously. In their experiment, the animated condition contained more interactivity than the static condition. In fact, the dynamic part mostly came from the interaction of learners with the material. In their first experiment (individual learning), the dynamic condition improved learning. In the second one (collaborative learning), this same condition induced lower scores than the static condition. As we see them, these results are very close to those from our previous experiment. The split-interaction hypothesis can explain these results since interaction with the dynamic display is positive for learning in individual condition, but not anymore when a co-learner also reclaims interaction and supplementary cognitive processes. Since these authors performed two different studies to differentiate individual and collaborative learning conditions we can not obtain the main effect of collaboration.

1.9 Research questions

We conducted this second study in order to verify the presence of an inhibiting effect between the presence of interface components and collaborative learning setting. This will help us define the inhibiting effect we were confronted with in our previous experiment and interpreted as a split-interaction effect. At the same time we deepened the interaction participants had with the material and created a more realistic permanence of information. Interactivity and control over an animation have already been studied but never in conjunction with collaborative learning setting.

Moreover, we wanted to explore the effect of visuo-spatial and verbal abilities in learning from multimedia instruction. Several studies from Hegarty and colleagues already showed the importance of visuo-spatial abilities while learning dynamic processes through static images (Hegarty et al., 2003; Hegarty & Sims, 1994): participants with better spatial abilities showed better retention and transfer performances than participants with low spatial abilities. We wanted to confirm these findings but also to extend them to verbal processing capabilities. Since a great part of the mental model is of semantic or descriptive nature (Schnotz & Lowe, 2003), the ability to process such kind of information should have comparable effects on the quality of the dynamic mental model obtained. Since such individual differences have an influence on the quality of the mental model obtained, instructional designers should find ways to reduce the importance of these differences. Control over the animation could allow less spatial learners to process information at their own speed and produce a high quality mental model.

2 Method

2.1 Participants and design

Participants were 86 students from the university of Geneva (age M = 22.3; SD = 2.9), 39 men and 47 women. Due to technical problems, 9 participants were removed from the sample. Participants were randomly assigned to one of 4 experimental conditions (2x2 between subject design). The first factor was learning setting (individual/collaborative). Participants were watching the animated graphic alone (individual) or in group of two (collaborative). Participants in collaborative condition had to speak together between the animation steps and were on their own when answering the tests. The second condition was called control (high/low). Participants with high control could stop the animation and watch any part at any time. Participants with low control had only 12 pre-defined moments at which the animation paused and they had to click to see the next sequence (participants with control also had the 12 forced pause).

2.2 Material

The learning material was an animation presenting the geological phenomenon of rift and subduction. It used dynamic pictures and synchronized audio commentary to explain how oceans and mountain form. The total time of the material was 3.32 minutes and was subdivided in 12 parts according to instructional units. The duration of one sequence varied between 6 and 28 seconds. In the low control condition, the animation paused automatically after each sequence and a “next” button appeared. When feeling ready to continue, participants had to click on it to run the next sequence. After the last sequence, the animation started again from the beginning. In the high control condition, several supplementary controls were accessible to the participants. They could stop and restart the animation at any time, go forward and backward quickly or by sequence. They also could use the cursor showing their position in the animation to directly go to a specific point (see figure 4 for a screenshot in high control condition, the animation is also available in annexe 1). In both conditions, the learning time was fixed to three times the total duration of the animation, i.e. 10.36 minutes (participants could not quit before the end of that time). A small clock displayed the remaining study time and participants were alerted when only 3 minutes were left. The animation also stopped at the 12 pre-definite steps and participants a written prompt asked participants to “explain themselves the changes”.

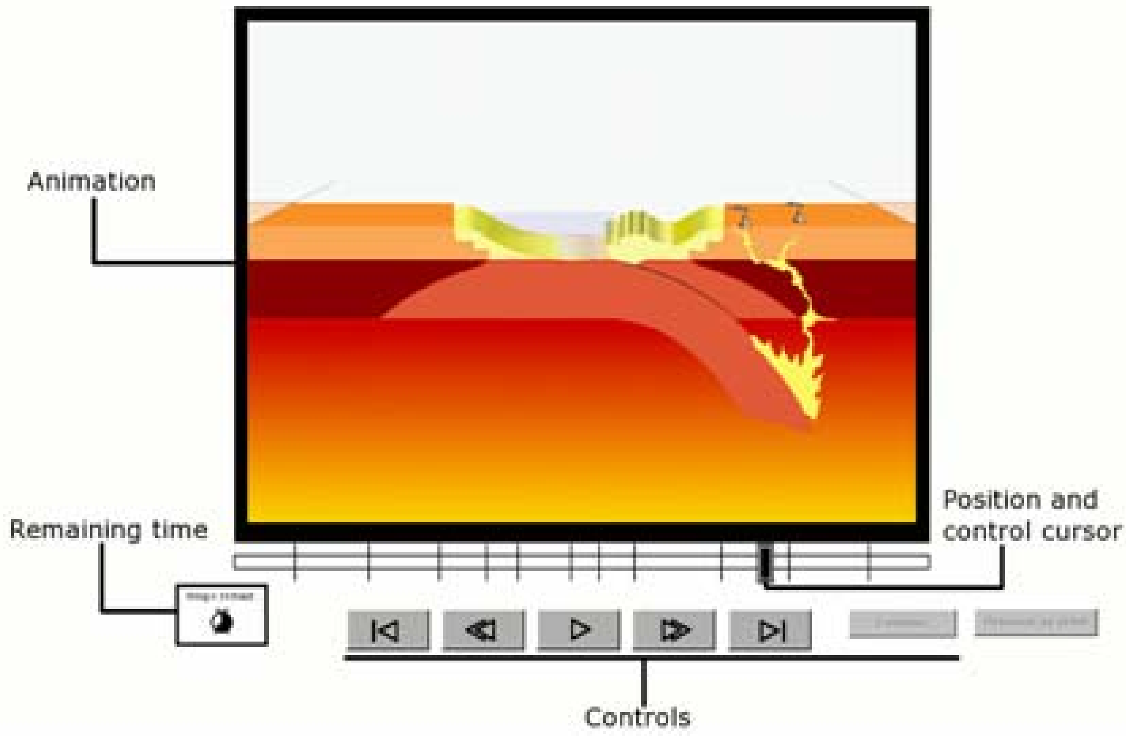

Figure 4: Commented screenshot of the material in high control condition

2.3 Procedure

All participants started the experiment with a short test on their previous knowledge (five multiple-choice questions to verify they are novices in the domain). After a short introduction to the phenomenon, participants were invited to study the animation for 10.36 minutes. Participants in collaborative condition were set in front of the same computer during the whole study time. After every sequence the animation automatically stopped and a written instruction requested participants to elaborate on the changes. This instruction was designed in order to force reflection and mutual explanations in the collaborative condition and to create a reflective and recapitulative opportunity for individual learners. The time elapsed during pauses, both between sequences and during participant-initiated pauses (in the high control condition), was measured and called “elaboration-time”. Duration of pauses and their position relative to the animation were also recorded in order to understand how participants explored the material. Moreover, in collaborative condition, spoken and screen interactions were recorded.

When the learning time was over, participants in collaborative condition were separated. All participants were invited to fill a simplified version of the NASA-tlx test (Hart & Staveland, 1988) to assess their perceived cognitive load. Five scales were kept from the original Nasa-tlx: mental demand, temporal demand, performance, effort and frustration. Participants had to assess each scale by moving a cursor; scores were noted between 0 and 100. After that, participants were asked to freely recall the phenomenon by writing a short paragraph and also had to answer to 16 multiple choice questions. According to the mental model paradigm (Schnotz & Lowe, 2003), two types of question were designed. The response to nine questions had been explicitly showed in the animation or said through the commentary. These questions were called "retention" and involved remembering details from the material. The seven remaining questions involved a deeper understanding of the phenomenon. These were called "inference" questions. In order to compare the performances at both tests, we used the percent of correct answers instead of raw scores. For every question participants also had to give their certainty level in the answer they gave. This was done by moving a cursor on a 0 to 100 scale. The time necessary to answer each question was recorded as well. Finally, two cognitive tests were administrated. Firstly, the paper folding test (Ekstorm et al., 1976) was presented on a computer, one item at a time in order to measure response times. Secondly, similarity test from the WAIS-R intelligence scale (Wechsler, 1981) was administrated by one of the experimenters. Figure 5 shows a summary sketch of the procedure. The complete digitalized experimental material is available in annexe 2.

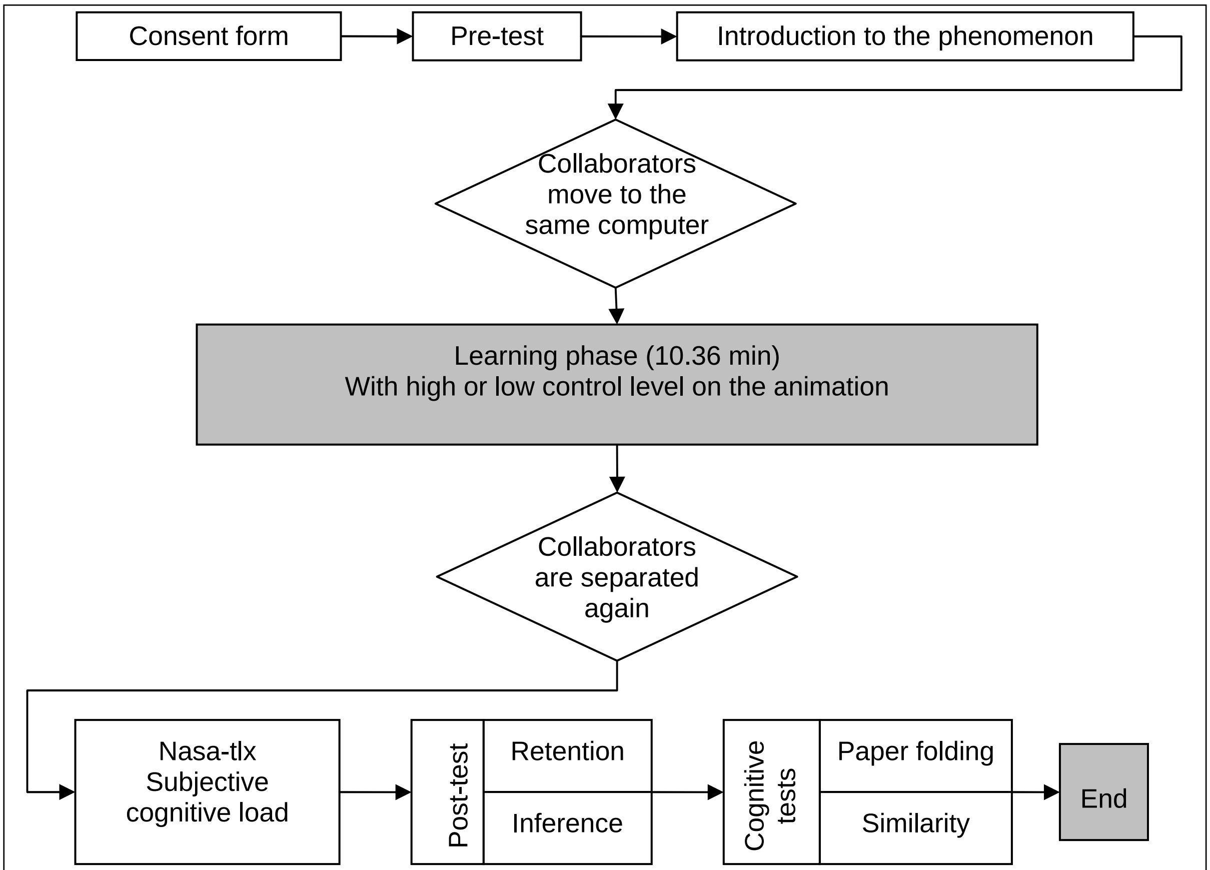

Figure 5: Sketch of experimental procedure

2.4 Data processing

Nine variables were retained to account for the learning performance. All of them were depending of the two main assessments in the post-test phase: the retention and inference questionnaires. These values were expressed in proportion of correct answers. The total of proportion of correct answers (retention + inference) was also used. From these three first dependent variables we constructed six others by using the answer time and certainty level. Data from the two cognitive tests were also processed. All the data are available in annexe 4.

2.4.1 Correctness of answers weighted by the answer time

To account of the answer time at the same time of the given answer, without giving too much power to extreme values, we used the following formula: $\sum \frac{S_{[-1;1]}}{\log(t)}$

Where S is 1 if the answer was correct, and -1 if the answer was incorrect; t is the answer time in seconds. This calculation was done for all three basic performance variables and then standardized. These variables are further referenced as retention_time, inference_time and total_time. Of course, the answer time was also used as is.

2.4.2 Certainty and correctness of answers

To account for the certainty of the answer we used a scale developed by Leclercq & Gilles (1994) designed to reward more participants self-evaluating correctly (see annexe 3 for the detailed scale). For each question, participants had the highest number of points if they answered correctly, with a bonus proportional to their level of certainty. If they answered wrongly they received very few points if they were also very uncertain. But a wrong answer with a high level of certainty would lead to a score loss. The calculation was done for all three basic variables and then standardized. These variables are further referenced as retention_certainty, inference_certainty and total_certainty.

2.4.3 Paper folding score and answer time

For every paper folding answer, the score was calculated using $S_{[-1;1]}*t$. where S is 1 if the answer was correct or -1 if it was incorrect, and t is the time in seconds to answer. Asking participants to answer quickly to the paper folding test and taking their answer time in consideration was done to reduce the effect of overall intelligence factors or resolution strategies, unlinked with visuo-spatial skills.

2.4.4 Exploration of the material

In order to observe how participants explored the experimental material, we created several indicators. They are explained further in the result part but the number of views of each sequences, the order of viewed sequences, the duration of each sequence and the associated forced of voluntary pauses were the basic observations used to create our indicators.

2.5 Hypotheses

We expect a main effect of learning setting: participants in collaborative learning setting will also show better retention and inference results, report lower perceived cognitive load and get lower answering times than participants in individual learning setting. When working in group, participants will use grounding and artefacts to construct a shared representation of the problem. This will be helpful for both learners to build their individual mental model.

We expect no main effect of control since we expect an interaction effect between control and learning setting factors. Because of split-interaction effect, participants in collaborative setting will have higher retention and inference scores than individual learners only in low control condition. Pairs in high control conditions will devote cognitive resources both to interact with the peer and with the computer. Therefore pairs in high control condition will not have higher scores than individual learners.

Accordingly a simple effect of learning setting in low control condition is expected. With low control, pairs will perform higher in retention and inference test than individual learners (they also will report lower perceived cognitive load and answer more quickly). Inversely, with high control, pairs will have lower learning performance scores than individual learners (with higher cognitive load and slower answering times).

According to the literature, individual spatial and verbal skills will be correlated with good performances. Moreover, we can expect that higher verbal skilled participants also benefit more from collaborative setting.

We expect participants with low spatial abilities to take more advantage of the control condition. Without control, participants with low paper folding scores will have lower learning outcomes than high performers. With control this difference should disappear.

3 Results

Because of the high quantity of results, we chose to present them in five parts. The first part will describe the direct effect of our main variables on learning and subjective load. The second part will determine if the experimental conditions induced mediatory variables such as differences in the material exploration. The third part will consider the influence of these mediatory variables on the learning performance and declared subjective cognitive load. The fourth part will depict the effects of individual cognitive skills on the performance. To finish, a fifth part will investigate the potential influence of cognitive skills on the exploration. See figure 6 for an organisational diagram or the results. Data used are available in annexe 4, the file containing spss analysis is available in annexe 5.

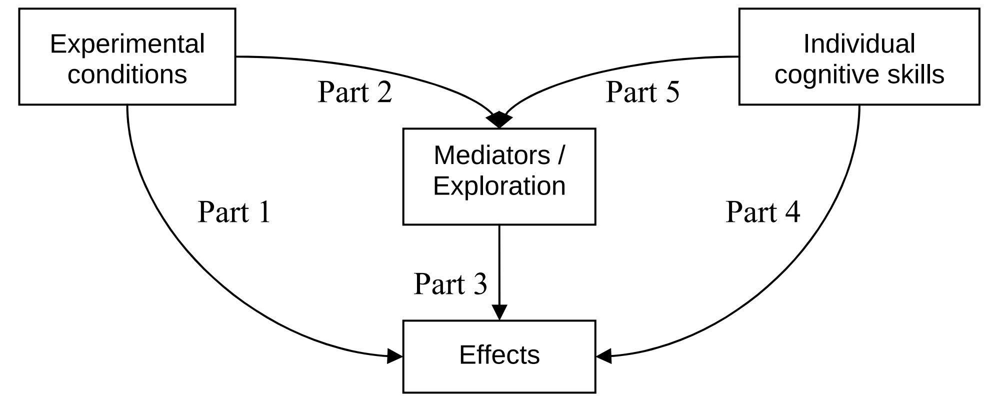

Figure 6: structure and presentation of the results

3.1 *Effects of the experimental conditions*

3.1.1 Learning performances

Nine dependant variables were used to assess the level of understanding of participants: retention score, inference score, total score, the same three weighted with answering time (retention_time, inference_time ant total_time), and again weighted with certainty level (retention_certainty, inference_certainty and total_certainty).

The manova using our two experimental conditions on these nine dependent variables showed few significant results: No effects of the control condition were significant. And only the total weighted by certainty was significant depending on the learning setting factor ($F_{(1;73)} = 5.46$; MSE = .96; $p < .05$; $d = 0.53$). Participants in collaborative condition had a lower total_certainty score than individual participants. Cohen's *d* is used to account of the effect size (Cohen, 1988).

A significant interaction between control and collaboration condition was present on the retention score ($F_{(1;73)} = 4.31$; $MSE = .03$; $p < .05$). As depicted in figure 7, participants learning in pairs with low control showed lower retention performances than all other groups.

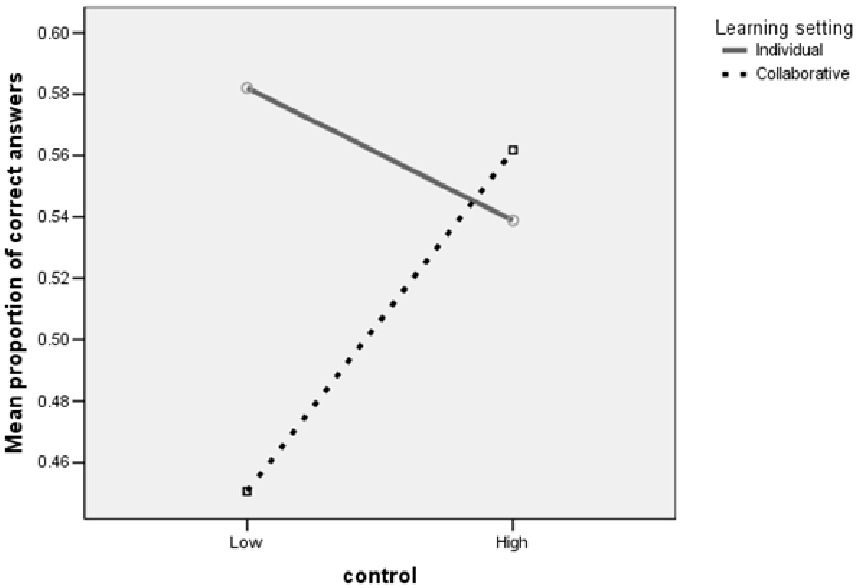

Figure 7: interaction between learning setting and control on retention score.

3.1.1.1 Simple effects

We ran another analysis of variance to investigate the simple effect of the learning setting in low control condition. The analysis showed a great deal of differences between individual and collaborative setting when the control was low: retention scores were significantly higher in individual condition than in collaborative condition ($F_{(1;39)} = 7.99$; $MSE = .02$; $p < .01$; $d = 0.92$), the total scores showed the same effect ($F_{(1;39)} = 6.18$; $MSE = .01$; $p < .05$; $d = 0.86$).

More detailed results:

These results were also present for scores weighted with response time, both for retention_time ($F_{(1;39)} = 7.94$; $MSE = .74$; $p < .01$; $d = .91$), and total_time ($F_{(1;39)} = 6.13$; $MSE = .71$; $p < .05$; $d = .79$). And again for retention_certainty ($F_{(1;39)} = 4.46$; $MSE = .99$; $p < .05$; $d = .67$) and total_certainty ($F_{(1;39)} = 6.85$; $MSE = .78$; $p < .05$; $d = .83$).

Another result showed that participants in collaborative setting had lower retention scores with low control than with high control. This result was significant for retention_time ($F_{(1;34)} = 4.33$; MSE = 1.01; $p < .05$; *d* = .69), the difference on retention score was marginal but we accepted it as significant ($F_{(1;34)} = 4.03$; MSE = .03; $p = .053$; *d* = .67).

3.1.1.2 Covariance

The main analysis of variance was performed again using the pause time as a covariance (time elapsed without the animation running), and a third time using the similarity score. Results were similar to the main manova.

More detailed results:

Using the pause time as a covariance, total_certainty was significant on the learning setting factor (F(1;72) = 4.93; MSE = .92; p < .05), again individual participants performed better than collaborative ones. The interaction was significant over retention_time (F(1;72) = 4.08; MSE = .96; p < .05), with the similar effect: collaborative participants with low control had lower scores than the other groups.

Using the similarity score as a covariance, the total score as well as total_certainty are significant for the learning setting factor (total: F(1;71) = 4.46; MSE = .01; p < .05. total_certainty: F(1;71) = 7.12; MSE = .84; p < .01). Again, participants learning individually had higher scores than participants learning collaboratively. The interaction was significant on retention_time (F(1;72) = 4.08; MSE = .87; p < .05). Collaborative participants with low control had lower scores than the other groups.

3.1.2 Subjective cognitive load

An analysis of variance over our five nasa-tlx scales showed no significant results. Our four experimental groups declared no different cognitive load. Subsequent analyses of variance for simple effects of our factors were also not statistically different.

3.1.3 Relationship between learning performance and perceived cognitive load

The Nasa-tlx “performance” score (indicating the participant’s feeling of success on the task) was negatively correlated with retention score ($r = -.35$; $p < .01$) and with total score ($r = -.31$; $p < .01$). Indicating that participants better remembering the material also knew they understood more (a high nasa-tlx performance score meaning a perceived disastrous understanding). This result is also present for retention_time ($r = -.37$; $p < .01$) and total_time scores ($r = -.36$; $p < .01$).

The Nasa-tlx "effort" score was also correlated with retention ($r = -.28$; $p < .05$) retention_time ($r = -.28$; $p < .05$) and total_time ($r = -.24$; $p < .05$). Although these correlations were light, they reflected that participants who felt the task needed more effort were also the best to retain elements. Nevertheless, this can be linked with the precedent result as the Nasa-tlx effort and performance scales are positively correlated ($r = .51$; $p < .001$)

A subsequent regression analysis showed that the Nasa-tlx performance score is negatively linked with the retention score (β-.34; $t = -2.27$; $p < .05$) and with total score (β-.32; $t = -2.08$; $p < .05$). The effect of effort score is too small to be significant on a regression analysis.

3.1.4 Response time

The time participants needed to answer the questionnaires can also be a good indicator of their learning performance. A more accomplished mental model would require less time to be activated than a mental model less organized and less linked to previous knowledge. This dependent variable is significantly different depending on the learning setting ($F_{(1;73)} = 6.98$; $MSE = 28620.00$; $p < .05$; $d = .60$). Participants in collaborative setting needed less time to answer the questions than individual learners. Effects on control and interaction were not significant.

The same analysis concerning the mean response time of correct answers reported no significant results for our factors.

3.2 Mediators depending on experimental conditions

Facing the deceiving results of our main analysis we decided to explore the activity of participants in the diverse experimental condition. If our factors did not change the level of understanding of our material, it can have changed the way participants studied it. We also need to know if participants with control over the material really used the controls. We supposed that our experimental conditions changed the way participants studied the experimental materials in several ways.

3.2.1 Time passed with the animation playing or paused.

The total amount of time at the disposal of participants to watch and reflect upon the material was limited and controlled (three times the animation's duration). Participants divided their time between watching the playing animation and being in pause. An analysis of variance using the time of animation at play as a dependent variable showed a significant effect on the learning setting factor ($F_{(1;73)} = 70.00$; MSE = 5388.45 $p < .001$; $d = 1.9$). Participants in individual condition spent more time playing the animation than participants in collaborative setting. This also means that participants in collaborative condition spent more time with the animation paused than individual participants.

3.2.2 Visualizing the way learners explored the material.

In order to understand if all participants studied at the material the same way or if different ways of exploring it existed, we started with two types of visualisations of the learners' exposition to the different parts of the material.

The number of times a participant looked at a sequence is already a good indicator of its exploration of the material. In figure 8 we can see how many time participants in the four experimental conditions looked at every sequence. Of course participants with low control looked more at the first sequences than at the lasts (they could not control what they wanted to see next and had to watch the sequences in order until the time was up). Participants with high control also looked more at sequences 8 to 10 than the others (sequences relatively hard to understand involving the beginning of a subduction). Participants in individual conditions looked overall more sequences than collaborating pairs. This is linked with the fact that participants in pairs spent more time with the animation paused (needed more time to talk and to mutualise).

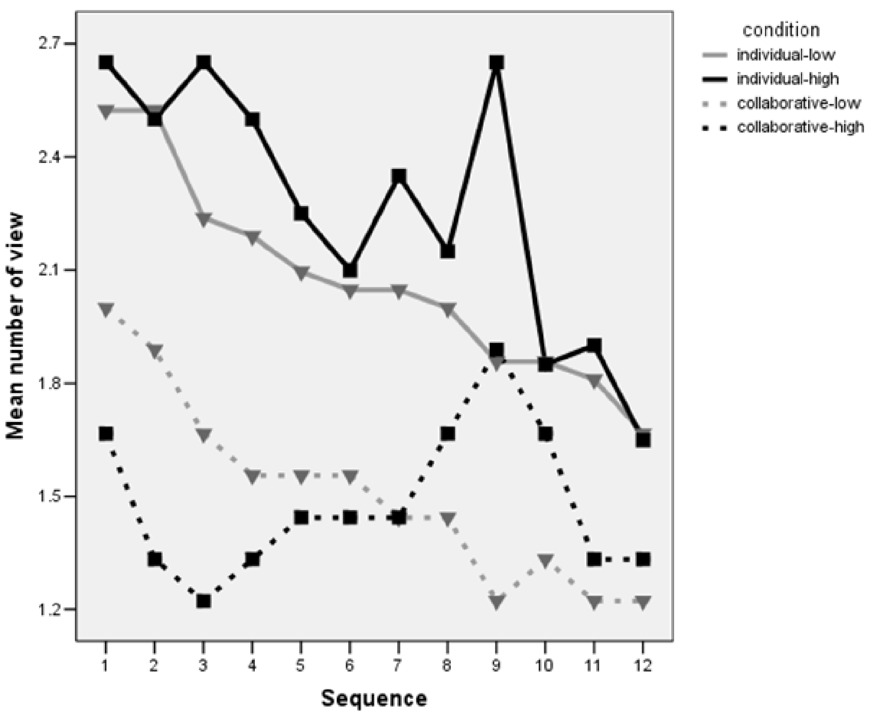

Figure 8: number of views of each sequence depending on the experimental condition

The number of times participants looked at a sequence seems to be different between experimental conditions. Nevertheless if the count is important, the order of exploration is also crucial. To visualise this information, we simply drew a graphic of the sequence corresponding to every pause made by the participant (figure 9).

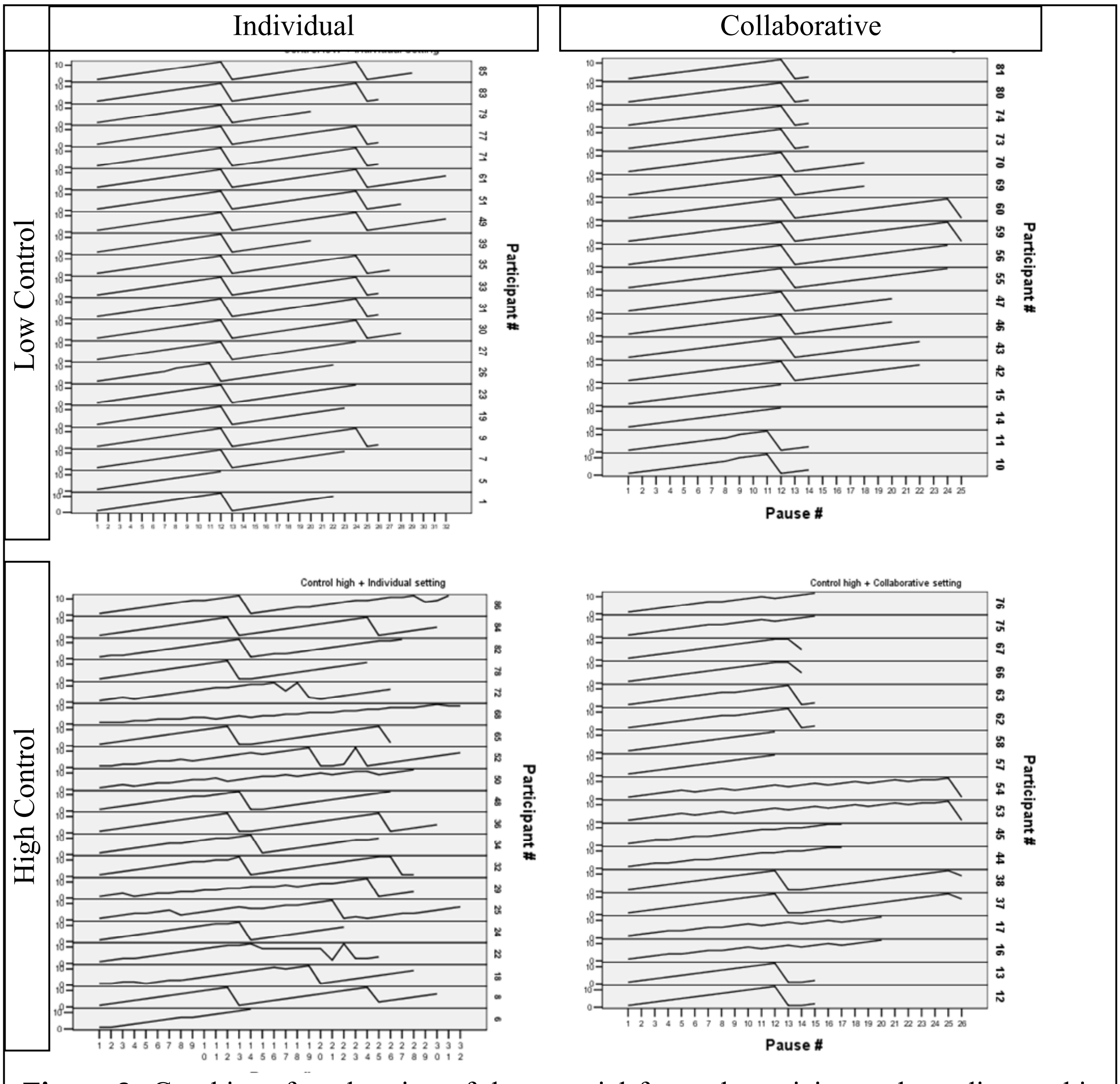

Figure 9: Graphics of exploration of the material for each participant, depending on his experimental condition. Each graphic represent an experimental condition, each line represents a different participant. Abscissa represents the participant's pauses, in order. Ordered represents the sequence number after which the pause occurred. The more the pattern is uneven, the less the participant followed sequences in order. Duration was constant, but as time (or pause duration) is not depicted here, it explains why all lines do not have the same length (all participants did not make the same number of pauses).

These graphics convinced us of the presence of different exploration behaviours between our participants. In order to describe these differences and to evaluate their link with our experimental condition, we developed and used different indicators.

3.2.2.1 *“Global” and “analytic” patterns*

The patterns of exploration presented in figure 9 allow us to divide participants in two groups depending on the way they explored the material.

- The first group is called “global” and regroups participants willing to see the entire animation first, and then to deepen their understanding by starting again or looking at specific sequences. These participants want a global view of the phenomenon. They typically spent less time in pause, at least before having seen the whole material once. A participant viewing all the sequences several times, making short pauses and looking at the sequences in the order is classified as “global”. Participant n°61 in low control + individual condition is typically in this category.
- Participants in the “analytic” group want to understand every sequence before getting to the next. They build their knowledge on solid bricks and typically spend more time reflecting after each sequence or look at it again before they get to the next one. Participants looking several times the same sequence before looking the next one, making long pauses between steps (i.e. with a short line in figure 7) are “analytic” ones. Participants n°53 and 54, a pair in high control condition, are good examples of “analytic” participants.

We tagged every participant as “analytic” or “global”. These categories refer more to exploration behaviours than meta-cognitive strategies. It is not likely that participants developed elaborated approaches to the content. Despite their apparent link to individual differences research domain, these categories refer only to observed behaviour. We do not claim to describe a cognitive state or capacity here.

When all the participants were tagged, we calculated a khi2 on the repartition of these approaches depending on the experimental conditions. Exploration patterns were different than the theoretical distribution both for the control factor (khi2 = 5.76; $p < .05$) and for the learning setting factor (khi2 = 18.07; $p < .001$). In the low control condition we observed more “global” participants, and in the high control condition the proportion of “analytic” participants was higher. But there were also more “global” in individual condition and more “analytic” in the collaborative condition). So depending on the experimental condition, participants explored the materials differently. Table 1 summarizes this distribution according to experimental conditions. Table 2 also presents total scores depending on the “global” or “analytical” approach.

3.2.2.2 "Duration of first passage" pattern

Derived from the global versus analytical approach, we decided to measure the time necessary for a participant to watch the whole twelve sequences. Participants taking more time to understand each sequence before getting to the next one would need more time than participants interested in seeing the whole phenomenon before studying it more closely. As the result was a quantitative variable we would have been able to perform an anova if its distribution was normal, which was not the case. Nevertheless, we split the participants in two groups depending on the median of duration of first vision (fast or slow) and processed a khi2 analysis.

Participants in individual conditions were mostly quick to see the entire material once, on the contrary participants in collaborative condition were more often slow to make their first view (khi2 = 21.86; $p < .001$). The distribution was not different from theoretical strength concerning the control condition (khi2 = 2.19; *ns*). Again, the experimental conditions changed the way participants explored the material.

Nevertheless, most of the participants categorised as "global" were also "fast" to view the material once, and most of the "analytic" participants were also "slow" to view the material once (khi2 = 48.47; $p < .001$).

3.2.2.3 "Exposal time" approach

Using the total time spent playing the animation; we median-split the population. They were tagged "high exposal" if they spent more time playing the animation and "low exposal" if they played it the less. We looked at the repartition of these participants in the diverse groups and saw that the collaboration factor changed the way participants would use the animation (khi2 = 31.24; $p < .001$). Participants in individual setting were for the most in the "high exposal" group and participants in collaboration were for the most in the "low exposal" group.

Again, this way to distinguish participants' explorations is very similar to the "global/analytical" (khi2 = 35.60; $p < .001$), participants in "low exposal" being for the most "analytical" and "high exposal" ones being "global". This is also true as compared to the "duration of first vision" groups (khi2 = 42.19; $p < .001$). Participants being less exposed to the animation are also slow to see it once (they spend more time in pause between sequences), and off course the most participants highly exposed to the animation are quick to see it entirely the first time. The duration of first vision is also strongly negatively correlated with the exposal time ($r = -.70$; $p < .001$).

3.2.2.4 Levenshtein linearity approach

The Levenshtein distance (Levenshtein, 1966) is an algorithm giving the minimum number of transformations needed to change one character string into another (in terms of addition, substraction and permutation of characters). The application here was to transform the exploration pattern of participants into a string (each letter corresponding to one sequence and each occurrence to one pause, the order is preserved). We used the Levenshtein distance to compare each exploration pattern to the pattern the more linear possible with the same number of pauses (letters in the string). This way, we obtained an index of linearity for the participant's exploration. The lower the Levenshtein distance, the more linear the exploration was. The spss script used is available in annexe 6.

As the distribution of this variable was far from normality we could not process any variance analysis. Nevertheless, we calculated khi2 as for the other exploration indicators. Of course, participants with low control could only have a linear approach so the significant effect was expected ($khi2_{(77)} = 54.87$; $p < .001$). On the collaboration factor, the difference was not significant ($khi2_{(77)} = .01$; *ns*). As all participants in low control condition were unable to have a strategy other than linear we split again the sample with the same variable but using only participants in high control condition. With this more specific analysis, the effect of the learning setting factor was still not significant ($khi2_{(38)} = .92$; *ns*). Hence, the linearity of exploration depends only of the control factor (i.e. the possibility of being non-linear) and not of the presence of a co-learner.

This distribution is comparable with the analytic/global distinction ($khi2_{(77)} = 5.76$; $p < .05$), a majority of participants classified as "global" were also "linear" and most of "analytic" participants were also "non-linear". Nevertheless, these categories are not comparable to "duration of first passage" ($khi2_{(77)} = 2.19$; *ns*) nor "exposal time" approaches ($khi2_{(77)} = .01$; *ns*). However, the Levenshtein distance is positively correlated with the duration of first vision ($r = .39$; $p < .001$) and with the exposal time ($r = .29$; $p < .05$). The more linear the exploration, the shorter the time spent to see all the sequences once, nothing surprising.

3.2.2.5 Homogeneity within groups

We attempted to verify the intra-group homogeneity of exploration patterns through a Kendall omega in every condition. But as the number of pauses changes between participants we could not process this type of analysis and could not respond to this question.

Table 1 summarizes the results obtained through the different approaches. The conclusions are rather consistent and the distinction “global-analytical” is very persuasive.

Table 1

Number of participants, mean and standard deviation for total score in the four experimental conditions, depending on the identified exploration behaviour. N.B. the 3 participants in low control condition classified as non-linear are due to drops in the recordings. The median-split number 4 is not informative.

Exploration pattern		Level of control		Learning setting	
		Low	High	Individual	Collaborative
		Number of participants			
1	Global	24	13	29	8
	Analytical	15	25	12	28
2	Fast 1st vision	23	16	31	8
	Slow 1st vision	16	22	10	28
3	Low playtime	20	18	8	30
	High playtime	19	20	33	6
4	Linear	36	3	21	18
	Non-linear	3	35	20	18
Among high control only					
5	Linear	n/a	n/a	8	10
	Non-linear	n/a	n/a	12	8

3.3 Learning effects depending on mediator variables

Participants explored the material differently, and we saw in last part that the way participants explored the materials could be due to their experimental condition. The next question would be to understand if the way participants explored the materials induced differences in their understanding. Table 2 displays total learning scores for participants depending on their experimental condition and on their exploration pattern (based on the global/analytical distinction).

Table 2

Number of participants, mean and standard deviation for total score in the four experimental conditions, depending on the identified exploration pattern.

Exploration pattern	Learning setting: Individual						Collaborative								
	Level of control: Low			High			Level of control: Low			High					
	Total Score			Total Score			Total Score			Total Score			Total Score		
	N	*M*	*SD*	*N*	*M*	*SD*	*N*	*M*	*SD*	*N*	*M*	*SD*	*N*	*M*	*SD*
Global	18	.50	.10	11	.45	.16	6	.40	.09	2	.38	.09	37	.46	.12
Analytical	3	.56	.17	9	.53	.11	12	.43	.11	16	.48	.14	40	.48	.13
Total	21	.51	.11	20	.49	.14	18	.42	.10	18	.47	.14	77	.47	.13

Considering the different approaches to understand the exploration patterns, we had one common problem: participants were unevenly distributed in the experimental conditions depending on these variables. This is natural as conditions by themselves seemed to induce the exploration patterns. Some groups would count less than five participants. In these circumstances it was not very judicious to process an analysis of variance with three variables.

However we processed analysis of variances using only the exploration pattern as a factor, knowing that doing so would increase the unexplained variance between groups. The risk taken was of Type II, to fail to see an effect. Analyses were run on the nine performance variables and the five Nasa-tlx scales but no effect was significant for any of the five approaches. None of the identified exploration patterns could prove an effect on the performance or on the perceived cognitive load.

But there was one exception, when splitting the participants of the high control group depending on their linearity value (obtained through Levenshtein distance). In this case, for high control participants only, having a non-linear approach is linked with better performances. In particular, the total score is higher for non-linear participants than for linear ones ($F_{(1;36)} = 6.40$; $MSE = .11$; $p < .05$; $d = .84$). This result is also valid for total_time ($F_{(1;36)} = 6.89$; $MSE = 6.96$; $p < .05$; $d = .85$) and total_certainty ($F_{(1;36)} = 7.86$; $MSE = 6.91$; $p < .01$; $d = .90$).

Instead of using median split to create new factors, some of our indicators could be used as quantitative variables. The time of first vision, time of playing animation and the Levenshtein distance are all continuous variables. If their distributions were far from normality and did not allow us to use them as dependent variables in an anova, these indicators still gave us some information through correlations. Only the Levenshtein distance showed significant correlations, they were positive with the total score ($r = .23$; $p < .05$), retention_time ($r = .23$; $p < .05$), total_time ($r = .25$; $p < .05$), total_certainty ($r = .25$; $p < .05$) but also with mental demand ($r = .24$ $p < .05$) and temporal demand ($r = .26$; $p < .05$) from the nasa-tlx.

3.4 Effects of individual cognitive skills

The paper folding and similarity tests show strong correlations with several performance variables. In particular, the similarity score is positively correlated with retention ($r = .34$; $p < .01$) and total scores ($r = .35$; $p < .01$). The paper folding score (weighted with response time), was also correlated with retention ($r = .25$; $p < .05$), and total scores ($r = .27$; $p < .05$).

Further linear regressions showed that the retention score was increasing proportionally with the verbal skills ($\beta = .34$; $p < .01$), but also with the spatial ability ($\beta = .25$; $p < .05$). The same result was found for the total score, both for similarity ($\beta = .35$; $p < .01$) and paper-folding ($\beta = .27$; $p < .05$).

More detailed results:

The similarity score was correlated as well with retention_time ($r = .33$; $p < .01$), total_time ($r = .33$; $p < .01$), retention_certainty ($r = .30$; $p < .01$) and total_certainty ($r = .29$; $p < .05$).

The paper folding score was correlated with retention_time ($r = .24$; $p < .05$), total_time ($r = .25$; $p < .05$) and total_certainty ($r = .24$; $p < .05$).

3.4.1 Median split: skilled versus not-so-skilled

After two median split depending on similarity and paper-folding score, we distributed participants according to two new factors: visual skills (high/low) and verbal skills (high/low). An anova using these two factors was performed and the visual skills factor showed several significant effects. Participants with high visual skills performed better at the retention test than participants with lower visual skills ($F_{(1;72)} = 4.64$; MSE = .03; $p < .05$; $d = .56$). The effect is also visible on total ($F_{(1;72)} = 5.29$; MSE = .02; $p < .05$; $d = .64$), retention_time ($F_{(1;72)} = 4.55$; MSE = .90; $p < .05$; $d = .60$) and total_time ($F_{(1;72)} = 5.04$; MSE = .89; $p < .05$; $d = .63$).

No effect of the similarity groups was significant but the interaction showed an effect on retention_certainty ($F_{(1;72)} = 6.02$; MSE = .92; $p < .05$) and total_certainty ($F_{(1;72)} = 4.32$; MSE = .89; $p < .05$).

We were not able to perform a more complete analysis due to the very small number of participants in the groups when using our main factors. Nevertheless and despite the lack of explained variance with only these two factors, several effects of individual cognitive skills were underlined.

3.4.2 Regressions in the diverse experimental conditions

In order to investigate the effect of individual cognitive abilities in our diverse experimental conditions, we computed four different linear regressions in each of our conditions. Similar results would have meant that the two aspects were not interacting with each other. Table 3 regroups the regression results for both cognitive skills indicators in the four experimental conditions.

The positive relationship between retention and paper folding was present only for pairs with high control ($\beta = .51$; $p < .05$). This was also the case for total score ($\beta = .52$; $p < .05$).

The regression between retention score and similarity test is significant in the collaborative-low control condition ($\beta = .48$; $p < .05$) and in individual-high control ($\beta = .51$; $p < .05$). The result can also be accepted as significant in collaborative-high control condition ($\beta = .45$; $p = .06$).

Concerning total score in relationship with the similarity test, only the regression in collaborative-high condition is significant ($\beta = .51$; $p < .05$).

Table 3

Linear regression results between performance variables and cognitive skills in every experimental condition. Significant results are grayed. Retention and total scores were overall significant in a regression with the paper folding test.

Performance score	Learning setting											
	Individual						Collaborative					
	Level of control						Level of control					
	Low			High			Low			High		
	Similarity test											
	B	*β*	*p*	*B*	*β*	*p*	*B*	*β*	*p*	*B*	*β*	*p*
Retention	.57	.01	.95	.10	.51	.02	.03	.48	.04	.12	.45	.06
Inference	.05	.39	.09	.49	-.09	.72	.30	.10	.67	.11	.39	.11
Total	.65	.24	.32	.23	.26	.31	.19	.38	.09	.12	.51	.03
	Paper Folding test											
	B	*β*	*p*	*B*	*β*	*p*	*B*	*β*	*p*	*B*	*β*	*p*
Retention	.57	.24	.29	.54	-.06	.82	.45	.01	.96	.58	.51	.03
Inference	.40	.25	.28	.42	.59	.56	.37	-.25	.32	.36	.27	.27
Total	.50	.36	.11	.49	.06	.82	.41	-.68	.51	.48	.52	.03

3.5 Mediators and individual cognitive skills

In order to investigate whether individual cognitive skills could induce a way of exploring the material, we compared our four exploration approaches with the two individual cognitive skills indicators.

3.5.1 Global or analytical approach

Comparing the repartition of participants depending on their exploration pattern and their classification as high or low on paper folding test, we found a completely even distribution ($khi2_{(77)} = .01$; *ns*). The repartition on the similarity level showed also no significant difference with the theoretical numbers ($khi2_{(77)} = .01$; *ns*). The way to explore the material was not dependent on one of the individual cognitive skill measured here.

3.5.2 First vision duration approach

This indicator was not correlated with the paper folding score (r = .13; *ns*) nor the similarity score (r = .10; *ns*). We also looked at the distribution of participants after a median split using this variable. The distribution was not different as the theoretical one both compared with the visual skills level ($khi2_{(77)}$ = 1.05; *ns*) and the verbal skills level ($khi2_{(77)}$ = .20; *ns*).

3.5.3 Exposition time approach

Again, no correlation was found between this indicator and the paper-folding score (r = .01; *ns*), nor the similarity test (r = -.07; *ns*). The distribution was not different that the theoretical one as compared with the visual skills level ($khi2_{(77)}$ = 1.05; *ns*) or the verbal skills level ($khi2_{(77)}$ = 1.86; *ns*).

3.5.4 Levenshtein distance approach

Still no significant correlation was present comparing this variable with paper folding (r = .13; *ns*), nor similarity test (r = -.08; *ns*). And with this indicator also, the distribution of participants after a median split was not dependent of the visual skills level ($khi2_{(77)}$ = .01; *ns*) nor the verbal skills level ($khi2_{(77)}$ = 1.32; *ns*).

Considering only participants with a high level of control (the only ones able not to be linear), did not change the analysis. Participants were not differently linear depending on their visual skills level ($khi2_{(38)}$ = 3.80; *ns*) or their verbal skills level ($khi2_{(38)}$ < .01; *ns*).

3.6 Results summary

As the results are numerous and complex, we constructed the diagram showed in figure 10 to summarize the principal results.

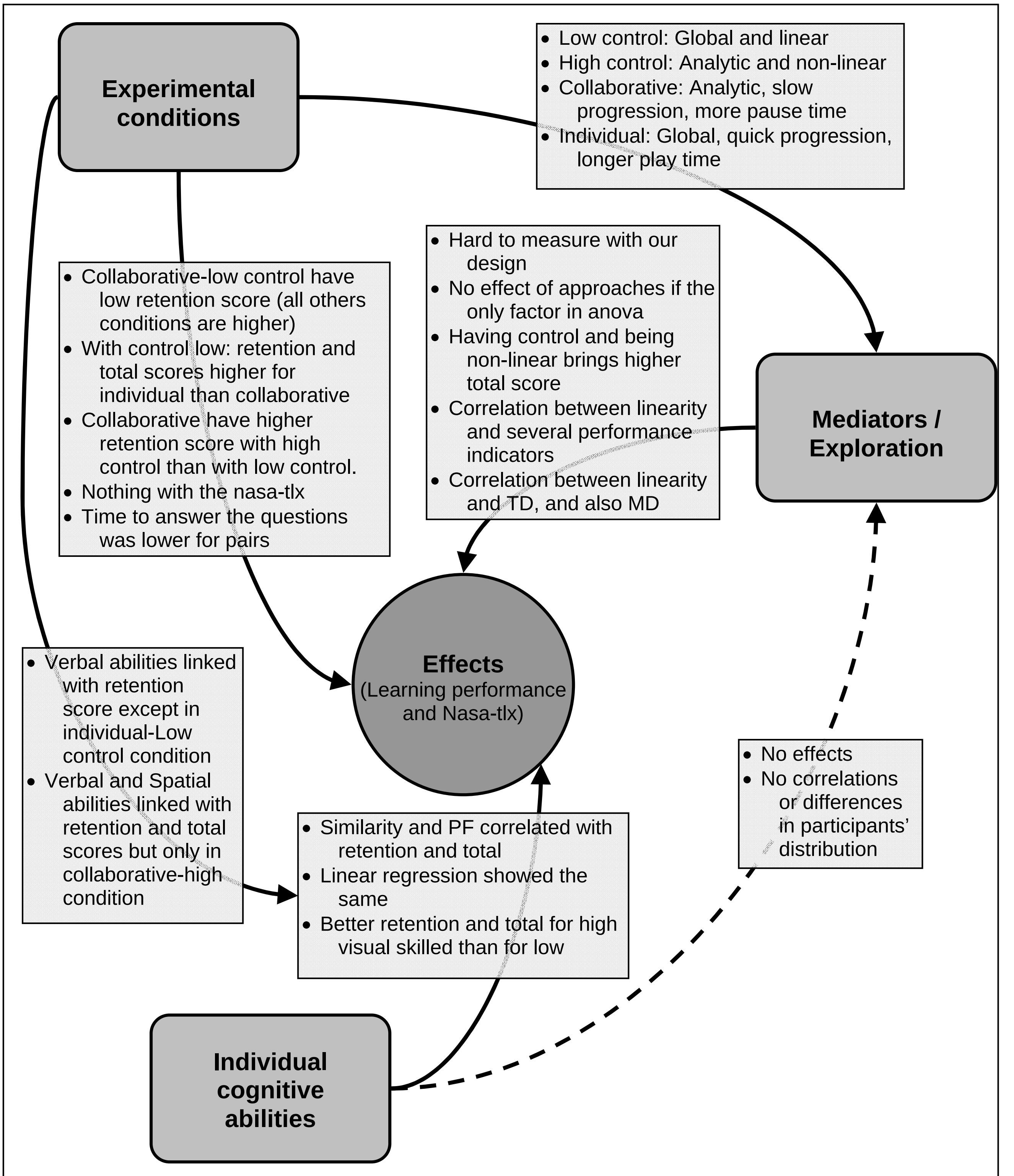

Figure 10: summary of results

4 Discussion

In our discussion, we will separate aspects depending on our experimental conditions, on the mediators and on individual cognitive abilities. After that, a broader discussion will take place, linking these aspects together.

4.1 Collaboration and control

The two factors had no direct influence on any of the learning performance scores, nor on the subjective cognitive load. These results do not confirm our hypothesis and question the results obtained previously (Rebetez et al., submitted). In our previous experiment, participants using dynamic presentations had better inference scores when learning in collaborative than in individual setting. In this study, such results could not be replicated.

4.1.1 No effect of the learning setting

Our hypothesis concerning the learning setting factor was not verified. We expected an improvement of the learning performance due to the construction and maintain of a shared representation in collaborative learning setting. However, no effect was observed on the learning setting factor. This is also clearly opposed to our previous results with the same material (Rebetez et al., submitted). In our previous experiment, participants in a condition corresponding to low control obtained higher inference results when learning in pairs than individually. The biggest difference we have between the two experiments to explain this result is the duration of the learning phase and the exposition to the animation. Previously, participants could only study the animation once but could take the time they wanted to elaborate and discuss between sequences. Here, participants had a limited and controlled time to study the material (superior to what participants used in the first experiment). Besides, they could study the animation several times. These design differences were not controlled and seemed to have an influence on the other conditions.

The study time is an issue in collaborative learning setting; pairs have several mutualisation and verbalisation tasks to put in place during the session. In the previous experiment, participants in collaborative setting spent more time between sequences and an overall higher study time than individual learners. In this experiment, pairs spent more time in pause between (or in the middle of) sequences than individual learners. This was obviously necessary to manage their interaction and maintain a shared representation. These activities need time, and this time is needed in order to build a good and efficient collaboration. In the present study, participants had enough time (more than what they used in experiment one). Nevertheless, time was limited. This time constraint may have been the change that impaired the building of an efficient collaboration (Dillenbourg, 2002).

We did not replicate Schnotz, Böckheler & Grzondziel (1999) results either. In their second experiment, participants in pairs had lower scores with animated pictures. Here no differences were found. Nevertheless, the conditions were different as the learning material consisted of an interactive presentation and not a controllable animation. Moreover, the experimental setup of these authors did not allow them to directly compare individual and collaborative learners.

4.1.2 No effect of control

We expected no main effect of this condition but a positive simple effect of control for individual learners. A high level of control could benefit to individual learners since they would be able to process the different parts of the dynamic material at their own pace, thus improving their mental model construction. In this study, no differences in learning performance were observed between participants with high or low control over the animation. This result contradicts the study of Schwan & Riempp (2004) using nautical knots and Mayer & Chandler (2001) using an animation about lightings. However, in Mayer & Chandler (2001) participants with control only had pre-determined pauses, their only control possibility was to press play when they wanted to see the next animation. In our study, participants with low control already benefited of pauses, and of the same minimal level of control to choose when the animation was to start again. The better explanation in the light of these other studies is that giving low or full control on the animation is not really different in terms of advantage for the learner. The mere fact of having a discontinuous animation could be enough to compensate the supplementary cognitive process due to the dynamic format. This could explain why we saw no differences between our control conditions. A new experiment comparing several levels of controls (including a total absence of control) should be carried out in order to verify this assumption.

4.1.3 Interaction result and split-interaction hypothesis

An interaction was found between the two factors. All participants seemed to have comparable retention results, except for collaborative learners with low control which remembered fewer elements. No split-interaction was observed in this setting. Furthermore, this result is clearly opposed to our split-interaction hypothesis since pairs in low control should have obtained higher results than pairs in high level control. Additionally, in previous experiment, the split-interaction effect was showed on the inference score and not the retention one. In this experiment, no effect was ever found using the inference score.

As the present results did not support our split-interaction hypothesis we need to question its exact conditions of emergence. After that, we will discuss its existence and other possible explanations for the results of our previous experiment.

4.1.3.1 Why no split-interaction was observed?

Like for the absence of effect between individual and collaborative learners, the interaction result can be due to the limited time in sessions. This difference may have changed the way the collaborative learners behaved to the point that they were not able to build an effective collaborative learning situation. This argument finds support in the fact that no effect was found for the learning setting factor. In this case, split-interaction might have been avoided since the interaction with the peer was not as high as needed and thus maybe inefficient.

A second difference from previous experiment comes from the nature of the interaction with the computer interface. With permanent snapshots participants already had some kind of permanent information always on sight. These permanent reminders could have acted as memory cues, helping participants to remember what they already saw and to elaborate further explanations. With a high level of control, participants really had to go back and actively and intentionally search for the information. We thought the computer interaction would be higher and induce even more split-interaction. This was not the case.

We described two differences between this experiment and the previous one (time and control). Both of these differences can theoretically explain why we did not observe the same interaction effect in this experiment than in the previous one. We can criticise our experimental design since it imprudently modified several elements of the paradigm. This problem does not permit us to state about the existence or not of a split-interaction effect. Moreover, we formulated another hypothesis to explain our previous results, the socio-cognitive underwhelming.

4.1.3.2 Socio-cognitive underwhelming

This alternate hypothesis is based on computer-supported collaborative learning theories (Dillenbourg, 1999; Roschelle, 1992; Roschelle & Teasley, 1995) and gains inspiration from Lowe’s underwhelming concept (Lowe, 2003). In our previous experiment, the permanence of snapshots allowed the learner to offload their working memory. We can imagine that participants used only this cognitive support, without engaging in collaboration with their peer. If the participants did not engage in collaborative learning processes due to interface features, they could not create and maintain a shared representation of the problem at hand. This ended in an underwhelming of socio-cognitive involvement; participants were together but did not take advantage of it, thus only the drawbacks of collaboration load were presents to the learner. This second hypothesis can explain the results of previous experiment.

Nevertheless, in this experiment, a socio-cognitive underwhelming hypothesis would have awaited the same results as the split-interaction hypothesis. The present experiment was not designed to decide which one of these explanations is the more convincing. For now we can only say that whatever the effect present in previous experiment was, it seems to be fragile, only specific learning conditions see it happen.

4.2 Approaches to explore the material

Main effects of our experimental conditions may vary depending on how participants actually used the provided tools and resources. From the different data collected about the activity of learners when using the materials and the diverse approaches to define their explorations we found several effects. The “global” and “analytical” approaches are the most interesting. Although the denomination is far from perfect because it appeals to a kind of individual or planned strategy, which is not the case.

As showed in the results, for every approach we looked at, we saw differences of distribution depending on our experimental conditions. As our factors were provoked and participants randomly assigned to a group it is likely that our experimental conditions induced the exploration patterns we observed.

4.2.1 Exploration depended on the experimental condition

The low control condition induced a global and linear approach. Of course participants could only have a linear approach since in this condition they could not choose the order of the sequences.

In the high control condition, most participants were more likely to adopt an “analytic” and “non-linear” behaviour when exploring the material. Those both tags fit well together since a typical analytic approach would be to watch the same sequence until it is perfectly understood and switch to the next. However, taking time between sequences to recapitulate would also count as an analytical approach while being completely linear (as did some participants).

In the collaborative condition, pairs were more likely to be "analytic", slow to see the entire animation and using their time to take breaks between sequences, certainly to discuss and maintain a shared representation. In the individual condition, participants were more likely to be "global", quick to see the material once, and their time was used to watch sequences.

Looking at both factors together, we see that individual learners (with or without control) were mostly using a "global" approach, were quick to see the whole animation, spending most of their time in watching the sequences. Collaborative learners (with or without control) were more likely to have an "analytical" approach, slow to see the whole animation, talking or thinking about each sequence before going to the next (low playtime). Of course, only participants with control could adopt a non-linear exploration, but not all of them did it.

Participants studied the material differently, depending on their experimental condition. The way they explored, depending on the condition, is not surprising. Nevertheless, our conditions did have an effect on the way participants learned the material.

4.2.2 Does the way the material was explored change the understanding?

As our experimental condition did have very few effects on the learning performance, one can ask if participants learned differently depending on the way they explored the material.

This question is difficult to answer solely on the basis of the results of this experiment. Because exploration patterns were identified post-hoc and categorisations found not to be evenly distributed in our conditions (and less than five subjects in some groups). This did not allow us to run statistical analyses using all our factors and exploration approaches. However, we ran anovas using only one exploration pattern approach at a time, without our experimental factors. No effects were found with any approach on the learning performances or subjective cognitive load. Nevertheless, using only one factor limited the potentially explained variance between our groups and might explain why no significant effect was found. But as the classification of participants depending on our approaches was relatively close to our experimental conditions, we can doubt that a specifically designed experiment would find different results. We showed that if an effect of exploration (at least as we defined it) exists and we did not see it, it certainly is of low strength.

Nevertheless, some effects of the mediators were found. A non-linear approach rather than a linear one when disposing of high control is related to higher total scores. As well, the linearity coefficient is related with several performance indicators and also subjective cognitive load (mental demand and temporal demand): participants with less linear approach (for those it was possible), slightly better learned from the material. They also felt the material more demanding. We can explain this through the involvement of learners in the material and in the learning activity. Recent results showed positive effects of invested effort (defined as a motivated involvement) on retention (Gendolla & Richter, 2005). Positive effect of “activity” on multimedia messages were also reported (Mayer, 2001; Palmiter & Elkerton, 1993). Participants with high control who decided to really be active on the material and to use the controls were the more active, and maybe the more motivated. This resulted in higher perceived cognitive load as the involvement in the task was higher and as the interface controls were more used (correlation between linearity and temporal demand, and also with mental demand). Moreover, some learning performance results were higher.

The high control condition created a suitable ground for non-linear and “top-down” exploration of the material. As we saw in previous part, the experimental conditions have something to do with the effective exploration of the materials. However, the manner the learner uses the condition is important to its comprehension. Giving control has no effects per se; the efficient use of controls (here pointed out by non-linearity) can have effects on learning. But, with low control, no chance is given to be non-linear. In collaborative condition, low control even lowered retention scores as compared to high control condition. The absence of control seemed to prevent collaborative learners to build and maintain an efficient shared representation. As past informations could not be displayed easily to create grounding, collaboration mechanisms were harder to deploy. This would explain the interaction result on our main anova.

4.3 Individual cognitive skills

Like for mediators and exploration patterns, the individual cognitive skills were assessed during the experiment and participants were attributed a score and a post-hoc group. This leaded to difficulties in assessing the precise effect of these variables in conjunction with our experimental conditions. Indeed, when using four factors, some groups were too small to run an analysis of variance. However analysis could be made on three aspects. Firstly, we investigated direct effects of individual cognitive abilities on learning performance an perceived cognitive load. Secondly we looked at these same effects depending on the experimental condition. Third, we asked if individual abilities affected mediators or exploration decisions.

4.3.1 Cognitive abilities and learning

Using correlations and linear regressions, we saw a positive relationship between both similarity and paper-folding scores on retention and total scores. Using median split to create groups of high and low able learners, an anova confirmed better retention and total score for higher spatial learners than lower spatial learners. Thus as already observed by Hegarty and collaborators (Hegarty & Sims, 1994; Hegarty & Steinhoff, 1997), and in accordance with our hypothesis, paper folding scores were positively related to performances. As we used answering time to weight paper-folding answers this result shows an effect of visual abilities, not related to intelligence or higher problem resolution abilities.

Our hypothesis about verbal abilities was also confirmed. Efficient mental model creation needs both verbal and pictorial information (Pavio, 1986; Schnotz & Bannert, 2003). Thus, efficient verbal integration and processing abilities helps the comprehension of the material and the creation of the mental model (assuming the similarity test assesses a cognitive ability close to the one needed for the task).

These results reinforce the text and picture integration model (Schnotz & Bannert, 2003). We used materials involving both graphical and verbal information. According to the model, these contents needed to be processed, organised and integrated in their own modality to induce efficient comprehension. In this point of view, individual visual and verbal processing abilities are inevitably important to integrate text and pictures. As our results showed it, both are linked with better retention. If we can regret the absence of effect on the inference score (which showed very few effects in this experiment), retention and total scores speak in favour of this model.

4.3.2 Did experimental conditions interact with cognitive abilities?

As we expected, Verbal and spatial abilities were positively linked with retention and total scores. However, this was only the case in collaborative-high condition. It seems that top-down effects were likely to occur only in the richest condition. We can only speculate on an explanation but a convincing one might be in terms of condition's potential of benefit. As we saw it in the results, these conditions did not improve comprehension in the present experiment. It is likely that participants could not avoid the inhibiting effects of load due to collaboration increased interaction with the screen. Only high abilities participants could manage the supplementary processing.

Another result showed that verbal abilities were positively linked with retention score, except for individual learners with low control. In this condition, the setting is the less rich of our experiment, the potentials for deepening comprehension were lower but so was the potential extraneous cognitive load involved. So the need for quickly and efficiently process verbal information was less critical in this condition. This can explain why learners with high or low verbal abilities performed similarly.

These results are very interesting because they make us doubt of overall importance of individual abilities when learning from multimedia. We saw here that their effects are only cloistered in specific conditions, strongly depending on learning conditions to emerge. They have an effect only when higher demands are imposed to the learner.

4.3.3 Did individual abilities determine the way people explored the material?

In order to understand if individual variables like visual and verbal processing abilities would have an influence on the way participants explored the materials, we used khi2 and correlation analysis. The results showed both aspects were unrelated. The way participants looked at animations did not depend on their spatial or verbal processing abilities.

The conclusion is that the exploration patterns were not likely to be "top-down", they were definitely not strategies nor competence-guided exploration. As participants were all novices in the domain, knowledge-based "top-down" processes were unlikely to happen. However individual abilities are not on the same level and could have influenced learners on the way they would apprehend the materials, but this is not the case here. Of course we can discuss the pertinence of our two cognitive ability measures as predictors of multimedia exploration behaviour. But, as they had some effects on learning performance and nasa-tlx results they might as well be involved in the material exploration.

A new experiment could be carried out in order to further define the importance of verbal and visual individual abilities when learning from multimedia materials. A pre-test should be used in order to select participants before the experiment depending on their abilities. Four groups with a similar number of participants could be formed. This paradigm would especially be interesting in conjunction with collaborative learning setting where the composition of groups could be based on participant's abilities (complementary or not).

5 Conclusion

The main results for this experiment were: 1) the interaction between our two factors; pairs with low control had lower learning performances than the others. 2) The level of control had no effect on individual learners. 3) The control allowed higher understanding when it was used but it also added to subjective cognitive load. 4) The exploration behaviours did not seem to depend on the individual cognitive abilities.

The split-interaction, as well as the socio-cognitive underwhelming hypotheses, were not supported by the empirical results. If they can not be totally rejected due to some arguable effects described previously, they suffer severe flaw. Even if these effects exist they are not very strong nor easily observable in other contexts. Moreover, the result of the interaction between our two factors contradicts our hypothesis about these effects.

Our factors had few direct effects on learning performance but strongly influenced the way participants explored the materials. This influence of experimental conditions is even stronger as we saw none from individual cognitive abilities. This strongly speaks in favour of bottom-up and design-guided exploration of the learning material for novice learners. Even participants abilities, which were linked to learning performance and are not knowledge based, had no influence on exploration behaviour. This supports theories of strong influence of delivery features on the participant's processing of information (Lowe, 2003; Schnotz & Bannert, 2003).

Nevertheless, identified mediators had very few effects on learning performances. As a matter of fact, no strong effects were found since the only result concerns the linearity among participants in high control conditions. To sum up, our conditions had few effects on learning performance but induced different uses of the experimental material. But these different uses had few to no effects on learning performances. So, our conditions well and truly changed the way participants worked on the material, but this had no results, why? Were the participants always "overwhelmed", or "underwhelmed", in all conditions so that despite different behaviours very few understanding occurred at all? Another explanation would be that our material was designed for a straightforward, linear approach (like in our previous experiment). Thus, information was coming to the learner in an instructionally designed progression. There was no real need to have a nonlinear or an analytical approach to better understand the material (like it could have been to learn nautical knots maybe). The efficiency of the control condition could depend on the conceptual mapping involved in the material (Bétrancourt, Bauer-Morrison, & Tversky, 2001). Further experiments involving control should take the instructional design and conceptual mapping into account. Moreover, three levels of control should preferably be used, in addition to high and low control, a condition without control might reveal useful.

Concerning the learning setting factor, this experiment certainly flawed the quality of the collaborative condition by controlling study time. This was a very difficult design issue since with high level of control; the time of study had to be constant in order to compare participants. However, further experiments involving both control and learning setting should maintain constant only the time spent in playing the animation and not the whole study time. Concerning this, we know no experimental study which compared time constraints when collaboratively learning from multimedia materials. Using a common material and manipulating the real or perceived time constraints to study it could help us explain the present results, but also to describe the best conditions for collaborative learning.

Since no split-interaction was observed, this study can not really answer the question it asked and that conditions of its potential appearance might also have been missed. Other studies involving computer supported collaborative setting should be aware of the concept of split-interaction and of socio-cognitive underwhelming. More data are actually needed to define existence and conditions of expressions of these two potential effects. A new experiment should be carried out in order to distinguish split-interaction and socio-cognitive underwhelming. The activity proposed to participants could be manipulated in order to force collaboration or not. For example the learning activity could be similar to the present experiment (learn together), or force collaboration (write a text or draw a concept map together). If the interaction effect is present in both conditions, the split-interaction effect would better explain the results. If the interaction effect is not present only when collaboration is forced, the socio-cognitive underwhelming would describe the results more accurately.

The present study showed, through numerous analyses, that learning with multimedia material is a complex phenomenon. It clearly underlined the design and bottom-up based approach of novice learners. It also showed that if participants explored the material differently it will not likely induce differences in their understanding. Delivery features are important but instructional design might be even more. Factors like control and learning setting only lightly influenced the understanding of our material. Factors like conceptual mapping or level of abstraction might reveal much more powerful to improve multimedia learning.

6 References

Ainsworth, S., & Van Labeke, N. (2004). Multiple forms of dynamic representation. *Learning and Instruction, 14*(3), 241-256.

Atkinson, R. C., & Shiffrin, R. M. (1968). Human memory: A proposed system and its control processes. In K. W. Spence & S. J. T. (Eds.), *The psychology of learning and motivation: advances in research and theory* (Vol. 2, pp. 89-195). New-York: Academy Press.

Baddeley, A. D. (1986). *Working Memory*. Oxford: Clarendon Press.

Baddeley, A. D., & Hitch, G. J. (1974). Working Memory. In G. A. Bower (Ed.), *Recent advances in learning and motivation* (Vol. 8, pp. 47-90). New York: Academic Press.

Bétrancourt, M., Bauer-Morrison, J., & Tversky, B. (2001). Les animations sont-elles vraiment plus efficaces? *Revue d'intelligence artificielle, 14*(1-2), 149-166.

Bétrancourt, M., & Realini, N. (2005). *Le contrôle sur le déroulement de l'animation.* Paper presented at the Journée d'Etude sur le Traitement Cognitif des Systèmes d'Information Complexes, JETCSIC'2005, Nice.

Bétrancourt, M., & Tversky, B. (2000). Effect of computer animation on user's performance: a review. *Le travail humain, 63*(4), 311-330.

Catrambone, R., & Fleming Seay, A. (2002). Using animations to help students learn computer algorithms. *Human Factors, 44*, 459-511.

Clark, H. H., & Brennan, S. A. (1991). Grounding in communication. In L. B. Resnick, J. M. Levine & S. Teasley (Eds.), *Perspectives on socially shard cognition*. Washington: APA Books.

Cohen, J. (1988). *Statistical power analysis for the behavioural sciences* (2nd ed.). Hillsdale, NJ: Lawrence Earlbaum Associates.

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

Dillenbourg, P. (2002). Over-scripting CSCL: The risks of blending collaborative learning with instructional design. In P. A. Kirschner (Ed.), *Three worlds of CSCL. Can we support CSCL* (pp. 61-91). Heerlen: Open Universiteit Nederland.

Dillenbourg, P., & Bétrancourt, M. (2006). Collaboration Load. In J. Elen & R. E. Clark (Eds.), *Handling complexity in learning environments: research and theory*. Amsterdam: Elsevier.

Ekstorm, R. B., French, J. W., Harman, H. H., & Dermen, D. (1976). *Manual for kit factor-referenced cognitive tests* (Educational testing services ed.). Princeton, New Jersey.

Gendolla, G. H. E., & Richter, M. (2005). Ego involvement and effort: Cardiovascular, electrodermal, and performance effects. *Psychophysiology, 42*, 595-603.

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

Hegarty, M., Kriz, S., & Cate, C. (2003). The roles of mental animations and external animations in understandig mechanical systems. *Cognition and Instruction, 21*(4), 325-360.

Hegarty, M., & Sims, V. K. (1994). Individual differences in mental animation during mechanical reasoning. *Memory and cognition, 22*(4), 411-430.
Hegarty, M., & Steinhoff, K. (1997). Individual differences in use of diagrams as external memory in mechanical reasoning. *Learning and Individual Differences, 9*(1), 19-42.
Hutchins, E. (1995). *Cognition in the wild*. Cambridge: MIT Press.
Jermann, P. (2004). *Computer support for interaction regulation in collaborative problem-solving.* University of Geneva, Geneva.
Johnson-Laird, P. N. (1983). *Mental models: Towards a cognitive science of language, inference, and consciousness*. Cambridge: Harvard University Press.
Kaiser, M., Proffitt, D., Whelan, S., & Hecht, H. (1992). Influence of animation on dynamical judgements. *Journal of Experimental Psychology: Human Perception and Performance, 18*, 669-190.
Kalyuga, S., Chandler, P., & Sweller, J. (1999). Managing split-atention and redundancy in multimedia instruction. *Applied Cognitive Psychology, 13*, 351-371.
Leclercq, D., & Gilles, J.-L. (1994). *GUESS, un logiciel pour entraîner à l'auto-estimation de sa compétence cognitive.* Paper presented at the colloque QCM et questionnaires fermés, Paris: EISEE.
Levenshtein, V. I. (1966). Binary codes capable of correcting deletions, insertions and reversals. *Soviet Physics-Doklady, 10*, 707-710.
Lowe, R. K. (1999). Extracting information from an animation during complex visual learning. *European journal of psychology of education, 14*, 225-244.
Lowe, R. K. (2003). Animation and learning: selective processing of information in dynamic graphics. *Learning and Instruction, 13*, 157-176.
Lowe, R. K. (2004). Interrogation of a dynamic visualisation during learning. *Learning and Instruction, 14*, 257-274.
Mandl, H., & Levin, R. (1989). *Knowledge acquisition from text and pictures*. Amsterdam: Noth holland.
Mayer, E. R. (1989). Models for understanding. *Review of Educational Research, 59*(1), 43-64.
Mayer, E. R. (1999). *The promise of educational psychology* (Vol. 58). Amsterdam: Noth Holland.
Mayer, E. R. (2001). *Multimedia learning*. Cambridge: Cambridge University Press.
Mayer, E. R. (2003). The promise of multimedia learning: using the same instructional design methods across different media. *Learning and Instruction, 13*, 125-139.
Mayer, E. R., & Chandler, P. (2001). When Learning is just a click away: does simple interaction foster deeper understanding of multimedia messages? *Journal of Educational Psychology, 93*(2), 390-397.
Mayer, E. R., Heiser, J., & Lonn, S. (2001). Cognitive constraints on multimedia learning: when presenting more material results in less understanding. *Journal of Educational Psychology, 93*(1), 187-198.
Moore, J. L., & Rocklin, T. R. (1998). The distribution of distributed cognition: multiple interpretations and uses. *Educational Psychology Review, 10*(1), 97-113.
Narayanan, N. H., & Hegarty, M. (2002). Multimedia design for comnuication of dynamic information. *International journal of human-computer studies, 57*, 279-315.
Paas, F., Renkl, A., & Sweller, J. (Eds.). (2004). *Advances in cognitive load theory: Methodology and instructional design [Special issue]* (Vol. 32).
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, 13-18 April). *Assessing dynamics in computer-based instructions.* Paper presented at the CHI.

Pavio, A. (1986). *Mental representations: a dual coding approach*. New York: Oxford University Press.

Perret-Clermont, A. N. (1993). What is it that develops? *Cognition and Instruction, 11*, 197-205.

Rebetez, C. (2004). *Sous quelles conditions l'animation améliore-t-elle l'apprentissage?* Unpublished Master thesis, Unversity of Geneva, Geneva.

Rebetez, C., Sangin, M., Betrancourt, M., & Dillenbourg, P. (submitted). Learning from animation enabled by collaboration. In P. Gerjets & K. Scheiter (Eds.), *Cognitive and socio-motivational aspects in learning from dynamic visualizations, Special Issue of Instructional Science.*

Rieber, L. P. (1991). Animation, incidental learning and continuing motivation. *Journal of Educational Psychology, 83*, 318-328.

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

Roschelle, J., & Teasley, S. (1995). The construction of shared knowledge in collaborative problem solving. In C. E. O'Malley (Ed.), *Computer supported collaborative learning* (pp. 69-97). Heidelberg: Springer-Verlag.

Schnotz, W., & Bannert, M. (2003). Construction and interference in learning from multiple representation. *Learning and Instruction, 13*, 141-156.

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*, 117-123.

Schnotz, W., & Rasch, T. (2005). Enabling, facilitating and inhibiting effects of animations in multimedia learning: Why reduction of cognitive load can have negative results on learning. *Educational Technology Research and Development, 53*(3), 47-58.

Schwan, S., & Riempp, R. (2004). The cognitive benefits of interactive videos: learning to tie nautical knots. *Learning and Instruction, 14*(3), 293-305.

Stahl. (2002). *Contributions to a theoretical framework to CSCL.* Paper presented at the CSCL, Boulder, Colorado.

Sweller, J. (2003). Evolution of human cognitive architecture. In B. H. Ross (Ed.), *The psychology of learning and motivation* (Vol. 43, pp. 215-266). New-York: Academic Press.

Sweller, J., Chandler, P., Tierney, J., & Cooper, M. (1990). Cognitive load as a factor in the structuring of technical material. *Journal of Experimental Psychology: General, 119*, 176-192.

Sweller, J., van Merrienboer, J. J. G., & Paas, F. G. W. C. (1998). Cognitive architecture and instructional design. *Educational Psychology Review, 10*(3), 251-296.

Thompson, S. V., & Riding, R., J. (1990). The effect of animated diagrams on the understanding of a mathematical demonstration in 11 to 14 years-old pupils. *British Journal of Educational Psychology, 60*, 93-98.

Tversky, B., Morrison, J., B, & Betrancourt, M. (2002). Animation: can it facillitate? *International journal of human-computer studies, 57*, 247-262.

Wechsler, D. (1981). *Wechsler adult intelligence scale - revised (WAIS-R)*. Cleveland: The psychological corporation

7 Acknowledgments

This experiment was funded through a FNS grant (#11-68102-02).

Mireille to correct the first version during her vacation time with her friends (I hope it will not be too sunny), and of course for multiple help, support, tips and references. Annelore, who agreed to have a busy week during the redaction phases of this document. Mirweis, who first formulated the socio-cognitive underwhelming hypothesis. Stephane and Stephane for endless bomberman cool-down sessions (I still see bombs when I close my eyes). Fatec, the lab's aibo, for keeping me company during the night. The jury, to read and analyse this work during sunny weather. And myself for not reading the diploma's regulation and forgetting the necessity of an oral exam and board review (that's why I did not present this work in march 2006).

N.B. some parts of the introduction are rewordings from of other works I wrote:

Rebetez, C., Sangin, M., Bétrancourt, M., & Dillenbourg, P. (submitted). Learning from animations is improved by collaboration. Submitted to *Instructional Sciences* in december 2005.

Rebetez, C., Bétrancourt, M., Sangin, M., Dillenbourg, P., Mollinari, G. (2006). Handling complexity in multimedia learning: effects of control and collaboration when studying animation, EARLI SIG Conference: Instructional design & Learning and instruction with computers, 21-23 june 2006, Leuven (Belgium).

8 Annexes

Most of the annexes are available from the cd-rom included in the paper version of this master thesis. If you read an online version, please send an email to cyrilrebetez at gmail dot com.

8.1 Animated pictures used

Disponible on the cd-rom

8.2 Complete experimental material

Disponible on the cd-rom

8.3 Certainty scale

From (Leclercq & Gilles, 1994), also found at http://www.smart.ulg.ac.be/smart/ee/guess/.

Voici les consignes et le barème des tarifs en cas de réponse correcte et incorrecte :

		Vous obtiendrez en cas de réponse	
Si vous estimez que votre réponse a une probabilité d'être correcte comprise entre	Ecrivez	**Correcte :**	**Incorrecte :**
0 % et 25 %	**0**	**+ 13**	**+ 4**
25 % et 50%	**1**	**+ 16**	**+ 3**
50 % et 70 %	**2**	**+ 17**	**+ 2**
70 % et 85 %	**3**	**+ 18**	**0**
85 % et 95 %	**4**	**+ 19**	**- 6**
95 % et 100 %	**5**	**+ 20**	**- 20**

Ce tarif est calculé de telle façon que ceux qui s'auto-estiment correctement (sans trop de sur ou sous-estimations) récoltent le plus de points.

8.4 Raw data of the experiment

Disponible on the cd-rom

8.5 Statistical analysis

Disponible on the cd-rom

8.6 Levenshtein distance script

Disponible on the cd-rom