Chapter for the Book 'Dealing with Complexity in Learning Envionments'

Collaboration Load

Pierre Dillenbourg [1] & Mireille Betrancourt[2]

[1] CRAFT, Ecole Polytechnique Fédérale de Lausanne , Switzerland

[2] TECFA, Université de Genève, Switzerland

Abstract

Does collaboration increase or decrease cognitive load during learning? On one hand, collaboration enables some degree of division of labour that may reduce cognitive load. On the other hand since interacting, expressing thoughts, monitoring another's understanding, grounding, etc., are mechanisms inducing some extraneous cognitive load, they may create cognitive overload and impede learning mechanisms. However this additional load may explain why collaboration sometimes leads to knowledge construction. This trade-off between productive versus counter-productive load is not specific to collaborative learning. It is also present in individual learning, namely questioning guided-discovery learning methods. This contribution explores the concept of cognitive load in collaborative situations. We raise more question than provide answers. What constitutes collaboration load, i.e. which mechanisms triggered during collaborative learning more often than during individual learning, contribute to increase cognitive load? In collaborative learning software, which interface features and tool functionalities increase or decrease the different costs factors (verbalization, grounding, modelling…)? We explore these questions and illustrate our arguments with three studies on computer-supported collaborative problem solving. We also consider how the collaboration

load may be tuned through the design of computer-supported collaborative learning environments.

1.   Introduction

This chapter addresses the notion of 'complexity' in collaborative learning, especially computer-supported collaborative learning (CSCL). Collaborative learning may simply be the joint use of learning material such as textbooks or drill-and-practice software. However, due to its socio-constructivist roots, CSCL research often addresses group learning in complex environments. Most empirical studies actually investigate collaborative problem solving. CSCL environments often include a problem space and a social interaction space (chat, forum, argumentation tools…). The problem space can be a computerized environment (a microworld, a simulation…) or not (paper readings, field trips, physical experiment,…). Hence, learning in a CSCL environment cumulates the complexity of computer-based constructivist environments with the complexity of computer-mediated communication. While these two spaces have often been computationally separated (two different windows on the screen), recent CSCL environments integrate them computationally, for instance by relating utterances to task objects (Zahn et al; 2004)

The interaction with any computerized learning environment imposes an additional cognitive load, especially at the beginning of its use, as developed in chapter ??? (this volume). In CSCL, this 'computer interaction' additional load is increased with the 'social interaction' load or collaborative load (i.e. the need to manage interactions with the other group members). On one hand, this increased complexity may interfere negatively with learning processes, consuming some of the individuals' cognitive resources. On the other hand, it may be beneficial to learning. Complexity increases in group learning because additional cognitive mechanisms are triggered by collaborative interactions. Those mechanisms (e.g., explanation, argumentation) may generate the learning effects that are expecting from collaborative situations. In other words, complexity has advantages and drawbacks. The question of

collaboration load is a particular instance of a trade-off that exists in any learning situation; there is no learning without some cognitive load, but there is no learning with too much cognitive load either. In light of the latter, this contribution begins with a short review of this trade-off as it has been investigated in learning in general (Section 2).

This question of collaboration load and its effect on learning has been largely unexplored. We do not report on results of empirical studies that were specifically targeted to measure collaboration load, instead this paper addresses three main questions.

1. What constitutes collaboration load, i.e. which mechanisms triggered during collaborative learning more often than during individual learning contribute to increase cognitive load? Is it the need to verbalize one's own thoughts? Is it the effort to understand one's team mates and, more globally, to construct a shared understanding of the task at hand? Is it the need to maintain some kind of representation of a partner's goals, knowledge and actions? While collaboration enables division of labour does it decrease cognitive load? In section 3, we try to disentangle some of the factors that come into play when estimating cognitive load.

2. Do CSCL environments influence collaboration load? Different media have an impact on the cost factors. Which interface features and tool functionalities increase or decrease the different costs factors (verbalization, grounding, modelling…)? Reviewing all the features of human-computer interaction that have an effect on collaboration load would be beyond the scope of this paper. Section 4 illustrates some tool features (persistency of information, mutual awareness) that were revealed by empirical studies conducted with CSCL environments. These studies were not designed for measuring collaboration load, but nonetheless provided insights on this topic.

3. What are the implications for CSCL designers? Section 4 stresses that the collaboration load is not an intrinsic feature of CSCL environments but depends on specific features of the CSCL environments. Section 5 describes some properties that designers may use to "tune" the collaboration load induced by a specific environment.

We are unable to provide definite answers to any of the questions above. This contribution does however explore these questions, disentangle factors and raise sub-questions that could initiate further research in this avenue.

2. Cognitive load and learning

Students use the same brain when they learn alone or in groups. Some people would even claim that one never learns alone. For these reasons, our analysis of cognitive load in collaborative learning is first situated with more general debate on the cognitive load involved in learning.

Since long ago, there has been a discrepancy between the psychological and the educational perspective on the cognitive load factors in a learning task. Since the constructivist approach in the sixties, educational scientists consider that conceptual learning occurs only if the learning task requires learners to engage in intensive cognitive processes. Conceptual learning, also considered as 'deep learning' is characterized by the transformation of learners cognitive structures in a way that the acquired knowledge, procedures or schemata could be used in other situations or domains (De Corte, 2003). On the contrary, surface learning, or rote memorization, enables the learner to apply the learned schema only to similar situation. Whereas practices to improve surface learning are quite well known and used in educational situations, methods that improve conceptual learning are still under experimental investigation. Learning at a conceptual level means tremendous changes in the learners'

cognitive schemes or conceptions. Learning tasks and practices that engage learners in rich and complex interactions with the learning environment, such as inquiry learning or discovery learning, have been shown to be situations in which deep learning can occur (Schnotz, Vosniadou, & Carretero, 1999). Collaborative learning belongs to the pedagogical practices that generate a rather heavy cognitive load.

There are however some psychologists who claim that learning can only occur if the cognitive resources required to process a learning task are maintained below a 'reasonable' level. This claim is based on the current view on the cognitive system as described in Baddeley's model (Baddeley, 1997, 2000). In this view, the cognitive system consists of two processing components: A long-term memory (LTM) in which knowledge is stored permanently, and a working memory that processes the information sent by the perceptual system on the basis of the knowledge stored in LTM. Many experimental results comforted the assumption that working memory is limited in capacity (Baddeley, 1997). The consequences of cognitive processing limitations on learning have been formalized and extensively investigated by the tenants of the Cognitive Load Theory (Paas, Renkl, & Sweller, 2004; Sweller, 1988). According to the Cognitive Load Theory, deep learning is described as the acquisition of cognitive schemata that enable categorizing the problem, choosing the correct procedures to apply and regulating problem solving. The construction of such schemata is cognitively demanding. Consequently, the processing of the learning task itself competes with the construction of cognitive schemata if the learning task is too demanding. A large body of research, has investigated the effect of the format of instruction on learning. For example, it was demonstrated that a multimedia instructional material in which mutually referring verbal and graphic information are displayed separately on the page is detrimental to learning compared to material in which graphic and verbal information are spatially integrated (Sweller et al., 1990). According to the authors, the separated display forced learners to

repeatedly shift their attention from one source to the other and thus increased the cognitive resources that should be dedicated to mentally integrating the two sources of information. The cognitive overload induced by the 'split-attention effect' would explain the learning impairment. Principles for designing effective design material have been derived from the research and are still the object of thorough investigation (Mayer et Moreno, 2002; Sweller, 2003).

Recent results showed that some factors such as expertise can tremendously increase processing capacities of the working memory (Ericsson & Kintsch, 1995). When dealing with new elements that have not been previously learned, there is no schema in long term memory that can indicate how the elements should be processed and all the burden falls on the working memory. Conversely, for well-learned material and activity, schemas stored in long-term memory take in charge the coordination and the combination of elements, allowing huge amounts of information to be processed in working memory (Sweller, 2003). As a consequence, some instructional guidelines that have been proved effective for novice learners are not applicable for more advanced learners. For example, Kalyuga, Chandler and Sweller (2000) showed that the modality effect, according to which it is better to present verbal information in auditory mode when the material also involves graphical information, could be reverted by expertise. While novice learners benefited more from the audio-visual material compared with the visual-only material, this advantage disappeared after a few training sessions. As expertise increased, a visual-only presentation was superior to an audio-visual presentation, particularly when the text information was removed (graphic presentation only). Guidance provided by text information was necessary for novices but was redundant for experts who had a schema available to process the material.

2.1. Intrinsic and extrinsic load

Current developments of Cognitive Load Theory consider two sources of load when learners have to process instructional material in order to achieve a learning task (Paas, Renkl, & Sweller, 2004; Sweller, 2003):

- Intrinsic load refers to the load required to process the instructional task. It is related to the complexity of the content itself and particularly to the degree of interactivity between elements, which impacts the number of elements that must be held in working memory simultaneously;

- Extrinsic load refers to two sub-categories of load:

    - Germane load promotes the construction of the cognitive schema, which is the ultimate goal of deep learning;

    - Extraneous load refers to the additional load that is influenced by the format of instruction (material presentation or structure of the learning task) and that does not contribute to learning.

Deep learning can occur only if cognitive resources are sufficient to cover the processing requirements. In other words, cognitive overload may explain why some learning situations fail to induce deep learning. Extraneous load should thus be reduced to its minimal by adequate presentation format and learning task.

These educational and psychological views on cognitive load in learning seem contradictory; educational scientists aim to design cognitively demanding learning tasks, psychologists care to minimize the cognitive resources engaged in the learning task. The 'goal-free effect' (Sweller, van Merriënboer, & Paas, 1998) is an excellent example of this tension; deep learning is improved when learners are not provided with the final goal but only with intermediate goals. While this seems contradictory to the self-monitored approach that claims

that explicitly stating learning objectives increases learning outcomes (Tourneur, 1975), the two perspectives may not be as contradictory as they appear at first sight. Firstly, the notion of germane load, recently taken into consideration by the cognitive load model, acknowledges that cognitive load can be beneficial to learning, provided that this load is allocated to the construction of cognitive schemata rather than to the processing of extraneous information. Furthermore, the cognitive load concept in psychology has more than one component: In subjective scales (i.e., NASA-TLX; Hart & Staveland, 1988), cognitive load refers to the cognitive effort but also to frustration and stress. Effort is not always painful or unpleasant. Effort can be an excellent motivator to proceed, as the sensation of 'flow' in game situations and can be turned to a learning motor (Rieber, 1996).

2.2. *Measurement of cognitive load.*

As useful as it can be to provide instructional guidelines, the cognitive load model raises difficulties regarding assessment and measures. In most studies, cognitive load was assessed by self-reporting indicators and as a relative measure to distinguish between conditions. Gerjets, Sheiter and Catrambone (2004) used a self-reporting scale adapted from the NASA-TLX to investigate the processing of multimedia instructional materials. They found no clues in the cognitive load estimation to understand the results. More complex indicators, based on both subjective evaluation and performance (scores, time) were identified to constitute a reliable estimate of the mental efficiency of instructional methods. (Paas, Tuovinen, Tabbers & van Gerven, 2003). Physiological measures (i.e., heart beats, electro-dermal reactions, pupil dilatation, blinking, neuroimaging techniques,) can be regarded as more direct evidence for cognitive load, but they are difficult, if not impossible, to apply in ecological learning situations. Less indirect than learning outcomes or subjective evaluation, but easier to handle in a deep learning situation, the dual task paradigm has scarcely been used in instructional studies (Brünken, Steinbacher, Plass & Leutner, 2002), The dual task

methodology consists in measuring how reaction times to a secondary task vary over different treatment conditions in the primary task. The dual-task paradigm can effectively distinguish between load in different sensory modalities (auditory vs. spatial, e.g. Brünken, Plaas & Leutner, 2004) or processing modes (verbal vs. visual, e.g. Gyselink, Ehrlich, Cornoldi, de Beni and Dubois, 2000). However, the methodological challenge remains to design 'pure'secondary tasks: spatial tasks often involving visual processing (Gyselink et al., 2000) and conversely. Verbal material may involve auditory load by way of the auditory loop even when presented in written mode. As a consequence, the effect of instructional formats on secondary tasks may be confusing and does not necessarily assess well established findings in instructional design (Brünken et al 2004). Finally, one should keep in mind that the maximal cognitive load is a versatile value, varying over people, time and context. Working memory processing capacity depends on the level of expertise (Ericsson & Kintsch, 1995), individual abilities (Gyselink et al, 2000), metacognitive processes (Valcke, 2002) and level of involvment in the task. It is therefore difficult to discriminate between a cognitive load level that is manageable and beneficial to learning and the overload level that is detrimental to learning.

In addition to methodological toughness, a more fundamental issue is being able to measure the desired variable. In subjective evaluation, do the students express workload or cognitive load? Workload is the student's perception of his or her amount of work. It is quite different from the cognitive load that refers to the limited capacity of working memory. Actually, these two concepts are easily differentiated at the theoretical level, but more difficult to dissociate at the empirical level. As discussed earlier, measures of cognitive load in instructional design studies often involve self-evaluation scales or questionnaires. In this case, we wonder if students are able to distinctively perceive their workload and their cognitive load, as will be explored later on (section 4.1). Besides, no method to date permits

identifying the 'nature' of the cognitive load that is being measured. How can we tell extraneous cognitive load, detrimental for the construction of knowledge, and germane cognitive load, an indicator that deep learning is occurring, apart?

2.3. Minimal versus optimal collaborative load?

The notion of collaborative effort has been addressed in the study of dialogue. Clark and Wilkes-Gibbs (1986) analyzed the effort necessary for two people to understand each other. They stressed that what is important is not individual effort by the receiver of a communicative act, but the overall 'least collaborative effort'. Thereby, they mean that the cost of producing a perfect utterance may be higher than the cost of repairing the problems that arise. For instance, subjects are less careful about adapting utterances to their partner when they know they can provide feedback on their understanding (Schober, 1993).

The notion of least effort fits with the economy of discussion in everyday life situations, when partners naturally minimize efforts to reach mutual understanding. This also holds for two students who have to work together. However, studies of collaborative learning reveal that collaboration leads to learning if peers are engaged in intensive interactions such as argumentation or explanation. What produces learning is the 'effort after shared meaning' (Schwartz, 1995). For instance CSCL designers deliberately form pairs with conflicting opinions (See section 5), because this conflict solving situation will require peers to produce a higher effort to build a shared solution. In other words, we face the same argument on the status of cognitive load in learning in groups as in learning individually.

We hence need to discriminate between the positive and negative effects of cognitive load, the later being implicit in the term 'overload'. Misunderstandings increase the effort of knowledge elicitation and may hence be positive for learning. In contrast, too many misunderstandings

would of course spoil collaboration. We therefore use the notion of **optimal collaborative effort** (Dillenbourg, Traum & Schneider., 1996) that is the equivalent of the notion of 'germane load': The interactions that enable students to co-construct knowledge are not effortless; they require some effort. For instance, Webb (1991) discriminated two levels of elaboration in explanations: the low elaborated explanations were not predictive of learning gains while elaborated explanations which have a higher cost, led to learning gains. The goal of the CSCL designer is to tune this collaboration load within an acceptable range, i.e., above a floor threshold below which too few cognitive processes are triggered but below a ceiling threshold (overload) above which collaboration becomes painful or unmanageable... Any collaboration load that goes beyond the optimal collaborative effort, makes collaboration unnecessarily difficult, and could be associated to extraneous cognitive load.

3.   What constitutes collaboration load?

We now attempt to disentangle the different factors that impart on collaboration load. First, we address the load reduction that may be produced by the division of labour and then we review mechanisms that increase cognitive load: the verbalization of thoughts, the construction of a shared understanding and the maintenance of a representation of the other team members. The ideal collaborative learning situation would minimize extraneous load (by load reduction mechanisms) and generate germane load by rich social interactions.

3.1.   The benefits of division of labour

The benefits of division of labour have been illustrated in situations where the task regulation could be performed by a subject other than the one carrying out the task operations. Even though group members act together on a task, often one partner takes responsibility for the low level aspects of the task while the other focuses on strategic aspects (Miyake, 1986). Blaye et al. (1991) showed that this mutual regulation was progressively internalized as self-

regulation skills. If mutual regulation is an intermediate phase in the acquisition of self-regulation skills, one may hypothesize that it is easier. The division between cognitive and metacognitive layers of the task may lead to an individual offload; the cognitive load of mutual regulation (A does the task and B monitors A) can be hypothesized to be lower than the cognitive load of self-regulation (A does the task plus A regulates A).

Some pedagogical methods for group learning use division of labour for reducing cognitive load. When students have to learn about complex issues, a common collaborative script (see section 5) is to ask them to adopt controversial roles. For instance, student A plays the role of an engineer who designs a new product, while B is a financial manager and C is the lawyer. Each student has to handle his or her own set of arguments, detecting when they are relevant. If one student would have to conduct the same reasoning individually, he/she would need to handle multiple sets of arguments, i.e., checking for each argument produced if there is an existing counter-argument. This recursive self-refutation process is expected to induce a higher cognitive load than mutual argumentation.

This task/meta division occurs in collaborative sessions rather than in cooperative settings: When team members split the task into independent subtasks, it does not change the cognitive load, but only the workload. If, instead of writing a full report, I have to write half of it, the cognitive load of writing a sentence at a given time remains the same, what changes is the time I will spend on the whole report.

We conducted a study (Rebetez et al, 2004) in which we compared individuals and pairs during a learning task. Pairs reported a lower cognitive effort than individuals. The students were asked to study some multimedia material together; a situation that afforded no division of labour. The explanation of the perception of a lower load has to be explained by a lower workload or by other factors. One may be the 'social facilitation effect' (Michaels et al., 1982), which explains performance increases by the mere presence of others, despite any

interaction. Social facilitation may impact on the subjective feeling of effort but not on the very notion of cognitive load.

3.2.   The cost of verbalization

Even if our own reasoning is structured by language, turning ideas into sentences consists in an additional process that is not free of charge. Verbalizing one's thoughts requires a metacognitive activity that can be harmful to the task performance particularly for procedural and automatic tasks. Biemiller and Meichenbaum (1992) showed that students who found a task was cognitively demanding had little resources left for following their 'think aloud' instructions. Verbalization implies reflective mechanisms (being aware of one's own knowledge) plus the voicing internal reasoning aloud. Pure verbalization is cognitively demanding because of the necessity to apply discourse linearization processes (Levelt, 1989). For instance, during a highly automated task, verbalization induces a cognitive overload that is perceptible through pauses, hesitations and slowing down of the task achievement (Hoc & Leplat, 1983). In addition, it has been shown that thinking-aloud methods, often used in expertise modeling and usability studies, change the task itself. Verbalization during the course of action induces more planning than when the action is performed regularly (Ericsson & Simon, 1980).

Since verbalisation induces some cognitive load, it raises the same argumentation as depicted in section 3. Is it beneficial or detrimental to learning? Webb (1989, 1991) found that the effects of verbalization depend on the degree of elaboration of the explanation produced during collaborative learning. This effect may be related to the 'self-explanation-effect', which refers to learners learning more from examples when they are asked to explain these examples to themselves (Chi, Bassok, Lewis, Reiman, & Glaser, 1989). The learning situation typically entailed both declarative and procedural knowledge that the learner should acquire in order to be capable of solving application and transfer problems. Using thinking-aloud

protocol, the studies showed that the difference between good and poor learners could be explained by quantitative and qualitative differences in the explanation produced. For example, not only did good learners produce more self-explanations during examples studying, but they also tried to find out how each information was derived from the other. The self-explanation effect is not due to a simple externalization process but rather to deep cognitive processing, consisting in drawing inference from the example and the conceptual knowledge already acquired (Van Lehn, Jones, & Chi, 1992). These cognitive effects of explanations do not come for free, they reflect intense cognitive processing and the load necessary to construct an explanation. Since collaborative learning induces, although not systematically, the need to explain to each other usually in order to make joint decisions, one can infer that at least the same load as self explanation is incurred.

3.3. The cost of grounding

Dialogue is of course more than verbalization. The construction of a mutual understanding or grounding (Clark & Brennan, 1991) requires additional mechanisms; when A tells something to B, B listens and interprets what A says, A monitors what B has understood what A meant while B provides A with some cues of his (or her) understanding (backchannel). A may repair his/her utterances on the basis of this backchannel or A may anticipatively tailor his/her utterances to what he/she expects B to understand (audience design). This formal analysis of dialogue seems to indicate a huge cognitive load, much higher than what we experience in daily conversations. Actually, we use a range of default reasoning mechanisms (e.g., A agrees with me unless he/she explicitly disagrees) that reduces the cognitive load of everyday dialogue. We nonetheless experience this load in specific situations where communication is made difficult by the channel (e.g., bad phone conversation), the content (a complex domain), a difference of referential background (e.g., cultural differences), etc.

Among the different sub-processes of grounding, let us focus on the process of tailoring utterances to the receiver. Mechanisms of 'audience design' (Lockridge & Brennan, 2002) are salient for instance if one compares explaining how to reach an address to a foreigner versus to somebody who knows the city well. Horton and Gerrig (2005) explored the memory demands induced by the 'audience design' mechanisms. Their experiment shows that audience design was more likely to occur in conditions where to the subjects received some help for remembering what the listener had seen before. In other words, audience design participates to cognitive load.

One could expect that explaining to somebody else generates a higher cognitive load (and hence higher learning outcomes) than self-explanation since explaining to someone else requires both the mechanisms of the audience design and of constructing an explanation (self-explanation). We explored this hypothesis with different levels of interactivity in the explanation process (no listener, silent listener, interactive listener), but we found no clear evidence that explaining to someone is more or less effective than self-explanation (Ploetzner et al., 1999) . Does it mean that there is no pure self-explanation, i.e., that any self-explanation experiment actually includes some listener: the experimenter does not directly interact with the subject but the subject knows that the experiment will listen the recording later on? Or, does it mean that the additional cognitive mechanisms involved in dialogue (but not in self-explanation) are not cognitively demanding enough? As we said earlier, this chapter raises more questions than it provides answers; the cognitive load/benefits of audience design are largely unexplored.

It is only for the sake of disentangling mechanisms that we dissociated the effort that A makes to be understood by B and the effort that B makes to understand A; mutual understanding is of course a joint effort where both the emitter and listener contribute. The process of constructing a shared understanding is referred to as the grounding process in

psycholinguistics. Clark and Brennan (1991) studied how the cost of grounding varies from one medium to another, where the concept of 'cost' covers both cognitive load and physical workload. They discriminate several sub-costs: the production costs refer to the effort for articulating or typing the message; the formulation costs refer to how easy is it to decide exactly what to say; the reception costs are concerned with the effort necessary for listening to or reading the message, including attention and waiting time; the understanding costs are those necessary for interpreting the message in context; the start-up costs refer to how partners initiate a conversation; the delay costs are those necessary for making the receiver wait during formulation. This set of factors also includes the asynchrony costs (for instance, not being able to tell what is being responded to), the speaker change costs, the fault costs and the repair costs. Each of these factors varies with the medium, not only globally (e.g., chat versus voice) but also in a very specific way (two different chat systems may vary the costs). The study 2 (Section 4.2) provides more details in the cost of grounding across different media.

The sum of these costs corresponds to the global effort required to build a shared understanding. For instance, in a remote audio-conferencing environment, the lack of visual cues requires devoting more attention to turn taking than in face-to-face conversation. Another example is that, in face-to-face dialogue, misunderstandings are detected even before an utterance is completed by the emitter, while in chat environments, the receiver does not see the utterance until it is completed and sent. Computer-mediated communication may induce high costs. The relationship between these costs and cognitive load is not simple. Some of these costs may slow down dialogue or increase the physical load, but not necessarily increase the cognitive load. Nevertheless, all together, these costs increase the complexity of the learning task for novice users who have to simultaneously learn the domain and the communication environment (Hron & Friedrich, 2003). This extraneous load concerns not

only the usability of the tool but also the acquisition of specific conversation rules. For instance, chat users will learn that adding "[…]" at the end of a turn means that the message will continue in the next turn ('delay costs').

Common sense tells us that these extrinsic costs will quickly decrease as learners become familiar with the system. This is only true if the system is used on a regular basis. Actually, most empirical research on complex CSCL environments consists of short duration experiments in which this decrease cannot be rigorously assessed. Most long term experiments tend to use more standard environments, such as forums.

Finally, some of the factors that increase extraneous load may not be detrimental for some learning tasks. For instance, the time ones apparently wastes typing sentences in web-based forums is also a time that can be used for reflection. Learners benefit from having more time to reflect on their answers. in other words, any equation such as "the closer to face-to-face, the lower the cognitive load" would fail to account for the very adaptive nature of humans and the cognitive off-load generated by some software features (Dillenbourg, 2005).

3.4. The cost of modeling

The construction of a shared understanding requires that each partner build some representation of the other partners' beliefs, knowledge or goals. We refer to this process as **mutual modeling**, a facet of intersubjectivity (Bromme, 2000; Wertsch, 1985). By using the term 'model', we do not imply this is a detailed or explicit representation of the partner. Simply stated, if A wants to (dis-)agree with B, A needs some representation of B's position; if A wants to repair B's misunderstanding, A needs some representation of what B understood. Mutual modeling is, like the grounding process, very functional with its degree of precision depends on the task. For instance, it has to be extremely high when two pilots are discussing

the track where their aircraft should land, but can be much lower if they are discussing about the last party they went to (down to what politeness allows).

This mutual model is not constructed in a vacuum but is based on initial representations. Common grounds are initialized by the assumptions people make about their partners from cues such as their community membership (age, culture, profession...) and from co-presence (e.g., common grounds include any event A and B attended together) (Clark & Marshall, 1981). Several scholars studied how this initial model imparts on communication, namely because it can easily be manipulated. For instance, Slugoski et al. (1993) told some subjects that their (fake) partners received the same information as the subjects themselves and told other subjects that their partners received a different piece of information. They observed that the subjects adapted their explanation to their partners by focusing on items that this fake partner was supposed to ignore. The mutual modeling part can be seen as the diagnosis part (or 'communality assessment – Horton & Gerrig, 2005) of the audience design process; A needs some information about B to tailor his/her explanation to B. Similarly, Brennan (1991) showed that the subjects used different initial strategies in forming queries depending on who they were told their partner was. Other simpler inference mechanisms such as default reasoning rules (e.g., B agrees with me unless he disagrees) are developed according to the conversational context.

In summary, the need for modeling a peer's knowledge ranges from global clichés to detailed information. Subsequently, the mechanisms for modeling range from default assumptions to elaborated inferences. Hence, the cognitive load and the benefits of the mutual modeling process will be very different in various collaborative contexts. We nonetheless hypothesize that, when triggered, the mechanisms for monitoring how one's partner understands the task contribute to deepen one's own understanding. Mutual modeling is some kind of "thinking in stereo", looking at the knowledge from two different angles.

4.   Do computerized environments influence collaboration load?

We did not conduct studies that specifically aimed at measuring collaboration load. However, over the last years, we carried out experiments on CSCL environments that provided us with some preliminary elements for understanding collaboration load.

4.1.   *Cumulating load from learning material and collaboration*

This research project originated from a study of Schnotz, Boeckeler and Gzrondziel (1999) who found evidence for the collaboration load assumption. In a first experiment, learners had to individually study a hypertext explaining time zones on earth with either interactive and animated graphics (simulations) or static graphics. They found that learners studying individually with the interactive graphics performed better than learners studying with static graphics when answering factual questions, but not when answering comprehension questions. One hypothesis was that learners with the animated graphics condition looked at the simulation passively, whereas learners with the static graphics condition mentally performed the time zone simulation. The tendency of subjects to 'underprocess' animated learning content has been found in other studies and is referred to as the 'underwhelming' effect of animated graphics (Lowe, 2004). In a second experiment, learners were grouped in pairs to study the same instructional material. Pairs learning with the animated material had poorer performance to both kinds of questions than pairs with the static graphics condition. According to Schnotz et al. (1999), learners in pairs did not benefit from the simulation because they had to allocate cognitive resources to co-ordinate their learning with the peer, in addition to process the visual display. The authors inferred that the sum of the load induced by visually complex graphics and the load involved by managing the collaboration could lead to cognitive overload and hence impair the learning process. As complementary evidence for this explanation, the subjects reported that they had less peace to think deeply in the

collaborative condition. However, as the two studies were two distinct experiments, no direct statistical comparison was possible between the individual and collaborative situations regarding performance.

An alternative explanation of Schnotz et al. (1999) results is that pairs could not benefit from the dynamic graphics since they could not base their shared understanding on a stable graphic representation. Since animations display rapid and transient information, they do not provide external support to the referring objects. Deictic gestures are a key mechanism in the construction of common grounds. By inhibiting them, animated pictures might also inhibit the most interesting interactions of collaborative learning, leaving pairs only with the drawbacks. Since the persistency of representation is essential for grounding (see section 4.2), displaying permanent static snapshots of the critical steps of the animation while it is running would be a good way to ensure persistency of the depiction in addition to the dynamic visualization of the phenomenon.

We carried out an experiment to investigate this hypothesis, using a factorial design with three factors: visualization (static or animated); permanence (presence of static snapshots or not); and a learning situation (individual or collaborative) (Rebetez, Sangin, Bétrancourt, & Dillenbourg, 2004). The participants had to study two animations explaining astronomic and geologic phenomena. Then, they had to answer, individually in all conditions, to retention and transfer questions. Retention questions involved recalling information that was provided in the instructional material whereas transfer questions required learners to draw inferences from the material by mentally simulating the phenomenon. Describing our results in a detailed way goes beyond the scope of this chapter. However we found three interesting results that pertain to the notion of collaboration load.

First, in the individual situation, the animated visualization improved performance compared with static graphics for retention questions only, while no difference was found on transfer

questions. In contrast, in the collaborative situation, the animated visualization was beneficial both for retention and transfer questions, compared with static graphics. Second, a series of five scales adapted from the NASA-TLX was used as a self-reporting measure of cognitive load (cognitive effort, time pressure, mental demand, performance satisfaction and frustration). Learners in pairs reported significantly less cognitive effort than learners in individual situations. The other scales pertaining to cognitive load followed the same trend but not significantly (see Figure 1). In other words, not only did pairs have better performance than individual learners but they also evaluated their cognitive load as lower than individual learners. These two results contradict Schnotz's et al. (1999) 'collaboration load' hypothesis.

<insert here figure 1 >

Our third main result is however intriguing. When static snapshots were provided alongside the graphics, performance increased for individuals but decreased for pairs. This is surprising since these snapshots aimed at supporting grounding mechanisms during group discussions. Our tentative explanation is that this condition produced a **'split-interaction effect'**. This effect appears when collaborative learning is impaired by interferences between the two concurrent modes of interactions; on the one hand, the interactions between the users and the system, and, on the other hand, the interactions among users. These results are even more surprising if one considers that the difference of interactivity between the two conditions was limited to being allowed to view snapshots or not and that those who were allowed to do so, did not do it very frequently. Actually, the observed effects are not as much due to a lack of cognitive resources as to a split-attention effect between interacting with the simulation and interacting with the peer.

This hypothesis relies on recent results (Wilson & Peruch, 2002) indicating that the users' focus of attention is more influential on learning achievement than users' physical interaction with the material. In two first experiments, the authors compared learners who actively explored a virtual spatial environment to learners who just looked at someone exploring for them. Surprisingly, results showed that active and passive learners were not systematically different regarding their learning of the environment (including navigation measures). In a third experiment, Wilson and Peruch (2002) added instructions for the active and passive learners: In one condition, learners were told to pay attention to objects in the environment and in another condition, they were told to pay attention to the spatial layout. Again the active or passive factor did not yield any consistent difference, but the instructions had a significant effect on the remembering of objects or layout accordingly. The authors proposed that cognitive interactivity is determined by the focus of attention and is not heavily affected by behavioral interactivity. This might explain why, in our study, the snapshots caused a split-interaction effect in the collaborative condition even if learners did not interact heavily with the device. The snapshots would have attracted learners' attention, creating cognitive interactivity, which affected their processing even though there was hardly any physical interaction.

4.2. Grounding across different media

The second study (Dillenbourg & Traum, to appear) revealed that the medium features that influence collaboration load are not those one initially expects. For instance, studies showed that collaborating with a video-conferencing system is not necessarily more effective than with an audio-conferencing system. A chat (synchronous text-based communication) enforcing turn taking rules and similar to face-to-face meetings is not more effective than a chat with its peculiar turn taking habits, etc. (Dillenbourg, 2005). The reported study aimed at analyzing the grounding mechanisms in a multimodal collaborative environment (Dillenbourg

& Traum, to appear). Twenty pairs had to solve an enigma problem using a chat and a whiteboard (a drawing software where they can jointly construct and edit graphical representations). We estimated the grounding effort by counting the rate of acknowledgement, i.e., the ratio between the number of acknowledged interactions and the total number of interactions. In typed interactions, peers acknowledge 41% of the utterances of their partners, on average. The rate for the spoken conversation pairs however, was 90%. This comparison is slightly awkward since the acknowledgment rate is dependent on the way speech is segmented into utterances. Nevertheless this difference tells us something about the cost of grounding in a chat environment. Because this cost is high, subjects are selective in the type of information that justifies a grounding act. This difference of acknowledgement concerns the physical workload (the time and effort for typing) rather than the cognitive load (the resources used to acknowledge). Two other findings are more relevant for the appraisal of collaboration load.

First, the rate of acknowledgement was 26% for information that subjects simply retrieved from the game and transmitted to each other (hereafter 'facts'), for instance "Hans is the barman", while the rate of acknowledgement was 46% for the information they inferred from the situations (hereafter 'inferences') such as "Hans had a motive to kill Lisa". Syntactically, a sentence such as "Hans is the barman" is identical to "Hans is the killer". What is different is the role of such a sentence in the joint problem solving process, namely the probability of disagreement. Inferences such as "X has a good reason to kill" are personal interpretations of facts and hence more likely to be points of disagreement. If the acknowledgment rate of such utterances varies, it implies that grounding is sensitive to the status of these utterances within the problem solving process. These findings contribute discriminate grounding at the utterance level and grounding at the knowledge level (Dillenbourg & Traum, to appear). In terms of collaboration load, the latter matters more than the former. The cognitive load

necessary to remember what the partner knows, what he agreed upon or what he might disagree with – the process of mutual modeling –was described in section 3.4. In a sentence such as "He is really suspicious", the mutual modeling required for grounding at the utterance level (what does my partner refers to by "He"?) is less demanding than the mutual modeling required for grounding at the knowledge level (What does he mean by "really suspicious"?).

Second, we expected the whiteboard to be used to clarify their verbal interactions in the chat, in the same way we draw a schema on the napkin for explaining what we mean. This experiment revealed the opposite relationship; the whiteboard was the central place of interaction, and the chat interactions were mostly accessory to it. The chat was mostly used to ground short notes posted on the whiteboard and to discuss non-persistent information, namely the strategy (e.g., "let do this now"). The whiteboard was used for gathering all important information and for organizing it (namely, structuring information by suspects and discarding suspects proved to be innocent) (see Figure 2). In other words, the pairs jointly maintained a representation of the state of the problem on the whiteboard. The whiteboard was used as (external) working memory, i.e., the place to store a (shared) representation of the problem state. As the notion of cognitive load is intrinsically related to the limits of our working memory, the notion of collaboration load should be related to the capacity of maintaining this shared working memory. However, the effect of such external working memory on individual cognitive load is complex; it is not because pairs share a physical representation of the problem state that they do not also build a mental representation that remains subject to the limitations of working memory. This relationship between external group memories and collaboration is further discussed in section 6.

<insert here figure 2 >

4.3. Supporting mutual modeling

This third study aimed at investigating mutual modeling mechanisms. The collaborative task was a game called SpaceMiners (Figure 3) (Nova et al., 2003). This 3D game involves two players in space missions where they have to launch drones in order to collect asteroids full of minerals and bring them to space stations. The drones' direction is modified by the planet's gravity and by some objects that the teams drop in space. The teams' task was to collect the largest amount of minerals located in asteroids and to bring them to the space station on the left.

We measured mutual modeling by interrupting group activity at 3 pre-defined points in time and by asking them to select from a list what they would do next and what they think their partner would do next. The accuracy of mutual modeling was estimated by the overlap between what A says (s)he will do and what B says A will do (hereafter MM-rate). We manipulated the degree of mutual modeling by using the availability of so-called awareness tools (Gutwin & Greenberg, 1998) as the independent variable. These tools are software components that inform users, within multi-users environments, of what the other users are doing, where they are located, what they are looking at, and so forth. The awareness tools implemented in SpaceMiners informed A about what B was looking at and where B was going to set some objects.

The awareness tools led to higher group performance, which is consistent with previous findings (Gutwin &Greenberg, 1998), but did not improve the accuracy of the mutual model; pairs with the awareness tools were not higher at MM-rate than pairs without awareness tools. However, a two-way analysis of variance conducted on contrasted groups (post-hoc split) showed that pairs in the awareness condition who spent more time using the awareness tools reached higher levels of mutual modeling than the others.

However, a fourth study led to contradictory results (Nova, Dillenbourg, & Girardin, 2005). Teams of 3 subjects had to walk across the EPFL campus in order to find a virtual object. Each of them carried a laptop (a tabletPC) displaying a map of the campus, their own position and a proximity sensor (telling them how far the object to be found is). In the first condition, the subjects had an awareness tool; they could see the position of their team mates on the map. In the second condition, they had no awareness tool. In both conditions, they could also draw annotations on the map. The performance of the groups (the length of the path of each group member before finding the object) was equivalent in both conditions. We measured mutual modeling by asking the subjects to draw their own path and the path of their partners on a paper map of the campus. The groups without the awareness tool drew more accurate paths than the groups with the awareness tool. The explanation is that the lack of the awareness tool led them to draw more annotations on the map, annotations that constituted a more explicit act of mutual information.

In summary, the awareness tool partly off-loaded groups with the burden of mutual modeling, which led to more accurate and to less accurate mutual models in the third and fourth studies respectively. These results illustrate our earlier discussion on the relationship between learning outcomes and cognitive/collaboration load and our notion of 'optimal collaborative effort'. Reducing load is not always more effective; some collaborative load is part of the natural mechanisms that makes collaboration produce learning.

5.   How can CSCL designers tune collaboration load?

These four studies show that there is no intrinsic cognitive load to collaborative situations in general because there are many of ways of collaborating, multiplied by many of ways for supporting collaboration in computerized environments. The cognitive load depends on the nature of the task, on the composition of the group, on the features of CSCL environments

and on the interactions among these three factors. Hence, the key issue is how to design a CSCL environment in a way that increases or decreases extraneous collaboration load. Research in human-computer interaction has devoted a lot of attention to the cognitive load induced by any interface. We do not address these general issues of interface design here, but focus on factors that specifically concern the cognitive load of CSCL environments. Empirical studies have not yet provided us with a set of rules that designers should apply. Instead, we describe the design space, i.e. the features that designers should consider when thinking in terms of collaboration load.

- As mentioned earlier, CSCL environments vary by the extent to which each team member is informed of what the others are doing. The design of shared editors was based on the WYSIWIS principle: What you see is what I see; all group users have the same view (Stefik et al., 1987). This principle must be relaxed when the group includes many members or when the task is so complex that group members have to work at different times on different parts of the task. In these cases, users need different views. In order to sustain group coordination, designer created 'awareness tools' (Gutwin & Greenberg, 1998), that were addressed in section 4.3. A key question in the design of environments for collaborative problem solving is: Which elements of information should be provided by the awareness tools in order to decrease extraneous collaboration load but without inhibiting the mutual modeling process that is necessary for learning? This question is similar to the one a teacher asks when he decides that his students could use their pocket calculators in order to off-load some parts of the computation while not affecting notions they are supposed to learn.

- Another feature of CSCL environments that we mentioned earlier is the persistency of display; jointly constructed schemata, when they are persistent,

partly off-load memory by providing group members with a shared representation of the problem state. This is not basically different from face-to-face collaborative settings where group members draw a schema on paper or construct an artefact together. However, an interesting difference is that a shared representation contains its own history. Some environments will colour objects according to the author or will fade out colours with time. Other environments enable the user to 'undo' a certain number of recent changes or even to scroll back in the history of previous problem states. This history results from a combination of automatic and explicit "save as" recordings. We have no empirical data regarding how much this history off-loads the metacognitive processes of the team (perceiving their own strategy, identifying loops, errors…) but it certainly extends the group memory.

- Many CSCL environments record the history of interactions, which is another form of the persistency principle. This history may be a flat list of interactions as in a chat environment, a WIKI or a blog. It may reflect the structure of conversation as in "threaded" environments (Reyes & Tchounikine, 2004) or even the rhetorical structure of the argument (as in Belvedere – Suthers et al., 2001). The interface reifies the dialogue to the group members themselves. It can be argued that the long term memory off-load generated by an interaction history may indeed increase working memory load. Searching through a long list of messages is intensive. However, all chat users know how useful it is to scroll up a few lines to clarify an utterance. We saw in the experiment mentioned in section 4.2 that even novice chat users were able to participate in multiple parallel synchronous conversations because the tool partly offloads the need to memorize the conversational context for each interlocutor. In a CSCL

environment, the communication among team members becomes a substance that both the system and team members may process. One research direction is to improve the analysis of this substance by the system (Barros & Verdejo, 2000; Constantino-González & Suthers, 2001; Inaba & Okamoto, 1996). Another research direction is to provide group members with some graphical representation of their interaction histories and let them interpret these representations. We refer to these representations as **group mirrors** (Dillenbourg et al., 2002; Jermann, 2004; Zumbach et al., 2002). Jermann (2004) investigated the hypothesis that these mirrors would off-load the process of self-regulation. The task given to the pairs was to tune the lights of several crossroads to optimize the flow of cars through a city. In a first set of experiments, he observed that the most effective pairs were those who discussed their options before tuning the lights. In such a dynamic system, a simple approach by trial-and-error leads to low performance. Therefore, this information was encompassed into a group mirror (figure 3). Experiments show that pairs do interact differently when provided with this feedback, although this difference does not lead to higher group performance. The question here is whether the cognitive load provided by the external regulation is higher than the new cognitive load induced by reading and interpreting the group mirror.

<insert here figure 3 >

- Another feature of a CSCL environment is the use of semi-structured communication interfaces, as illustrated in Figure 4; which is usually a text-based communication tool embedded into a series of interface items (Baker & Lund; 1996; Soller, 2002; Suthers et al., 2001; Veerman & Treasure Jones,

1999). These tools partly offload the physical load of typing utterances by offering predefined speech act buttons (e.g., "I agree"). The purpose of these interfaces overall is to favour the emergence of specific interaction patterns. This may also support the grounding mechanisms by making explicit which utterance is being acknowledged, which object is being referred to or which type of speech act is being uttered (Reyes & Tchounikine, 2003). Do these tools off-load the students working memory? This is certainly not the case at the beginning. Dialogue moves are usually performed in an implicit way. Forcing students to express their next dialogue move constitutes another example that increases cognitive load, which is didactically suitable but may nevertheless be detrimental to collaborative learning.

<insert here figure 4 >

The features we listed refer mainly to software tools. Another trend in CSCL research concerns the design and experimentation of 'scripts'. **A collaboration script** (Aronson et al., 1978; O'Donnell & Dansereau, 1992) is a set of instructions prescribing how the group members should interact, how they should collaborate and/or how they should solve the problem. The collaboration process is not left open but structured as a sequence of phases. Each phase corresponds to a specific task where group members have a specific role to play. A well-known script is the 'reciprocal teaching' approach set up by Palincsar & Brown (1984). One peer reads a paragraph of text and the other questions him/her about his/her understanding. The roles are shifted for the next paragraph. We developed and investigated several scripts (Dillenbourg & Jermann, in press), such as the ArgueGraph scripts aimed at raising conflict solving interactions among group members or the ConceptGrid script, a JIGSAW-like script, aimed at triggering explanations. Do scripts increase or decrease

collaboration load? Any script has to be "played", i.e. the student has to remember what to do, when, how, etc. which increases the load. However, some scripts may reduce cognitive load since they reduce the need for mutual modeling; when roles are clearly defined, the learner may play his or her role and spend less energy for a close coordination with other roles. Actually, the scripts we designed deliberately increase the collaboration load. In order to trigger argumentation, the ArgueGraph script forms pairs with student having conflicting opinions, i.e., it increases the effort necessary to reach consensus. The ConceptGrid script provides groups of four students with four different subsets of knowledge, thereby increasing the explanation effort they have to engage in to build a concept grid. Scripts are ways for designers to tune the collaboration load.

6.   Conclusion: Group cognitive load

The discussion so far focused on the load that collaboration imposes on individuals. A distributed cognition theories (Pea, 1993; Hutchins, 1995; Salomon, 1993) perspective would be to consider the cognitive load for the group as a whole. This alternative was briefly mentioned in section 4.2 when we discussed that at the group level jointly constructed representations play the role that working memory plays at the individual level; maintaining and updating representations of the state of the problem. This distributed cognition viewpoint also concerns the mutual modeling process. It may be the case that team members do not build a representation of their partners' mental states but instead a representation of the interaction process at the group level. Instead of modeling who knows what, who does what, who said what, the team members could maintain a representation of what the team knows, did or said. We refer to this as the **group model** instead of the mutual model.

These two visions of teams, as collections of individuals or as larger units, have been opposed for the sake of argument, but the real challenge is to understand how they articulate with each

other. Let us take a simple example; a knot in my handkerchief to remind me to buy bread is expected to offload my memory. Actually, the situation is slightly more complex; I still have to remember that this knot means "buy bread". In our study, peers co-constructed a visual and physical representation of the task that included information beyond the capacity of working memory. However, they still needed, in order to take decisions, some kind of mental representations of this external representation (e.g., a guy with a red cross meant "this person is not guilty" for many teams). We know that information may stay in the working memory for longer periods by using an articulory loop (repeating it) or using knowledge structures in long-term memory (Ericsson & Kintsch, 1995). It may be that the shared visual representation plays a similar role, providing group members with a continuous reactivation of the elements to be maintained in the working memory.

The notion of memory at the group level is clearly different from the notion of working memory of individuals. It has a physical counterpart (usually some artefact), it has a larger capacity and it is more visual than auditive. In other words, group memory could be conceived as the equivalent, at the group scale, of the concept of long-term working memory at the individual scale. It extends individual and collective cognitive capacities by offloading, organizing and updating information available to the group. In this context, collaboration load could be defined as the effort engaged in by team members to co-construct a long-term working memory by incrementally grounding the role of each piece of information with respect to the problem solving process.

7.   Acknowledgements

8.   References

Aronson, E., Blaney, N., Sikes, J., Stephan, G., & Snapp, M. (1978). *The Jigsaw Classroom.* Beverly Hills, CA: Sage Publication.

Baddeley, A. (1997). *Human memory: Theory and practice*. London: Lawrence Erlbaum.

Baddeley, A. (2000). The Episodic Buffer: A New Component of Working Memory?" *Trends in Cognitive Sciences*, 4, p. 417-423.

Baker, M.J. & Lund, K. (1996) Flexibly structuring the interaction in a CSCL environment. In P. Brna, A. Paiva & J. Self (Eds), *Proceedings of the European Conference on Artificial Intelligence in Education*. Lisbon, Portugal, Sept. 20 - Oc. 2, pp. 401-407.

Barros, B. & Verdejo, F. (2000) Analysing student interaction processes in order to improve collaboration: The DEGREE approach. *Journal of Artificial Intelligence in Education*, 11, 211-241.

Biemiller, A., & Meichenbaum, D. (1992) The nature and nurture of the self-directed learner. *Educational Leadership*, 50(2), 75-80.

Blaye, A., Light, P., Joiner, R., & Sheldon, S. (1991) Collaboration as a facilitator of planning and problem solving on a computer based task. *British Journal of Psychology*, 9, 471-483.

Brennan, S. E. (1991) Conversation with and through computers. *User Modeling and User-Adapted Interaction*, 1, pp. 67-86.

Bromme (2000) Beyond one's own perspective: The psychology of cognitive interdisciplinarity. In P. Weingart & N. Stehr, (Eds), *Practicing interdisciplinarity*. (pp. 115-133)Toronto: Toronto University Press.

Brünken, R., Plass, J.L. & Leutner, D. (2004). Assessment of cognitive load in mutlimedia learning with dual-task methodology: Auditory load and modality effect. *Instructional Science*, 32, 115-132.

Brünken, R., Steinbacher, S., Plass, J.L. & Leutner, D. (2002) Assessment of cognitive load in multimedia learning using dual-task methodology. *Experimental Psychology*, 49, 1-12.

Chi, M. T. H., Bassok, M., Lewis, M.W. Reiman, P. & Glaser, R. (1989). Self-explanations: How students study and use examples in learning to solve problems. *Cognitive Science, 13,* 145-182.

Clark, H.H. & Marshall, C.R (1981) Definite reference and mutual knowledge In A. K. Joshi, B. L. Webber, and I. A. Sag (Eds), *Elements of Discourse Understanding*. Cambridge University Press.

Clark, H.H. & Wilkes-Gibbs, D (1986). Referring as a collaborative process. *Cognition*, 22:1–39.

Clark, H.H., & Brennan S.E. (1991) Grounding in Communication. In L. Resnick, J. Levine & S. Teasley (Eds.), *Perspectives on Socially Shared Cognition* (127-149). Hyattsville, MD: American Psychological Association.

Constantino-Gonzales, M. A., Suthers, D., Icaza, J. (2001). Designing and Evaluating a Collaboration Coach: Knowledge and Reasoning. In J. D. Moore, C. L. Redfield, & W. L. Johnson (Eds.) *Artificial Intelligence in Education: AI-ED in the Wired and Wireless Future* (10th International Conference on Artificial Intelligence in Education), Amsterdam: IOS press, May 19-23, San Antonio Texas, pp. 176-187.

De Corte, E. (2003). Designing learning environment that foster the productive use of acquired knowledge and skills. In E. De Corte, L. Verschaffel, N. Entwistle & J. van Merrienböer (eds.) *Unravelling basic components and dimensions of powerful learning environments.* (pp 21 - 33). Pergamon: Elsevier Science Ltd.

Dillenbourg P. & Traum, D. (to appear) The complementarity of a whiteboard and a chat in building a shared solution. *Journal of Learning Sciences*.

Dillenbourg, P. & Jermann, P. (to appear). SWISH: A model for designing CSCL scripts. In F. Fischer, H, Mandl, J. Haake & I. Kollar (Eds) *Scripting Computer-Supported Collaborative Learning – Cognitive, Computational, and Educational Perspectives* . Computer-Supported Collaborative Learning Series, Springer

Dillenbourg, P., Traum , D. & Schneider D. (1996) Grouding in multi-modal task oriented collaboration. *Proceedings of the European Conference on Artificial Intelligence in Education*, Lisbon, Portugal, September, pp. 415-425.

Dillenbourg, P. (2005) Designing biases that augment socio-cognitive interactions. in R. Bromme, F. Hesse and H. Spada. (EDS) *Barriers and biases in computer-mediated knowledge communication* (pp. 243-264) Computer-Supported Collaborative Learning Series, Springer

Dillenbourg, P., Ott, D., Wehrle, T., Bourquin, Y., Jermann, P., Corti, D. & Salo, P. (2002). The socio-cognitive functions of community mirrors. In F. Flückiger, C. Jutz, P. Schulz and L. Cantoni (Eds). *Proceedings of the 4th International Conference on New Educational Environments*. Lugano, May 8-11, 2002.

Ericsson, K. A., & Kintsch, W. (1995). Long-term working memory. *Psychological Review, 102*, 211-245.

Ericsson, K. A., & Simon, H.A. (1980). Verbal reports as data. Psychological review., 87 (3), 215-251.

Gerjets, P. Scheiter, K. & Catrambone, R. (2004). Designing instructional examples to reduce intrinsic cognitive load: Molar versus modular presentation of solution procedures. *Instructional Science*, 32, 33-58 (a la place de Scheiter et al)

Gutwin, C. & Greenberg, S. (1998) The effects of workspace awareness on the usability of real-time distributed groupware. *Research report 98-632-23*, Department of Computer Science, University of Calgary, Alberta, Canada.

Gyselink, V., Ehrlich, M.-F., Cornoldi, C. de Beni R. & Dubois, V. (2000). Visuospatial working memory in learning from multimedia systems. *Journal of Computer Assisted Learning, 16*, 166-176.

Hart, S. G., & Staveland, L. E. (1988). Development of NASA-TLX (Task Load Index): Results of experimental and theoretical research. In P. A. Hancock & N.

Meshkati (Eds.), *Human Mental Workload* (pp. 139-183). Amsterdam: North Holland.

Hoc, J.M., & Leplat, J. (1983). Evolution of different modalities of verbalization. *International Journal of Man-Machine Studies*, 19, 283-306.

Horton, W.S. and Gerrig, R.J (2005) The impact of memory demands on audience design during language production.. *Cognition, 96, 127-142.*

Hron, A., & Friedrich, H. F. (2003). A review of web-based collaborative learning: Factors beyond technology. *Journal of Computer Assisted Learning, 19*, 70-79.

Hutchins, E. (1991) The Social Organization of Distributed Cognition. In L. Resnick, J. Levine and S. Teasley. *Perspectives on Socially Shared Cognition* (pp. 383 - 307). Hyattsville, MD: American Psychological Association.

Hutchins, E. (1995). How a cockpit remembers its speeds. *Cognitive Science*, 19, 265-288.

Inaba, A. & Okamoto, T (1996) Development of the intelligent discussion support system for collaborative learning. *Proceedings of Ed-Telecom* '96. (pp 494-503), Bostoo.

Jermann, P. (2004) Computer Support for Interaction Regulation in Collaborative Problem-Solving, Unpublished doctoral thesis. Faculté de Psychologie et des Sciences de l'Education de l'Université de Genève, Switzerland.

Kalyuga, S., Chandler, P., & Sweller, J. (2000). Incorporating learner experience into the design of multimedia instruction. *Journal of Educational Psychology, 92, 1 11.*

Levelt, J.M.W. (1989). Speaking : from intention to articulation. Cambridge, M.A.: the *MIT Press*.

Lockridge C.B. & Brennan, S.E. (2002), Addressees' needs influence speakers' early syntactic choices, *Psychonomic Bulletin & Review, 9 (3), 550-557*

Lowe, R. K. (2004). Interrogation of a dynamic visualization during learning. *Learning and Instruction*, 14, 257-274.

Mayer, R. E. & Moreno, R. (2002). Aids to computer-based multimedia learning. *Learning and Instruction, 12*, 107 – 119.

Michaels, J.W., Blommel, J.M., Brocato, R.M.,, Linkous, R.A. and Rowe, J.S. (1982) Social facilitation and inhibition in a natural setting. *Replications in Social Psychology* **2,** pp 21-24.

Miyake, N. (1986) Constructive Interaction and the Iterative Process of Understanding. *Cognitive Science*, 10, 151-177.

Nova, N., Girardin, F. & Dillenbourg, P.(2005) 'Location is not enough!': an Empirical Study of Location-Awareness in Mobile Collaboration *IEEE International Workshop on Wireless and Mobile Technologies in Education*, Tokushima, Japan.

Nova N., Wehrle, T., Goslin, J., Bourquin, Y. & Dillenbourg, P. (2003). The Impacts of Awareness Tools on Mutual Modelling in a Collaborative Video-Game. In Proceedings of the 9th International Workshop on Groupware, Autrans France, September 2003.

O'Donnell, A. M., & Dansereau, D. F. (1992). Scripted cooperation in student dyads: A method for analyzing and enhancing academic learning and performance. In R. Hertz-Lazarowitz and N. Miller (Eds.), *Interaction in cooperative groups: The*

theoretical anatomy of group learning (pp. 120-141). London: Cambridge University Press.

Paas, F., Renkl, A., & Sweller, J. (Eds.). (2004). Advances in cognitive load theory: Methodology and instructional design [Special issue]. Instructional Science, 32.

Paas, F., Tuovinen, J.E., Tabbers, H., Van Gerven P.W.M. (2003). Cognitive Load Measurement as a Means to Advance Cognitive Load Theory. Educational Psychologist, 38 (1), 63-71.

Palincsar A.S. and Brown A.L. (1984) Reciprocal Teaching of Comprehension-Fostering and Comprehension-Monitoring Activities. *Cognition and Instruction,* vol.1, n°2, pp. 117-175.

Pea, R. (1993) Practices of distributed intelligence and designs for education. In G. Salomon. (Ed). *Distributed cognitions. Psychological and educational considerations* (pp. 47-87) Cambridge, UK: Cambridge University Press.

Ploetzner R., Dillenbourg P., Praier M. & Traum D. (1999) Learning by explaining to oneself and to others. In P. Dillenbourg (Ed) *Collaborative-learning: Cognitive and Computational Approaches* (pp. 103-121). Oxford: Elsevier

Rebetez, C., Sangin, M., Bétrancourt, M., & Dillenbourg, P. (2004). Effects of collaboration in the context of learning from animations, In *Proceedings of the EARLI SIG meeting on Comprehensionof Texts and Graphics: Basic and applied issues* (pp 187-192). September 2004, Valencia (Spain).

Reyes P., Tchounikine P. (2003), Supporting emergence of threaded learning conversations through augmenting Interactional and Sequantial Coherence, In: *International Conference on Computer Supported Collaborative Learning* (CSCL, best PhD student paper award), 2003, Bergen (Norway), p. 83-92.

Rieber, L. P. (1996). Seriously considering play: Designing interactive learning environments based on the blending of microworlds, simulations, and games. *Educational Technology Research & Development, 44*, 43-58.

Salomon, G. (1993). No distribution without individual's cognition: a dynamic interactional view. In G. Salomon. (Ed). *Distributed cognitions. Psychological and educational considerations* (pp. 111-138) Cambridge, USA: Cambridge University Press.

Schober, M.F. (1993) Spatial perspective-taking in conversation. *Cognition*, 47, 1-24.

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

Schnotz, W., Vosniadou, S. & Carretero, M. (Eds.) (1999*). New perspectives on conceptual change*. Oxford: Elsevier.

Schwartz, D.L. (1995). The emergence of abstract dyad representations in dyad problem solving. *The Journal of the Learning Sciences*, 4 (3), pp. 321-354.

Slugoski, B.R., Lalljee, M., Lamb, R. & Ginsburg, G.P. (1993) Attribution in conversational context: Effect of mutual knowledghe on explanation giving. *European Journal of Social Psychology*, 23 (219-238).

Soller, A. (2002). *Computational analysis of knowledge sharing in collaborative distance learning*. Unpublished Doctoral Dissertation. University of Pittsburgh, PA.

Stefik, M., Bobrow, D,G., Foster, G., Lanning, S. & Tatart, D. (1987) WYSIWIS Revised:

Early Experiences with Multiuser Interfaces. *ACM Transactions on Office Information Systems*, 5(2), 147-167, April.

Suthers, D., Connelly, J., Lesgold, A., Paolucci, M., Toth, E., Toth, J., and Weiner, A. (2001). Representational and Advisory Guidance for Students Learning Scientific Inquiry. In Forbus, K. D., and Feltovich, P. J. (2001). *Smart machines in education: The coming revolution in educational technology*. Menlo Park, CA: AAAI/Mit Press, pp. 7-35.

Sweller, J. (1988). Cognitive load during problem solving: Effects on learning. *Cognitive Science, 12*, 257-285.

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, P. and Cooper, M. (1990). Cognitive load and selective attention as factors in the structuring of technical material. *Journal of Experimental Psychology: General, 119*, 176-192.

Sweller, J., van Merriënboer, J.J.G., & Paas, F. (1998). Cognitive Architecture and Instructional Design. *Educational Psychology Review, 10*, 251-295.

Tourneur, Y. (1975), *Recherche en Education Effets des Objectifs dans l'Apprentissage*, Edité par la Direction Générale de l'Organisation de L'Enseignement., Bruxelles, Belgique.

Valcke, M. (2002). Cognitive load: Updating the theory? *Learning and Instruction 12*, 147–154.

VanLehn, K., Jones, R. M., & Chi, M. T. H. (1992). A model of the self-explanation effect. *Journal of the Learning Sciences*, 2, 1-59.

Veerman, A.L. & Treasure-Jones, T. (1999) Software for problem solving through collaborative argumentation. In P. Poirier and J. Andriessen (Eds) *Foundations of argumentative test processing* (pp. 203-230). Amsterdam: Amsterdam University Press.

Webb, N. M. (1989). Peer interaction and learning in small groups. *International Journal of Educational Research, 13,* 21-40.

Webb, N.M. (1991) Task Related Verbal Interaction and Mathematical Learning in Small Groups. *Research in Mathematics Education*. 22 (5) 366-389.

Wertsch, J.V. (1985) Adult-Child Interaction as a Source of Self-Regulation in Children. In S.R. Yussen (Ed).*The growth of reflection in Children* (pp. 69-97). Madison, Wisconsin: Academic Press.

Wilson P.N. & Peruch P. (2002). The influence of interactivity and attention on spatial learning in a desktop virtual environment. *Current Psychology of Cognition, 21*, 601-633.

Zumbach, J., Mühlenbrock, M., Jansen, M., Reimann, P. & Hoppe, H.U. (2002) Multi-dimensional tracking in virtual learning teams: An exploratory study. In G. Stahl (Ed.), Computer support for collaborative learning: foundations for a CSCL community (pp. 650-651). Mahwah, NJ: Lawrence Erlbaum Associates.

Zahn, C. , Barquero, B., & Schwan, S. (2004). Learning with hyperlinked videos – design criteria and efficient strategies of using audiovisual hypermedia. Learning and Instruction, 14, 275-291.
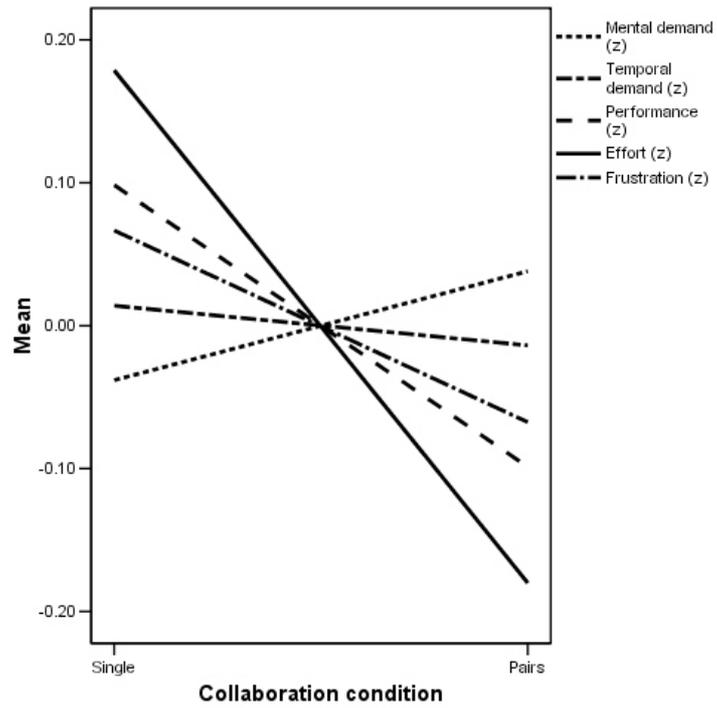
Figure 1. Self-reported five-scale measure of cognitive load in pairs (right) and individual (left) learning situations (using Z scores for the sake of comparison)

Room 6:
-left her room at 7:30
-SAW HEIDI VERY UPSET ABOUT HANS
-went out with Heidi all the night
-her mother his upset (she's the daughter of Salève's)
-don't know very well the victim, BUT KNOWS THAT HANS WAS CHASING AFTER THE VICTIM AND THAT HEIDI WAS UPSET ABOUT THAT
-Profession: student
-Age: 20   no gun

Room 7:
-at the bar with Hans
-then at 7.45 went out with Lucie
-then ate at together at Pizzeria and came late in the night
-thinks the victim was a snobish buisnesswomen
-profession: chambermaid   no gun
-age: 58 37

Room 8:
-to bar from 7.30 till 9 with Heidi first, then with the Colonel
-go to bed after the bar
-profession: ski instructor   no gun
-age: 56 38
-was very close with the victim

Oscar:
- says he was in kitchen all night until 1010, but rolf said that he couldn't find him there, so IL MENT!

Room 5:
-at bar from 8:00 to 9:00 with Hans Wenger
-in the other case, in his room
-don't know the victim
-HE SAWS THE VICTIM HAVING SOME HEATED WORDS WITH LISA JONES
-HIS GUN WAS STOLEN WHEN HE WAS IN THE BAR
-Profession: retired Colonel   no gun
-Age: 68

Bar:
Painting:
- looks like early painting of Magritte

Jacques:
- spent night in bar
- Hans and Heidi had drink at 7.30
- Heida mad at Hans and left with sister
- many other people came but spent night with english girl
- at 10 his mother comes in with people from restaurant
- heard that victim was not nice, but doesn't know her personally, and thinks victim knows his father
- gun belongs to Colonel
- Colonel a bit aggressive
- thinks jacket is Hans'

Giuzeppe:
- went to resto about 6, after victim tried to make phone call, victim left at 9
- went to bar with loretan's when restaurant closed, and saw Hans looking suspicious in corridor
- doesn't like hans, thinks it is his jacket in room 4

Room 1:
Rolf:
- rolf + claire dinner with vesuvios last night
- at 8.30 Salèves couldn't be found (at that time, rolf left restaurant momentarily)
- when restaurant closed went to bar until Gieuzeppe discovered murder
- worked in same insurance company and office as victim, but she was promoted and changed office
- ask colonel about gun   no gun

Claire:
- when restaurant closed, went to room briefly, before going to bar
- doesn't like victim, she took the job her husband was supposed to get

Private residence:
- victim signed insurance contract to insure the   1000000 Frs.
painting for oscar saleve

Room 3:
- lisa spent night in bar
- didn't know victim well
- victim seemed disappointed when discovering that painting in bar was fake (bad attitude?)

Room 4:
- red swiss ski jacket
- old swiss army gun   no gun
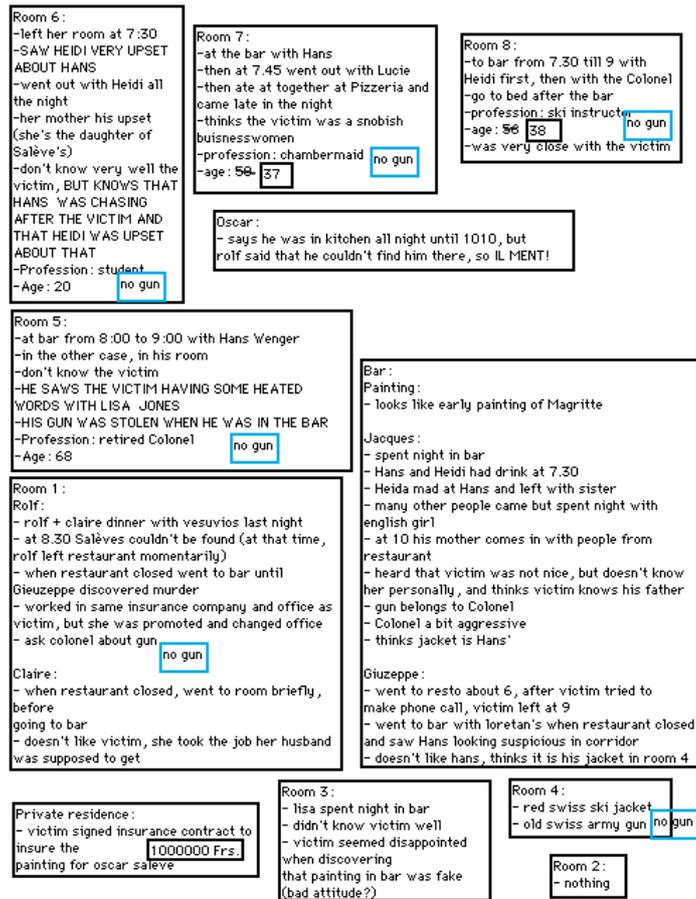
Room 2:
- nothing

Figure 2: Group working memory: Constructing shared and persistent representation of the problem state
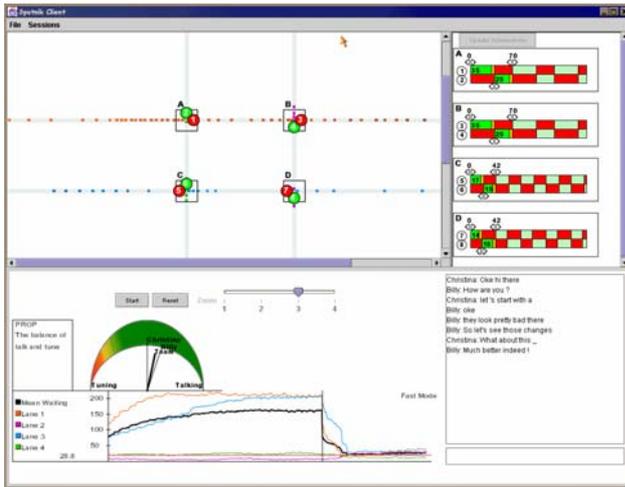
Figure 3: COTRAS (Collaborative Traffic Simulator). The group mirror is the red-green meter with 3 arrows, one for each user and one for the group. (Jermann, 2004)
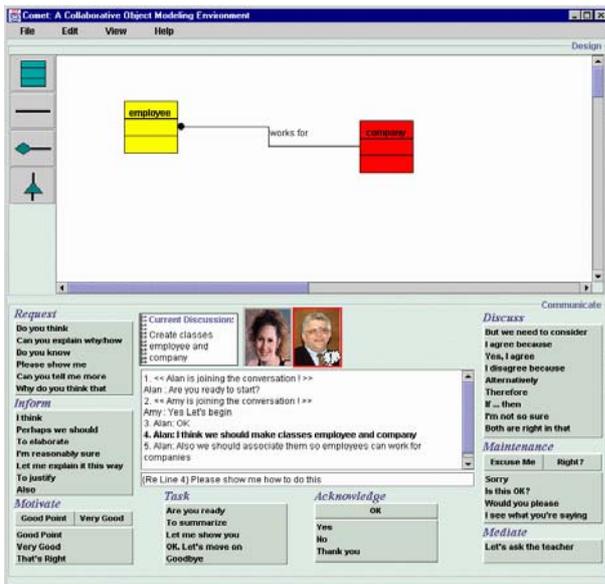

Figure 4: Example of semi-structured interface; buttons in the bottom part offer pre-defined communication acts and sentence openers (Soller, 2002)