M.J. BAKER, M. QUIGNARD, K. LUND, A. SÉJOURNÉ

COMPUTER-SUPPORTED COLLABORATIVE LEARNING IN THE SPACE OF DEBATE

CSCL Conference 2003

Abstract. Cutting across the common distinction between learning to argue and arguing to learn, this research is concerned with *arguing to learn argumentative knowledge*, or broadening and deepening understanding of the space of debate. A secondary school experiment compared broadening/deepening of understanding of the space of a debate on genetically modified organisms, using either a CHAT tool, or else an argument-diagram tool (DREW: Dialogical Reasoning Educational Web tool) that was designed as a medium for interactive debate. Although there was no significant difference between the quality of students' texts before and after debating across the conditions, a new interaction analysis method ("Rainbow") revealed differences in students' expression and elaboration of arguments.

1. INTRODUCTION

Research on argumentation-related learning is commonly based on a dichotomy between acquisition of the abstract and transdisciplinary skill of argument, or *learning to argue* (e.g. Voss & Means, 1991) and acquisition of specific or abstract knowledge in a particular domain, as a result of arguing, or *arguing to learn* (e.g. Andriessen & Coirier, 1999). The research presented here cuts across this divide: it is concerned with *arguing to learn argumentative knowledge*, or *broadening and deepening understanding of the space of debate.* Students broaden their understanding of a space of debate when they are better acquainted with societal and epistemological points of view, their associated arguments and value systems; they deepen it when they are able to go deeper into argument chains, to elaborate upon the meaning of arguments, and to better understand the notions involved.

It is of course not easy to create (computer-supported) collaborative situations that favour such debates and associated learning. Important factors include the choice of a suitable topic, intersubjective /interpersonal differences in groups, and of course facilitating or inhibiting characteristics of CSCL interfaces (see e.g. Golder, 1996; Quignard & Baker, 1999). Since students' knowledge is, by hypothesis, undergoing co-construction in such situations, we should expect a *cooperative exploration of a dialogical space of debate* (Nonnon, 1996; Baker, 1999) rather than an adversarial confrontation of established viewpoints.

We describe new Internet tools and pedagogical situations that were developed for favouring students' understanding of the space of debate, within the European Union "SCALE" project (Internet Support for Collaborative Argumentation-Based LEarning; IST-1999-10664, http://www.euroscale.net). An experiment carried out at secondary school level compared students' debates about Genetically Modified Organisms (henceforth GMOs), using either 1) a CHAT tool, or else 2) the

"DREW" (Dialogical Reasoning Educational Web tool) argument-graph tool, together with the CHAT tool (required for non-argumentative interaction management). The DREW argument-graph tool was specifically designed as a *medium* for educational debate, rather than as a focus for it (c.f. de Vries, Lund & Baker, 2002). Here we focus on methods for analysing the quality of students' understanding of the space of debate, as manifested in their texts and computer-mediated interactions.

2. DESIGN OF A CSCL SITUATION FOR ARGUING TO LEARN ARGUMENTATIVE KNOWLEDGE

Designing an educationally-relevant CSCL situation for arguing to learn argumentative knowledge requires satisfying many constraints, for example: 1) the topic debated must be integrated into the curriculum; 2) the students must be trained on the tools; 3) appropriate teaching materials need to be designed (one does not debate 'from scratch'); 4) the sequence of tasks must enable the target knowledge to be elaborated; 5) the task sequence must be feasible within the school timetable, and, of course, 6) CSCL tools must be designed to satisfy these other constraints.

2.1. Teaching materials and task sequence

In order to choose a topic for debate that was integrated into the French national curricula, we reviewed official programmes (see http://www.education.gouv.fr/sec/), and on this basis chose *GMOs.* This is dealt with as *a scientific question* in the Life and Earth Sciences curriculum, and as *a societal question* in the Civic Education curriculum.

In collaboration with French and Economics teachers, we developed teaching materials (texts) on GMOs that were practically readable within school timetable constraints, and that presented a wide and balanced set of arguments and points of view (breadth of the space of debate), together with information about key concepts such as "gene" (depth of space of debate). Several primary sources (notably websites) were used, that corresponded to clearly identifiable 'voices' (Bakhtine, 1929/1977) of *social actors* implicated in the question of GMOs (e.g. "Limagrain", a major grain producer, the French Research Ministry, and Greenpeace). Within each 'voice', several *epistemological* points of view are represented (e.g. scientific, agronomic, economic and ethical). Societal and epistemological points of view are distinct since, for example, each social actor presents the 'facts' differently.

The task sequence shown in Figure 1 aims to prepare students for debating, in terms of acquiring appropriate knowledge of the domain, the tools to be used, argumentation itself, and consolidation of knowledge co-constructed in the debate.

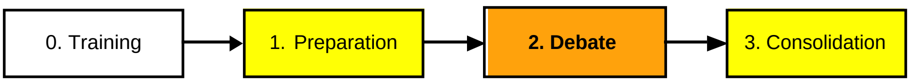

Figure 1. Generic task sequence for Collaborative Argumentation-Based Learning

The training phase (0), of 2 hours' duration, comprised a short introduction to argumentation notions and techniques that would be necessary during the debate phase (2), including use of Toulmin-like diagrams (Toulmin, 1958) to represent theses, pro and contra arguments. In addition, students were trained on the DREW interface tools. During the preparation phase (1), students were given the teaching materials on GMOs to read (during their own time), together with a table for noting arguments, structured according to social actors and epistemological points of view (see above). At the beginning of the part of phase 1 that took place in class, after re-familiarising themselves with the dossier on GMOs, the students were asked individually to write a short text presenting their own opinions, and associated arguments, on the question: *"Should the production of GMOs be allowed or not?"*. The pedagogical rationale of this sub-task was rendering explicit, reflecting upon and restructuring argumentative knowledge in preparation for debating. The ensuing debate phase was carried out either using a CHAT interface, or else using the CHAT in combination with the DREW argument graph (see below). In both cases, students were asked to each express their opinions and provide arguments for them, then to explore and deepen the question together in order to subsequently enrich their individual texts. Within the last 10 minutes of the debate, the students were asked to sum up their points of agreement and disagreement. The pedagogical rationale of this phase was that by interacting together, students would deepen and broaden their understanding of the space of debate, by various means: acquisition of new arguments from their partners, refinement of their own understanding by expressing arguments and by understanding criticisms of them, negotiating refined meanings of key concepts (such as the notion of genetic modification).

In the final phase of the task sequence (3), students return to individual work, and are asked to improve their individual texts, *in the light of the discussion* that had just taken place. This task was intended to help students to integrate the knowledge they had acquired during the debate, and as a result of it.

2.2. The DREW argument-grapher: a tool for graphical argumentation dialogue

The principal interface of the DREW CSCL environment is shown in Figure 2. The CHAT window, with the trace of the interaction is on the left and the argument graph window on the right. In this first version of the software, our aim was to produce a graph that is as simple as possible: boxes for arguments/theses, and only two types of argumentative links ("+" and "–", or else undefined: "?"), whose interpretation is left to the students and their teacher in a given session (Quignard, *in press*). A more important feature concerns the fact that the students are able to express their opinions — "in favour" and "against" — for any element of the argument graph (each person's opinion appears in a different colour). In order to highlight differences of opinion, and to focus discussion upon them, boxes with respect to which opposed opinions have been expressed appear in a 'crushed' form. In this respect, the DREW argument graph tool differs from several others, such as the argument graphs in Suthers' "Belvédère" system (e.g. Suthers & Hundhausen, 2001), since it is intended to be more a *medium* through which argumentation *dialogue* can occur, than as a third-party 'object' to be commonly constructed.

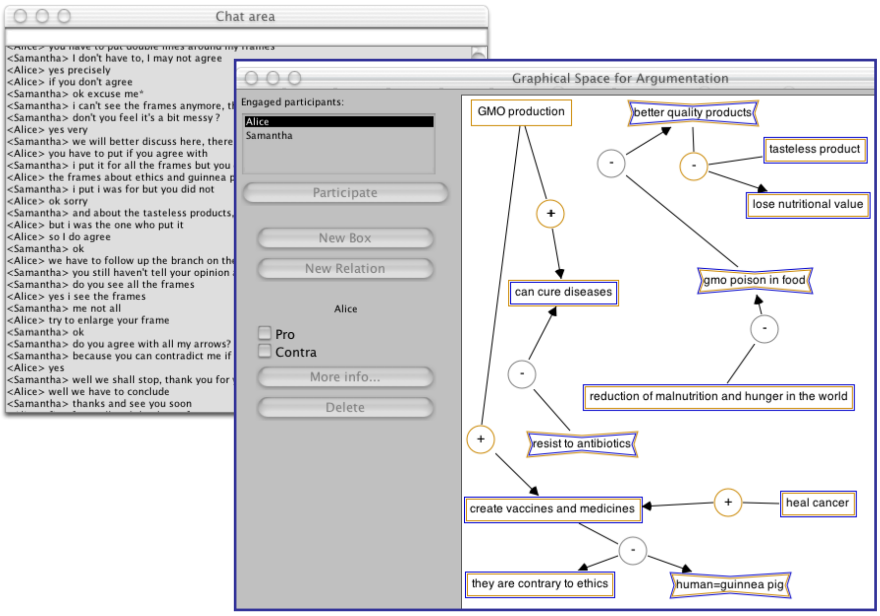

Figure 2. CHAT and Argument Grapher of the DREW interface.

In initial versions of Belvedere, the argument diagrams contained many different types of nodes and links (e.g. *Principle*, *Theory*, *Hypothesis*, *Claim* for nodes, and *Supports*, *Explains*, *Predicts*, *Conflicts*, *Justifies*, *Undercuts*, *Causes*, for links). However, it was found that students spent most of their type arguing about the meaning of these elements, rather than reasoning in the scientific domain itself, so — as with DREW — the diagrams were simplified in later versions.

3. AN EXPERIMENT: MULTI-REPRESENTATIONAL COLLABORATIVE LEARNING IN THE SPACE OF DEBATE

In November 2001 we carried out an experiment in a secondary school in Lyon, using the teaching materials and task sequence described above, together with the DREW CSCL environment. Firstly, we aimed to determine the extent to which the teaching materials and task sequence would in fact enable students to deepen and broaden their understanding of the space of debate. Secondly, we wanted to determine the extent to which such understanding would be influenced by the use of an argument-graph drawing tool, in comparison with CHAT interactions.

In the experiment, phase 0 lasted for one session of 2 hours; phases 1 to 3 together lasted a second session of 3 hours. The CHAT condition involved 21 students from a single class and the "CHAT+GRAPH" condition, 28 students from a different class. In each case, the students were randomly grouped into dyads (and one triad in the CHAT condition), since we wanted to eliminate the possible effect of dyad constitution (c.f. Quignard & Baker, 1999).

With respect to our second objective, we hypothesised that students using the argument graph would acquire deeper and broader understanding of the space of debate than students using CHAT alone. Although verbal interaction (CHAT condition) is an effective means of negotiating meaning, due to its intrinsic or strategic indeterminacy (Edmondson, 1981), we hypothesised that this effect would be outweighed by the fact that diagrammatic representations are more determinate, and thus more memorable (e.g. Ainsworth, Bibby & Wood, 1999). The data collected from the experiment consisted of students' individual texts, produced before the debate, and then revised after it, together with automatic traces of the interactions themselves.

We measured the contribution of the students' discussions (CHAT or CHAT with the graph tool) to improved understanding of the space of debate by evaluating the differences between individuals' texts produced before and after the discussion, using a newly devised method called "QED" ("*Qualité de l'Espace du Débat*", or "Quality of the Space of Debate").

The first step of analysis involves segmenting the text into (counter-)arguments, with respect to a principal thesis (e.g. "GMOs should be allowed"), and identifying the student's general opinion with respect to that thesis (e.g. "against", "in favour", "neither for nor against"). Each segment must then be classified as a pro or a contra argument with respect to the thesis, then classified in terms of one of a list of epistemological points of view (e.g. "economic", "ethical", "agronomic", …), and finally, its degree of elaboration must be assessed. Thus classified, the text is evaluated according to a complex weighted sum of the following factors that correspond to a good, wide, elaborate and coherent space of debate: *richness* (number of arguments), degree of *elaboration*, *balance* (pro vs. contra arguments), *coverage* of different viewpoints, and *coherence* (between expressed opinion and expressed pro or contra arguments). In its present version, the QED method only takes argumentative content into account, to the detriment of text structure, for mainly practical reasons: 98 texts had to be analysed, and four other partners in the SCALE project also used the method for comparable experiments.

Results of analyses of students' texts using the QED method are shown in Figure 3. Our first question was: *do students' QED scores improve significantly, in both conditions*? A paired-samples t-test was done on QED scores before and after discussion (pretext and posttext). Results show a significant difference, $t(48) = -4.61$, $p < .001$, with a higher QED score after discussion ($M = 42.92$) than before ($M = 34.92$). This means that students' texts showed a higher quality of space of debate after discussion, irrespective of the experimental condition.

Our second question was: *do students perform better in the chat-only condition than students in the chat-graph condition*? The repeated measures show that there was in fact no effect of the interaction between condition and QED-scores, $F(1.47) = 0.25$, $p > .01$. This result indicates that the increase of QED-scores was the same for both students in the chat-condition and students in the chat-graph-condition.

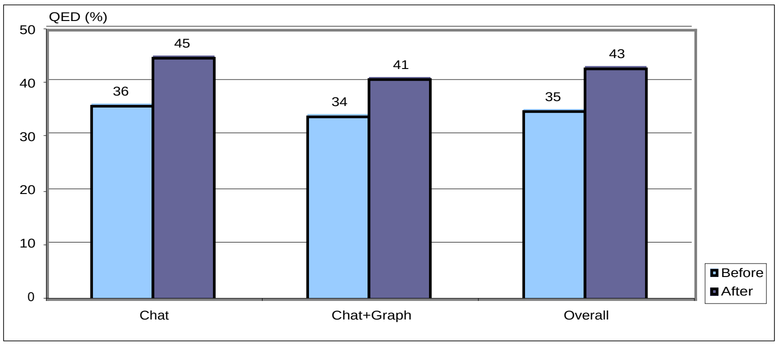

Figure 3. Graph of QED scores with different tools (Chat without and with Graph) before and after the interaction. The last column gives the overall results.

Our results thus show that students' knowledge of the space of debate for GMOs improved significantly during the experimental task sequence, but that the use of an argument graph for communication, as compared with a CHAT interface, made no significant difference with respect to this improvement.

4. ANALYSING THE SPACE OF DEBATE IN INTERACTIONS

It is possible that students' texts before and after debate do not provide sufficient indication of the degree of argumentative understanding that was elaborated in interaction. For this, we needed to look at interactions themselves. We analysed the students' interactions in both conditions using a method that was devised in the SCALE project, called "Rainbow". It comprises seven sets of main *functional* categories (Figure 4), each of which is ascribed a different colour (red, orange, yellow, green, blue, indigo, violet) to enable data visualisation.

Definition of each category and application to the corpus raises many methodological issues that we cannot deal with here. Clearly, category 7 "Explore and Deepen" (understanding of the space of debate), is at the heart of our pedagogical objectives. There are three main ways in which students can explore and deepen:

1) by expressing a (counter-)argument that builds on an already expressed (counter)argument, i.e. 'going deeper' in the argument tree;

2) by discussing meaning of argumentative relations, such as questioning or supporting that link (c.f. "backing" or "warrants" in Toulminian models);

3) by discussing the meaning of a notion underlying an argument, or by performing a discursive operations on an argument (see below).

We also carried out a parallel analysis of (sub-)topics debated for functional categories 6 and 7, in order to understand which of them were deepened/broadened.

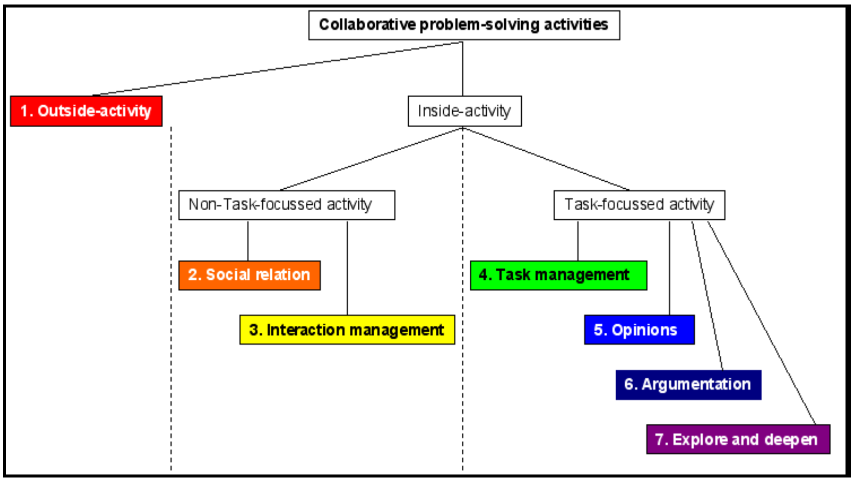

Figure 4. Principal categories of the rainbow functional analysis

Table 1 shows an analysed extract from a CHAT interaction. Prior to this extract, “Christine” (students’ names are changed, preserving gender) had argued generally in favour of GMOs, and Anne had countered her arguments.

Table 1. Extract from a CHAT interaction (translated/transliterated from French)

N	*T(h:m:s)*	*Loc*	*Interaction*	*Rainbow analysis*
46	09:44:03	Christine	but tell me why you're against then explain why to me?	5. Opinions
47	09:44:26	Anne	because it's bad for the human organism	6. Argumentation
48	09:44:55	Christine	reply to me	3. Interaction management
49	09:45:11	Anne	and then if we start with plants in 10 years or less it will be humans turn	6. Argumentation
50	09:45:38	Christine	to be modified?	7. Explore/deepen
51	09:46:02	Anne	er yes perhaps we'll even be cloned	7. Explore/deepen
52	09:46:19	Christine	yes it's true me you know I'm dead against cloning anybody	5. Opinions
53	09:46:33	Anne	errmm yes me too	5. Opinions
54	09:48:07	Christine	why are you against GMOs? isn't there the least positive argument for you?	5. Opinions
55	09:48:33	Anne	ermm perhaps but nothing's proved	6. Argumentation
56	09:48:46	Anne	for vaccinations nothing's proved	7. Explore/deepen
57	09:50:08	Christine	it's clear that they're only hypotheses for the moment	7. Explore/deepen

From the point of view of category 7, two *discursive operations* are interesting here. Firstly, Anne makes a *conceptual association* (c.f. Baker, 2002) between genetic modification and cloning of human beings (lines 49-52). Secondly, Anne performs an *epistemological operation*, relativising scientific claims in favour of GMOs (lines 55 and 57), and evoking a societal "principle of precaution".

Basic results of comparative analysis of average frequencies of Rainbow categories, of CHAT and CHAT-GRAPH interactions, are shown in Figure 5 (average frequencies of Rainbow categories in the CHAT and CHAT-GRAPH conditions).

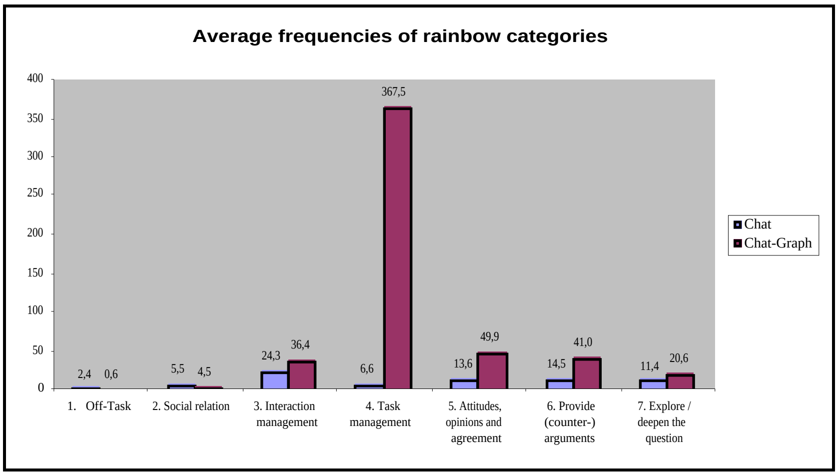

Figure 5. Comparative analysis of CHAT and CHAT-GRAPH interactions, using Rainbow.

The analysis was carried out jointly by the authors. For all categories, the CHAT-GRAPH interactions contain more analysed segments for comparable durations of tasks (17 minutes on average in the CHAT condition, 26 minutes on average for the CHAT-GRAPH condition). On average, students produce twice as many segments in argumentation and explore/deepen categories (6 and 7) in the CHAT-GRAPH condition (Freq. = 61.6), in comparison with CHAT alone (Freq.= 25.9). Category 4, Task management, is very high for the CHAT-GRAPH condition, principally because of the prevalence of "move-box" operations, that involve repositioning argument boxes on the interface. A large proportion of category 7 for the CHAT-GRAPH condition is actually carried out on the CHAT interface. Further analyses revealed that the average frequency of category 7 is similar during use of CHAT alone and during use of a CHAT with the GRAPH. However, the task sequence and the CHAT-GRAPH tools favour argumentation (cat. 6: Freq. = 36.6 with GRAPH, 14.5 with CHAT alone) and expression of opinions (cat. 5: Freq.= 41.7 with GRAPH and 13.6 with CHAT alone). In addition, the CHAT in the CHAT-GRAPH condition also plays its role in interaction management.

5. DISCUSSION

Why was there no significant difference between improvement of QED scores across the two conditions, despite their very different types of interactions? Although our hypotheses were not confirmed by QED analyses, they received support from interaction analysis.

1) Several problems with the argument graph interface emerged: students spent much time re-arranging the diagram in a relatively restricted screen space, and coordination problems occurred (one student could continue to edit an argument that had already been deleted by his partner). The former has now been facilitated by adding scrollable windows and the latter has been corrected

2) Differences in transitions between types of representations (diagrammatic and textual) from debate to consolidation, across conditions could have influenced the extent to which students were able to integrate knowledge elaborated in the interaction into their individual texts. In the CHAT condition, the transition was "text → text", and in the CHAT-GRAPH "diagram → text".

3) There might have been significant differences had textual structure been taken into account in QED scores. Comparing interactions is also difficult: is typing an argument really comparable to creating, naming and linking a box?

4) The students almost never elaborated the content of arguments, but rather shifted to CHAT. This problem could be addressed by changing task instructions ("Elaborate your arguments!").

5) The sharing out of activity across GRAPH and CHAT tools is significant, particularly with respect to the explore/deepen category.

6. CONCLUSION

Our aim was to understand how to design situations — teaching materials, task sequences and Internet tools, viewed as an integrated whole — for a specific form of collaborative argumentation-based learning related to broadening and deepening the space of debate. Although we have obtained mixed results with respect to the contribution of an argument graph tool (DREW), our results show at least that it is in fact possible to create situations in which students will broaden and deepen their understanding with respect to subjects taught in school, in argumentative interactions across Internet, provided that teaching materials, tasks and tools are appropriately designed. Apart from improving interface ergonomics, future research will concentrate on determining which of the different uses of interactive argumentation diagrams are most effective in such situations: are such diagrams better used as a *focus* for discussion, as *tools for analysing* discussions or, as we have attempted to determine, as *media* for discussion (or a complex combination of all three)?

7. ACKNOWLEDGEMENTS

This article is dedicated to the memory of Anne-Marie Chevalier, teacher, colleague and friend, in whose class we carried out the experiment reported here.

Michael Baker, Kristine Lund, Arnauld Séjourné: Groupe de Recherches sur les Interactions Communicatives, C.N.R.S. & Université Lumière Lyon 2 (France) — {Michael.Baker, Kristine.Lund, Arnauld.Sejourne}@univ-lyon2.fr
Matthieu Quignard: C.N.R.S. & LORIA, Laboratoire lorrain de Recherche en Informatique et ses Applications, Nancy (France) — Matthieu.Quignard@loria.fr

8. REFERENCES

Ainsworth, S.E., Bibby, P.A., & Wood, D.J. (1999). Analysing the costs and benefits of multi-representational learning environments. In M.W. van Someren, P. Reimann, H.P.A. Boshuizen, & T. de Jong (Eds.), *Learning with multiple representations* (pp. 120-134). Oxford: Elsevier Science.

Andriessen, J. & Coirier, P. (Eds.) (1999). *Foundations of Argumentative Text Processing*. Amsterdam: University of Amsterdam Press.

Baker, M.J. (1999). Argumentation and Constructive Interaction. In J. Andriessen & P. Coirier (Eds.), *Foundations of Argumentative Text Processing* (pp. 179–202). Amsterdam: University of Amsterdam Press.

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

Bakhtine, M. (1929/1977). *Le Marxisme et la Philosophie du Langage*. Paris: Minuit. [Ist edition: Voloshinov, V.N., Leningrad, 1929].

de Vries, E., Lund, K. & Baker, M.J. (2002). Computer-mediated epistemic dialogue: Explanation and argumentation as vehicles for understanding scientific notions. *The Journal of the Learning Sciences*, 11(1), 63-103.

Edmondson, W. (1981). *Spoken Discourse: A model for analysis.* London: Longman.

Golder, C. (1996). *Le développement des discours argumentatifs.* Lausanne: Delachaux & Niestlé.

Nonnon, E. (1996). Activités argumentatives et élaboration de connaissances nouvelles: le dialogue comme espace d'exploration. *Langue Française,* 112, 67-87.

Quignard, M. & Baker, M.J. (1999). Favouring modellable computer-mediated argumentative dialogue in collaborative problem-solving situations. In S.P. Lajoie & M. Vivet (Eds.) *Artificial Intelligence in Education* (pp. 129-136). Amsterdam: IOS Press.

Quignard, M. (in press). A Collaborative Model of Argumentation in Dyadic Problem-Solving Interactions. In F. van Eemeren et al. (Eds), *Proceedings of the Fifth International Conference of the International Society for the Study of Argumentation* (ISSA'02). Amsterdam: Sic Sat.

Suthers, D. & Hundhausen, C. (2001). Learning by Constructing Collaborative Representations: An Empirical Comparison of Three Alternatives. In P. Dillenbourg, A. Eurelings & K. Hakkarainen (Eds.), *European Perspectives on Computer-Supported Collaborative Learning* (pp. 577-584). Maastricht: Universiteit Maastricht Press.

Toulmin, S.E. (1958). *The Uses of Argument.* Cambridge: Cambridge University Press.

Voss, J.F. & Means, M.L. (1991). Learning to reason via instruction in argumentation. *Learning and Instruction,* 1, 337-350.