Co-construction of knowledge in computer supported collaborative argumentation (CSCA)

Gellof Kanselaar, Jerry Andriessen, Gijsbert Erkens, Jos Jaspers, Maaike Prangsma
Department of Educational Sciences, Universiteit Utrecht
Arja Veerman, TNO Human Factors, Soesterberg

We have been experimenting with electronic conferencing (CMC), Computer Supported Collaborative Learning (CSCL) and Computer Supported Collaborative Argumentation (CSCA) at the Department of Educational Sciences of Utrecht University for a period of nearly 10 years now. We will first present a couple of studies with students in our own department and then with some studies with students in upper secondary schools in the Netherlands. We will end with a discussion about the use of computer tools to improve discussion and argumentation in education.

Introduction

In academic and in secondary education, there is wide interest in using Internet and electronic communication applications for educational purposes. Such applications offer not only advantages of time and/or place, but also of flexibility of information exchange and options for electronic communication. Information can be easily stored, presented and accessed in multiple formats (e.g. text, graphics). Communication within communities of education (students, teachers, tutors etc.) can be facilitated by the use of computer-mediated communication (CMC) systems (e.g. discussion forums, chat box, e-mail, newsgroups). The question we address here, is how those new communication tools can provoke discussion and argumentation to improve learning in education.

Electronic discussions to support learning in higher education[1]

In academic education, students have to deal with abstract, ill-defined and difficult to understand knowledge as well as with open-ended problems. Collaborative learning is one of the pedagogical methods that can stimulate students to discuss such information and problems from different perspectives, and to elaborate and refine these in order to re- and co-construct (new) knowledge or to solve the problems. In such situations, argumentation is considered to be one of the main mechanisms that can promote collaborative learning. (Veerman, Andriessen, & Kanselaar, 2000.)

However, little is known about the effective use of educational technology to support collaborative learning in academic education, particularly considering the role of argumentation. (Kanselaar, & Erkens, 1996; Kanselaar, de Jong, Andriessen, & Goodyear, 2000).

In the present chapter, collaborative learning through argumentation in computer supported collaborative learning (CSCL) environments is examined. The purpose is twofold: to increase knowledge about the effective use of educational technology to support collaborative learning in education and to contribute to a better understanding of the role of argumentation as a mechanism for collaborative learning itself.

The research is framed by socio-constructivist learning theory, the nature of academic education and current technology. The following general research questions are addressed:

1. How can collaborative learning situations be arranged that provoke and support students' argumentation, examining contextual aspects affecting argumentation (the role of student, peer student, tutor, task, instruction and medium)?

2. How can student groups' argumentative discussions be characterized in relation to collaborative learning processes in CMC systems?
3. How can students' computer supported collaborative learning be enhanced by providing pedagogical support or electronic facilitation at the user-interface?

To study the first question, the research started with a search for principles that provoke argumentation in collaborative learning situations. Effects of important contextual aspects were specifically looked at, such as the role of the student, peer, tutor, task, instruction and medium. To study the second and third research question, process analyses were applied to assess student groups' argumentative discussions in relationship to collaborative learning in different CMC systems, with and without various forms of pedagogical support.

Study 1: Tutoring sessions

The study on collaborative argumentation in F2F tutoring sessions was implemented as an open-ended problem-solving task as part of an 8-week course on developing CBL programs. After an introduction to the course and a presentation of some principles for instructional design, 23 students organized themselves in 11 units of one, two or three students each (2*1 student; 6*2 student; 3*3 students). They all reached a comparable level in Educational Sciences and had to work together during the whole course. Their first group assignment was to construct learning goals for a CBL program they had to design. To this end they were instructed to use concept map techniques (see Figure 1) and to construct a plan on paper (size A1/ A2), in which their learning goals were described, interrelated, organized and justified. In the second week of the course, this plan was subject to a one-hour tutoring session, in which a tutor evaluated the feasibility of the learning goals by discussing the students' assumptions. Neither the students nor the tutor were instructed on what to say or how to act during the tutoring sessions. We expected argumentation to occur, provoked by the tutor questioning and critiquing the students' work.

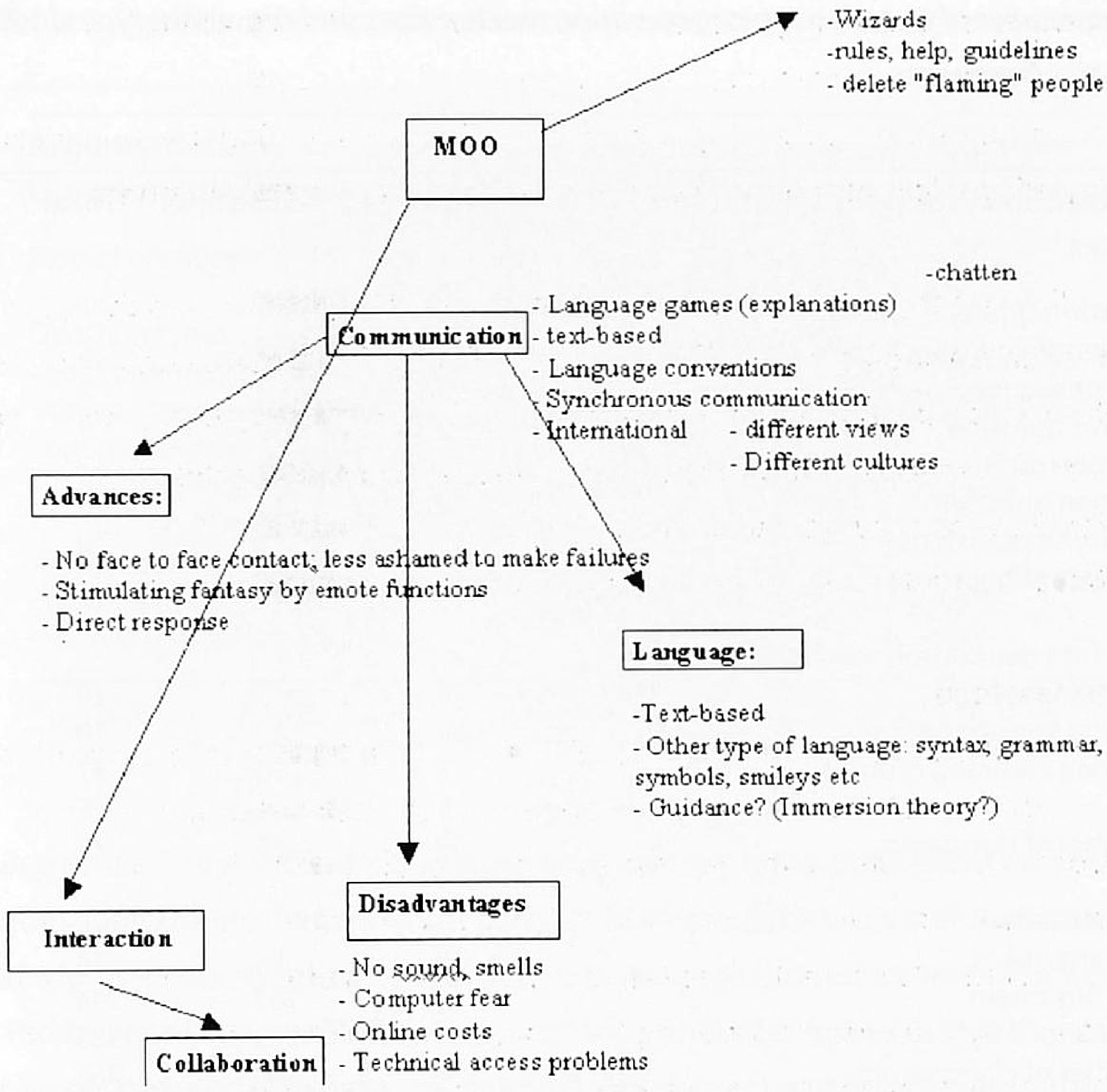

Figure 1. Example of a concept map, subject to one of the tutoring sessions.

The data analysis revealed that the tutoring sessions hardly contained argumentative discussion. From the 11 one-hour tutoring sessions analysed, only 25 argumentative fragments (a few minutes each) could be gathered in which in sum 72 questions were asked (mean .3 questions per fragment) and 767 elaborations were stated (mean .31 arguments per fragment). In Table 1, questions and arguments are split into different categories for analyses. Considering multiple perspective taking, students explored on average 2.1 different perspectives, the balance between positively and negatively oriented arguments was on average 0.4.

Table 1. Frequencies of question types, generation mechanisms and argument types in 25 argumentative fragments.

Categorisation system for separated utterances	Tutor	Students	Total
(A1) Question types			
1. Goal-oriented question	5	-	5
2. Cause-consequence question	2	-	2
3. Evaluative question	9	-	9
4. Other open questions	13	-	13
5. Verification question	12	11	23
6. Other closed questions	11	9	20
Total number of questions	52	20	72
(A2) Question generation mechanisms			
1. Inferring knowledge	20	-	20
2. Correcting knowledge	20	-	10
3. Monitoring common ground	7	16	37
4. Other...	5	4	5
Total number of questions	52	20	72
(B) Argumentation			
1. Neutral argument	20	18	38
2. Positive argument	161	183	344
3. Negative argument	203	182	385
Total number of elaborations	384	383	767

In Table 1 it is shown that the tutor, who aimed primarily at correcting and inferring knowledge, asked most questions (72%). Students mainly asked verification and other types of closed answer questions, aimed at reaching common ground. Considering argumentation, students and tutors contributed about the same number of neutrally, positively and negatively related arguments. The tutors contributed more positive arguments than negative ones; the students used both types of arguments comparably often.

Within the argumentative fragments (including both student(s) and the tutor), since all categories had a skewness and/or kurtosis > 1, the relationship between question asking and argumentation was measured by Spearman's nonparametric measure of correlation. In Table 2 it is shown that question asking in general and some specific question types and generation mechanisms were significantly related to the incidence of arguments, particularly to negatively oriented arguments. In addition to open questions, verification and cause-consequence types of questions, questions aimed at correcting knowledge and at monitoring common ground showed significant relationships.

Table 2. Significant relationships between question asking and argumentation.

Argumentation / Question asking	Elaboration (Σarguments)	Negative arguments
Sum of questions	0.40*	0.46*
Question types:		
- goal-oriented question	-	0.41*
- cause-consequence question	0.42*	0.49*
- open question	0.46*	-
- verification question	0.41*	-
Generation mechanisms:		
- correcting knowledge	0.40*	0.57*
- monitoring common ground	0.46*	0.43*

* $p < 0.05$

In this study, the students and tutor acted according to different roles. The tutor asked most questions, the students tried to find out what the tutor meant and what he wanted the students to do. The students proved to be strongly biased towards the tutor's evaluations and to adopt his plans without engaging in further discussion. This may be explained by the role of the tutor, which was by default to evaluate the students' plans. The students appeared to be not very well prepared to defend their point of view. They only made rough sketches for their conceptual design. It seemed they did not feel highly committed and/or self-confident to defend their work. Our assumption is that the dominant position of the tutor, combined with the insecurity of the students, did not provoke students to critically question ideas and to engage more often in argumentative discussion.

To overcome some of these problems, we decided to organize a more collaborative learning situation (F2F). First, we wanted the students to show a stronger attitude in defending positions. Therefore, we decided not to bring the tutor back on stage and, moreover, we provided the students with predefined conflicting claims and a competitive role (to win the discussion!). Second, students were provided with instruction on critical question asking (particularly asking open questions and verification questions aimed at correcting and inferring knowledge). We expected the students to be encouraged to resist critique, to discuss their positions and to engage in multiple-sided argumentation.

Study 2: Collaborative learning sessions

Based upon the results of the first study, we developed in this second study two short (open-ended) argumentative discussion tasks as part of an introductory course on Educational Technology for undergraduate students. For several years, this course has focused on the book 'Rethinking University Teaching' (Laurillard, 1993). The book is used as a theoretical framework to discuss affordances of media applications in higher education (hypertext, simulations, CMC, etc.). It centres on a discussible 'conversational framework' (Bostock, 1996) that describes crucial activities necessary to complete the learning process in teaching-learning dialogues.

In order to develop insight into the implications of this framework, we developed two 10-minute argumentative discussion tasks in the third week of the course. Students with a comparable background in Educational Sciences (n = 24) were required to prepare two chapters of 'Rethinking University Teaching' (Laurillard, 1993) at home about the role of (1) feedback and (2) tutoring strategies in tutoring sessions. Before the meeting, they had to take an individual knowledge test, in which they were tested on their knowledge about the concepts and activities mentioned in the conversational framework. During the meeting, they were randomly paired and instructed to competitively discuss a protocol of a tutoring dialogue. In two 10-minute sessions they were asked to defend controversial claims in relationship to the protocol and to win the discussion. The claims were:

- (Ad 1) "The tutor provides feedback that improves / does not encourage the student's learning process"
- (Ad 2) "The tutor's strategy improves / does not encourage the student's learning process"

In the study 12 student pairs engaged in the first discussion task without additional instruction on question asking. In the second task they switched partners randomly and were provided with additional instruction on question asking, as shown in Figure 2.

Instruction on question asking

- Ask each other questions
- Don't ask too many questions at once
- Ask questions, such as:
  - Is it true that...?
  - Why...?
  - Can you explain...?
  - Can you give an example...?
  - Can you give a characteristic...?
  - What can you infer on...?
  - What do you mean...?
  - Is that true...?
  - ...

Figure 2. Instruction on question asking.

We analyzed the data in the same way as in the formerly described study on the F2F tutoring sessions. First of all, we gathered 24 collaborative learning sessions, which could all be considered as complete argumentative fragments. We gathered 13 sessions without instruction on question asking and 11 sessions with instruction. One of the student pairs accidentally did not receive instruction in the second session. Secondly, within these argumentative fragments we analyzed separated utterances on question types, question generation mechanisms and argument types. In addition, we measured argumentative fragments on multiple perspective taking, balanced argumentation, the total number of elaboration and the sum of questions asked (see Table 3).

The 24 argumentative fragments included in sum 1422 elaborations (mean . 58 per fragment) and 156 questions (mean .6 per fragment). Compared to the study on tutoring sessions, twice as many questions were asked and the sessions were twice as elaborated in 1/6 of the time available (10 minutes versus 1 hour). Students asked each other mainly verification questions, but also many open, goal-oriented and cause-consequence questions. Questions were mainly aimed at inferring knowledge and at monitoring common ground. Considering multiple perspective taking and balancing the argument, an independent T-test showed no differences between conditions (T(df = 22) = 0.99; p = 0.33 resp. T(df = 22) = 0.26; p = 0.80). Overall, students explored on average 3.5 different perspectives; the balance between positively and negatively oriented arguments was on average 0.5. This means that on average for every 5 positively oriented arguments, 10 negatively oriented arguments were stated (or the other way around). Some fragments included more positively oriented arguments whereas others contained mainly negatively oriented arguments. This explains why the total number of positive and negative arguments is comparable (672 versus 651) whereas the balance is far below 1.

Table 3. Frequencies, means (0), standard deviations (s.d.) and p-values considering question types, question generation mechanisms and argumentation in the instructional condition versus the control group.

	Instruction (n = 11)			**Control (n = 13)**			
	freq.	*M*	*SD*	*freq.*	*M*	*SD*	*p-value*
(A1) Question types							
1. Goal-oriented question*	12	1.1	1.0	10	0.8	1.2	0.28
2. Cause-consequence question*	13	1.2	1.0	7	0.5	1.4	0.04
3. Evaluative question*	1	0.1	0.3	0	0.0	0.0	0.73
4. Other open questions	23	2.1	1.8	18	1.4	1.3	0.39
5. Verification question	25	2.3	1.9	36	2.8	2.3	0.61
6. Other closed questions*	7	0.6	1.0	4	0.3	0.5	0.65
Total number of questions	81	7.4	3.5	75	5.8	4.6	0.41
(A2) Question generation mechanisms							
1. Inferring knowledge*	44	4.0	2.1	32	2.5	3.0	0.04
2. Correcting knowledge *	8	0.7	1.0	14	1.1	1.5	0.87
3. Monitoring common ground	28	2.6	2.3	29	2.2	2.0	0.82
4. Other...*	1	0.1	0.3	0	0.0	0.0	0.73
Total number of questions	81	7.4	3.5	75	5.8	4.6	0.41
(B) Argumentative fragments							
1. Neutral argument*	58	5.3	6.5	41	3.2	3.0	0.46
2. Positive argument	306	27.8	14.7	366	28.2	12.0	0.43
3. Negative argument	308	28.0	15.7	343	26.4	13.9	0.69
Total number of elaborations	672	61.1	30.8	750	57.7	25.7	0.36

* skewness > 1 and/or kurtosis > 1; categories are included in a nonparametric test (Mann-Whitney)

Against expectations, hardly any differences were found between the instructional conditions of question asking (see Table 3). Student groups that were not provided with instruction on question asking produced a comparable number of questions and arguments as the instructed student groups. In addition, they also produced comparable types of questions and arguments considering means and standard deviations. Using a nonparametric Mann-Whitney test for measuring differences between groups showed that instructed students only scored higher on asking cause-consequence questions and questions aimed at inferring knowledge. However, means are small and standard deviations are quite high (especially in the control group), therefore, we have to consider these findings as tentative results. Considering the lack of (strong) variation between conditions and our interest in the overall relationship between question asking and argumentation, we continued our analyses by measuring correlations between question asking and argumentation. Categories with a high skewness and/or kurtosis were included in Spearman's nonparametric correlation measurement. Significant results are shown in Table 4.

Table 4. Significant relationships between questions asked, the number of elaboration and types of arguments across the 24 collaborative learning sessions.

Argumentation / Question asking	Elaboration (Σarguments)	Negative arguments	Positive arguments
Number of questions	0.42*	0.44*	-
Question types:			
- goal-oriented question	-	-	0.44*
- verification question	0.43*	0.41*	-
Generation mechanisms:			
- monitoring common ground	0.42*	0.41*	-

In Table 4, comparable to Study 1 it is shown again that question asking is related to argumentation. In this study, goal-oriented questions, verification questions and questions aimed at monitoring common ground were related to different types of argumentation. Remarkably, most question generation mechanisms were aimed at inferring knowledge (49%) and another 14% was aimed at correcting knowledge. However, no relationships were found between these types of question generation mechanisms and (types of) argumentation.

We have interpreted the results as follows. Students established a strong motivation to engage in discussion. They focused on winning the argument that started from the predefined conflicting claims. Although the students proved to be able to ask every type of question without instruction on question asking (which suggests they suppressed that ability in the first study), argumentation was not related to the question generation mechanisms of correcting and inferring knowledge, only to monitoring common ground. As a result, it appeared that questions were asked and arguments were given just to keep the discussion going, without arriving at new insights or conclusions. The knowledge tests showed that students did not prepare themselves enough (only 36% of the students passed the test) and the impression was that students just followed the assignment to competitively engage in argumentative discussion, considering the conflicting claims and aiming at winning the discussion.

We concluded that affordances and drawbacks of competitive instruction and providing students with predefined conflicting claims are not yet clear. In order to stimulate students to put effort in critical argumentation, to care for their own learning and to increase commitment to the collaborative learning task, we decided to organize a third study in which the students predefined their own conflicting claims as part of their task preparation. In addition, students were asked to submit a joint product to the tutor as a result of their discussion. We designed the study in a CMC environment in which students could graphically represent and organize arguments generated by themselves and their peers. Instruction on competitive versus consensual behaviour was varied in order to assess effects on collaborative argumentation.

Study 3: Electronic collaborative learning sessions

In the third study we questioned how instruction on competitive versus consensual behaviour affects students' engagement in (a) collaborative argumentation and in (b) argumentative diagram construction. We integrated this study in a course on developing CBL programs (the same course as described in Study 1). After an introduction and the presentation of some principles for instructional design, the students grouped themselves into 7 pairs in order to construct learning goals for a program they had to design. After jointly defining these goals, students were asked to produce conflicting claims on three pedagogical aspects that they thought would meet their learning goals:

- what pedagogical strategies to use
- how to sequence learning activities
- what programming tool to use

The three self-defined conflicting claims then had to be discussed electronically by using the Belvédère system, a synchronous network-tool developed by the Learning Research and Development Center at the University of Pittsburgh (Suthers, et al., 1995). Among many other applications, Belvédère can be used for constructing argumentative diagrams online with individuals or groups of students of any size (see Figure 4). The working screen of the program displays private and shared windows. To communicate with a partner the student has a text-based chat box in which multi-line messages can be created and sent. Messages will then be displayed, linked with the writer's name, in the shared chat-history. Adding data into the diagram is constrained; students must use the predefined set of boxes ('hypothesis', 'data', 'unspecified') and links ('for', 'against', 'and'). These are shown in the menu bar in Figure 3.

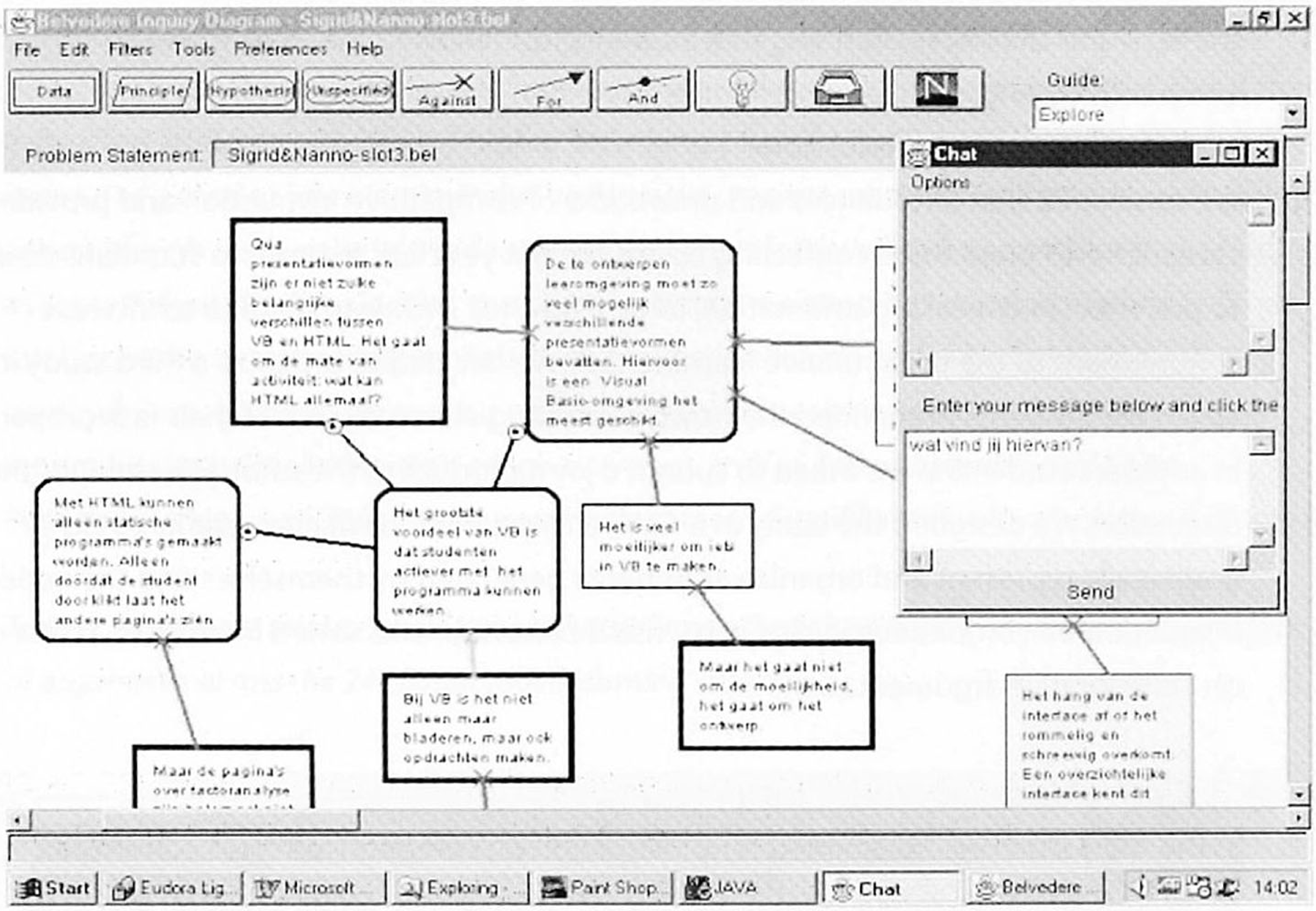

Figure 3. Interface of the Belvédère system.

Each Belvédère session took about one hour to complete. Besides basic instruction on the technical use of Belvédère, all students received guidelines on how to engage in argumentative discussions (see Figure 4). We figured that perceiving the discussible nature of the task would support students' mental preparation.

Guidelines to engage in collaborative argumentation

- Be critical in argumentation
- Use task-related arguments
- Detect feigned arguments (e.g. based on misinterpretation)
- Be co-operative (e.g do not use pressure, aggression)
- Use multiple perspectives in argumentation
- Elaborate on arguments
- Ask 'open' and verifiation question

Figure 4. Basic guidelines to engage in collaborative argumentation.

In addition, students were provided with instruction on competitive or consensual behaviour. In the competitive condition students were instructed to behave as competitively as possible, even when they were already convinced by the other party (considering the conflicting claims). They were told to end the argument only when getting stuck or not being able to think of any more (pro or contra) arguments. Then, a (sub) conclusion could be drawn and a next discussion could be provoked. In the consensual condition students were instructed to behave competitively considering the conflicting claims but to arrive at consensus as soon as being convinced by their peer student. We expected the students in the competitive mode of instruction to elaborate more and to take more perspectives into account than students in the consensual mode of instruction. Due to politeness strategies, the latter group was expected to agree on each other too soon, for example when arguments were still doubted or disbelieved.

We analyzed the data in the same way as the formerly described studies on F2F tutoring sessions and collaborative learning sessions. In analyzing the gathered dialogues and diagrams, it appeared that the third pedagogical aspect for discussion ('What programming tool to use') was not appropriate since students had too little experience in programming. As a consequence, we deleted the (partly) produced chats and diagrams produced on this issue.

Each of the 14 text-based chat discussions left in the study could be recognized as an argumentative fragment (7 student pairs * 2 Belvédère sessions). Within these fragments, we analyzed separated utterances on question types, question generation mechanisms and different types of arguments. The chats were subsequently measured on multiple perspective taking and balanced argumentation. In addition, the Belvédère diagrams were analyzed on the number of neutrally, positively and negatively oriented arguments in relationship towards the self-defined claims, and on multiple perspective taking and balanced argumentation.

The results showed, first of all, that during the Belvédère chat discussions 80 questions were asked and 410 arguments were stated. Students mainly asked verification questions and open answer questions. They aimed at inferring knowledge, correcting knowledge and monitoring common ground. The students stated more positively oriented than negatively oriented arguments (see Table 5). Students explored about two different perspectives (for and against the self-defined claim) and competitively instructed students showed a trend to a higher balanced argumentative chat discussion than students aimed at reaching consensus (mean balance = 0.65 respectively 0.43). However, an additional and independent T-test showed that this difference was not significant ($T(df = 12) = 1.55$; $p = 0.15$).

Table 5. Frequencies, means (0), standard deviations (s.d.) and p-values considering question types, generation mechanisms and argumentation within the Belvédère chat conditions of competition versus consensus.

	Belvédère chat discussions						
	freq.	*M*	*SD*	*freq.*	*M*	*SD*	*p-value*
(A1) Question types							
1. Goal-oriented question*	4	0.6	0.8	10	1.4	1.4	0.26
2. Cause-consequence question*	2	0.3	0.5	6	0.9	1.9	0.90
3. Evaluative question*	2	0.3	0.8	3	0.4	0.5	0.54
4. Other open questions*	10	1.4	1.5	12	1.7	1.5	0.71
5. Verification question*	13	1.9	2.0	13	1.9	2.3	0.81
6. Other closed questions*	1	0.1	0.4	4	0.6	1.1	0.62
Total number of questions	32	4.6	2.8	48	6.9	4.9	0.30
(A2) Question generation mechanisms							
1. Inferring knowledge*	14	2.0	1.8	13	1.9	1.7	1.00
2. Correcting knowledge *	8	1.1	2.2	13	1.9	0.9	0.07
3. Monitoring common ground*	6	0.9	1.6	15	2.1	3.0	0.39
4. Other...*	4	0.6	1.1	7	1.0	1.4	0.62
Total number of questions	32	4.6	2.8	48	6.9	4.9	0.30
(B) Argumentative fragments							
1. Neutral argument*	11	1.6	3.3	30	4.3	2.9	0.54
2. Positive argument	117	16.7	14.1	102	14.6	6.9	0.72
3. Negative argument*	102	14.6	15.4	48	6.7	5.0	0.23
Total number of elaborations	230	32.9	28.4	180	25.7	9.5	0.54

* skewness > 1 and/or kurtosis > 1; categories are included in a nonparametric test (Mann-Whitney)

Considering question asking, hardly any significant differences were found. However, student groups that were instructed to aim at reaching consensus asked more questions (particularly more goal-oriented and cause-consequence questions) than competitively instructed students. In addition to aiming at inferring knowledge, they also aimed at correcting knowledge and monitoring common ground. Competitively instructed students aimed their questions mainly at inferring knowledge. However, they stated more arguments, negatively oriented arguments in particular.

Considering the lack of (much) variation between conditions and our interest in the overall relationship between question asking and argumentation, we continued our analyses by measuring correlations between question asking and argumentation. Categories with a high skewness and/or kurtosis were included in a Spearman's nonparametric correlation measurement. We found an overall and significant relationship between inferring knowledge and the number of negatively and positively oriented arguments triggered ($r = 0.58; p < 0.05$ resp. $r = 0.64; p < 0.05$). Other relationships found were between open question types and the sum of arguments stated ($r = 0.63$; $p < 0.05$) and between verification questions and the number of neutral arguments ($r = 0.60$; $p < 0.05$).

Looking at both the results triggered between conditions (Mann-Whitney test; T-test) and across conditions (correlation test) leaves us to think that aiming questions at correcting knowledge and monitoring common ground may be less important to provoke argumentation (particular negative types of arguments) than questions aiming at inferring knowledge.

The Belvédère diagrams contained in sum 181 arguments (see Table 6). Comparable to the chat discussions, in both conditions the diagrams included about two different perspectives (for and against the claim). We found that, first of all, competitively instructed student groups generated 30% more arguments than student groups aimed at reaching consensus (107 against 74 arguments). Using independent T-test measurements revealed that this could be due particularly to the production of negatively oriented arguments ($T(df = 12) = 2.93$; $p < 0.05$). An interesting finding, which is in line with the results concerning the chat discussions, is that competitively instructed student groups balanced their argumentative diagram significantly better than student groups aimed at reaching consensus (balance = 0.72 resp. 0.53; $T(df = 12) = 2.41$; $p < 0.05$).

Table 6. Argumentation in the Belvédère diagrams.

(B) Argumentation Belvédère diagrams	Consensus (n=7)			Competition (n=7)			
	freq.	*M*	*SD*	*freq.*	*M*	*SD*	*p-value*
1. Neutral argument	-	-	-	-	-	-	-
2. Positive argument	49	7.0	3.3	57	8.1	2.5	0.48
3. Negative argument	25	3.6	1.8	50	7.1	2.7	0.01
Total number of elaboration	74	10.6	4.5	107	15.3	4.5	0.08

To summarize, the study showed that students only engaged in argumentation when they were sufficiently prepared for the assignment. Their lack of knowledge about programming inhibited them to discuss the third pedagogical aspect. The argumentative task design appeared to be successful, encouraging students to engage in argumentation based upon their self-defined conflicting claims and the requirement to submit a jointly constructed product (the diagram). The combination of competitive instruction and the graphical overview showed by the Belvédère system may have supported the argumentative process to be multiple sided, balanced and elaborated on different positions. Students provided with competitive instruction produced better balanced argumentative chat discussions than students aimed at reaching consensus: they considered multiple perspectives and elaborated on both positive and negative sides of the argument. Shown as a trend in the chat discussions, these findings were significant in the diagrams: the constructed diagrams showed an even better balanced argumentation than the created chat discussions.

Discussion about collaborating students

Academic education can be viewed as an ongoing process of negotiation in which there is an important role for collaborative argumentation. In collaborative argumentation, multiple perspective taking and elaboration may stimulate learning through knowledge transforming.

The studies described in the preceding paragraphs explored the question of how to provoke collaborative argumentation in educational situations with university students, considering contextual features such as the role of the tutor, (peer) student, task, instruction and medium. We conducted three studies in sequence, two in F2F situations, with and without a tutor, and a third study by using Belvédère, an electronic CMC system for synchronous communication and graphical argument construction. Although the studies differed on many aspects such as task design and instruction, in all studies we collected argumentative fragments that were analyzed similarly on question asking and argumentation. This allows us to make some comparisons among the three studies.

We found most of our students were not well prepared. Given this situation, we looked for facilitating factors that engaged students in fruitful collaborative argumentation. The results of the studies indicate some principles on how to provoke collaborative argumentation in academic learning sessions. In producing self-defined (competitive) stances and being asked for a joint product, students can be encouraged to participate in collaborative argumentation. To support students' multiple perspective taking and elaboration, additional instruction can be given by providing students with basic guidelines for argumentation, competitive instruction on argumentative behaviour and tools to organize arguments graphically. Belvédère provides students with a shared overview of their arguments. In addition to text-based and permanent argumentation, showing arguments graphically in a 'persistent' window may stimulate students to balance positively and negatively oriented arguments and elaboration.

In general, question asking showed to be important in relationship to collaborative argumentation, specifically considering some types of open answer questions, verification questions and questions aimed at correcting and inferring knowledge and at monitoring common ground. It appeared that open-answer question types aimed more at triggering elaborations and negative arguments whereas verification questions seemed to be more important in order to establish common ground. We think the role of verification questions in particular and the importance of seeking common ground deserve more prominence in today's research about question asking and learning (e.g. King, 1990; Hakkarainen, 1995; Erkens, 1997). Furthermore, we revealed that students were perfectly capable of asking effective question types without additional instruction, as long as this ability was not suppressed by the presence of a tutor.

While the metaphor of education as argumentation points out the importance of focusing on the learner within the educational situation, it also asks us to consider the other party – the educator. Engaging the learner in collaborative argumentation gives the learner an opportunity to articulate his or her understanding and offers the educator the best chance of intervening (Petraglia, 1997). In our first study, however, we found our students to be strongly biased towards the tutor's beliefs and values, which inhibited their argumentation. In our second study no tutor was involved. We found our students to engage in argumentation easily, but not to ask or answer questions in connection to each other. The discussions, as a result of this, were rather superficial. In the third study, Belvédère supported students in discussing questions in relation to their answers or solutions. However, the program only provides students with an overview and is not able to correct students in for instance discussing misconceptions. At some point in the discussion, some sort of tutor support may still be useful.

To support collaborative argumentation for learning purposes, we first of all have to know more about the specific needs students have in different educational situations considering collaborative argumentation. Secondly, we have to think about how to adapt the tutor's role or other forms of (electronic) support in relationship to the role of entities already provided by the task environment (task features, instruction, electronic characteristics of CMC systems, etc.). The need for research results specifying how to provide effective support increases, especially for collaborative argumentation in electronic environments. Since educators discovered the Internet, CMC systems have increasingly been implemented for educational purposes. We think these systems include promising characteristics to support learning through collaborative argumentation and, therefore, this issue deserves further empirical study.

Interactive learning in secondary education[2]

Secondary school students in The Netherlands – as a result of recent changes in the curriculum of the final years (the 'studyhouse') – are doing increasingly independent research in preparation for college studies. The focus has shifted towards working actively, constructively and collaboratively, as this is believed to enhance learning. We have developed a groupware computer environment that supports collaborative writing that should fit well within this curriculum, as the Information and Communication Technology (ICT) involved can emphasize both the constructivist and collaborative aspects through its active and interactive nature. The purpose of our research is to investigate the effect of the computer supported writing environment and its tools on the final written product through differences in the participants' collaboration processes. The study discussed here deals with the influence of CMC (Computer Mediated Communication) tools on the argumentation process in writing an argumentative text.

Argumentation and Collaboration

One of the main principles of constructivist learning theory is the negotiated construction of knowledge through dialogue. Such learning through negotiation can consist of testing understanding and ideas against each other's as a mechanism for enriching, interweaving and expanding understanding of particular phenomena. Active engagement in collaborative argumentation during problem solving fits this principle by giving prominence to conflict and query as mechanisms for enriching, combining and expanding understanding of problems that have to be solved (Savery & Duffy, 1995). After all, as Von Glaserfeld (1989) has noted, other people are the greatest source of alternative views to challenge our current views and hence to serve as the source of cognitive conflict that stimulates learning.

Knowledge is actively constructed, connected to the individual's cognitive repertoire and to a broader, often team-based and interdisciplinary context in which learning activities take place (Salomon, 1997). Constructivism seems to be influenced not only by a Piagetian perspective on individual cognitive development through socio-cognitive conflict, but also by the socio-cultural approach emphasizing the process of interactive knowledge construction in which appropriation of meaning through negotiation plays a central role (Greeno, 1997). From a constructivistic perspective, collaborative argumentation during problem solving can be regarded as an activity encouraging learning through mechanisms such as externalizing knowledge and opinions, self-explanation, reflecting on each other's information and reconstructing knowledge through critical discussion.

We consider an argument to be a structured connection of claims, evidence and rebuttals. A minimal argument is a claim for which at least doubt or disbelief is expressed (van Eemeren, Grootendorst & Snoeck Henkemans, 1995). Such doubt or disbelief can be expressed by an individual (if working alone) or by a partner in an argumentative dialogue. In response to such doubts a complex structure may be produced potentially including features such as chaining of arguments, qualifications, contraindications, counter-arguments and rebuttals. Hence the argument is the product, the structure linking claims, the evidence or rebuttals. The process by which the argument is produced we refer to as argumentation.

Our interest lies in argumentation structures that are built by groups of students involved in collaborative problem solving and writing. During problem solving we expect students to make various claims about the domain and the potential solutions. It is possible that during the problem solving no doubt is expressed regarding claims and solutions and hence no argument emerges in the dialogue. However, such a situation seems unlikely and we believe would not produce the best solution to the problem. Certainly if the students have not produced reasons to support the claims and solutions during the problem solving process itself then we have no reason to believe that they will be able to produce such reasons at a later date. Therefore we believe that students should be encouraged to use argumentation processes to build argument structures during problem solving.

The TC3 environment

In the COSAR project (Erkens, Prangsma, Jaspers, & Kanselaar, 2002) we developed the groupware program TC3 (Text Composer, Computer supported & Collaborative) with which the students carry out the main writing task. This environment is based on an earlier tool called CTP – Collaborative Text Production (Andriessen, Erkens, Overeem, & Jaspers, 1996), and it combines a shared text editor, a chat facility, and private access to a notepad and to information sources to encourage collaborative distance writing. The participants worked in pairs within TC3, each partner working at his/her own computer, and wherever possible partners were seated separately in different classrooms. The main screen of the program displays several private and shared windows. The basic environment, shown in Figure 5, contains four main windows:

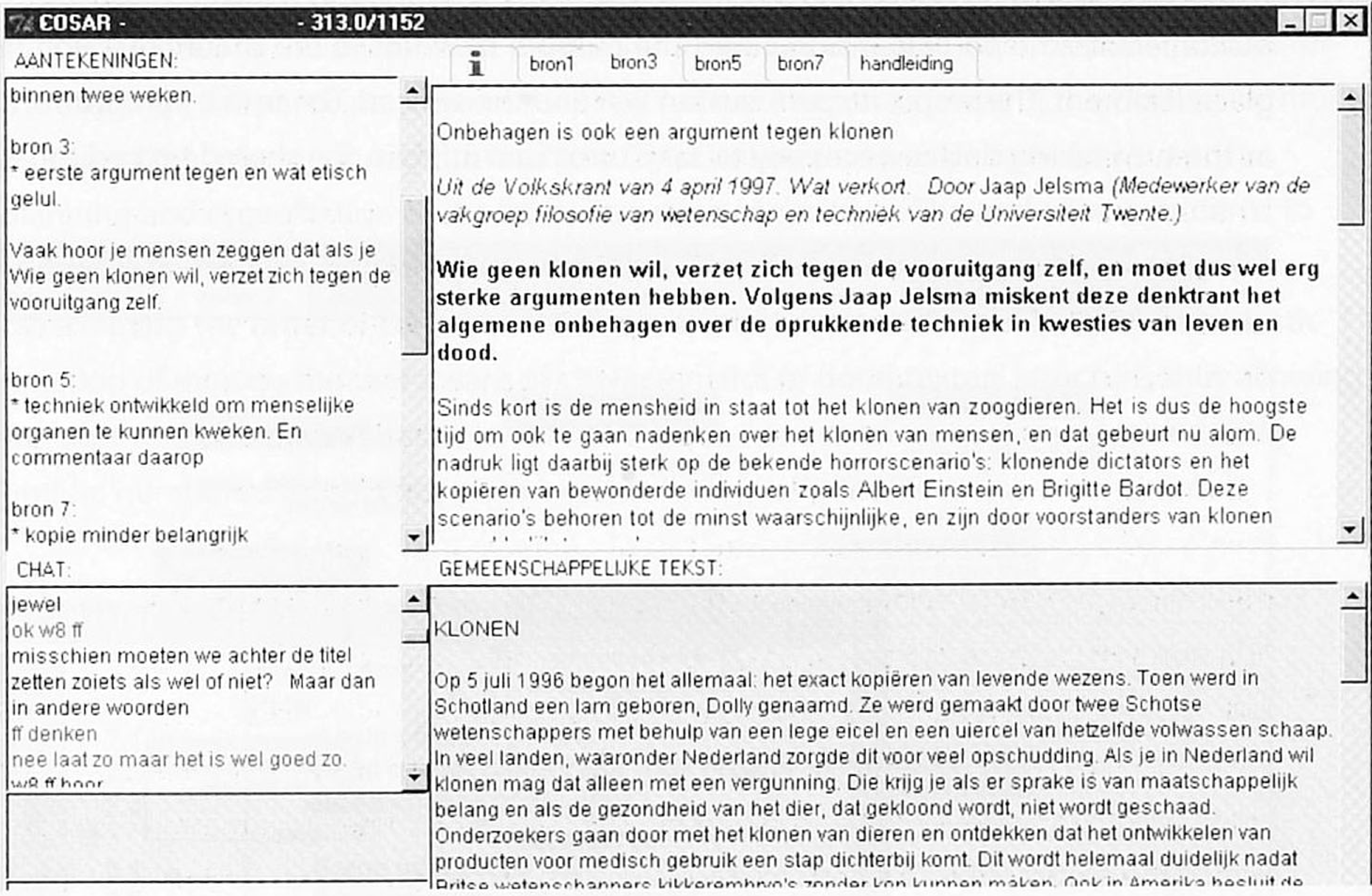

Figure 5 The layout of the interface of the TC3-basic environment

- The upper half of the screen is private and the lower half is shared.
- *Information* (upper right window): This private window contains tabs for the assignment, sources and TC3 operating instructions. Sources are divided evenly between the students. Each partner has 3 or 5 different sources plus one – fairly factual – common source. The content of the sources cannot be copied or pasted.
- *Notes* (upper left window): A private notepad where the student can make non-shared notes.
- *Chat* (lower left, 3 small windows): The student adds his/her chat message in the bottom box: every letter typed is immediately sent to the partner via the network, so

that both boxes are WYSIWIS: What You See Is What I See. The middle box shows the incoming messages from the partner. The scrollable upper chat box contains the discussion history.

- *Shared text* (lower right window): A simple text editor (also WYSIWIS) in which the shared text is written while taking turns.

Text from the private notes, chat, chat history and shared text can be exchanged through standard copy and paste functions. To allow the participants to focus more on private work or on the collaboration, three layout buttons were added in the left-hand corner: the middle layout button enlarges the private windows, the rightmost button enlarges the shared windows, and the leftmost layout button restores the basic layout. The buttons search; mark and delete (*zoek*, *markeer* and *wis*) can be used to mark and unmark text in the source windows and to search through the marked texts. The number of words (*aantal woorden*) button allows the participants to count the number of words in the shared text editor at any given moment. The stop (*stoppen*) button will end the session. The traffic light button serves as the turn taking device necessary to take turns in writing in the shared text editor.

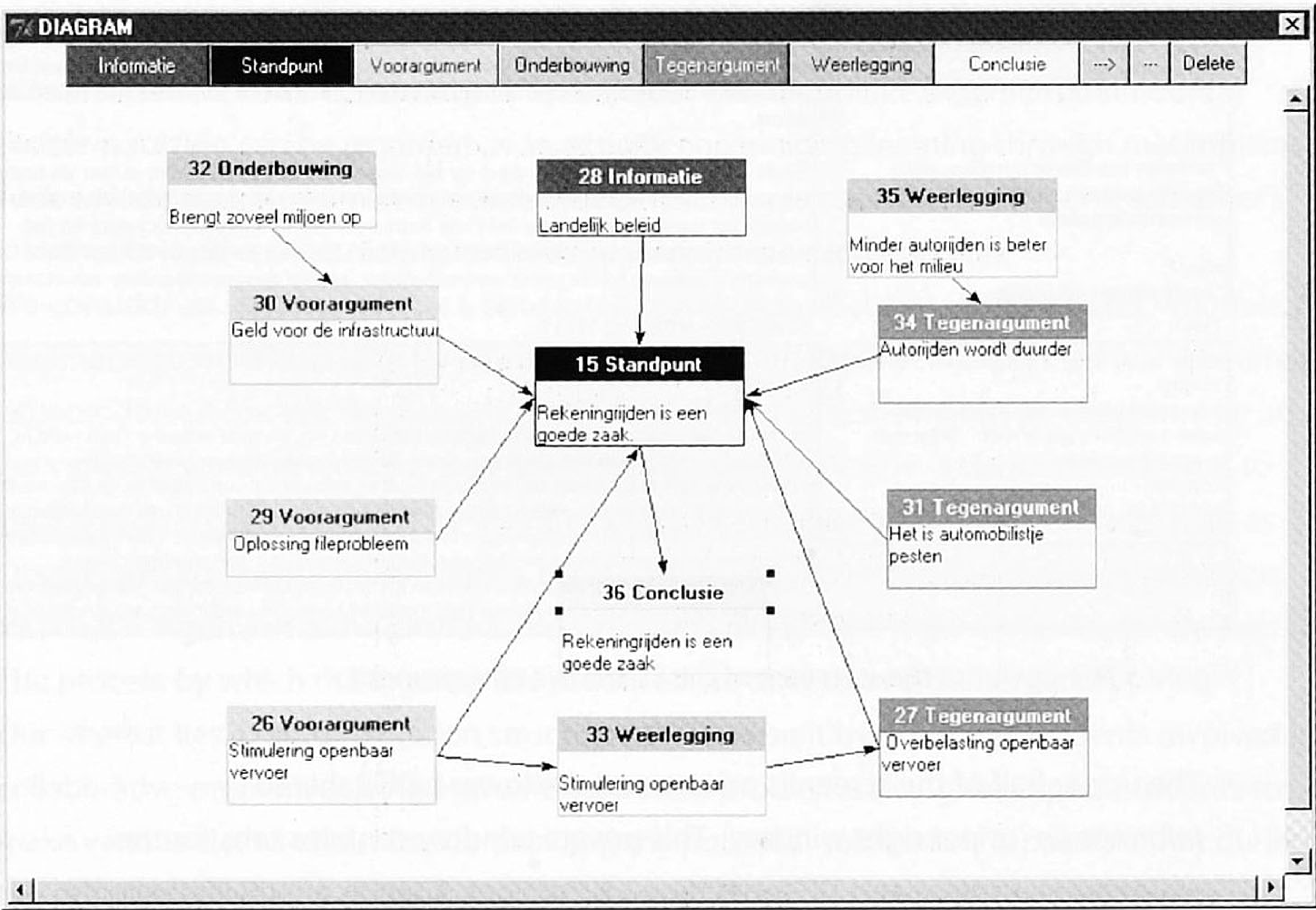

Figure 6 The Diagram window in the TC3 program

In addition, two planning modules were developed in the TC3 program for the experimental conditions: the Diagram and the Outline. The Diagram (see Figure 6) is a tool for generating, organizing and relating information units in a graphical knowledge structure comparable to Belvédère (Suthers, Weiner, Connelly, & Paolucci, 1995; Suthers, & Hundhausen, 2001). The tool was conceptualized to the students as a graphical summary of the information in the argumentative essay. Students were told that the information contained in the Diagram had to faithfully represent the information in the final version of their essay. We hoped that this requirement would help students to notice inconsistencies, gaps, and other imperfections in their texts, and encourage them to review and revise. In the Diagram, several types of text boxes can be used: information (Informatie), position (Standpunt), argument pro (Voorargument), support (Onderbouwing), argument contra (Tegenargument), refutation (Weerlegging), and conclusion (Conclusie). Two types of connectors were available to link the text boxes: arrows and lines. The Diagram can be used to visualize the argumentative structure of the position taken.

The Outline (see Figure 7) is a tool in the TC3 program for generating and organizing information units as an outline of consecutive subjects in the text. This tool was conceptualized to the students as producing a meaningful outline of the paper, and as for the Diagram, the participants were required to have the information in the Outline faithfully represent the information of the final text. The Outline tool was designed to support planning and organization of the linear structure of the texts. The tool allows students to make an overview or hierarchical structure of the text to be written. This should help in determining the order of content in the text. In addition, the Outline tool has the didactic function of making the user aware of characteristics of good textual structure, thus allowing the user to learn to write better texts. The Outline has a maximum of four automatically outline numbered levels. Both planning windows are WYSIWIS.

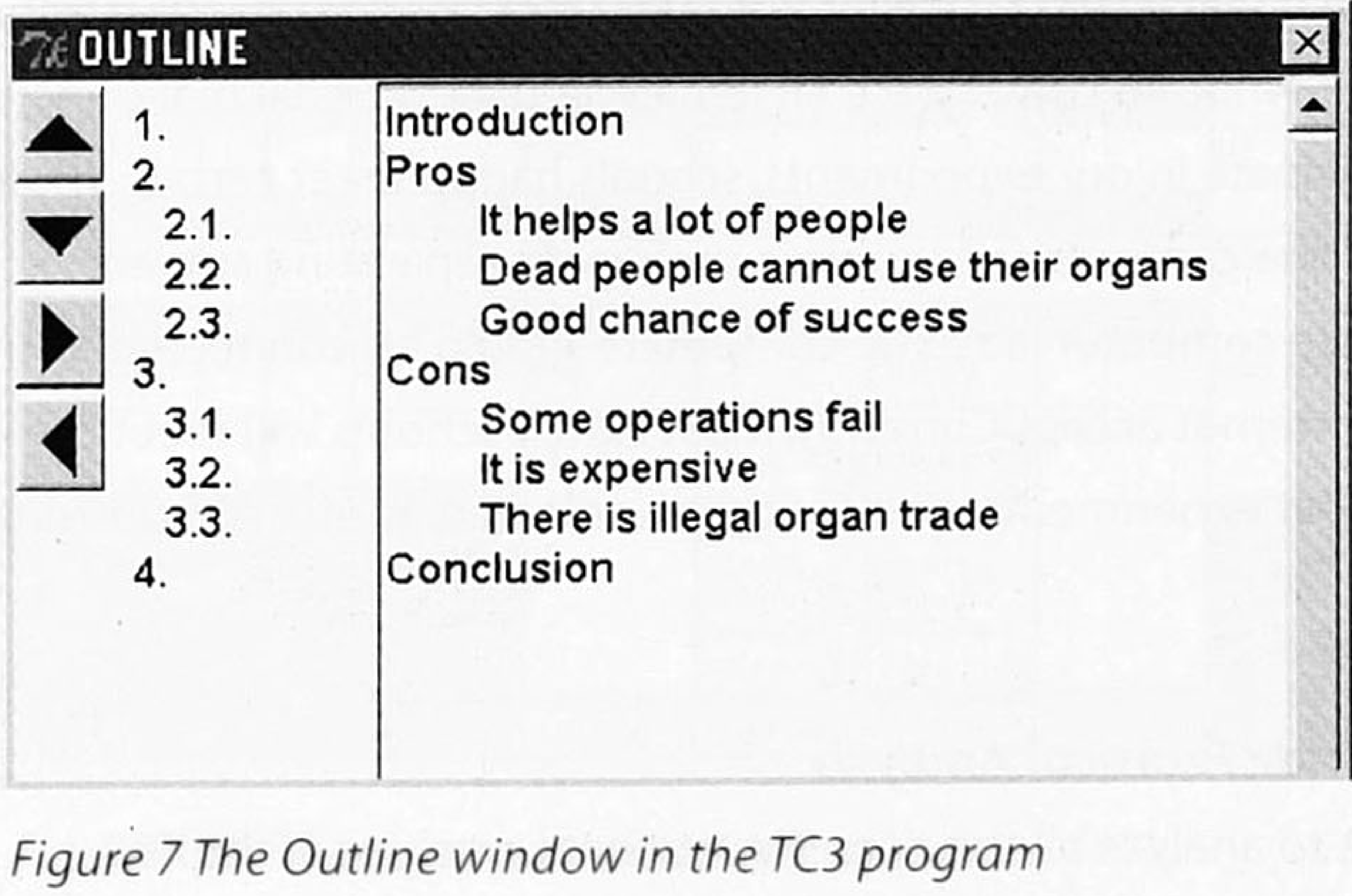

Figure 7 The Outline window in the TC3 program

We expected to find that the effects of the Diagram would mainly concern the consistency and completeness of the argumentation of the text (Veerman & Andriessen, 1997). Using the Outline may result in a better and therefore more persuasive argumentative structure and a more adequate use of linguistic structures such as connectives and anaphors (Chanquoy, 1996).

The writing task

The main task in this study was a collaborative writing task. The assignment was to write an argumentative text of 600 to 1000 words in Dutch on cloning or organ donation. For organ donation each partner had five private sources plus one common source, so there were eleven sources in total. The sources were taken from the Internet sites of Dutch newspapers. The assignment was to convince the Minister of Health, Welfare and Sport of the position they had taken. For cloning the partners each had three sources and one common source, so there were seven sources in total. In all groups, partners were seated in separate computer rooms, to encourage them to communicate only through TC3. Naturally, we could not prevent communication during breaks and between sessions. The students received teacher grades for their texts as part of their normal curriculum.

Subjects

Our participants were 290 Dutch students aged 16 to 18 from six secondary schools in the Netherlands. The assignment was completed during one to six sessions. The initial sample was about 50 % larger: dyads who were partially absent during the experiments were excluded from the final sample, as were dyads caught using sources other than those given, communicating through mobile phones or chat programs external to TC3, as well as a few dyads who logged on under each others names to read their partner's information sources. The analyzed samples included 151 girls and 139 boys. All students participated twice, but only 74 of the students are included twice in the data. Of these, 32 worked with different partners, and 42 of the students who participated twice worked with the same partner. The students worked in pairs that were put together randomly. Mixed gender dyads comprised 58 pairs of the total sample, while 46 dyads were all female, and 41 were all male.

In order to be able to participate in our experiments, schools had to meet certain criteria. The experiments required one computer per subject preferably separating subject collaborating in two or more computer labs. The computers should be connected to the school network and have Internet access. Currently, most Dutch schools will meet these criteria, but at the start of our experiments mostly schools selected as official frontrunners responded.

MEPA: A tool for Multiple Episode Protocol Analysis

We use the program MEPA to analyze all the data the students produce in the TC3 environment. The purpose of MEPA[3] (Multiple Episode Protocol Analysis), a program for protocol analysis, is to offer a flexible environment for creating protocols from verbal and non-verbal observational data, and annotating, coding and analyzing these. Examples of suitable data within education are class discussions, collaborative discussions, teaching conversations, thinking-aloud protocols, e-mail forums, electronic discussions and videotape transcriptions.

The program is multifunctional in the sense that it allows for development of both the coding and protocolling systems within the same program, as well as direct analysis and exploration of the coded verbal and non-verbal data using several built-in quantitative and qualitative methods of analysis. In its current version, MEPA can execute frequency and time-interval analyses; construct cross-tables with associative measures; perform lag-sequential analysis, interrater reliability, visual, word frequency and word context analyses; and carry out selecting, sorting and search processes. Also, some aids for inductive pattern recognition have been implemented. MEPA uses a multidimensional data structure, allowing protocol data to be coded on multiple dimensions or variables. To minimize the work associated with coding protocols and to maximize coding reliability, MEPA contains a module that can be used to program complex structured if-then rules for automatic coding. Figure 8 shows a screen dump of the MEPA program.

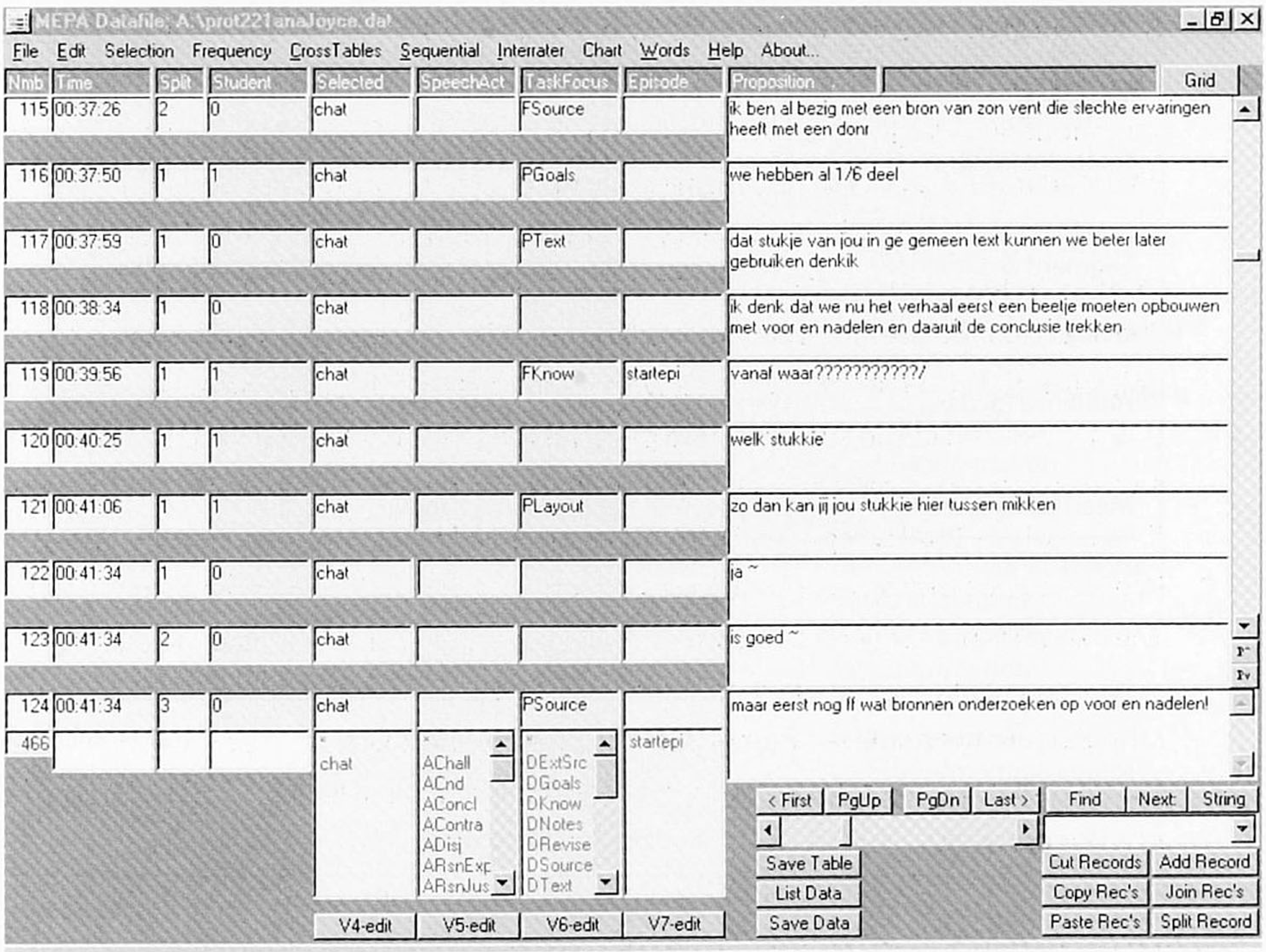

Figure 8 MEPA program for protocol analyses

The writing product: Analysis of the argumentative texts

Each of the 145 student pairs produced one argumentative text, and these were analyzed on several dimensions. As a preparation for the final assessment, the texts were imported in MEPA, with a single sentence – defined by a period – per line. The sentences with potential multiple argumentative functions were split into smaller units using an automatic splitting filter, so that the constituents of sentences such as "Cloning is good, but it can also have side effects" could be properly coded as position and argument contra. The sentences were split automatically where necessary on the basis of argumentative and organizational markers, such as but, however, although, therefore, unless. Before coding, the experimenters manually divided the texts into segments, largely based on the existing paragraph structure. The final argumentative texts were scored on five variables.

Table 7 Description of text quality measures

Variable	Description
Textual structure	The formal structure of the text as defined by introduction, body, and conclusion.
Segment argumentation	The quality of the argumentation within the paragraphs.
Overall argumentation	The quality of the main line of argumentation in the text.
Audience focus	The presentation towards the reader and the level of formality of the text.
Mean text score	The mean of the four scores above.

Table 8 Communicative functions and Dialogue acts in chat discussions.

Communicative function	Dialogue act	Specification	Explanation
Argumentatives	Reason		Ground
	Contra		Counterargument
Argumentative task focus	Conditional		Condition
	Then		Consequence
	Disjunctive		Disjunctive
	Conclusion		Conclusion
Responsives	Confirmation		Confirmation of information
	Deny		Refutation of information
Reaction, or response to an elicitative	Acceptation		Acceptation of information, without confirming or refuting the information
	Reply	Confirm	Affirmative response
		Deny	Negative response
		Accept	Accepting response
		Statement	Response including a statement
		Performative	Response containing an action performed by saying it
Informatives	Performative		Action performed by saying it
Transfer of information	Evaluation	Neutral	Neutral evaluation
		Positive	Positive evaluation
		Negative	Negative evaluation
	Statement		Statement
		Action	Announcement of actions
		Social	Social statement
		Nonsense	Nonsense statement
	Task		Task information
Elicitatives	Question	Verify	Yes/no question
Questions or utterances requiring a response	Set	Set	Question/multiple choice
		Open	Open question
	Proposal	Action	Proposal for action
Imperatives		Action	Order for action
Commanding utterances		Focus	Order for attention

The chat protocols were not analyzed at a propositional level, like the argumentative texts, but rather at an episode level based on the task oriented collaboration process. The chat protocols were manually divided into episodes of different Task act categories. Whenever the focus of the discussion changed within a particular type of Task act, a new episode was started as well. In addition, MEPA automatically coded a new episode whenever the partners had not used the chat window for more than 59 seconds.

Design

For answering the research question about the influence of the different CMC tools on the interaction and argumentative text, the Internet mediated writing environment TC3 (Text Composer, Computer supported & Collaborative) was used. All communication and activities during the collaboration are logged automatically in a chat and activity protocol. To this basic environment of TC3, three tools are added to support knowledge construction during collaborative writing:

Organizer — A tool for generating, ordering, and relating information units in a graphical knowledge structure (Diagram, see Figure 6).

Lineariser — A tool for linearizing information units as an outline of consecutive topics in the text (Outline, see Figure 7).

Advisor — A help facility that gives advise on how to use the organizer and/or lineariser.

The effects of the organizer will be related mostly to the consistency and completeness of the knowledge structure in the text (Veerman & Andriessen, 1997). The effects of the lineariser will be related mostly to the persuasiveness of the argumentation and the adequate use of language in the shape of connectives and anaphora (Chanquoy, 1996). We expected these effects to take place especially when both organization and linearization are supported, and explicit attention is paid to translating the conceptual structure into the linear text. The main indicators of this are increasing attention to the opposite position, and the use of counterarguments.

In order to compare the effects of the planning tools on the process of collaborative argumentative writing a (quasi) experiment was set up varying the different combinations of planning tools. The effect of the tools on collaborative writing is investigated in the experimental conditions shown in Table 9.

Table 9 Experimental design.

Condition		Tools
C	Control	TC3 basic
D	Diagram	Basic + Organizer
DA	Diagram Advisor	Basic + Organizer + Advisor
DO	Diagram Outline	Basic + Organizer + Lineariser
DOA	Diagram Outline Advisor	Basic + Organizer + Lineariser + Advisor
O	Outline	Basic + Lineariser
OA	Outline Advisor	Basic + Lineariser + Advisor

It was not possible logistically to assign students to the experimental conditions at random, so we assigned entire classes to the conditions. To control for school effects, classes from different schools were assigned to each condition. To control for differences in writing and argumentation skills two pretests were administered before executing the writing task. The randomly assigned pairs were asked to write an argumentative text of about 600 to 1000 words defending a position on cloning or organ donation. The shared text had to be based on information sources given within the groupware program. The experiment was executed in two separate studies: the Control group and the experimental groups.

Results

Structural characteristics of the chat dialogue

This section[4] contains a description of the results for the structural characteristics of the dialogue in terms of communicative functions and dialogue patterns within the collaboration dialogues, and the relationship between these features and the final product, the argumentative text. Table 10 shows the distribution for the five communicative functions for the Control group and for each experimental condition.

Table 10 Distribution of communicative function in the dialogue in percentages

	Total *Mean*	**C** *M*	**D** *M*	**DA** *M*	**DO** *M*	**DOA** *M*	**O** *M*	**OA** *M*
Argumentatives	9.80	8.98	10.74	10.51	9.72	9.03	10.70	9.04
Elicitatives	20.55	20.46	1.26	20.39	20.92	19.30	20.11	21.30
Imperatives	7.93	8.06	6.40	6.36	7.68	10.74	9.18	9.18
Informatives	37.66	38.65	36.04	38.28	37.93	33.94	36.50	40.22
Responsives	24.06	23.84	25.56	24.45	23.75	26.99	23.51	20.26
Total number of contributions	425.37	421.15	312.59	441.81	518.00	460.27	401.72	385.91
N (dyads)	145	39	17	26	23	11	18	11

	Total *SD*	**CD** *SD*	**DA** *SD*	**DO** *SD*	**DOA** *SD*	**O** *SD*	**OA** *SD*	*SD*
Argumentatives	2.82	2.89	2.98	2.77	2.39	2.89	2.51	2.71
Elicitatives	3.97	5.08	3.72	4.04	3.27	1.84	2.74	4.20
Imperatives	3.45	3.31	1.82	1.69	2.50	4.99	5.08	2.69
Informatives	5.65	6.09	4.73	5.51	5.64	3.02	6.11	4.67
Responsives	4.71	4.33	4.48	5.08	3.62	5.18	5.23	3.48

See first column in Table 8 for description of categories and Table 9 for description of conditions

The distribution for all groups together shows that Informatives occur most frequently (37,66%), followed by Responsives (24,06%). Argumentatives make out an encouraging 10% of the communicative functions, and Imperatives are the least frequent with 8%. Compared to the other conditions, the Control group uses significant fewer Argumentatives, especially in comparison to the Diagram, Diagram-Advisor and Outline conditions. Imperatives are more frequent in the Diagram-Outline-Advisor condition, but less frequent in the Diagram and Diagram-Advisor conditions. The Diagram-Outline-Advisor condition also used fewer Informatives, and the Outline-Advisor group used relatively few Responsives.

Table 11 shows the mean percentages of the specific Dialogue acts. In general, the distributions within the communicative functions are very similar for all conditions, so we will only discuss the total sample here. Within the Argumentatives, the relatively most frequent Dialogue act is Contra: counterarguments (3,86%). This is a nice surprise, as relatively novice writers are usually thought to use counterarguments quite sparsely. The verifying question is relatively most frequent in the Elicitatives (9,53%), followed by proposals (5,75%) and open questions (4,66%). Urging the partner to take action or fulfill a task is the more frequent Imperative with 4,91%, although asking for attention follows closely behind at 3%. Task information is exchanged relatively often (Statement Info 26%), while evaluative informatives are used less frequently (3,74%). Finally, within Responsives the most frequent Dialogue acts are Confirmation (13,46%) and plain replies (Reply Statement 3,78%).

Table 11 Mean percentages of Dialogue act

	Total *Mean*	C *M*	D *M*	DA *M*	DO *M*	DOA *M*	O *M*	OA *M*
Argumentatives								
Conclusion	1.43	1.28	1.80	1.29	1.52	1.48	1.57	1.20
Conditional	1.36	1.32	1.45	1.34	1.51	1.20	1.52	0.98
Contra	3.86	3.39	4.17	4.32	3.50	3.65	4.49	3.93
Disjunctive	0.69	0.59	0.78	0.77	0.80	0.63	0.70	0.52
Reason	1.66	1.59	1.80	1.91	1.66	1.26	1.60	1.61
Then	0.80	0.81	0.74	0.88	0.73	0.81	0.82	0.82
Elicitatives								
Proposal Action	5.75	5.49	6.03	6.08	5.71	4.80	6.14	5.86
Question Open	4.66	4.69	5.03	4.14	5.04	5.17	4.63	4.00
Question Set	0.60	0.56	0.80	0.40	0.56	0.52	0.79	0.75
Question Verify	9.53	9.72	9.41	9.77	9.61	8.81	8.55	10.69
Imperatives								
Action	4.91	5.19	3.98	3.94	4.59	6.61	5.85	5.02
Focus	3.02	2.87	2.42	2.42	3.09	4.13	3.33	4.16
Informatives								
Evaluation Negative	0.55	0.93	0.35	0.41	0.37	0.32	0.48	0.55
Evaluation Neutral	0.35	0.41	0.36	0.31	0.46	0.20	0.21	0.38
Evaluation Positive	2.84	2.92	2.52	2.99	2.66	2.38	2.63	3.93
Performative	0.97	0.94	1.02	1.11	0.76	0.62	1.11	1.28
Statement Info	26.00	25.63	25.40	27.03	26.40	23.97	25.37	28.08
Statement Action	4.99	4.58	5.03	5.17	5.36	5.12	5.50	4.27
Statement Nonsense	0.67	1.51	0.21	0.48	0.52	0.23	0.27	0.29
Statement Social	1.27	1.72	1.17	0.79	1.39	1.11	0.93	1.44
Responsives								
Acceptation	1.39	1.79	1.32	1.37	1.20	0.98	1.22	1.25
Confirmation	13.46	14.09	13.21	13.80	12.53	16.07	13.63	9.89
Deny	1.61	1.96	1.48	1.27	1.63	1.88	1.26	1.59
Reply Accept	0.14	0.16	0.22	0.15	0.13	0.12	0.05	0.11
Reply Confirm	3.03	2.09	3.96	3.37	3.31	3.62	3.01	2.98
Reply Deny	0.61	0.41	0.81	0.67	0.65	0.64	0.69	0.67
Reply Performative	0.04	0.04	0.11	0.06	0.01	0.01	0.03	0.03
Reply Statement	3.78	3.30	4.46	3.77	4.29	3.68	3.61	3.77

The total percentage per column is 100%; see second column of Table 8 for categories

Transitions between Dialogue acts

Figure 9 and Figure 10 show the transition diagrams made by the MEPA program for the Control and the Diagram condition. We will discuss the other transition diagrams too, but they are not shown here. The transition diagrams result from lag-sequential analyses (Wampold, 1992). In lag-sequential analysis the number of transitions of one event to the next (lag = 1) are tested for significance with regard to the expected number of transitions of that type based on the distribution of probability. In the diagrams, only the significant transitions are shown, with the width of the arrows indicating the level of significance. A large number of different transitions in the diagrams points towards unstructured dialogues: the fewer arrows, the more structured the dialogues were for that condition. A relatively high number of autocorrelations – indicated by the circular arrows – also indicates relatively unstructured dialogues. For readability reasons, a number of categories were merged in these analyses.

The Control group with only the TC3 basic environment, shown in Figure 9, differs from the experimental conditions with extra tools: this group shows a lot more different significant transitions between the Dialogue acts. The Control group displays relatively more different patterns than the experimental groups, and 8 out of 19 of its Dialogue acts show autocorrelations, which means that the dialogue is less structured in the Control group. Possibly, the planning tools in the experimental groups stimulate structuring of the dialogue.

All transition diagrams show one typical pattern in particular: the arrows from open questions (EliQstOpn) and verifying questions (EliQstVer) to statement replies (ResRplStm). Although the obvious answer to a verifying question would be a denying or accepting reply (ResDen or ResAcc) in all seven conditions verifying questions are relatively often answered with an elaborated statement.

Another characteristic pattern is the strong presence of argumentative sequences throughout the conditions (see upper half in Figure 10). Only the Diagram-Advisor condition differs on this point, as it shows fewer transitions between argumentatives than any other condition. The Diagram-Advisor condition generally differs from the other experimental conditions in its transitions. There are more significant transitions and these transitions are different from the ones that occur in the other experimental groups. For example, argumentative conclusions (ArgCcl) are followed significantly by social statements (InfStmSoc), conditionals (ArgCon) are followed significantly by imperative actions (ImpAct), and there are relatively many transitions to accepting responsives like mmm or oh (ResAcc). Just like the Control group, the Diagram-Advisor condition contains relatively many autocorrelations.

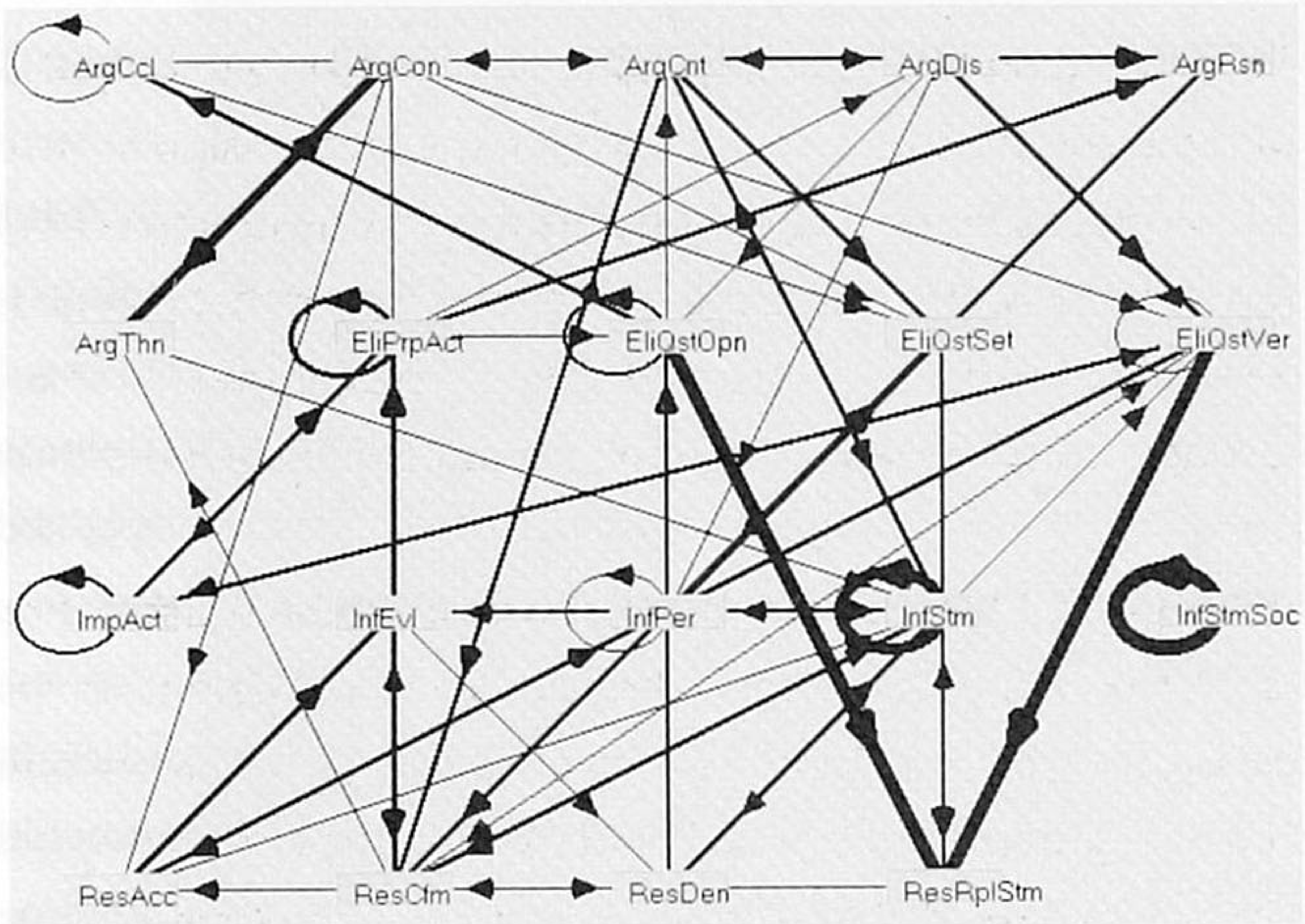

Figure 9 Transition diagram for the Control group

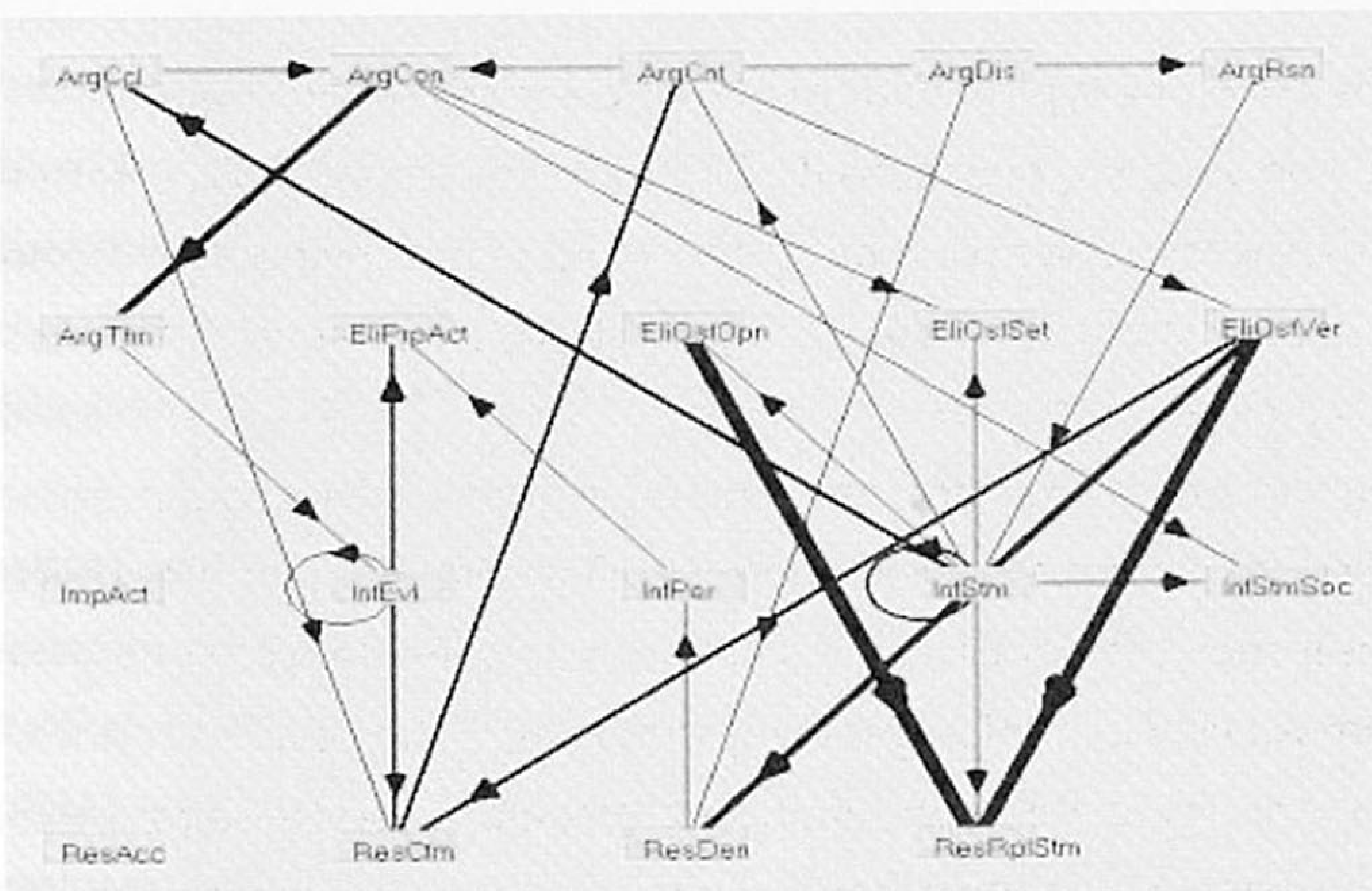

Figure 10 Transition diagram for the Diagram condition

The transition between 'if'-argumentatives (ArgCon) and 'then'-argumentatives (ArgThn) is not significant for the Outline and Outline-Advisor conditions, whereas the transition is significant in the Control group and the conditions with the Diagram. Possibly, the diagram stimulates the use of if-then patterns, whereas the Outline suppresses these patterns.

Conclusions and Discussion

Education can be viewed as an ongoing process of argumentation (Petraglia, 1997). It is the process of discovering and generating acceptable arguments and lines of reasoning underlying scientific assumptions and bodies of knowledge. In collaborative learning, students can negotiate different perspectives by externalising and articulating them, and learn from each other's insights and different understandings. Thus, through negotiation processes, including argumentation, they can (re-) and co-construct knowledge in relationship to specific learning goals.

The present research suggests that the role of argumentation needs to be reconsidered. Across studies, 'direct' forms of argumentation (challenges, counter-argumentation) did not relate well to the production of constructive activities, a measure to define learning-in-process. This may be explained by the paradox that students should have a well-established understanding of knowledge in order to take firm positions. However, their knowledge is under discussion and subject to the learning process itself. Therefore, offering support to students to challenge and counter each other's information may not be the most fruitful approach. However, information checking did show to be important, which was regarded as an 'indirect' form of argumentation. The more information was checked, the more constructive activities were produced. Students can be provoked to critically check each other's information through instruction and task design.

With regard to computerized learning environments, the research indicates that students particularly need facilitation in co-ordinating electronic and text-based communication and in keeping track of the main issues while producing networked-based discussions. Technical disturbances and a loss of thematic focus easily occur, especially in synchronous CMC systems, and have a negative effect on collaborative learning processes. Additional tools to keep a (graphical) overview of the issues at hand can be helpful, such as the diagram construction tool provided by the Belvédère system and the TC3 program.

In research by Erkens (1997), focusing, checking and argumentation were revealed as essential factors in collaborative learning processes. In addition, parallel studies aimed at argumentation, epistemic interactions and grounding processes contributed to gaining more understanding of the mechanisms that can support collaborative learning through (electronic) dialogue. We presented some results of a couple of studies in this chapter to explore those relations more in depth.

We found that the Diagram, Outline and Diagram-Advisor conditions all have a positive affect on the number of Argumentatives. This suggests that the moderate availability of extra tools has a positive influence on the number of arguments in the chat and also on some aspects of the quality of the argumentative text.

The transition patterns show that the experimental groups are more structured in their direct communication than the Control group. This suggests that the planning tools

(Diagram, Outline, Advisor) stimulate a more structured dialogue. The same difference in the structure of dialogues can be observed when comparing high scoring and low scoring dyads. This leads us to conclude that the experimental condition (extra tools) has a direct effect on text quality, but also through the communicative function in the chat dialogues.
Some of the results in the last study (Erkens, Prangsma, Jaspers, & Kanselaar, 2002) are not simple to interprete.

The analyses of chat dialogues about the Diagrams suggest that for some participants this tool did not serve as a basis for discussion or a tool for idea generation, as it was intended, but rather functioned as a visual representation. The correspondence of arguments between Diagram and the final text reveals a discrepancy between the two: only about a third of the arguments are found both in the final text and the Diagram. Although the use of wholly original arguments seems to be slightly positively related to text quality, these are hardly used, and most of the arguments are taken directly from the given sources.

With respect to these results, the study 3 (see paragraph 2.3) (Veerman, 2000) can be mentioned in which students used the Belvédère environment to chat electronically and to visualize their discussion about a computer-based design by the use of an argumentative diagram construction tool. It showed that the students only gained from the Belvédère environment, when they linked their chat discussions closely to their diagrams. A significant relationship was found between the amount of overlapping information between chats and diagrams, and the amount of constructive activities produced. However, student groups varied in linking information between chats and diagrams. This appeared to depend heavily on student groups' task approaches and preparation activities.

We also found that using the private –hence non-collaborative– notes window (the upper left window in TC3) is detriment to the quality of the collaborative product. This confirms our idea that collaboration is necessary on all subtasks, including planning, idea generation, coordination and information processing.

We also found that explicit argumentation on content, coordination, and metacognitive strategies is related positively to text quality, whereas argumentation on technical aspects of the task and on non-task related topics is related negatively to text quality. The relation between non-task chat and text quality is negative throughout the groups, although the relation is the most clear for the Control group.

When we compare the Diagram (Figure 6) with the Outline (Figure 7), the Outline tool was more successful. Availability and proper use of this planning tool have a positive effect on the dialogue structure, and on the coordination processes of focusing and argumentation, as well as on text quality. The Diagram often functions as a visual representation, and not as a basis for discussion or a tool for idea generation. When a diagram reflects the discussion itself, it can be a valuable starting point for writing the text, and of benefit to textual structure. Students don't have much experience with the use of Diagram tools. Perhaps a different approach to the task instruction – for example by giving the students time to practice using the complex Diagram tool – could encourage the students to use the tool as it was intended, and thus lead to different results.

Much is possible in electronic learning environments, but so far not enough is known about the relationships between collaborative learning, argumentation and educational technology. This research has shown that such relationships are neither simple nor very predictable. Hence, much more research is needed that examines the role of (interactive) mechanisms such as argumentation and focusing in relationship to features of CSCL situations.

In this contribution we showed some of the characteristics of current electronic discussions in education. We think the discussions that we obtained rate among the more successful implementations of CMC in this context. Nevertheless, we feel much more can be achieved if a carefully designed approach to educational practice in terms of educational scenarios is implemented.

Our ideas as they have been presented still remain at the level of intentions. Even in transmission not much collaboration is designed in practice. Many exceptions can be found in O'Donnell & King (1999), and these examples show the many things students have to learn before effective collaboration is achieved. In our scenarios not all details of the complex processes students are engaged in are well known or enough articulated.

The most important message of this chapter is that new learning has to be designed and needs careful study. Changing educational practice has to be an engineered approach, in which goals change as a function of the scenario users are engaged in. If the goal of education is personal understanding, maybe any scenario will do as long as it is properly designed. Results in this case will depend on the appropriate interplay of individual and task situation characteristics. However, if the goal of education is shared understanding, transmission is not good enough. Design of learning arrangements in which awareness of collaboration is raised and encouraged are then a necessary requirement.

Notes

[1] Studies reported by Arja Veerman in her PhD. thesis in chapter 2 (Veerman, 2000)

[2] This study is reported in more detail in Erkens, Prangsma, Jaspers, & Kanselaar, 2002.

[3] MEPA was developed as a general program for protocol analysis and is being used in several research projects at Utrecht University, as well as abroad. For further information, please contact G. Erkens (G.Erkens@fss.uu.nl).

[4] The analyses of the Dialogue Acts were done by Floor Scheltens as part of her Master's thesis

References

Andriessen, J., Erkens. G., Overeem, E., & Jaspers, J. (1996, September). *Using complex information in argumentation for collaborative text production*. Paper presented at the UCIS '96 conference, Poitier, France.

Bostock, A. J. (1996). A critical review of Laurillard's classification of educational media. *Instructional Science, 24*, 71-88.

Chanquoy, L. (1996, October). *Connectives and argumentative text: a developmental study*. Paper presented at the First International Workshop on Argumentative Text Processing, Barcelona, Spain.

Erkens, G. (1997). *Cooperatief probleemoplossen met computers in het onderwijs: Het modelleren van cooperatieve dialogen voor de ontwikkeling van intelligente onderwijssystemen* [Cooperative problem solving with computers in education: Modelling of cooperative dialogues for the design of intelligent educational systems]. Unpublished doctoral dissertation, Utrecht University, Utrecht.

Erkens, G. (1997). *Coöperatief probleemoplossen met computers in het onderwijs: Het modelleren van coöperatieve dialogen voor de ontwikkeling van intelligente onderwijssystemen.* [Co-operative problem solving with computers in education: Modelling of co-operative dialogues for the design of intelligent educational systems]. Utrecht, The Netherlands: Brouwer Uithof.

Erkens, G., Prangsma, M.E., Jaspers, J.G.M., & Kanselaar, G. (2002). *Computer supported collaborative and argumentative writing*. Utrecht : Utrecht University, ICO-ISOR Onderwijsresearch.

Glasersfeld, E. von (1989). Cognition, Construction of Knowledge and Teaching. *Synthese, 80*, 121-140.

Greeno, J.G. (1997). Response: On claims that answer the wrong question. *Educational Researcher, 20*, 5-17.

Hakkarainen, K. (1995, August). *Collaborative inquiry in the computer supported intentional learning environments (CSILE)*. Paper presented at the 6th European conference for Research on Learning and Instruction, Nijmegen, The Netherlands.

Kanselaar, G., & Erkens, G. (1996). Interactivity in co-operative problem solving with computers. In S. Vosniadou, E. DeCorte, R. Glaser & H. Mandl (Eds.), *International perspectives on the design of technology-supported learning environments* (pp. 185 - 202). Mahwah, New Jersey: Lawrence Erlbaum.

Kanselaar, G., Jong, T. de, Andriessen, J. E. B, & Goodyear, P., (2000). New technologies. In P .R. J. Simons, J. L. van der Linden, & T. Duffy (Eds.), *New learning* (pp. 49 - 72). Dordrecht: Kluwer Academic Publishers.

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

Laurillard, D. (1993). *Rethinking university teaching: A framework for the effective use of educational technology*. London: Routledge.

O'Donnell, A. M. & King, A. (Eds.) (1999). *Cognitive perspectives on peer learning*. Mahwah, NJ: Lawrence Erlbaum Associates.

Petraglia, J. (1997). *The rhetoric and technology of authenticity in education*. Mahwah, NJ: Lawrence Erlbaum.

Salomon, G. (1997, August 26-30). *Novel constructivist learning environments and novel technologies: Some issues to be concerned*. Invited key-note address at the EARLI conference, Athens.

Savery, J.R., & Duffy, T.M. (1995). Problem based learning: An instructional model and its constructivistic framework. *Educational Technology, 35*, 31-38.

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: Proceedings of thé First European Conference on Computer-Supported Collaborative Learning* (pp.577-584). Maastricht, The Netherlands, University of Maastricht.

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

Van Eemeren, F.H., Grootendorst, R., & Snoeck Henkemans, A.F. (1995). *Argumentatie*. [Argumentation]. Groningen: Woltersgroep, The Netherlands.

Veerman, A.L. (2000). *Computer-Supported collaborative learning through argumentation. Doctoral dissertation*. Enschede: Print Partners Ipskamp.

Veerman, A.L., & Andriessen, J.E.B. (1997, September, 4-6). *Academic learning & writing through the use of educational technology*. Presented at the conference on Learning & Teaching Argumentation, Middlesex University, London.

Veerman, A.L., Andriessen, J.E.B., & Kanselaar, G. (2000.) Enhancing learning through synchronous discussion. *Computers & Education*, *34*, (2-3), 1-22.

Wampold, B. E., & Margolin, G. (1982). Nonparametric strategies to test the independence of behavioral states in sequential data. *Psychological Bulletin*, *92*, 755-765.