COMPUTER SUPPORT FOR COLLABORATIVE AND ARGUMENTATIVE WRITING

G. Erkens
M.E. Prangsma
J.G.M. Jaspers
G. Kanselaar

Onderwijskunde
ICO-ISOR Onderwijsresearch
Universiteit Utrecht

Heidelberglaan 2
Postbus 80140
3508 TC Utrecht

telefoon: 030 - 253 49 40
fax: 030 - 253 23 52
e-mail:owksecr@fss.uu.nl
http://www.fss.uu.nl/owk

ISOR-rapportnummer 02.01

ISSN 0924-0217
ISBN 90-6709-053-0

NWO-project 575-33-008
“COmputerondersteund Samenwerken bij ARgumentatieve tekstproductie (COSAR)”

Druk: Drukkerij Zuidam & Uithof B.V. - Utrecht

CONTENTS

PREFACE

This book gives a report of the COSAR project, a research project into the opportunities for computer support of planning in argumentative collaborative writing. The project was aimed at the senior years of the highest level of Dutch secondary schools, and ran for three years. The results gave us insight in the planning process in collaborative writing, as well as in the possibilities of support through computer tools like diagrams and outlines. The COSAR project has also resulted in a successful groupware program, that will be used in a number of follow-up research projects.

Special thanks to Hermi Schijf, who was the main project researcher for almost the first two years. In carrying out the research we were assisted by a number of students from the Department of Educational Science. First of all, we would like to thank our graduating MA students, all of whom assisted in gathering and coding the data. Joyce van Berlo wrote her Master's thesis on the Task acts in this group. Floor Scheltens wrote her Master's thesis on the Dialogue acts. Paulien Honkoop wrote her Master's thesis on the Diagram tool. Johan Theil wrote his Master's thesis on the social aspects of the chat collaboration. Part of Chapter 6 was realized with the assistance of students Tobi Boas, Chris Phielix, Jan-Willem Schoonhoven and Nicolette van der Meijden as part of a second-year research course. Many thanks as well to Jean-Claude Wippler from EQUI4 Software for his contribution to programming the TC3 software.

We would like to thank everyone in the six participating schools – the students, the teachers, the system administrators, and headmasters – for taking part in the project: RSG Broklede in Breukelen, ORS Lek & Linge in Culemborg, the Griftland College in Soest, the Niels Stensen College in Utrecht, the Minkema College in Woerden, and KSG De Breul in Zeist.

The COSAR project was funded by NWO, the Netherlands Organization for Scientific Research, under project number 575-33-008. We would like to thank NWO for giving us the opportunity to do this research.

SUMMARY

The objective of the research project 'Computer Support for Collaborative and Argumentative Writing' (the COSAR project) was to study the relation between the collaborative process and support of the planning process in argumentative writing. Subject of our investigations are students in the 'studiehuis' – a recent innovation in the Dutch secondary school curriculum. Groupware was developed – called TC3: Text Composer, Computer supported & Collaborative – that allows collaborative writing by pairs of students, with or without support by specially designed planning tools for organization and linearization (the Diagram and the Outline). The TC3 environment offers the students a shared text editor, access to Internet based information sources, a private notepad, and a chat facility. The Diagram is a shared planning tool with which students can organize the elements of their argumentation in a graphical map. With the Outline students can order the paragraphs for their text in a linear structure.

Chapter 1 presents a review of the relevant literature on collaborative argumentative writing, on coordination processes in collaborative writing, and on groupware. Planning is a complex part of argumentative writing tasks, and coordination of activities is a vital aspect of collaboration. Computer software can be designed to support the coordination and planning processes.

The research methods are described in Chapter 2. We analyzed the chat and activity protocols of 145 dyads working collaboratively on planning and writing an argumentative text on cloning or organ donation. Pretests were administered to control for linearization and argumentation skills.

Although the underlying study is strongly process-oriented, the products of the collaborative work are not overlooked. The main results for the argumentative texts are discussed in a separate chapter (Chapter 3), and the results for the planning tools – the Diagram and the Outline – can be found in the chapter on the groupware support of organization and linearization (Chapter 5). We found little difference between the experimental conditions, but the data do show a slight positive effect of the use of the linearization tool – the Outline – on the argumentative and structural quality of the final text.

Activities of collaborative knowledge construction are discussed in terms of the use of the software tools – both the basic tools and those designed especially to support planning – and in terms of task discussion in the chat (Chapter 4). The results show that the types of constructive activities are different in different phases of the writing process. In addition, we found some interesting differences between the control group and the planning tool conditions.

Several aspects of coordination processes are discussed, including structural characteristics of the collaborative dialogues, symmetry of contribution to the chat by the collaborating partners, and the specific coordination processes of checking, focusing, and argumentation (Chapter 6). There are some interesting differences between conditions, and between dyads writing ‘good’ and ‘bad’ argumentative texts. In general, coordination processes were found to play an important role in collaborative argumentative writing.

As a long-term objective of this type of research is to develop and improve software for educational purposes, the participants were asked to give feedback on the task and the TC3 program (Chapter 7). Analysis of the evaluation showed that though most students thought there was room for improvement, the majority was reasonably positive about the groupware environment presented to them.

CHAPTER 1 INTRODUCTION

1.1 Purpose and rationale

Secondary school students in The Netherlands – as a result of recent changes in the curriculum of the final years (the 'studiehuis') – 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 task related activities and deliberation between participants on the quality of the final text.

1.2 Theoretical background and context

Collaborative argumentative writing
Writing clearly is an open task. Writing texts of any length is a complex process consisting of several interrelated sub processes, each with its own dynamics and constraints (Alamargot & Chanquoy, 2001; Rijlaarsdam & Van den Bergh, 1996). We conceptualize writing argumentative texts mainly as a knowledge construction (Galbraith, 1999) and problem solving task. This task requires that information is generated, collected, selected, related, and organized into a consistent knowledge structure. In addition, the writer must find a persuasive line of argumentation to convince the reader. For successful completion of the task social, cognitive, rhetorical, and cultural skills are called for.

The main advantage of collaborative writing when compared to individual writing, is the possibility of receiving and giving immediate feedback. According to Stein, Bernas and Calicchia (1997) argumentation itself facilitates learning because it necessitates searching for relevant information and using each other as a source of knowledge. In addition, the discussions generated by the argumentation task make the collaborators verbalize and negotiate, among others, purpose, plans, concepts,

and doubts. Collaborating writers need to test their hypotheses, justify their propositions, and clarify goals. This may lead to increased awareness of and more conscious control over the writing and learning processes (Gere & Stevens, 1989; Giroud, 1999).

Theories of writing (Hayes & Flower, 1980; Hayes & Nash, 1996) generally distinguish three types of activities within the writing process: planning (generating, organizing and linearizing content), formulating or translating (writing the text), and revising. For planning an argumentative text, arguments need to be generated and ordered based on ones position and the demands of the audience. Unlike in storytelling, the order of the content of an argumentative text does not inherently follow from the order in which events take place (McCutchen, 1987). During planning activities, ideas will probably be conceived and organized from a perspective other than time – for instance, in argument clusters. The contents need to be linearized (ordered) before the ideas can be rendered into text, and again when the contents are reorganized. Linearization, therefore, is an important activity in argumentative writing (Levelt, 1988). Research at our department showed that an explicit division between idea organization and linearization during planning leads to improved quality of the argumentative text (Coirier, Andriessen & Chanquoy, 1999). Converting the conceptual representation of ideas into linear text turned out to be a crucial problem for the novice writer of argumentative texts. Our computer environment endeavors to support students during these two activities by providing tools for conceptual organization and linearization and by offering help on using these tools for planning.

Prior research often focused on preplanning. Preplanning refers to planning activities that occur before writing the text. It has been shown that preplanning can have a favorable effect on the quality of the text (Andriessen, Coirier, Roos, Passerault & Bert-Erboul, 1996; Wesdorp, 1983). At the same time, we know that inexperienced writers rarely preplan (Alamargot, 1997; Bereiter & Scardamalia, 1987). Moreover, because of their insufficient knowledge of the issues involved, when preplanning does occur in novices it is more likely to be like superficial brainstorming: simple content activation based on the terms used in the assignment. Bereiter and Scardamalia (1987) found this to be true for children. Torrance, Thomas, and Robinson (1996) likewise found little idea generation based on rhetorical demands during preplanning for adult undergraduates (relative novices), whose idea generation rather fitted a simple content activation model.

Also, the number and originality of ideas in the draft did not correlate with time spent preplanning.

When lacking preplanning skills, support of online planning becomes especially important for inexperienced writers. Online planning consists of the monitoring activities that occur during writing based on set goals, ideas, expectations and strategies (Van der Pool, 1995). These activities direct the process of knowledge construction during writing. Online planning activities, unlike preplanning, are generally linked more strongly to the local organization of the text. Expert preplanning deals with broader issues like setting goals and determining overall organization and genre. In prior research, the transition between preplanning processes and writing the actual text was found to be a stumbling block. Kozma (1991), Scardamalia and Bereiter (1985), Bereiter and Scardamalia (1987), and Schriver (1988) all found positive effects of teaching preplanning on the amount and/or the quality of preplanning, but not on the quality of the final text. The problem could lie in the linearization or the translation process, both transitional processes.

In collaborative writing, reflecting on such transitions becomes a natural process, because by writing a shared text, the partners will have to agree on both the content and the organization of the text. In addition, the use of sources needs to be coordinated and discussed. In prior research, where college undergraduates selected arguments and produced an argumentative text while collaborating in a groupware environment, differences in the argumentative discussion were found to correlate with the representation of the source material. In a task where the arguments were presented as pictures, more inferences were needed to deduce the usefulness of the information. The students discussed more new arguments in the chat discussion and in their common argumentative text (Andriessen, Erkens, Overeem & Jaspers, 1996). Thus, the constructive activities of organizing, linearizing as well as translating to the common text have to take place in mutual deliberation, necessitating verbalization and reification of ideas. This negotiation, leading to shared knowledge construction, takes place in the collaboration dialogue between the partners (Erkens, Andriessen & Peters, 2002). We expect to find that more mutual coordinating activities in the dialogue result in a more consistent shared knowledge structure, and in a better mutual problem solution, that is, a better argumentative text (see also Baker, 1999). Furthermore, support of content generation, organizing and linearization should make these planning activities explicit and negotiable.

Coordination processes in collaborative learning

In natural educational settings we can specify a collaborative learning situation as one in which a small group of two or more students work together to fulfill an assigned task within a particular domain of learning in order to achieve a joint goal (Cohen, 1994). In natural collaboration, the collaborating partners must have a common interest in solving the problem at hand. Furthermore, they should be mutually dependent on the information, resources, tools and cooperative intention or willingness of the partner(s) to reach their (shared) goals. Collaboration can only be fruitful and be searched out in a natural way if the participants have complementary abilities, information and willingness (Erkens, 1997). Mutuality is a necessary condition for natural collaborative learning situations: a positive interdependence and equal opportunity of participation in the interaction. Under these conditions of mutuality, coordination of task strategy and of the constructive activities to achieve a shared understanding of the problem are crucial aspects of collaborative learning. In earlier research we found that this coordination is realized by a complex interaction between task related strategies, cooperative intentions and communication processes during collaboration. In the collaborative learning situation the learning results will be influenced by the type of task, the composition of the group, the common goal or task product, the complementarity in expertise of the participants, the resources and tools available, and the educational climate. In order to achieve the common goal the collaboration partners will have to coordinate their activities and their thinking. Real collaboration requires a common frame of reference in order to be able to negotiate and communicate individual viewpoints and inferences. Furthermore, shared understanding of the problem at hand – a joint problem space (Roschelle & Teasley, 1995) – or a collective landscape of concepts (Andriessen, Erkens, Peters & Van de Laak, in preparation) must be constructed and a problem solving strategy has to be agreed upon.

After reviewing research on the learning activities that may be stimulated by the dynamics of the interaction between the participants in the collaborative learning situation, we now think that there may be three main processes: activation of knowledge and skills, grounding or creating a common frame of reference, and negotiation or the process of coming to agreement. Specific activities can be distinguished within these three processes:

1. Activation of knowledge and skills
    a. Initiating (taking initiative in the task interaction)
        i. degree of participation
        ii. proposing topics of discussion (task strategy)
    b. Articulation of knowledge and information
        i. explicating & verbalizing
        ii. organizing & structuring
    c. Exchanging knowledge and information
        i. sharing information and resources
        ii. seeking or asking for information

We assume that the collaborative learning situation – simply by its shared goal directness and the interactive situation – stimulates processes of taking initiative in the interaction, encourages the verbalization and thus (re)structuring of knowledge (situated articulation, Brown, Collins and Duguid, 1989) and promotes the exchange of information and resources (Teasley and Roschelle, 1993). In short, collaboration stimulates the activation and exchange of task-related knowledge and information and thus stimulates a shared task orientation.

2. Grounding (creating a common frame of reference)
    a. Tuning
        i. adapting to the level of understanding of the partner
    b. Checking
        i. checking exchanged information in relation to the existing knowledge structure
    c. Focusing
        i. mutual control of focus and topic of discussion
    d. Co-construction
        i. complementing knowledge and skills of the partner

Collaborative communication requires that the students acquire a common frame of reference to allow them to communicate and negotiate their individual viewpoints and inferences. Grounding is a process characterizing all communication (Clark & Schaeffer, 1987; Clark & Brennan, 1991). For communication to be successful, we need to make sure that we understand each other. By back channeling (confirming, nodding, acknowledgements etc.) communicating participants can signal their understanding. By tuning, participants try to adapt to the perceived level of understanding of their collaborative partners. By focusing, students try to maintain a shared topic of discourse and to repair a common focus if they notice a focus divergence. By checking new information with regard to the knowledge that was (co)constructed, the students guard the

coherence and consistency of their collective knowledge base (Erkens, in preparation). In co-constructing the participants collaboratively add to this shared knowledge base by complementing each other's contributions (Van Boxtel, 2000). In short, by processes of grounding students maintain the consistency of their collective, commonly understood knowledge base of concepts and relationships between them. Mutual understanding is a necessary condition for communication and hence for collaboration. However, understanding each other's perspective is not the same as agreeing on one perspective.

3. Negotiation and coming to agreement
    a. Explanation
        i. elaboration, explanation and accounts
    b. Argumentation
        i. discussion, persuasion and criticizing
        ii. comparing and evaluating
    c. Coming to agreement
        i. deciding and according

In collaboration the participants also need to come to agreement about task strategies, relevant concepts and relationships. They try to change the other's viewpoint to arrive at the best way to solve the task at hand or at a definition of concepts acceptable for all. In this process they try to convince the other by elaborating on their point of view, giving explanations, justifications and accounts (Antaki, 1994). A process of explicit argumentation should lead to agreement on the task strategies to be followed and on the inferences to be drawn (Baker, 1999). Alternatives need to be deliberated and compared to each other, and a joint decision has to be made on which alternative to use (Di Eugenio, Jordan, Thomasson & Moore, 2000). In (neo-)Piagetian theory the resolution of the social-cognitive conflict between participants is seen as the most crucial factor for learning in collaborative learning situations (Doise & Mugny, 1984). In our opinion, however, it is in fact the paradox of collaborative learning: the assumption that students learn from arguing, criticizing and conflict versus the necessity of reaching consensus in order for collaboration to advance.

From our research we believe that the need to coordinate activities - in other words, to come to a common goal, a common task strategy and the construction of a shared knowledge base - is crucial for solving the collaborative task at hand. In the first place, this need for coordination stimulates the activation of knowledge and the initiative to share this private information or knowledge. Secondly, the

need to coordinate not only necessitates transfer of information, but also a common frame of reference in order to understand each other's perspective. Thirdly, agreement on a common line of reasoning should be reached. In fact this accounts for the difference between obtaining mutual understanding ("I understand what you mean") versus obtaining a common understanding ("I agree with what you mean"). While mutual understanding (grounding) can be seen as a 'cooperative' prerequisite for all communication, and thus also for collaborative learning situations, coordination of activities and agreement on a common line of reasoning is necessary for successful collaboration. Furthermore, one may assume that collaborating students will need to coordinate their activities on three levels of thinking and action: the task content level (concepts and procedural skills), the meta-cognitive level (task strategy and monitoring), and the socio-communicative level (interpersonal relations and interaction).

Groupware

Computer and Internet based environments seem especially suited for collaborative learning through the variety of possibilities they offer: they allow for integration of multimedia sources, data processing tools, and communication systems (not restricted by time or place) within a single workspace (Bannon, 1995; Van der Linden, Erkens, Schmidt & Renshaw, 2000). Computer Supported Collaborative Learning (CSCL) systems are assumed to have the potential of enhancing the effectiveness of peer learning interactions, which is considered vital for knowledge building in constructivism (Andriessen et al., 1996; Dillenbourg, 1999; Katz, 1995). When we speak of the role of computers in education, our focus is on the development of computer based multimedia environments: open learning environments that may give rise to multiple authentic learning experiences (The Cognition and Technology Group at Vanderbilt, 1994). The cooperative aspect is mainly realized by offering computerized tools that can be helpful for collaborating students in solving the task at hand (e.g., the CSILE program of Scardamalia, Bereiter & Lamon, 1994; the Belvedere program of Suthers, Weiner, Connelly & Paolucci, 1995). These tools are generally one of two types: task content related or communicative. Task related tools support task performance and the problem solving process (Roschelle & Teasley, 1995; Salomon, 1993; Teasley & Roschelle, 1993). Communicative tools give access to collaborating partners through Computer Mediated Communication (CMC) facilities like chat and discussion forums, but also to other resources, such as external experts, or information sources on the Internet. In this respect, the program functions as a communication medium (Henri, 1995). Programs that

integrate both tool types are generally known as groupware: they are designed to support collaborative group work by sharing tools and resources between group members, and by offering communication opportunities within the group and with the external world. The program we designed is also groupware.

The COSAR research project focused on the influence of collaborative writing strategies on the final product in collaborative writing. We are particularly interested in planning and coordination strategies. The following three research questions were addressed:

1. How do knowledge construction activities differ in different planning phases (before and whilst formulating the text)?
2. How does support of organization and linearization of knowledge through different ICT tools influence the consistency and coherence of argumentative texts in argumentative writing?
3. How do features of the planning process (organization and linearization) and the ICT tools relate to the coordination in the dialogue of collaborating students in terms of checking, focusing and argumentation?

1.3 Social and scientific significance

The project integrated three aspects of social significance: the introduction of ICT in education, writing as an important basic skill, and collaborative learning. For many teachers, collaborative learning proves difficult to realize within the recent changes in the curriculum (Bolhuis & Kluvers, 1996). The application and effectiveness of collaborative learning might increase if it is supported through electronic networks. Learning to write as preparation for the information society should play an important part in current education. Internet facilities like e-mail and groupware heighten the importance of collaborative writing.

The project had links with the program Internet for Education in Utrecht (Internet Voor Onderwijs (IVO)) in which Utrecht University, KPN (Dutch Royal Mail), and 20 schools cooperate to implement networking facilities within the schools, and to investigate the didactical use of the network infrastructure. The schools that took part in our studies are members of this program.

At the start of this project, knowledge of collaborative planning and writing within an ICT environment was limited, providing too little proof of its added value. With

the COSAR project we have endeavored to focus on the following relevant and innovative aspects:

- Concrete implementation of interactivity in a powerful learning environment
- Knowledge of text production through manipulation of support in co-construction of knowledge
- Distinction between different planning components in collaborative writing
- Function of coordination between collaborating partners in shared knowledge construction and planning
- Better insight in conditions determining success of collaborative writing.

1.1 Organization of the report

In the next chapter we will discuss the methods used to investigate the coordination between collaborating students during planning and writing the argumentative texts. In the chapters that follow we will present quantitative results on the three research questions. Chapter 3 is concerned with the quality of the final writing product. Chapter 4 deals with the first research question, on activities of knowledge construction. The second research question, on support of organization and linearization activities, is discussed in Chapter 5. Planning and coordination processes are discussed Chapter 6. Chapter 7 contains a discussion of the student evaluations. Chapter 8 deals with the practical and methodological issues encountered while carrying out the research. In the last Chapter the main implications of our findings are summed up and suggestions for further research are given.

CHAPTER 2 METHOD

2.1 The design

For answering the research question, an Internet mediated writing environment was developed that allows pairs of students to deliberate and collaborate in writing a text. This TC3 program (see 2.3 The apparatus: The TC3 environment) contains four windows containing the task and information sources, a private notepad, a chat facility, and a shared text editor. All communication and activities during the collaboration are logged automatically in a chat and activity protocol. To this basic environment, three tools can be added to support knowledge construction during collaborative writing:

Organizer	A tool for generating, ordering, and relating information units in a graphical knowledge structure (Diagram).
Lineariser	A tool for linearizing information units as an outline of consecutive topics in the text (Outline).
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 are investigated in the experimental conditions shown in Table 2.1.

Table 2.1: Experimental design.

	Condition	**Tools**
C	Control group	None
D	Diagram	Organizer
DA	Diagram Advisor	Organizer + Advisor
DO	Diagram Outline	Organizer + Lineariser
DOA	Diagram Outline Advisor	Organizer + Lineariser + Advisor
O	Outline	Lineariser
OA	Outline Advisor	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.

2.2 The participants

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.

The main task in this study was a collaborative writing task. Some of the students participated twice, but with different partners, and in different conditions. This meant that they wrote two texts – one on each topic – with at least two months

between the two tasks. The participants did not choose their own partners: pairs were assigned by the experimenter on the basis of the list of names provided by the teacher.

Table 2.2: Distribution of participants over conditions.

Condition	Number of dyads
Control	39
Diagram	17
Diagram-Advisor	26
Diagram-Outline	23
Diagram-Outline-Advisor	11
Outline	18
Outline-Advisor	11
Total	145

2.3 Material

2.3.1 Pretests

The students were given two pre-tests before starting on their first writing task. The Wild Cat Test was used to determine writing skills, and the Underline Arguments Test was used for measuring the students' understanding of argumentative structure.

In the Wild Cat Test (Coirier, Favart, & Broggio, submitted) the participants were asked to compose a story from a set of eleven given sentences. The sentences, all stating facts about and characteristics of the wild cat were carefully worded so that they could be clustered logically. Both clustering and ordering of the sentences are assumed to be indicative of specific writing skills: composition and linearization. A number of linguistic and semantic measurements were also taken. The total score on the Wild Cat Test was used as a measure of the participants' writing skills. An English translation of the Wild Cat Test can be found in Appendix 2.

In the Underline Arguments Test (Oostdam, 1991) the participants were asked to underline the part of a three part clause that is both a claim and an argument. For example: "He has gone bankrupt twice, so I feel that we should not do business with him. We have to tell the management that we will not be using his services." The underlined clause is a claim in relation to the first clause, but an argument in relation to the final sentence. This test should give an indication of the

participants' understanding of argumentative structure. The Dutch version of the Underline Arguments Test can be found in Appendix 2.

In addition to the pretests, we asked the teachers to supply assessments of participants' writing and language skills. This proved to be difficult for the teachers, so we ended up working with the students' Dutch language grades from the previous school year. Unfortunately, very few teachers from the experimental groups managed to provide us with these data.

2.3.2 The apparatus: The TC3 environment

General description

In the COSAR project 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 2.1, contains four main windows:

Figure 2.1: The neutral layout of the interface of the basic TC3 environment.

- 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 window): 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 *zoek*, *markeer* and *wis* (search, mark and delete) can be used to mark and unmark text in the source windows and to search through the marked texts. The *aantal woorden* button allows the participants to count the number of words in the shared text editor at any given moment. The *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.

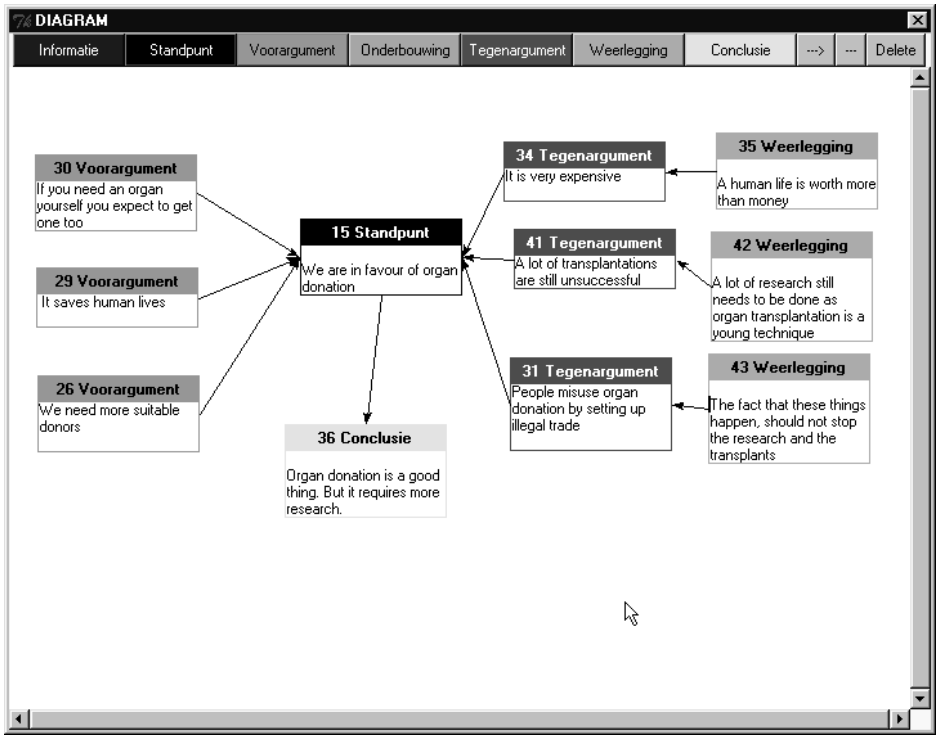

Figure 2.2: The Diagram window.

In addition, two planning modules were developed for the experimental conditions: the Diagram and the Outline. The Diagram (see Figure 2.2) is a tool for generating, organizing and relating information units in a graphical knowledge structure comparable to Belvedere (Suthers, Weiner, Connelly, & Paolucci, 1995). 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.

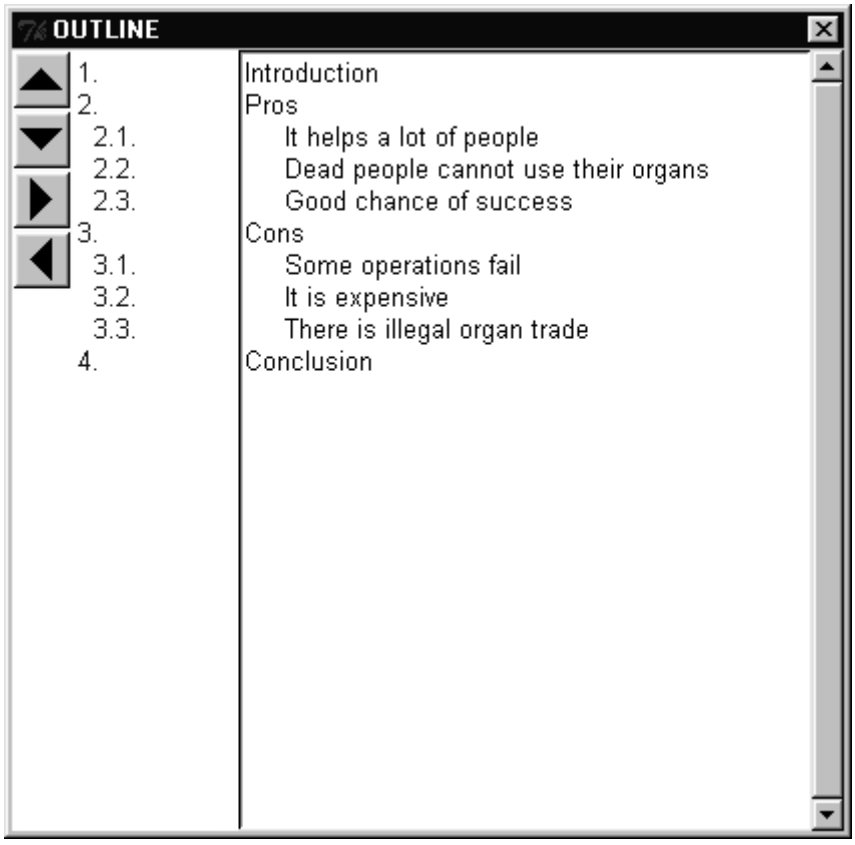

Figure 2.3: The Outline window.

The Outline (see Figure 2.3) is a tool 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.

Originally, we intended to add another module to the program: the Advisor. This module would advise and ask questions related to the consistency of the knowledge structure and the coherence of the contents of the Diagram and/or Outline, dependent on the phase in the writing process. The Advisor would also give advice on general writing problems. However, it proved to be too technically complicated and time-consuming to program this module. Instead, we added an Advisor tab sheet to the information window, and gave participants in the Advisor conditions extra instructions on the planning tools. The tab sheet gave tips and instructions for optimum use of the Diagram or Outline tool.

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). We hypothesize that these effects will be stronger when the Advisors support both the graphical/semantic and the linear/hierarchical organization, and when there is explicit help on pre and online planning and on translating organizational structures into linear text.

Table 2.3: Event types in the Integrated Activity Protocols.

Activity	Description
Name	Typing in the student's name at the initial start-up of TC3.
To task	Clicking into the window containing the assignment.
To manual	Clicking into the window containing the TC3 manual.
To Diagram tips	Clicking into the window containing the Diagram tips.
To Outline tips	Clicking into the window containing the Outline tips.
To source #	Clicking on a specific source tab in the information window.
Mark source	Clicking the mark source button to mark text in the information window.
To notes	Clicking into the private notes window.
To text	Clicking into the shared text editor.
To chat	Clicking into the chat window.
Chat	Entering chat into the chat history by pressing Enter.
To chat history	Clicking into the chat history window.
Turn-ask	Clicking on the traffic light to ask the partner for the turn to write in the shared text window and the planning tool(s).
Turn-give	Clicking on the traffic light to give the partner his/her turn to write in the shared text window and the planning tool(s). In the protocol, a turn-give is followed by a dump of the contents of the shared text, the private notes windows, and the planning tools at that time, and by an overview of added and deleted content since the previous turn-give.
Text	Shows the contents of the shared text window at the time of a turn-give.
Notes	Shows the contents of the private notes windows at the time of a turn-give.
Diagram	Shows the contents of the Diagram window at the time of a turn-give.
Outline	Shows the contents of the Outline window at the time of a turn-give.
Difference in text	Shows what has been added to (>) and deleted from (<) the shared text window since the previous turn-give.
Differences in notes	Shows what has been added to (>) and deleted from (<) the private notes windows since the previous turn-give.
Diagram open	Opening the Diagram window by clicking on the Diagram button.
To Diagram	Activating the inactive – but open – Diagram window.
Diagram close	Closing the Diagram window.
Diagram activities	Addition of Diagram activities below.
Diagram delete link	Deleting links between Diagram objects.
Diagram delete object	Deleting Diagram objects (text boxes).
Diagram new link	Adding new links between Diagram objects.
Diagram new object	Adding new Diagram objects (text boxes).
Diagram update object	Updating text in Diagram objects.
Outline open	Opening the Diagram window by clicking on the Outline button.
To Outline	Activating the inactive – but open – Outline window.
Outline close	Closing the Outline window.
Layout	Clicking on one of the three layout buttons to change the TC3 screen layout.
Word count	Clicking the word count button to count the number of words in the shared text.
Stop	Clicking the stop button to exit the program.

The program keeps a log file, saving type, time, content and position of all actions (keyboard strokes and mouse clicks) in the separate windows, and the chat discussion history. Specific content for deleting actions – by using the delete or backspace keys – was not recorded. Also, the cumulative contents of the chat history, shared text, private notes, Diagram and Outline windows were recorded at

each turn give and upon stopping the program. The log files can be used to replay all keystrokes and thus the full collaboration between the students. For our analyses, the log files were converted into activity and dialogue protocols, or Integrated Activity Protocols (IAPs). Special software was developed to pre-sort and interweave the data so that it could be imported into our coding and analysis program MEPA. A sample of an IAP can be found in Appendix 3. The IAPs recorded 35 different event types, as defined in Table 2.3. Unfortunately, due to technical difficulties, not all event types were logged during the first experimental survey. The following categories were not logged properly: name, layout, turn-ask, turn-give, stop (was not logged for the Control group either), and to-chat-history. Also, a maximum of two minutes at the ends of some of the protocols were not logged. Some of the turn-gives could be inferred from the rest of the protocol.

Technical aspects

The first version of the TC3 program (CTP Collaborative Text Production) was developed in 1996. The ideas stemmed from a longstanding interest in all forms of collaborative learning. In the experiments subjects were asked to collaboratively create a text based on different types of information. Few people really enjoy writing assignments. One of the most surprising results was the very favorable evaluation of this type of collaborative text production by the subjects (university students). This result was a major factor in pursuing further research into this type of groupware program. Over time the TC3 program has been redesigned and reprogrammed several times.

The TC3 program

The first version of TC3 was programmed in 1996 using Visual Basic (Microsoft). Although this version performed adequately, it was decided to develop a new version in Delphi, a Pascal variant (Borland). There were a number of reasons for this transition. Delphi allows generation of a standalone executable. The Visual Basic version required an installation procedure for every computer set to run TC3. At the time Delphi generated faster code and gave the programmer more control over network related issues.

When the COSAR project started it was decided to rebuild the entire program to meet the specific demands of the project. Since the TC3 program had been developed by just one programmer, we tried to hire an extra programmer to speed up the process and to be less dependent on one person. However, our attempts to hire extra programmers failed. Eventually we found assistance in a company called

EQUI4 software in Utrecht. This company introduced us to new technologies and new approaches to networked applications. The first major decision was to change the programming language. We chose TCL/TK, a scripting language developed by John Ousterhout at Berkeley University. This language is a merge between the Tool Command Language and the ToolKit, a set of tools to create user interfaces. A number of advantages are mentioned on the website dedicated to this tool:

- Rapid development: since TCL is a scripted language the program and variables can be inspected and changed at runtime.
- Graphical user interfaces: the ToolKit facilitates easy creating of sophisticated user interfaces.
- Cross-platform applications: programs run with little or no modification on Windows, Unix an Mac's.
- Extensible applications: A large number of additional libraries exist that can be added to programs.
- Flexible integration: TCL can be interfaced to nearly all programming languages.
- Ready for the enterprise: it is a mature system.
- Testing: because of the interpretative nature of the language testing is facilitated.
- Easy to learn: Anyone familiar with other programming languages will agree.
- Network-aware applications: Networking applications are easy to create and maintain.
- The Tcl community: A number of websites and newsgroups exist that create an effective platform for communicating problems and asking questions.

The TCL/TK language is open source; the source code is free and available for downloading. There is a large and active group of programmers that use this language. A substantial library of applications, libraries and tools exists that can serve as a framework for applications. This free tool has a level of support and documentation that many commercial packages can only dream of.

In order to minimize problems in schools we tried to keep the installation procedure very simple. The system had to be robust in the sense that errors should not result in the loss of work or data. The system consists of a server and a number of clients. The server coordinates the exchange of information between the clients and stores information and logfiles. No information is stored on the client computers. In the first version of TC3 the clients used a separate channel to exchange information (see Figure 2.4 left). This scheme was abandoned in the later release in favor of a scheme based on shared TCL arrays to communicate changes to the clients (see Figure 2.4 right).

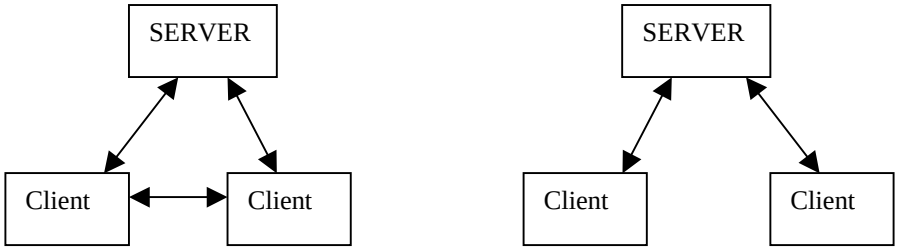

Figure 2.4: Left: initial version of TC3; Right: final version of TC3.

The schools

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.

The operating systems we encountered were Window95, Windows98, NT-workstation and Windows2000. All schools used Netware as their primary networking environment. Schools varied in the way they were connected to the Internet. In the Netherlands all primary and secondary schools will eventually be connected to the Internet by a government initiated project called “Kennisnet” (Knowledge net). Kennisnet will act as an Internet provider and all connected schools will be integrated into a national network of schools. This national network is isolated by means of a firewall from the rest of Internet mainly for security reasons. At the time of our experiments some of the schools were not yet connected to Kennisnet and used leased lines or ISDN connections.

In total, six schools participated in the COSAR project. After the Dutch language section of a school agreed to participate in the COSAR project we contacted the network administrator in order to install TC3 on the network. The network administrators in all schools were very cooperative. We usually installed the program two weeks before the start of the experiment in a particular school.

During our first tests we discovered that Kennisnet blocked the communication between the clients and our server at the university. While studying the documentation of Kennisnet we also found that schools depending on their requirements, were connected with different bandwidths. Since this would

introduce a large difference in the speed of communication between the clients and our server in the different schools we decided to hook up our server to the local network. This way we achieved a uniform response time for the different schools. We used a Linux server in the schools. Since TCL is platform independent the same version of the server will run on all versions of Windows and Linux.

The installation procedure of the program on the client computers allowed quick updating of the program in case of changes to the software. A directory was created on the schools server that contained a batchfile that copied the required files to a directory on the student computers. Then the program would be started from within that batchfile. A link would then be created in the menu system that executed the batchfile. Whenever we created a new version of our program only the files on the network needed to be replaced. Whenever students started the program a fresh copy of the necessary files was made from the network. The total amount that had to be copied per client was less than 1 MB.

Errors, mishaps and bugs

During our experiments we experienced a number of problems. We will provide a short description of the problems and the way we resolved these issues. In the first version of our program we used the FTP protocol to store and retrieve data on the server. During all our tests this worked flawlessly. In our first real experiment we arrived early in the labs to start up the computers and the TC3 program to save time. After the first break the students returned almost simultaneously. When they started the program many failed to connect to the server. We were forced to abandon that session. On inspection of the system logfile on the server the cause became clear. The system monitors the number of FTP connection made per minute. If this number reaches a certain level the system will assume a denial of service attack or a program going berserk. Then it will block any attempt to use the service for 10 minutes. The remedy was easy; the level can be set in the configuration file. Setting this number to a higher level solved the problem.

In one school we had great difficulty installing the system. After a day of searching we found that the school used a content filter for all Internet traffic on the network. This content filter replaced unwanted words by space characters. Two of the censored words were “chat” and “tequila”. These words were also used in the program as commands or as names of packages that were used by the program. On the receiving side these commands were replaced by spaces. The interpreter could

not understand the commands that were sent and terminated the program with an error message. The solution was to disable the content filter on the computers.

In one school there was a large difference in computer speed between the two computer labs we used. Several times the slower computers would fail to connect. We resolved this issue by first starting the program on the slower computers. When they were waiting for the other computers to connect, the program would be started on the faster computers. This way the connecting process worked fine. Before the next session we traced the error and removed this bug.

Whenever we set up our server in a school we require a unique IP number on that particular network. It is comparable to a phone number. Clients can contact the server through this number. It is absolutely essential that this number is unique on the network. In one of the schools after working for about one hour clients suddenly started to abort with error messages. After running some tests on the network it became clear that the IP number that was given to us was also in use by another computer that had been switched off during our tests. This computer received the same messages as our server and replied with errors because it was unable to respond in the correct way. The cure for this problem was to change the IP number to an unused one and restart the server and the clients.

After working for several hours in a school pupils experienced a sudden drop in response time of the program. This problem got worse and worse until clients started to abort with error messages. We went to see the network administrator to inquire about the state of the network. It turned out that the network administrator had started a process called “ghosting”. This is a process where a complete image of a harddisk is sent to several machines simultaneously in order to restore them to a particular configuration. This process sends tremendous amounts of information over the network. Since the amount of information one can send over a network is limited it effectively blocked all other network traffic. This disrupted the communication between our clients and the server to the extent that the clients aborted. Terminating the ghosting process solved the problem.

After TC3

During our experiments in the schools several teachers inquired whether they could continue to use the program. Our primary objective was to create a program that would enable us to answer the research questions. Since one of the programmers was always present during the experiments the user interface on the

server side is nonexistent. One of the schools was very anxious to obtain the program. We have provided them with the basic information needed to conduct a session in a 3-hour meeting. We have recently been informed that they have been successful in using the system. This means that the program can safely be used after a short instruction on installation. Several research projects in our department (the SCALE project and several PhD students) are using variants of the TC3 program in their research.

Many subjects in the school indicated that they perceived the TC3 program as a very useful tool. They had several suggestions to increase the utility of the program. The program should be accessible from any place at any time. They found it a nuisance that both students had to be logged on in order to change the text. They also wanted to be able to copy from the sources.

In a new project called PRO-ICT (NWO number 411-211-11) a more extended environment is being created. This program is designed to aid in the writing of research reports during the last two years of secondary school. Subject areas include history and geography. In this environment small groups of students can collaborate both synchronously and a-synchronously. Their teacher can monitor their progress and intervene by sending messages to the groups. The server will be accessible from within the schools as well as from their computers at home.

2.3.3 The 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.

2.4 Procedure

The research was done in two studies. The control study with the basic TC3 environment was conducted from October to December 1999 and from January to March 2000 in two schools. The experimental study, in which different combinations of the planning tools were added to the basic environment, was conducted in two surveys: from October to December 2000, and from January to March 2001. The Advisor was added only in the second experimental study, to all its conditions. In all studies, the students were first given oral instructions.

Table 2.4: The typical order of administration.

Activities	Duration
Pre-tests (only before first task)	15 minutes per test
Group instructions on task and software	10-20 minutes, depending on condition
Assign pairs and startup	10 minutes
Collaborative writing task	4-5 hours
Individual evaluation questionnaire	5-10 minutes

Scheduling constraints at the schools lead to differences in timetables between the groups. All groups started their first survey by taking the two pre-tests and receiving instructions on the task and the TC3 environment. The writing task was scheduled for one day or two consecutive days for some groups, but most groups completed their work during their Dutch classes, which meant that their 4-5 hours were spread over up to 6 sessions and several weeks.

2.5 Methods of analysis

2.5.1 MEPA: A tool for Multiple Episode Protocol Analysis

The purpose of MEPA (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[1]. Examples of suitable data within education are class discussions, collaborative discussions, teaching conversations, thinking-aloud protocols, e-mail forums, electronic discussions and videotape transcriptions.

[1] 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).

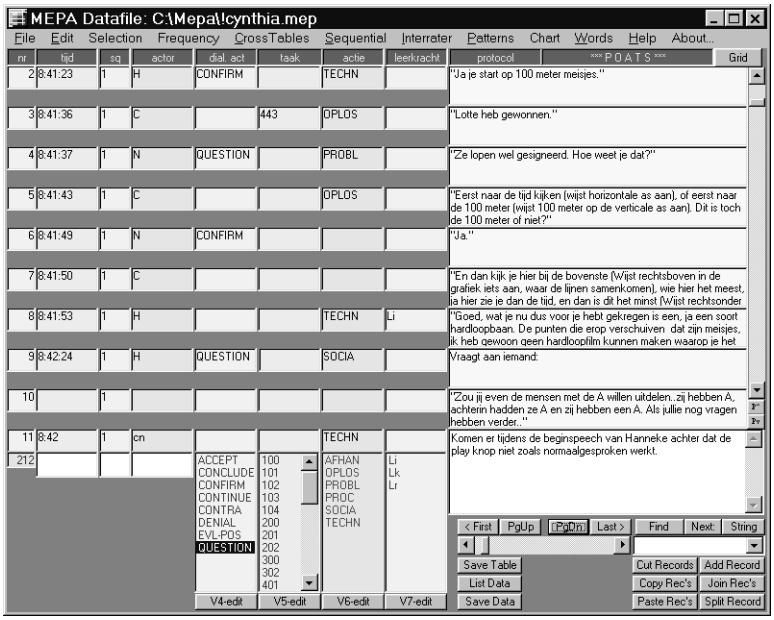

Figure 2.5: Screen dump of a MEPA file.

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, 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 structured if-then rules for automatic coding. Figure 2.5 shows a screen dump of the MEPA program.

2.5.2 Phases of the writing process

Some of the research questions were answered not only for the entire protocols, but also for three phases of the writing process. As it was impossible to pinpoint natural transitions of different activity phases in the writing protocols, the three

phases were determined as follows. There are two points in the writing process that can be clearly distinguished: the first draft and the final draft. The first draft in our definition is the point in the chat and activity protocol where the participants change turns after writing in the shared text for the first time – this is when the first draft is logged by TC3. The final draft is found at the last protocol line. We have used the first and final drafts as anchors to roughly mark out three phases of the writing process. The first phase refers to the chat before writing the first draft, and so reflects the preplanning phase. The middle draft – and so the transition between the second and the third phase – is found halfway the starting time of the second phase and the end time of the protocol. The third phase is from the middle draft up to the final draft. We expect the second phase to contain more activities related to formulating the shared text, and the third phase to contain more revision activities. The last two phases are always equally long, whereas the first phase usually lasts a lot shorter, as the participants tended to start writing quite soon after starting on the task.

2.5.3 Activity analysis: TC3 tool use

The TC3 program automatically logged all activities of the participants: every mouse click, and keystroke in every tool and window was saved in a chat and activity protocol. Unfortunately, due to a program bug not all categories were logged automatically during the second survey. This was discovered and corrected before the third survey, but for better comparison of groups from different surveys we only analyzed the categories that were logged in all surveys. These activity categories are explained in Table 2.5. The activities that were not logged properly were clicking the layout buttons, clicking into the scrollable chat history, and asking and giving turns through the traffic light. Although *layout* and *to chat history* were lost completely, most of the turn changes could be recovered by determining which partner wrote in the shared windows. After all, only the partner with the green traffic light would be able to do so.

Table 2.5: Activities logged in the protocols for frequency analyses.

Activity	Description
To chat	Clicking into the chat window.
Chat	Entering chat into the chat history by pressing Enter.
To source	Clicking into the source window.
Mark source	Clicking the mark source button.
To notes	Clicking into the private notes window.
To text	Clicking into the shared text editor.
To assignment	Clicking into the window containing the assignment.
To manual	Clicking into the window containing the TC3 manual.
To Diagram tips	Clicking into the window containing the Diagram tips.
To Outline tips	Clicking into the window containing the Outline tips.
Word count	Counting the number of words in the shared text with the word count button.
Stop	Clicking the stop button to exit the program.
Diagram open	Opening the Diagram window by clicking on the Diagram button.
To Diagram	Activating the inactive – but open – Diagram window.
Diagram close	Closing the Diagram window.
Diagram activities	Sum of all Diagram activities mentioned below.
Diagram delete link	Deleting links between Diagram objects.
Diagram delete object	Deleting Diagram objects (text boxes).
Diagram new link	Adding new links between Diagram objects.
Diagram new object	Adding new Diagram objects (text boxes).
Diagram update object	Updating text in Diagram objects.
Outline open	Opening the Diagram window by clicking on the Outline button.
To Outline	Activating the inactive – but open – Outline window.
Outline close	Closing the Outline window.
Total no. of acts	The total number of activities in the protocol.

Table 2.6: Duration variables for tool use analyses.

Activity	Description
In chat	Mean duration of a chat session.
Consists of	To chat; Chat
In source	Mean duration of a source reading session.
Consists of	To source; Mark source
To notes	Mean duration of typing or reading in the private notes window.
To text	Mean duration of typing or reading in the shared text window.
In instruction	Mean duration of an instruction reading session.
Consists of	To assignment; To manual; To Diagram tip;s To Outline tips
In Diagram	Mean duration of a Diagram session.
Consists of	Diagram open; To Diagram; Diagram delete link; Diagram delete object; Diagram new link; Diagram new object; Diagram update object
In Outline	Mean duration of an Outline session.
Consists of	Outline open; To Outline
Mean duration per activity	The mean time spent on each main interval activity.

In addition to the percentages we also analyzed the protocols for time spent on each tool, that is, the duration of the activities. Only the meaningful activities were included in these analyses, that is, simple mouse-click activities – e.g., *stop*, *word count* – were left out, while interval activities were included. The interval activities

were then grouped and reduced to six main categories: *in chat*, *in source*, *to notes*, *to text*, *in instruction*, and *in Diagram*. Table 2.6 shows the categories for tool use duration. The summary categories *in chat*, *in source*, *in instruction*, *in Diagram* and *in Outline* were calculated by adding up the durations of consecutive sub measures of the variable. For example, a chat session would start by clicking into the chat window, and could consist of multiple messages, marked by multiple strokes of the Enter key. The *to chat* duration and the consecutive *chat* durations were added up to form one chat session.

2.5.4 Chat analysis: Planning and executing through Task acts

As the only means of direct communication between the collaborating participants is the chat facility, the data captured in this window will no doubt contain valuable information about the writing process and the collaboration between the students. In addition to text content, the participants also discuss their writing strategies, such as planning and revision. This category of information was conceptualized as Task acts.

The Task act coding systems of Baltzer (1989) and Breetvelt (1991), and indirectly of Hayes and Flower (1980), lie at the basis of our system of analysis. We adapted the frameworks for analysis of collaborative data, as these models were originally intended for analysis of writing tasks for individuals. We had to take into account the influence of social communicative and coordinating aspects of the discussion: collaborating students do not just communicate task related information, but also try to get to know each other better and exchange non-task related information. In addition, our participants had to negotiate turn taking, and unlike Baltzer and Breetvelt we did not get our data from thinking-aloud-protocols, but from full written discussion protocols and additional information from the other TC3 windows. In this study, then, participants did not speak, but typed. As typing is generally slower than speaking, our protocols may contain less detail than thinking-aloud-protocols. However, as the activities were logged as well, the writing process can be reconstructed afterwards on the basis of the explicit communication in the chat.

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 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.

After dividing the chat protocols into episodes, the Task act coding plan was further developed. It did not prove possible to code the Task acts automatically in MEPA as we did for some of the other analyses. The Task acts were subdivided into 4 main categories:

- the planning level
- the executing level
- the writing level
- the non task level

The writing level consists of the actual writing and revision activities in the shared text editor, and is thus not included in our chat analyses of Task acts. The other three levels were further divided into 27 categories: 14 for planning, 11 for executing, and 2 for non task. The categories are described in Table 2.7. Task acts at the planning level refer to all utterances in which participants plan, propose or discuss future actions with regard to writing the text. In general, then, the planning level refers to metacognitive writing strategies. Task acts at the executing level are all utterances that are concerned with specific contents of writing. At the non task level, chat on technical aspects of the program is distinguished from socially oriented chat.

The categories Layout, Coordination, Alternate turn, and Experimenter do not exist at the Executing level, because these categories are not concerned with formulating. The first three of these only take place at the metacognitive level. Talking about remarks from the Experimenter is not concerned with formulating either, as it is simply taking note of utterances.

The category Execute Count does not exist at the planning level, because discussing the number of words in the shared text resembles a subcategory of Execute Goals. However, as discussing the word count is not very closely related to the content of the shared text (as opposed to Execute goals), but is task related and is not a planning activity, a separate category was made of it at the executing level.

Reliability analyses showed the Task act coding to be relatively reliable with a Cohen's Kappas between .57 and .64 and an interrater agreement percentage between 61% and 69%.

Table 2.7: Task act definitions: Planning, executing and non task.

Category	**Description**
Plan Advisor	Planning the use of the Advisor tab.
Plan turn alternation	Coordinating turn taking.
Plan coordination	Planning time and activities of interaction without going into detail.
Plan Diagram	Coordinating the use of the Diagram and asking for general feedback.
Plan Diagram layout	Planning the layout of the Diagram.
Plan external source	Planning the use of sources not given within TC3, without going into detail.
Plan goals	Discussing the task demands and goals.
Plan knowledge	Planning personal knowledge, experience, or opinions not stated in the sources, without going into detail.
Plan layout	Planning the layout of the argumentative text and the order of the units of information.
Plan notes	Planning and coordinating note taking without going into detail.
Plan Outline	Coordinating the use of the Outline and asking for general feedback.
Plan Outline layout	Planning the layout of the Outline.
Plan revision	Proposing and coordinating revision of the shared text.
Plan revision Diagram	Proposing and coordinating revision of the Diagram.
Plan revision Outline	Proposing and coordinating revision of the Outline.
Plan source	Planning the use of sources (including the assignment and given sources) without going into detail.
Plan text	Planning the main outline of the shared text without going into detail.
Execute Advisor	Discussing the contents of the Advisor tab.
Execute word count	Counting the number of words in the shared text.
Execute Diagram	Discussing specific contents of the Diagram.
Execute Diagram layout	Discussing specific layout of the Diagram.
Execute external source	Discussing specific contents of external sources.
Execute goals	Discussing the demands and goals for the contents of the shared text.
Execute knowledge	Discussing specific contents of personal knowledge, experience, or opinions not stated in the given sources.
Execute notes	Discussing specific contents of notes taken.
Execute Outline	Discussing specific contents of the Outline.
Execute Outline layout	Discussing specific layout of the Outline.
Execute revision	Discussing and executing revision of specific parts of the text.
Execute revision Diagram	Discussing and executing revision of specific parts of the Diagram.
Execute revision Outline	Discussing and executing revision of specific parts of the Outline.
Execute source	Discussing specific contents of sources.
Execute text	Discussing specific text and asking for feedback on contributions to the shared text.
Non task program	Discussing technical aspects of TC3 and use of the program.
Non task social	Discussing non task matters, mainly social talk.

2.5.5 Chat analysis: Coordinating through Dialogue acts

The Dialogue act coding indicates the communicative function of an utterance. The Dialogue acts were based on the VOS system (Erkens, 1997), and were mainly derived from discourse markers. Discourse markers are characteristic words showing the function of the phrase in a dialogue (Schiffrin, 1987). The coding system distinguishes between five communicative functions, that can be further subdivided into the Dialogue acts. Table 2.8 shows these communicative functions, and their Dialogue acts with specifications and explanations.

The Dialogue acts only concern the chat events that were coded as planning or formulating on the Task acts, that is, the non task chat episodes were excluded from the communicative analysis. This was done because we assumed that the structural features of the Dialogue acts within the non task episodes would not influence the final product. For example, the argumentative structure of social talk does not appear to influence the quality of the final text. Leaving the non task episodes in would result in a distorted image of the relevant coordination structures. A further explanation of the main communicative functions is given below.

Argumentatives are utterances indicating a line of argumentation or reasoning. Reasoning is used to clarify, but also to convince the partner. We distinguished 6 different argumentative types of Dialogue acts: Argumentative Reason indicates a reason, cause or ground; Argumentative Contra indicates an objection or counterargument; Argumentative Conditional indicates a condition or stipulation; Argumentative Then indicates a consequence or result; Argumentative Disjunctive indicates a disjunctive; Argumentative Conclusion indicates a conclusion.

Responsives are mostly answers to questions and proposals, but they can also be reactions to other utterances from the partner. Reactions to the partner can be affirmative (Responsive Confirm), negative (Responsive Deny) or accepting (Responsive Accept). Responses to proposals and questions can also be affirmative (Responsive Reply Confirm), negative (Responsive Reply Deny) or accepting (Responsive Reply Accept). In addition, responses to questions can be a statement (Responsive Reply Statement) or a performative – an action performed by saying it (Responsive Reply Performative).

Informatives serve to transfer information. Information can be transferred through a performative (Informative Performative) or through evaluative remarks that can be neutral (Informative Evaluative Neutral), positive (Informative Evaluative Positive) or negative (Informative Evaluative Negative). Information can also be transferred through a statement (Informative Statement); this statement can indicate an action (Informative Statement Action), but it can also be a social statement (Informative Statement Social) or contain nonsense (Informative Statement Nonsense).

Table 2.8: Dialogue acts.

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
	Evaluation	Neutral	Neutral evaluation
Transfer of information		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
		Set	Set question/ multiple choice
Questions or utterances requiring a response		Open	Open question
	Proposal	Action	Proposal for action
Imperatives		Action	Order for action
Commanding utterances		Focus	Order for attention

Elicitatives are questions or remarks requiring a response. We differentiated between three types of questions: open questions (Elicitative Question Open), set questions (Elicitative Question Set) and verifications – yes/no questions – (Elicitative Question Verify). The utterances that require responses are proposals for action (Elicitative Proposal Action).

Imperatives are commanding utterances. We distinguished between two types of imperatives: commends to take action (Imperative Action) and remarks to draw the attention of the partner (Imperative Focus).

The Dialogue act coding of the protocols was done automatically with the help of MEPA. In the program a filter file was made that could label the chat utterances with the Dialogue acts. A filter is like a sieve that sifts the protocols for typical words or phrases through if-then rules. For example, if a line in the chat protocol contains the word *because*, then it should be coded as 'argumentative reason'. The filter file for the Dialogue acts contained more than 700 of these if-then rules. The rules are applied to the protocol one by one in fixed order, so that a hierarchy could be imposed on the communicative functions: an argumentative is a more informative coding than an informative statement. Lines containing markers of multiple communicative functions could thus be coded as the most important of the possible Dialogue acts. Argumentatives came first in this hierarchy, followed by elicitatives, responsives, imperatives, and finally informatives.

Our Dialogue acts filter file could in theory be used for other types of discourse, although it was adapted to the sociolect of secondary school students. A separate section of the filter file also contains filters for the specific contents of the assignments: words like cloning and transplant were also coded automatically by adding a C for content to the general dialogue code.

With the filters, some 80 to 85 percent of the protocol lines were coded automatically. The remaining lines were coded as informative statements with a question mark (InfStm?), as we assume that this is the most probable code for non coded lines. A random check by a linguist and an educational scientist showed that over 90 percent of each protocol was coded correctly by the filters. The categories that proved to be the most faulty were the non-directed informative statements (InfStm?), the informative statements of content (InfStmC), and remarks and questions starting with *what* but ending without a question mark. This last category finally received the code EliQstOpn?. This rule was added to the final

version of the filter file. Approximately 75 percent of these three codes were manually changed to InfStm and EliQstOpn, respectively. The reliability of the automatic coding filters is naturally high, but the manual correction makes the procedure slightly less reliable.

Transition probabilities

To determine the structure of the dialogue, we used lag sequential analysis (see Wampold, 1992) in MEPA to make transition diagrams showing the significant consecutive Dialogue acts. For this analysis, the categories of Dialogue acts analyzed were reduced from 28 to 21 as shown in Table 2.9.

Table 2.9: The recoding for the transition diagrams.

Original code	**Changed to**
Responsive reply confirm	Responsive confirm
Responsive reply deny	Responsive deny
Responsive reply accept	Responsive accept
Responsive reply performative	Informative performative
Informative statement nonsense	Informative statement social
Informative statement action	Elicitative proposal action
Imperative focus	Imperative action

This analysis was used to compare the experimental conditions and the high and low performing dyads. For the condition analysis we randomly picked five dyads from each condition with a mean text score between 5.5 and 7.0. For the analysis of high and low performing pairs we selected dyads from the Control group only with a mean text score lower than 5.5 or higher than 7.0. We only used the Control group for this analysis, because this is the largest group, so sampling from it is more reliable than sampling from a smaller group. In addition, the analyses of dialogue structure per condition showed that the Control group had the least structured chat. Any differences between high and low performing dyads would therefore be most obvious for this group.

Equality of contribution

Equality of contribution to the chat dialogue is determined by comparing the proportions of contribution of the individual collaborative partners, resulting in a measure of asymmetry. A high score on this measure indicates strong asymmetry, or inequality of contribution. For this analysis, all Dialogue act codings are reduced to the five basic communicative functions – argumentative, responsive, informative, elicitatives, and imperative – and the percentages of contributions per communicative function and in total are determined for each participant.

Coordination processes

Focusing, checking and argumentation cannot be observed directly: these aspects of coordination are measured through indicators, represented by specific Dialogue acts. These Dialogue acts possibly indicate when a coordination process is taking place. The variables focusing, checking, and argumentation were obtained by adding up the percentages for the indicators in the chat protocol. The indicators are show in Table 2.10.

Table 2.10: Indicators for focusing, checking, and argumentation.

Focusing	**Checking**	**Argumentation**
Elicitative proposal for action	Elicitative question verify	Argumentative reason
Elicitative question open	Elicitative question set	Argumentative contra
Imperative action	Responsive confirm	Argumentative conditional
Imperative focus	Responsive deny	Argumentative then
	Responsive accept	Argumentative disjunctive
		Argumentative conclusion

2.5.6 Chat Analysis: Argumentation

In addition to the communicative function of argumentation, we wanted to take a closer look at the content of argumentative episodes. The episodes determined by content were coded with the categories shown in Table 2.11. An argumentation episode was assumed on the basis of two criteria:

- The episode must contain argumentation, that is, a students attempts to convince the partner with reasons.
- The partner responds to the reasoning at least once during the episode, that is, there must be some sort of dialogue.

The chat protocols of 17 dyads from the Control group were used for these analyses. An example of coded chat is shown in Table 2.12. The data were coded by four people, and the initial reliability results were very daunting: the Cohen's Kappas for two protocols were only .25. It was then decided to work with teams of two coders, where each protocol was coded by one person and checked and revised by the other. The interrater reliability between two teams then rose to Cohen's Kappas of .88 and .86 (so on two protocols coded two times two).

Table 2.11: Types of argumentative episodes in the chat.

Topic of the argumentation	Description
Content	The argumentation contains support or refutations of opinions on the topic (cloning or organ donation). The aim is to reach agreement on the main position taken in the shared text.
Coordination	The argumentation is aimed at establishing a division of tasks between the partners (who does what, and when?).
Metacognitive strategy	The argumentation is concerned with the characteristics of an argumentative text, and with the way in which it should be written ("How do we best compose this text?", "Where shall we put this?").
Technical aspects	The argumentation is aimed at solving problems with the computer hardware or software.
Miscellaneous topic	There clearly is an argumentative episode, but it does not fit in any of the above descriptions, and does not relate to the writing task as such. Often, the argumentation is concerned with non task matters ("What shall we do after school today?")

Table 2.12: Example of coded chat.

Line	Time	Actor	Episode	Protocol
127	00:37:12	1	Coordination	could you start with thinking of arguments pro and con?
128	00:37:27	0		ok
129	00:37:31	1		so we can exchange them later
130	00:38:36	1		would you please click the traffic light then??
131	00:38:37	0		if you type in the position and the introduction now, I will try to find arguments.
132	00:38:38	1	End	ythanl yi
133	00:39:56	0		by the way, how many words do we need?
134	00:42:23	1	Metacognitive	do you know how to formulate the position?
135	00:46:17	0		the position says whether we're pro or against, isn't it?
136	00:46:35	0		what you did looks fine to me.
137	00:46:47	1		yes, ~
138	00:47:11	1		but how do you formulate it. ~
139	00:47:11	1		we're not allowed to just say I'm pro or against, ~
140	00:47:11	1		we had to formulate it differently.
141	00:47:11	1		ok
142	00:47:27	1		against:
143	00:47:47	1	End	I have that as arguments contra, and you?
144	00:47:48	0	Metacognitive	so now we can start with the arguments
145	00:48:25	0		cloning people is ridiculous ~
146	00:48:49	0		because there are so may of them already.
147	00:48:56	0		what do you mean with general uneasiness.
148	00:49:05	1		if most people are negative about ~
149	00:49:27	1		cloning, ~
150	00:49:27	1		then why continue doing it
151	00:50:50	0		ok ~
152	00:50:54	1		I am busy
153	00:51:03	0	End	I guess we should type this in the text then.

2.5.7 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 original measurement system for the argumentative texts was based on Van Cuilenburg, Kleinnijenhuis and De Ridder (1988) and on Schellens and Verhoeven (1994). Unfortunately, due to a change in personnel, the texts of the experimental groups were coded and assessed by different researchers than the texts of the Control group. The new researchers could not reach satisfactory levels of interrater reliability for the original coding system, which forced us to simplify it. Although the coding system was changed, the assessment method was left intact, and T-tests of Control group texts coded in new system with their counterparts coded by the original researchers in the old system, showed no significant differences for the final assessment. The coding categories for the Control group are shown in Table 1.1 in Appendix 4, and those for the experimental groups are shown in Table 1.2 in the Appendix.

One of the aims of the coding and assessment was to determine the complexity of the argumentation. We found the original coding system too intricate for the texts we were coding. It assumed a high level of complexity in the texts and it assumed that learner writers use all argumentative functions and use them correctly. Also, the terminology was not completely unambiguous. The system was simplified to solve these problems, leaving the main ingredients of argumentation intact. In addition, the new system better reflects the possible structures of the Diagram tool, thus forming a closer match with the participants' frame of reference. Table 2.13 shows the changes to the coding system for the experimental groups.

Table 2.13: Transfer from Control group coding system to experimental group coding system.

Old system	New system	Justification of changes
Part of argument	-	This is a general term, and was not used for coding the Control group either.
Claim	Position, argument pro/contra	The old system did not allow us to distinguish the main position from the subordinate claims, whereas the assessment form did mention the overall position separately.
Part of argument & claim	-	By introducing argument pro/contra and slightly changing the definitions of support and refutation, the double coding was made redundant.
Conclusion	Conclusion	-
Solution	Conclusion	Solutions were rarely explicit in the texts. Where they were present implicitly, they functioned as conclusions.
Support	Support	In the new system, support only functions at the third level. At the second level it is replaced by argument pro.
Put in perspective	Refutation	Participants at this level rarely, if ever are this subtle in their argumentation. The code was therefore difficult assign and mostly superfluous.
Refutation	Refutation	-
Organizer	Organizer	-
Information	Information	-
Elaboration	-	Any part of argumentation that stretched over multiple phrases was continually coded as that part of argument.
-	Title	This was mentioned in the assessment form, but not in the old coding system.

After coding the argumentative functions of the phrases, the texts were assessed using the assessment form and instructions shown in Appendix 4. The assessment contained four summary measures: overall textual structure, quality of argumentation at segment level, quality of argumentation for the text as a whole, and audience focus. The scores were converted to a 10-point scale, and a mean was computed, functioning as an overall text quality score. The interrater reliability for the final measures was very high, with correlations between two independent raters for the five text scores on five texts ranging from .71 to 1.00 ($p < .01$).

Table 2.14: Descriptions 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.

2.5.8 The planning product: Analysis of the Diagrams

The Diagram tool was designed to give insight into the relationship between conceptualization in three phases of the writing process and the quality of the collaborative writing process, as measured by the final product – an argumentative text.

All changes made in the Diagram were logged in the protocols. Some of these, like the type of box and its contents, can be viewed and analyzed in MEPA. Other data, such as the relations between the different boxes, can better be viewed in TimeDump – the visual playback program designed for this purpose. This program allows us to view the Diagram at any given point in the writing process.

On the basis of thorough analysis of the information sources two automatic coding filters were developed – one for each topic – that could sift out the 160 different arguments – deduced from the sources – from the diagrams and the texts. New arguments that could not be traced back to the sources but that were used regularly by the participants were included in the filters and coded as generated by the participants themselves. The filter for organ donation coded 77% of the data, and the cloning filter 80%. Of course, 100% accuracy is not attainable, as participants occasionally generate wholly original arguments. The same filters were used for the diagrams and the texts. A sample of the filter for the topic Cloning can be found in Appendix 9.

After coding, the frequencies of the arguments were categorized on two dimensions: frequency of argument types and correspondence to the final text (see Table 2.15). The arguments from the given sources were divided into arguments on organ donation and on cloning, so we could determine the influence of the topic on the measures of interest.

Table 2.15: The two sets of argument labels for content analysis.

Frequency	**Correspondence**
Total no. of arguments	Arguments in diagram only
Self-generated arguments	Arguments in text only
Arguments from given source	Corresponding arguments
Arguments on organ donation	
Arguments on cloning	

The measure *arguments in diagram only* gives the number of arguments present in the diagram, but not in the text, whereas the measure *arguments in text only* gives the number of arguments found in the text that were not present in the diagram. The measure *corresponding arguments* gives the number of arguments that were found both in the diagram and in the text.

In addition to the MEPA analyses, the visual Diagrams were examined to get a view on their structure. This resulted in the variables mentioned in Table 2.16.

Table 2.16: The two sets of argument labels for structure analysis.

Frequency of object types	**General structure**
Argument contra	Total no. of elements
Argument pro	No. of content objects
Conclusion	
Information	
Position	
Refutation	
Support	
Arrow	
Line	

2.5.9 The planning product: Analysis of the Outlines

To answer the research question on linearization, the Outline products were analyzed for content and structure and compared to the final texts. Use of the Outline tool as found in the tool use analysis was also taken into account. We hypothesize that the use of the Outline enhances coordination in the collaboration between students during writing, and that effect of proper use of the Outline tool (that is, as a planning tool) on the final product will be positive for the persuasiveness of argumentation and adequacy of language use, for example through conjunctions and anaphora (Chanquoy, 1996).

A total of seven different measures were used to describe the outlines and to compare them to the argumentative texts. An overview of the measures and sub measures is given in Table 2.17. Making sure the two scoring systems were compatible, we adapted the coding instrument for the argumentative texts for identifying the structural complexity of the outlines (formal structure and argumentative structure). We developed two measures for determining the complexity of the Outline contents: abstractness of content and phrase complexity. In addition to the complexity of the outlines, the structural and content correspondence of the outlines to the texts was determined. The structural

correspondence was determined by comparing the order of the elements in the outlines to their order in the text. Content correspondence was assessed by determining whether all items from the outline were present in the text, and by determining whether all paragraphs from the text were present in the outline. A full description of the coding and assessment criteria is given in Appendix 5.

Table 2.17: Measures for outline complexity and correspondence to text.

	Structure	**Contents**
Complexity (vs. expert model)	1. Formal structure 1.1. Number of hierarchy levels 1.2. Number of organizational items per paragraph 1.3. Number of sub items per paragraph 2. Argumentative structure 2.1. Number of argumentative lines per paragraph 2.2. Variation in argumentative types	3. Formal content: abstract/mixed/concrete 4. Comprehensiveness: phrase complexity
Correspondence (vs. text)	5. Order correspondence	6. Item correspondence 7. Paragraph correspondence

2.5.10 Student evaluations

After finishing their assignment, the students were asked to fill out an evaluation form. The forms included general evaluative questions as well as questions adapted to the experimental conditions. The participants were asked to give their opinions on the writing assignment, their experiences with TC3 and the planning tools, to state their ideas on working collaboratively, and give suggestions for improvements of the TC3 program. An overview of the questions is given in Appendix 6.

CHAPTER 3 INFLUENCE OF CONDITION ON TEXT QUALITY

The COSAR research project was set up as a process-oriented study. However, the end results – the argumentative texts – are a vital ingredient for understanding the creative and collaborative processes that are our main interest. In this section the analyses of the argumentative texts are presented and discussed. Table 3.1 shows the means and standard deviations for all conditions separately and for the sample as a whole.

Table 3.1: Descriptive statistics for text quality per condition.

Condition	N (dyads)	Textual structure		Segment argumentation		Overall argumentation		Audience focus		Mean text score	
		Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
C	39	6.76	1.13	6.19	1.36	5.75	2.37	6.20	2.10	6.22	1.43
D	17	6.71	.97	5.63	1.34	6.81	2.29	5.81	1.84	6.29	1.09
DA	26	6.03	.82	5.49	1.34	6.41	2.07	6.01	1.64	6.00	1.01
DO	23	6.44	.83	5.64	1.32	6.16	2.25	6.20	1.60	6.17	1.03
DOA	11	7.15	.88	5.42	.84	5.76	1.69	5.57	1.00	6.19	.75
O	18	6.59	1.00	5.90	1.06	5.74	1.80	6.04	1.95	6.17	.96
OA	11	6.49	.83	6.34	.94	5.76	1.52	6.59	1.90	6.38	.74
Total	145	6.56	1.00	5.83	1.28	6.06	2.13	6.08	1.81	6.19	1.11

The table shows that the scores were quite close together for all groups. Independent samples T-tests showed no differences between the two topics – organ donation and cloning – and there were no significant gender differences either. The quality of the texts was not very high: average 6.2 on a scale of 1-10 is not very impressive.

As the participants in our research were relatively novice argumentative writers, their texts were of matching quality, although most of the text scores are above 6, so they would be a pass. The main imperfections we noticed whilst coding and assessing the texts were related to the clarity of the argumentation and the textual structure. Some students hardly used any argumentation, but only summed up the facts found in the sources without relating them to a position or to each other. If there was argumentation, this was often untrue, invalid, or insufficiently supported. In addition, students often started new paragraphs at illogical points in the argumentation, or were simply sloppy in placing paragraphs. Some examples are shown in Table 3.2. Two full texts are shown in Appendix 8.

Table 3.2: Examples of ill-placed paragraph marks.

Dyad 655	[NP] A donor transplant that did not go as desired, was the case of Sjef Timp: Sjef Timp received a cornea from a donor. After the transplantation the organ [NL] did not work properly, and the eye even started to reject the new cornea. 'It became,' he says, 'all very nasty and painful.' [NP] he ended up back in hospital and several doctors made sure during weeks of treatment that the cornea was saved. His eyesight has 'never been better'. All in all this transplant did have a happy ending after all. [NP] This happened among others because of the receiver of an organ is selected on medical criteria. … Proper matching can prevent the process of rejection in the receiver. [NP] The family of the donor can then be sure that the organ of their relative ends up in the right place.
Dyad 212	[NP] We want to draw your attention to donors, because they are important. [NP] Since one year the Ministry of Public Health has a Law on organ donation, called the WOD. The donor registry is part of this and the purpose of this law is to improve things around heart transplants. This law bans organ trade and states that organs and tissue must be distributed fairly. [NP] This law deals with two problems namely: …

Analyses of the pretests showed that there were no structural significant differences between participants from the different schools. The correlations between the pretest results and the text quality measures are given in Table 3.3. There is a clear positive relation between the score on the Underline Arguments Test and text quality, although the correlations are not very high. It seems that there is some influence of argumentative skill – as measured with this test – on argumentative writing. The results for the Wild Cat Test are less convincing, although there is a negative tendency for anaphor complexity, and a positive tendency for semantic clustering.

Table 3.3: Correlation pretests vs. text quality measures.

	Textual structure	**Segment argumentation**	**Overall argumentation**	**Audience focus**	**Mean text score**
Underline arguments	.07	.13*	.14*	.13*	.14*
Exactness characteristics	-.03	.06	-.01	-.03	-.02
Linearity	.00	.04	.08	.04	.06
Anaphor complexity	.01	-.06	-.19**	-.11	-.12*
Sentence complexity	.05	.00	-.07	-.04	-.05
Semantic clusters	.12	.09	.11	.12*	.14*
Total Wild Cat	.07	.05	-.03	-.01	.00

* $p < .01$; ** $p < .05$. N = 278 participants.

Initially, we also intended to use the teacher marks for the writing products in our analyses. However, it turned out that the teachers in the different schools used very different criteria for assessing the argumentative texts, both compared to our system and to each other. For example, most teachers included spelling mistakes, which we were not at all interested in. Only in one school the teachers could give us an overview of their assessment guidelines, and these turned out to be rather too severe to our – and the students' – liking. Among others, it involved deduction of 5 points per 'wrong' paragraph – but no definition of 'wrong' was given. Moreover, when the students wrote on organ donation, they did not receive a mark above 3 (on a scale of 1 to 10) if they focused on the pros and cons of a register of donors, even though this was well within the scope of the assignment. Eventually we decided not to use these data in our analyses.

We found a few differences in a multiple comparison analysis on the conditions, although only two of the five text quality measures showed significant differences: textual structure and segment argumentation. The significant results are shown in Tables 3.4 and 3.5. The values in the matrices are the mean values of the row variable minus the mean values of the column variable. It seems that the Diagram-Advisor group had slightly lower scores on textual structure and segment argumentation, especially in comparison with the Control group, the Diagram condition, and the Diagram-Outline-Advisor condition.

Table 3.4: Mean differences between conditions for textual structure (Bonferroni).

	C	**D**	**DA**	**DO**	**DOA**	**O**	**OA**
C	-		.73				
D		-	.68				
DA	-.73	-.68	-		-1.12		
DO				-			
DOA			1.12		-		
O						-	
OA							-

$p < .05$; only significant differences are shown. Values are row label minus column label.

Table 3.5: Mean differences between conditions for segment argumentation (Bonferroni).

	C	**D**	**DA**	**DO**	**DOA**	**O**	**OA**
C	-		.70				
D		-					
DA	-.70		-				
DO				-			
DOA					-		
O						-	
OA							-

$p < .05$; only significant differences are shown. Values are row label minus column label.

In general we can say that the planning tool conditions did not have a clear positive effect on the quality of the resulting texts in comparison to the texts written by the participants in the basic TC3 environment, the Control condition. In fact, the Diagram-Advisor condition seems to have had a negative effect. However, we must not confuse the availability of a planning tool with the proper use of it.

CHAPTER 4 ACTIVITIES OF KNOWLEDGE CONSTRUCTION

4.1 Tool use

One of our major interests was to investigate the effect of communication, writing and planning tools on the writing process. In this chapter we describe the results of several tool use analyses for the different conditions in an attempt to answer the matching research question and pinpoint potential differences between conditions: How do constructive activities differ in different planning phases: before and whilst formulating? To this question we add: What is the relationship between the activities in the three phases of the writing process and the quality of the resulting text? To answer these questions we analyzed the chat and activity protocols of 139 dyads and compared these to the each other, and to the quality of their final texts as measured by four scores and their mean, and we compared the activities in the three phases. In addition to the number of times each activity was performed, the mean duration of each activity was drawn from the protocols in MEPA.

We expected to find that the participants used the chat window frequently throughout the collaborative process, as this was their most direct means of communication within TC3. The source window (including marking the sources) and the different instruction windows – with the assignment, the program manual, and the planning tool tips – should be used less intensively towards the end of the process. During the earlier stages, the participants need those windows to get to grips with the topic, the task and the software. Due to the definition of the first phase – it ends where the first draft of the shared text is entered in the protocol – we expected very little shared text activity during this phase, and the same goes for counting the number of words in the shared text. We expected more intensive use of the planning tools during the first phase for the Advisor conditions, as they were more explicitly instructed to use the tools for preplanning. As the participants were required to make their planning product match their final text, we also expected to find more intensive use of the Diagram and Outline during the last phase. We did not have specific expectations for the other activities: the notes, window and stopping the program.

For the relations to text quality, we expected to find specific relations for several categories. We expected to find higher text scores for pairs making relatively more frequent use of the sources, marking the sources, or using the private notes

windows, in particular during the first phase. The same outcome was anticipated for pairs who frequently clicked into the shared text during writing and revision in the second and third phases. We predicted lower text scores for pairs who clicked into the chat window more frequently without entering chat, exited the program more regularly, or repeatedly counted the number of words in the shared text. We had mixed expectations for the other categories (*chat, to assignment, to manual*) and also for the total number of acts. In addition to our general expectations, we expected to find a positive effect on text quality for use of the Diagram and Outline tools in the conditions containing those tools, and for the use of the tips windows in the Advisor conditions, especially during the first phase of the writing process.

For readability reasons most of the tables referred to in this chapter are shown in Appendix 7. All references to tables within parentheses refer to this Appendix. The tables in the Appendix refer to the descriptive analyses of tool use in each phase, variance analyses with Bonferroni differences in tool use between the experimental conditions in total as well as per phase, and correlational analyses of tool use and text quality per phase.

4.1.1 Tool use differences between conditions and phases

Table 4.1 shows the mean percentages of the different types of tool use in all phases together, as well as the mean total number of acts in the different experimental conditions. The standard deviations for these means are shown in Table 4.2. Note that not all activities were logged for all dyads, as they were assigned to seven different experimental conditions, so the percentages should be looked upon with care. The analyses of tool use duration were less extensive than those for the tool use percentages, as these were only used to support (or refute) the findings from the percentage analyses.

As can be seen in Table 4.1 the chat window is the most used facility in TC3 in all conditions, with percentages ranging from 37 to 51 % of all activities. Reading the sources is also a frequent activity, ranging from 12 to 16 % in the experimental conditions and 8 % in the Control condition. It is striking that the percentage of writing in the shared text (*to text*) is relatively lower in the experimental conditions (5 to 9 %) than in the Control group (13 %). The Diagram conditions seem to use the Diagram more than the Outline groups use the Outline: about 6 % (*Diagram open* plus *to Diagram*) versus less than 4 % (*Outline open* plus *to Outline*).

Table 4.1: Means of tool use percentages in all phases for all conditions.

	Total	**C**	**D**	**DA**	**DO**	**DOA**	**O**	**OA**
N	139	33	17	26	23	11	18	11
To chat	13.11	16.42	11.72	11.07	12.04	12.37	12.67	13.91
Chat	41.92	51.05	37.26	37.20	41.13	37.36	40.05	42.19
To source	12.15	7.62	15.32	12.62	12.15	12.45	15.95	13.25
Mark source	2.10	2.86	1.45	1.81	1.29	.57	3.74	2.02
To notes	3.75	1.05	4.83	4.22	4.24	3.81	4.37	6.94
To text	8.16	13.25	5.88	4.98	6.48	6.82	8.07	8.96
To assignment	1.23	1.27	1.71	.87	1.28	1.04	1.22	1.37
To manual	.38	.27	.50	.30	.46	.18	.63	.31
To Diagram tips	.42			.44		.37		
To Outline tips	.44					.25		.63
Word count	4.10	5.82	3.66	3.70	2.72	3.78	4.29	3.46
Stop	.58	.39	.72	.92	.35	.63	.46	.79
Diagram open	4.37		3.84	5.28	3.88	4.05		
To Diagram	2.16		2.01	2.81	1.66	1.87		
Diagram close	3.15		3.19	4.66	1.63	2.70		
Diagram activities within Diagram	7.85		7.92	9.13	6.75	7.04		
Diagram delete link	.27		.26	.29	.28	.24		
Diagram delete object	.93		1.10	.82	1.05	.72		
Diagram new link	2.25		1.96	3.02	1.72	2.02		
Diagram new object	1.93		2.15	2.00	1.75	1.80		
Diagram update object	2.47		2.45	3.01	1.96	2.26		
Outline open	2.48				1.70	1.95	3.70	2.65
To Outline	1.13				.85	1.19	1.45	1.13
Outline close	2.17				1.39	1.56	3.40	2.41
Total no. of acts	955.17	821.82	750.29	1076.38	1161.23	1146.55	870.78	919.91

N = number of dyads. Conditions: C = Control; D = Diagram; DA = Diagram-Advisor; DO = Diagram-Outline; DOA = Diagram-Outline-Advisor; O = Outline; OA = Outline-Advisor.

Table 4.2: Standard deviations of tool use percentages in all phases for all conditions.

	Total	**C**	**D**	**DA**	**DO**	**DOA**	**O**	**OA**
To chat	2.72	2.31	1.42	1.59	1.73	.95	2.00	2.40
Chat	10.23	8.57	10.48	9.16	8.44	6.62	8.13	9.37
To source	5.11	3.06	5.67	4.11	4.08	2.59	5.94	3.89
Mark source	2.43	3.41	1.77	1.76	.97	.75	2.74	1.73
To notes	2.65	1.25	2.38	1.99	2.47	1.14	2.06	3.30
To text	3.56	2.62	1.80	1.31	1.37	1.15	1.76	1.59
To assignment	.83	.76	.98	.58	1.03	.77	.64	.77
To manual	.36	.29	.27	.35	.39	.13	.51	.15
To Diagram tips	.27			.29		.22		
To Outline tips	.30					.21		.25
Word count	2.29	2.41	2.04	1.83	1.18	1.75	2.54	2.17
Stop	.42	.51	.35	.36	.11	.30	.27	.35
Diagram open	5.17		1.81	3.13	8.37	3.49		
To Diagram	.96		.95	.90	.68	.76		
Diagram close	2.36		1.73	2.91	1.20	1.24		
Diagram activities within Diagram	3.24		4.25	2.83	2.92	1.68		
Diagram delete link	.31		.38	.25	.36	.22		
Diagram delete object	.92		.83	.68	1.31	.35		
Diagram new link	1.34		1.67	1.25	1.01	.77		
Diagram new object	.79		.95	.72	.82	.44		
Diagram update object	1.00		1.24	.86	.83	.57		
Outline open	1.62				.98	1.02	1.57	2.04
To Outline	.66				.50	.62	.71	.71
Outline close	1.60				.93	1.04	1.48	2.06
Total no. of acts	299.35	287.26	202.15	299.04	243.11	236.92	212.81	288.09

Table 4.3 shows the means and standard deviations for the duration of each interval activity in the three different phases of the writing process. The time spent in the chat is quite constant throughout the phases with about 21 seconds per event. The sources are not only visited less and less frequently as time goes by (compare Tables 1.1, 1.3, and 1.5), but the visits get shorter as well, from 27 seconds in the first phase to 18 seconds in the third phase. On the other hand, the private notes windows are used more intensively towards the end of the writing process, and so is the shared text. As we would expect, the different instruction tabs are visited the shortest during the final phase. The planning tool windows get more attention per event towards the end of the writing process, with durations ranging from 32 and 12 seconds in the first phase to 82 and 37 in the final phase. The mean duration per activity is the shortest in the first phase.

Table 4.3: Means and standard deviations of activity duration for all conditions in the three phases of the writing process.

	1st phase		2nd phase		3rd phase	
	Mean	**SD**	**Mean**	**SD**	**Mean**	**SD**
In chat	20.26	7.53	22.61	8.70	21.87	6.56
To chat	9.62	4.49	10.15	4.44	9.29	3.81
Chat	10.64	3.86	12.46	6.21	12.58	3.71
In source	27.30	17.02	26.75	21.17	17.80	14.88
To source	16.28	10.59	15.61	8.44	14.74	11.87
Mark source	11.03	11.78	11.14	18.95	3.06	10.09
To notes	18.46	39.51	23.23	20.53	24.64	24.41
To text	16.17	14.48	29.07	15.88	30.73	14.05
In instruction	26.70	21.75	22.13	35.69	16.48	19.95
To assignment	17.92	16.59	13.63	19.68	12.22	16.79
To manual	6.37	11.47	6.78	19.62	3.48	10.72
To Diagram tips	6.00	6.62	2.54	5.19	1.59	3.81
To Outline tips	5.11	4.80	6.59	12.04	2.21	3.84
In Diagram	32.55	30.70	44.94	33.62	81.86	42.35
Diagram open	8.04	6.39	11.64	10.98	15.49	9.60
To Diagram	3.24	4.61	3.28	5.42	9.23	12.39
Diagram delete link	1.45	3.76	1.86	3.80	3.57	7.51
Diagram delete object	2.64	4.83	5.57	9.83	6.15	10.05
Diagram new link	4.82	8.32	6.35	7.60	13.98	13.64
Diagram new object	7.24	8.47	10.23	10.56	22.30	18.70
Diagram update object	5.13	8.38	5.99	5.79	11.13	11.76
In Outline	11.97	13.89	18.78	16.71	37.45	36.79
Outline open	7.44	10.06	9.71	7.20	16.59	17.05
To Outline	4.53	7.94	9.06	13.85	20.86	31.73
Mean duration per activity	12.64	4.34	15.64	5.70	15.65	4.75

Mean duration in seconds.

As we expected to find significant differences in mean tool use percentages and in time spent per activity between the different conditions, variance and Bonferroni analyses were performed to test this hypothesis. We compared the tool use percentages for the Control group and the separate experimental conditions as well as the Control group and the experimental group as a whole. Significant differences were indeed confirmed for all three phases and for the protocols in their entirety. The basic tools that were available in all conditions show some interesting results, as do the Diagram, the Outline, and the Advisor for the conditions including these tools.

Chat

The chat activities – *to chat* and *chat* – show the strong predominance of the Control group throughout the phases that we expected for all basic tools: as the Control group does not have either planning tool nor any Advisors, they can divide

their activities between fewer tools, resulting in a higher mean percentage per tool. The mean time per event spent in the chat is consistent throughout the conditions (Table 2.2), with an average duration of about 21 seconds, as is shown in Table 4.3.

Shared Text

We expected to find similar results for *to text*, and even though their total number of activities was slightly smaller than in the experimental conditions, the Control group in fact managed to work on the text more often than any of the other groups (Table 2.1 and Table 3.6). On the other hand, the Control group spent less time per event in the shared text than the Diagram, Diagram-Advisor, and Diagram-Outline conditions during the third phase of the writing process (Table 9.4). Possibly, this is because these groups had some catching up to do after spending more time and energy on their planning during earlier phases.

Using the Private Notes and Sources

However, we did not find the same results for the Control group for all basic tools. The sources were found to be used less frequently in the Control group (Table 3.3), and so was *to notes* (Table 3.5). The relatively low percentage of *to notes* in the Control group might be explained from the fact that many participants in the experimental conditions used the notes window as a temporary text window whilst the partner was working on the Diagram or Outline. Otherwise, the alternatives for the participant with the red traffic light would be to read the sources (again), to chat, or to watch the partner write. After a turn change the notes text was copied and pasted into the shared text. This allowed the dyads to work on both collaborative goals simultaneously (both the text and the Outline and/or Diagram had to be completed). This is confirmed by the differences in duration: during the final phase, the experimental groups spent more time in the private notes window than the Control group (Table 2.2). Using the notes window as an alternative text window is a clever solution, but it does not encourage collaboration as the participants cannot see their partner's private window.

In general, the sources are consulted less frequently as the writing process advances (Table 10.1), and the duration of the events gets shorter as time goes by, as we can see in Table 4.3. It is unclear why the Control group should visit the sources less frequently than the other groups, as they do especially during the first phase (Tables 2.1 and 4.3). This does not mean that the Control group paid less attention to the source information, as the percentage says nothing about the time

spent in each tool. The T-tests for duration (Table 2.2), show that the Control group spent longer periods of time in the source texts than the experimental groups throughout the writing process. In other words, the experimental groups frequently switched to other tools, spending less time in a source per event. This difference can be explained from the fact that the Control group did not have any added tools, and thus could divide the same total time between fewer tools. The percentages for marking the sources (Tables 3.4 and 4.4) show a slight predominance of the Control group and the Outline condition during the first phase and in the entire protocols.

Total number of activities

We anticipated a lower mean number of activities for the Control group, as they simply had less to do: their assignment did not require them to make a diagram or outline in addition to the shared text. On the whole, this turned out to be true for the entire writing process, but not so much for the second and third phases, perhaps because the Control group chatted more frequently. Although the Control group chatted most frequently (Table 2.1), the experimental groups took slightly longer to type in their chat lines during the third phase (Table 2.2). In other words, the experimental chat was slower than the Control group chat. Possibly, the contents of the experimental group messages was more complex or more elaborate, as the task was more complex than for the Control group.

Using the Program Manual

As a result of the higher complexity of the TC3 environment in the experimental conditions, we expected the two Diagram-Outline groups to consult the program manual more frequently. At the same time, we expected the Advisor groups to consult the program manual less intensively, as all but four dyads in these conditions had worked with the program in an earlier survey. We found higher percentages in the protocols as a whole for the Outline condition compared to the Control group, the Diagram-Advisor condition, the Outline-Advisor condition and the Diagram-Outline-Advisor condition (Table 3.8). This confirms our assumption that the Advisor groups would need to consult the program manual less frequently. We similarly found higher percentages for the Diagram-Outline condition compared to the Diagram-Outline-Advisor condition. We also found higher percentages for the Diagram condition compared to the Control group and the Diagram-Outline-Advisor condition in the entire protocols. The first phase shows the same results for the Outline condition (Table 4.7). There are no significant differences in duration for this phase (Table 2.2). During the second phase, the

Diagram-Outline group consults the program manual more often than the Diagram-Advisor group (Table 5.7), and the Control group spends more time per event in the manual than the experimental groups (Table 2.2). The third phase shows no significant differences for the relative frequency of this activity, which is not surprising, as the students should have mastered the program by the time they reach the last stages of writing. However, the Control group did spend more time in the program manual than the experimental groups (Table 2.2).

Reading the Assignment

The assignment document is visited more intensively by the Control group in the second phase (Table 8.4) and by the Diagram group in the third phase – both in duration and percentage (Table 6.6 and Table 9.6) – when compared to each other and to the Diagram-Advisor, Diagram-Outline-Advisor, and Outline conditions. During the third phase, the Outline-Advisor condition clicked into the assignment more often than the Control group, and the conditions with the Diagram except the basic Diagram condition (Table 6.6). This means that the Control group felt the need to check task requirements during the first stages of writing, whereas the Diagram and Outline-Advisor groups felt the same need nearer the end of the writing process.

Using the Planning Tools

When we move on to the use of the planning tools, we see in Table 4.3 that the amount of time per event spent on these tools increases with each phase. Duration of Diagram activities goes from 32 to 45 to 82 seconds per visit, and for the Outline this goes from 12 to 19 to 37 seconds. When we look at the significant differences, the first thing we notice for the first phase is the predominance of the Diagram-Outline condition at the lower end of the scale and of the Diagram-Advisor condition, and to some extent the Diagram condition, at the upper end (Tables 4.9, 4.10, 4.11, 7.4, and 7.5). In other words: the Diagram-Outline group spent less time thinking about the next Diagram activity or performing the current one than the Diagram and Diagram-Advisor groups, and did so less frequently. This is hardly surprising, as the Diagram-Outline condition demanded that students divide their attention between two planning tools.

The same goes for the Diagram-Outline condition compared to the Outline and Outline-Advisor conditions in the first phase. Again, the Diagram-Outline group spent less time per event in the Outline tool (Table 7.5). The percentage analysis shows a different picture; here, the Outline condition shows a higher percentage than the Diagram-Outline and Diagram-Outline-Advisor groups for opening the

Outline, implying that the Outline group spent the most total time in the Outline of all Outline conditions during the first phase, whereas the opposite goes for the Diagram-Outline condition (Table 4.12). The second phase shows similar results for the Outline and Diagram-Outline conditions (Tables 5.13 and 8.6). The results for the Diagram-Outline-Advisor condition are similar to those for the Diagram-Outline group, though less convincing. The Outline-Advisor condition ends up in between the Outline group on the one side and the Diagram-Outline groups on the other: the Outline-Advisor students open the Outline less frequently than the Outline group, but more frequently than the Diagram-Outline students. They spent more time in the Outline per event than either of the Diagram-Outline groups. During the third phase, the Outline group still uses the Outline tool more frequently than the two Diagram-Outline conditions (Table 6.11), but there is no significant difference for the time per event in the Outline tool during this phase.

Use of the Advisor does not show very surprising results: the Diagram-Outline-Advisor condition shows lower mean percentages of *to Outline tips* during the second phase and in the entire protocols, and it pays significantly shorter visits to the tips window during the first phase.

Counting Words

The percentage analyses of the entire protocols for word count show that the Control group counts words more often than any of the experimental groups (Table 2.1). In addition, the Outline group counts words more frequently than the Diagram-Outline group (Table 3.9). The second phase shows the same tendency for the Control group, though not as strong as in the protocols as a whole; here, the Control group scores higher than the Outline, Diagram-Outline, and Diagram-Outline-Advisor conditions (Table 5.8). During the final phase, where we would expect students to count words more often in any case, the Control group counts words more frequently than the Diagram-Outline and Outline-Advisor groups only (Table 6.7). The Diagram-Advisor group and the Outline group also count words more frequently than the Diagram-Outline group, and in addition the Outline condition shows a higher percentage than the Diagram, Diagram-Advisor, and Outline-Advisor groups.

Stopping

Stopping the program happens most frequently in the Diagram, Diagram-Advisor, and Outline-Advisor conditions in the protocols as a whole (Table 3.10). The first phase shows only a few differences: the Control group stopped less frequently than

the Diagram, Diagram-Advisor, and Outline conditions, and the Diagram-Advisor condition also shows a higher stopping percentage than the Diagram-Outline condition during the first phase (Table 4.8). This last difference is also found in the second phase (Table 5.9). The third phase shows lower stopping percentages for the Control group, the Diagram-Outline condition, and the Outline condition compared to the Diagram, Diagram-Advisor, and Outline-Advisor conditions (Table 6.8). In addition, the Diagram-Outline-Advisor condition scores significantly lower here than the Diagram-Advisor condition.

Duration of Activities

The mean duration per activity is significantly lower for the Diagram-Outline-Advisor condition in the first phase compared to the Diagram, Diagram-Outline, and Outline-Advisor conditions and the Control group (Table 7.6). The students in the Diagram-Outline-Advisor condition had to divide their attention between the basic tools and all four extra modules. This naturally leads to a lower mean duration per activity, as the time students had to complete the assignment was the same for all conditions. During the second phase, the Control and Diagram conditions both had longer activity episodes than the Diagram-Advisor, Diagram-Outline, Diagram-Outline-Advisor, and Outline conditions (Table 8.7). This was anticipated for the Control group, as the students in this group had fewer tools at their disposal than the students in the experimental conditions. The results for the Diagram condition are most likely strongly influenced by the significantly longer chat activities, whereas the Control group total mean seems to be particularly influenced by reading the sources and the assignment.

4.1.2 Relation between tool use and text quality

Although the main research question dealt with in this chapter is not concerned with the final product – the shared text – it would still be very interesting to see whether differences in text quality could be related to different use of the software. This could help us in improving the program and in programming better new software for collaborative writing and learning. In this section, the correlations between tool use percentages and duration of activities, and the five text quality scores are discussed.

All phases together

Table 4.4 gives the correlations between the frequencies of the different activities in percentages and the text quality scores for the protocols as a whole.

Table 4.4: Percentage correlations all phases all conditions.

	Textual structure	**Segment argumentation**	**Overall argumentation**	**Audience focus**	**Mean text score**
To chat	.13*	.13*	-.15*	.09	.03
Chat	.10	.12*	-.03	.12*	.08
To source	-.07	.02	.01	-.08	-.02
Mark source	.09	.04	.01	.10	.06
To notes	-.14*	-.15*	.04	-.15*	-.09
To text	.16*	.23**	-.09	.12	.09
To assignment	.15*	-.01	-.02	.05	.06
To manual	.04	.11	.03	.02	.07
To Diagram tips	-.01	-.07	-.19	-.33**	-.21
To Outline tips	-.04	.54**	.24	.41*	.44**
Word count	.03	.03	-.02	-.04	-.02
Stop	-.03	-.09	.05	-.14*	-.05
Diagram open	-.22*	-.21*	-.07	-.26**	-.25**
To Diagram	-.05	.02	.07	.15	.08
Diagram close	-.14	-.14	-.07	-.02	-.11
Diagram activities within Diagram	.07	.00	.13	.00	.07
Diagram delete link	.13	.05	.08	.04	.11
Diagram delete object	.14	-.02	.06	-.05	.05
Diagram new link	-.06	-.01	.08	-.04	-.01
Diagram new object	.14	.01	.14	.00	.09
Diagram update object	.02	.02	.12	.10	.09
Outline open	-.07	.20*	.03	.18*	.13
To Outline	-.02	.18	-.08	.04	.03
Outline close	-.10	.18*	-.01	.19*	.10
Total no. of acts	-.07	-.25**	-.11	-.19**	-.18**

** $p < .01$; * $p < .05$.

Textual structure correlates positively with to chat, to text, and to assignment. It correlates negatively with to notes, and opening the Diagram. This means that, on the whole, it is better to focus on the collaborative dialogue and the product (both by checking the product requirements and by concentrating on the product itself), than to spend a lot of energy on writing private notes that the partner cannot see. Also, it seems better to leave the Diagram open in the background than to close and reopen it every time one needs to consult or change it.

Segment argumentation correlates positively with *to chat*, *chat*, *to text*, *to Outline tips*, and opening and closing the Outline. It correlates negatively with to notes, opening the Diagram, and the total number of activities. We draw from this that

the argumentation on the paragraph level benefits from using the Outline tool. Although this tool is meant to outline the argumentation on overall text level, by doing so it becomes clear to the students what each of the segments should contain. Not using the Outline when it is available results in neglect of argumentative structure on paragraph level. In addition, working in the collaborative windows – chat and text – results in discussion on and clarification of the argumentation, whereas by writing notes privately, thoughts and ideas are not shared and thus cannot be optimized by the partner. The negative influence of the Diagram may seem unexpected, but we should realize that the Diagram is meant to reflect the structure of the argument of the text as a whole, and not of each of the segments. Finally, the total number of actions influences the result negatively. In other words, switching between tools and crisscrossing the computer environment is detriment to the production of a high quality text, possibly because students do not focus efficiently on the task at hand, but let themselves be distracted by the multitude of possibilities.

Overall argumentation correlates negatively with *to chat*. Although clicking into the chat window results in worse overall argumentation of the text, entering chat into the chat history – so actual chatting – does not. This seems to confirm our conclusions for the total number of activities in the previous paragraphs: crisscrossing the environment and clicking from one window to another disturbs the writing process.

Audience focus correlates positively with *chat*, *to Outline tips*, and opening and closing the Outline. Paying attention to the communication with the partner results in a text that is better directed at its audience, and so does working on the Outline, and learning to work with it properly. Just like segment argumentation and textual structure, audience focus correlates negatively with *to notes,* opening the Diagram, and the total number of activities. In addition, it correlates negatively with *to Diagram tips* and *stop*. Again, activity in abundance leaves no room for working on the text. Also, as the notes are private, the partner cannot be tested out as a possible reader and audience, so that this function of collaborative writing is lost. Stopping the program more frequently is detriment to text quality. Reading the Diagram tips does not have the desired effect on text quality.

The mean of the four text quality scores correlates positively with *to Outline tips*. It correlates negatively with opening the Diagram, and the total number of activities. The last two finds are hardly surprising considering the results for the

separate text scores. The contrast between the negative tendency of using the Diagram tips and the positive correlation of consulting the Outline tips is an interesting one: the help texts were very similar, and so were the oral instructions. However, the two tools are of course very different: the Diagram is visual and quite abstract, whereas the Outline is much closer to the text. Also, students seemed to be more familiar with the concept of an outline than with the idea of a concept map.

The first phase

Table 4.5 gives the correlations between the different activities and the text quality scores for the first phase, in which an initial setup for the text should be made according to our hypotheses: preplanning. The correlations between duration of the activities and text quality in the three different phases of the writing process can be found in Section 11 in Appendix 7.

Textual structure correlates positively with marking source text, *to assignment*, and deleting Diagram objects. It correlates negatively with *to notes*. Although clicking into the sources in general does not correlate with textual structure during the first phase – and this is a little surprising, as we assumed the sources need to be read before the text can be written – highlighting specific parts of the source text has a positive effect on this text score, implying that distinguishing between main points and side issues pays off. Since the students need to understand the task in order to perform it properly, it is not surprising that *to assignment* has a positive effect during the first phase. Textual structure also correlates positively with *to assignment* duration: longer and more frequent visits to this window seems to have a positive effect on textual structure (Table 11.1). Using the private notes window during the initial phase of the collaborative process is detriment to the textual structure of the final text, possibly because frequent use of this window inhibits the collaboration: the partners cannot see each other's work.

Segment argumentation correlates negatively with *to Diagram tips*, *word count*, *stop*, and the total number of activities in this phase. Making an effort to understand the Diagram by reading the tips has the opposite effect of the Diagram's desired effect. Counting words during the first phase is naturally pointless, as there is (virtually) no text to be counted yet. Stopping the program during this phase interrupts the writing process, causing the students to lose track of their argumentation. They can be similarly distracted by the possibilities of the program itself: performing too many different activities does not leave room for

contemplation. Segment argumentation correlates positively with *to source* duration, to Outline tips duration, and to outline duration. This is in line with our expectations: it makes sense to read the sources before starting to write, as it helps to determine the exact topic, and a main outline of the argument.

Table 4.5: Percentage correlations 1st phase all conditions.

	Textual structure	**Segment argumentation**	**Overall argumentation**	**Audience focus**	**Mean text score**
To chat	-.01	-.03	-.12*	-.04	-.07
Chat	.06	.02	-.03	.02	.01
To source	-.12	.09	.04	-.01	.01
Mark source	.13*	.09	.03	.07	.08
To notes	-.14*	-.03	.03	-.05	-.04
To text	.00	.04	-.07	-.01	-.01
To assignment	.14*	.02	.01	.01	.06
To manual	.02	.09	.00	-.06	.01
To Diagram tips	-.22	-.29*	-.13	-.38**	-.36**
To Outline tips	.25	.26	.11	.05	.28
Word count	-.04	-.13*	-.07	-.07	-.09
Stop	-.10	-.24**	-.05	-.24**	-.18**
Diagram open	.00	-.11	.02	-.05	-.05
To Diagram	-.01	-.04	.02	.10	.03
Diagram close	-.08	-.08	-.05	.02	-.06
Diagram activities within Diagram	.07	-.03	-.02	.05	.02
Diagram delete link	-.07	-.13	-.13	-.15	-.16
Diagram delete object	.18*	.02	-.04	-.03	.04
Diagram new link	.00	-.05	-.07	-.05	-.06
Diagram new object	.07	-.01	-.01	.06	.04
Diagram update object	.04	-.03	.06	.17*	.08
Outline open	-.04	.17	-.05	.18	.09
To Outline	-.10	.14	-.20*	.06	-.06
Outline close	-.02	.17	-.05	.18*	.10
Total no. of acts	.04	-.13*	.03	-.05	-.01

** p < .01; * p <.05.

Overall argumentation correlates negatively with *to chat*, and activating the open Outline window. Frequently leaving the chat window and coming back to it implies that students had lots of short chat episodes. It is likely that these students did not go very deep into the argumentation as this will usually require more elaborate discussion. A similar explanation can be given for the negative influence of *to Outline*. Writing short bits at a time instead of focusing on the Outline for an extended period results in a less coherent framework for the overall argumentation of the final product. Overall argumentation correlates positively with in instruction duration, especially with reading the assignment and Outline tips. The assignment defines the end product as an argumentative text, and reading the assignment is

likely to have helped the students in understanding the task requirements. Spending time on the Outline during the first phase correlates negatively with overall argumentation.

Audience focus correlates positively with updating objects in the Diagram and closing the Outline tool. Improving the structure of the Diagram and thus of the argumentation should normally lead to a clearer discussion of the argument, and this is confirmed by the positive results for updating objects. On the other hand, audience focus correlates negatively with *to Diagram tips* and *stop*. Audience focus is influenced positively by spending more time reading the instruction texts in the information window, especially the program manual and the Outline tips. It also correlates positively with the mean duration per activity.

Mean text score correlates negatively with *to Diagram tips*, and *stop*. Reading the Diagram tips has the opposite effect from what we expected. It is no use stopping during this phase, and interrupting the planning process this early has a negative effect on text quality. The mean score correlates positively with *to assignment* duration and *to Outline tips* duration, as well as *to Diagram* duration.

The second phase

Table 4.6 gives the correlations between the different activities and the text quality scores for the second phase. This phase begins when students start writing the shared text. Table 11.2 in Appendix 7 shows the correlations between tool use duration and text quality for this phase.

To chat, and *to text* correlate positively with textual structure. It correlates negatively with opening and closing the Diagram window. It seems better to keep the Diagram window on screen than to close and reopen it. Frequent deliberation during this phase seems to be important for the collaborative writing process. This is also the time to write the bulk of the text, as the positive effect of to text shows. Textual structure correlates positively with in source duration (Table 11.2). Apparently, it is important to keep checking the sources even whilst writing the first draft of the essay. In this phase as well, using the Diagram for longer periods of time has a negative effect, as does using the Outline.

Table 4.6: Percentage correlations 2nd phase all conditions.

	Textual structure	Segment argumentation	Overall argumentation	Audience focus	Mean text score
To chat	.23**	.20**	.00	.10	.14*
Chat	.05	.08	-.08	.06	.00
To source	-.08	-.04	-.01	-.14*	-.08
Mark source	-.03	.01	-.10	.11	-.01
To notes	-.02	-.11	.09	-.19**	-.05
To text	.22**	.31**	.04	.15*	.20**
To assignment	.07	-.08	-.06	.03	-.01
To manual	.05	.09	.04	.08	.09
To Diagram tips	.22	.14	-.09	-.22	.01
To Outline tips	-.09	.24	.02	.45**	.25
Word count	.03	.13*	.04	.01	.05
Stop	.08	.01	.09	.01	.07
Diagram open	-.17*	-.14	-.05	.03	-.09
To Diagram	-.09	-.14	.00	.12	-.01
Diagram close	-.18*	-.14	-.05	.02	-.10
Diagram activities within Diagram	-.10	-.20*	-.01	-.02	-.09
Diagram delete link	.08	-.10	-.04	-.07	-.01
Diagram delete object	-.05	-.23**	-.01	.02	-.07
Diagram new link	-.09	-.17*	-.02	-.09	-.09
Diagram new object	-.11	-.15	.02	-.02	-.08
Diagram update object	-.10	-.18*	-.01	.06	-.05
Outline open	-.09	.11	.03	.14	.07
To Outline	.03	.13	.08	.09	.12
Outline close	-.10	.08	-.01	.13	.04
Total no. of acts	-.10	-.20**	-.13*	-.15*	-.18**

** p < .01; * p <.05.

Segment argumentation correlates positively with to chat, to text, and word count. It correlates negatively with the total number of activities in this phase, and with Diagram activities. Using the Diagram, then, has an overall negative effect on this text score. This is a disappointing result, as the Diagram is meant to specify the argumentative structure of the text. However, by its nature it does not help students to write good segments, but only supports development of argumentation on a higher level. Unfortunately, the results for overall argumentation do not show the intended positive influence of the Diagram. Focusing on the collaboration is important during this phase, as is shown by the positive effects of *to chat* and *to text*. Now is the time to start counting words, although the positive effect is not very strong. As in the first phase, performing too many different actions distracts the students from the task and is detriment to segment argumentation. Segment argumentation correlates positively with *in source* duration, *in instruction* duration, and mean duration per activity. There is a negative tendency for Diagram duration. Students are still learning to work with the program and keep checking

the assignment to make sure they are on the right way to a good text. Again, it is better to spend some more time in a window than to go window hopping. As we have noted before, the Diagram is not very helpful at all in constructing good argumentative paragraphs.

Just like segment argumentation, overall argumentation correlates negatively with the total number of activities in this phase. Spending more time in the Diagram shows a negative tendency for Overall argumentation.

Audience focus correlates positively with *to text* and *to Outline tips*. Audience focus correlates negatively with *to source*, *to notes*, and the total number of acts. If we consider the partner as part of the audience, and a guinea pig for testing the readability of the shared text, it is understandable that focusing on the private windows instead of the collaborative ones results in weaker audience focus. Audience focus correlates positively with *to source* duration and *to notes* duration. This is surprising, as we found negative correlations with to notes throughout for the percentages, including audience focus in the second phase. It seems that it is fine to use the private notes window for writing larger amounts of text, but not for scribbling down brief sentences at a time. We found mixed correlations for the different Diagram duration sub measures, while mean duration per activity again correlates positively with the text score.

Mean text score correlates positively with *to chat*, and *to text*. It correlates negatively with the total number of activities in this phase, but positively with the mean time spent on each activity. The basic collaborative windows – chat and the shared text – play an important role whilst writing the first draft of the essay, whereas too much activity in general is detriment to the final product. The mean score correlations with activity duration show that it is better in this phase not to spend too much time working on the diagram, whereas it is fine to spend more time in the sources.

The third phase

Table 4.7 shows the percentage correlations for the third and final phase of the writing process, where the draft from the second phase is extended and revised. Table 11.3 in Appendix 7 shows the correlations between tool use duration and text quality for this phase.

Table 4.7: Percentage correlations 3rd phase all conditions.

	Textual structure	**Segment argumentation**	**Overall argumentation**	**Audience focus**	**Mean text score**
To chat	.07	.08	-.15*	.12*	.00
Chat	.08	.13*	-.02	.17*	.10
To source	.02	-.02	-.02	-.04	.00
Mark source	-.04	.03	.03	.01	.02
To notes	-.14*	-.18**	.00	-.08	-.11
To text	.13*	.07	-.15*	.07	-.01
To assignment	.03	.09	.03	.11	.10
To manual	-.05	.06	-.05	.02	.00
To Diagram tips	.05	.15	-.14	.16	.08
To Outline tips	-.04	.09	.17	.17	.17
Word count	.01	-.10	-.02	-.09	-.07
Stop	-.07	-.04	.05	-.10	-.04
Diagram open	-.21*	-.15	-.08	-.27**	-.23**
To Diagram	-.04	.24**	.07	.02	.09
Diagram close	-.03	-.04	-.07	-.07	-.05
Diagram activities within Diagram	.09	.25**	.20*	.01	.17
Diagram delete link	.19*	.20*	.20*	.17*	.26**
Diagram delete object	.07	.05	.08	-.04	.05
Diagram new link	.02	.23**	.23**	.09	.18*
Diagram new object	.12	.18*	.13	-.04	.12
Diagram update object	.03	.27**	.11	-.05	.10
Outline open	.03	.15	.09	.10	.14
To Outline	.09	.10	-.07	-.06	.02
Outline close	-.05	.16	.06	.13	.11
Total no. of acts	-.07	-.15*	-.12	-.18**	-.17*

** p < .01; * p <.05.

Textual structure correlates positively with *to text* and deleting links between Diagram objects. It correlates negatively with *to notes* and opening the Diagram window. The focus of attention during this phase of the writing process should be on finalizing the shared text. By this time, it is really too late for writing private notes. The positive effect of deleting Diagram links may point out that it is important to keep adapting the diagram throughout the writing process and check the argument structure with the structure of the text continually. For this purpose, it is best to leave the Diagram window open at all times. Textual structure correlates negatively with *in source* duration and *to text* duration. By now, the

sources should have been read and incorporated in the text. If students still need to read or check source information at this stage, it is rather too late and it will decrease the quality of the textual structure.

Segment argumentation correlates positively with *chat*, *to Diagram*, and Diagram activities in general. It correlates negatively with *to notes* and the total number of activities. An overall positive influence of Diagram use can be observed for segment argumentation in this phase. This suggests that the Diagram is effective when used as a revision tool during the final stages of the writing process, rather than as a concept map for setting up the first draft. As in the second phase, collaborative discussion is a positive influencing factor for segment argumentation, whereas using the private notes window has the opposite effect.

Overall argumentation correlates positively with Diagram activities in general. It correlates negatively with *to chat*, and *to text*. As opposed to segment argumentation, the overall argumentation of the text should have been made clear during the second phase. This might explain why the focus on collaboration has the opposite effect here. On the other hand, the Diagram has the same positive influence on both levels of argumentation. Overall argumentation correlates positively with *to text* duration, so it helps to write during the final phase to improve the argumentation of the text. It also correlates positively with *to Outline tips* duration, which means that checking the requirements of the Outline is beneficial for text quality. However, Overall argumentation correlates negatively with in source duration, just like textual structure.

Audience score correlates negatively with opening the Diagram and the total number of acts. It seems better to leave the Diagram open and on screen and, again, not to develop too many different activities. The chat activities correlate positively with Audience focus, and this implies that coordination is an important factor during the completion of the text. Audience focus is influenced slightly negatively by using the Outline for longer periods: perhaps the planning tool distracts the students from presentation and formality.

The Diagram activities show a positive tendency for the mean text score. The mean score correlates negatively with opening the Diagram and the total number of activities. In general, opening the Diagram window has a negative effect on text quality. Activities within the Diagram, however, have an overall positive effect on the shared text. Just as in the second phase, performing too many different

activities decreases the quality of the text. The correlations for tool use duration and mean score show what we knew already: source information should be read early on in the writing process, not during the final stages. Taking time to go to the Outline tips has a positive effect on the mean score.

4.1.3 Conclusion

In general, then, the results for the activities in the different phases are quite different, especially for the planning tools: the Diagram and the Outline. Focusing on the text by clicking into the shared text window – and possibly writing in it – results in qualitatively better texts. This is in line with our expectations. Also, using the private – hence non-collaborative – notes window is detriment to the quality of the collaborative product. This confirms the idea that collaboration is necessary on all subtasks, including idea generation and information processing. Students who need to reopen the Diagram do not have the Diagram on screen at al times, and so they risk losing sight of the argumentative structure of their text. In general, there is a positive tendency on text quality for the percentage of Diagram use in the third phase, though the duration correlations give a more mixed picture. Reading the Outline tips, on the other hand, has a general positive relation with text quality. A multitude of activities may lead to distraction from the task and leaves less room for planning and writing.

4.2 Planning and executing: Task acts

4.2.1 Task act differences between conditions and phases

The question addressed in this section is what writing strategies are discussed by the participants and how these relate to the quality of the final product. Task acts are the types of writing activities and strategies refereed to by the participants in their chat in order to coordinate their actions in collaborate writing the argumentative paper. What is more, we were interested to see whether the presence of the different planning tools – the Diagram and the Outline – changed the distribution of strategies in the chat.

Table 4.8 gives the means, standard deviations of the Task act percentages in all phases for the Control group and the Experimental conditions, as well as the mean differences on T-tests between these two groups. In addition, the numbers of dyads

contributing to each type of Task act are given. The results for the separate experimental conditions are given in Table 1.1 in Appendix 10. As there were only very few significant differences between the experimental conditions, these will not be discussed here.

On the whole, planning was done more often than executing in both groups (47% and 70 % vs. only 37% and 18%), and non task related episodes occurred the least frequently (16% and 12%). In general, the distributions of episodes for the two groups over the three main Task act categories are very different: for the Control group it is 47:37:16, whereas for the Experimental groups it is 70:18:12. The Experimental conditions, then, have the inclination to plan more than their Control group colleagues. We would expect that this difference can be accounted for by the percentages for the categories related directly to the planning tools, such as Revision Diagram and Plan Outline. However, these Task act categories can only account for about 14% of the difference between Control group and Experimental conditions, whereas the significant mean difference for Total Planning is about 23%. We infer from this that the presence of the planning tools stimulates planning in general.

Within the planning categories, discussion about planning the text was done most frequently (13% and 19%), followed by planning coordination (10% and 14 %), and both of these categories occur more frequently in the chat of the Experimental groups than in the chat of the Control groups. After this, the two groups diverge: the Experimental groups spend relatively more chat episodes on planning knowledge (8%), whereas the Control group's next most frequent planning activity is planning turn alternation (6%). The latter is one of the few categories that shows a significantly higher percentage for the Control group than for the Experimental groups, and Plan knowledge shows a large significant higher percentage for the Experimental groups. In both groups, relatively few chat episodes were spent on planning the use of external sources, the layout, and the use of the private notes.

Table 4.8: Descriptive statistics Task act percentages in all phases for the Control group and the Experimental conditions and mean differences on independent samples T-tests.

	Control group			Experimental group			T-test
	N	Mean	SD	N	Mean	SD	Mean differences
Plan advisor				48	.01	.06	
Plan turn alternation	39	6.32	3.09	106	4.72	2.79	1.60**
Plan coordination	39	9.51	3.72	106	13.52	4.56	-4.01**
Plan Diagram				77	7.62	3.55	
Plan Diagram layout				77	.32	.54	
Plan external source	39	.66	.91	106	1.20	1.26	-.54**
Plan goals	39	1.85	1.15	106	1.59	1.21	
Plan knowledge	39	2.45	1.45	106	8.11	4.44	-5.66**
Plan layout	39	1.71	1.39	106	.44	.71	1.27**
Plan notes	39	1.54	1.37	106	1.41	1.25	
Plan Outline				63	4.04	2.31	
Plan Outline layout				63	.03	.11	
Plan revision	39	3.44	1.86	106	3.67	2.24	
Plan revision Diagram				77	.80	1.06	
Plan revision Outline				63	.30	.53	
Plan source	39	5.67	2.28	106	7.04	2.69	-1.38**
Plan text	39	12.98	3.98	106	18.80	4.43	-5.82**
Total percentage Plan	39	47.31	5.75	106	69.70	6.11	-23.33**
Execute advisor				48	.00	.00	
Execute word count	39	3.82	2.44	106	1.64	1.58	2.17**
Execute Diagram				77	1.39	1.69	
Execute Diagram layout				77	.03	.12	
Execute external source	39	.97	1.15	106	.26	.50	.71**
Execute goals	39	2.63	1.51	106	1.07	1.02	1.55**
Execute knowledge	39	5.17	3.51	106	4.86	2.93	
Execute notes	39	.41	.55	106	.03	.18	.38**
Execute Outline				63	.49	.96	
Execute Outline layout				63	.00	.03	
Execute revision	39	9.74	5.49	106	1.56	1.56	8.17**
Execute revision Diagram				77	.33	.54	
Execute revision Outline				63	.09	.32	
Execute source	39	4.83	3.22	106	2.16	1.58	2.67**
Execute text	39	9.49	3.52	106	4.98	3.17	4.51**
Total percentage Execute	39	37.05	8.24	106	18.18	5.14	18.86**
Non task program	39	3.11	1.75	106	3.82	2.38	-.71*
Non task social	39	12.54	6.15	106	8.28	5.08	4.25**
Total percentage Non task	39	15.65	6.83	106	12.10	5.45	3.54**

N is the number of dyads; ** $p < .01$, * $p < .05$. Only significant differences are shown.

The Executing activities show very little differences in the order of the categories, except for Execute revision and Execute knowledge. For the Control group, Revision is the most frequent Execute episode (10%) with 26% of the total percentage for Execute, whereas for the Experimental groups it only comes fifth

with only 2% of the total number of Task acts, and only 9% of all Execute episodes. Although the two groups do not differ significantly in talking explicitly about their knowledge of the topic (Execute knowledge), for the Control group it is about 14% of all Execute episodes, whereas for the Experimental conditions it is 27%. The two groups both show high percentages for Execute text, both compared to the entire number of Task act episodes (9% and 5%, resp.) and in relation to the total percentage of Execute episodes.

The Control group shows a significantly higher percentage of Non task episodes than the Experimental conditions, especially of social talk (13% vs. 8%). However, the Experimental group talked significantly more frequently about the TC3 software. About one third of all Non task episodes were spent on this by the Experimental groups, whereas the Control group devoted only one fifth of their Non task episodes to TC3. These results are hardly surprising, as the program was more complicated for the Experimental groups, and the Control group – not having to spend time on filling in diagrams or outlines – had more time to chat socially.

When we look at the discussion on the planning tools, the percentages for the Advisor are the most striking: they are virtually nonexistent. The Diagram is talked about relatively more frequently than the Outline, both in terms of planning and in terms of executing. Whereas the total percentage for Planning decreases towards the completion of the task, planning the Diagram and the Outline increases quite considerably in the third phase.

The descriptive statistics and significant differences between the two groups for the three phases are given in Tables 1.2, 1.3 and 1.4 in Appendix 10. In addition, the total percentages of Planning, Executing, and Non task episodes for the Control group and for the Experimental conditions are set out in Figure 4.1, Figure 4.2 and Figure 4.3. The tables and Figure 4.1 show that while the Control group plans relatively less frequently as the collaboration advances – from 51% in the first phase to 49% and 42% in the second and third phases – the Experimental groups plan slightly more frequently during the second phase compared to the first phase, and the percentage seems to stabilize in the third phase (from 67% to just over 70% in the second phase and just under 70% in the third phase). Although starting out planning the sources significantly more frequently, the Control group shows a significantly lower percentage of Plan source during the third phase.

Whereas the Experimental groups plan revisions significantly more often during the second phase, the Control group speaks of executing revision significantly more frequently throughout the writing process. The total percentage of Execute episodes increases throughout the collaborative process for both groups, but with a line from 23% to 35% to 41% the Control group shows a much stronger increase than the Experimental groups from 15% to 18% to 20%. Finally, the Non task episodes are most common during the first phase, with the percentage dropping rapidly in the second phase. For the Control group, this category seems to stabilize around 16%, whereas the Experimental group Non task chat drops a bit more in the third phase, to 10%. The Non task discussion of TC3 shows a significant difference between the two groups only for the first phase. This suggests that the problems with the software occurred mainly during the first phase, and that the students discussed the use of the program rather than technical hiccups, as these did not occur solely during the earliest phase.

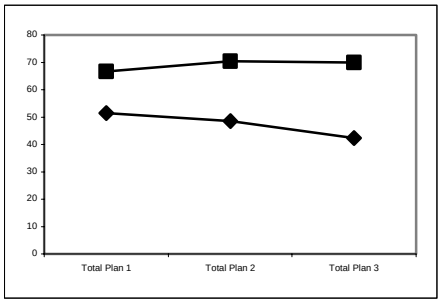

Figure 4.1: Chart showing the differences between phases for the Control group and the Experimental groups for the Total Plan percentage. ◆ Control group, ■ Experimental groups.

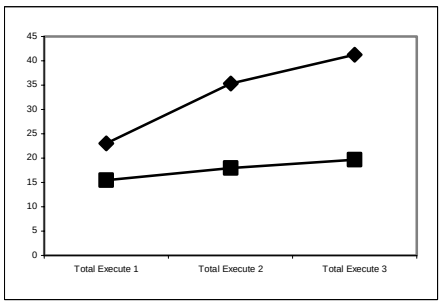

Figure 4.2: Chart showing the differences between phases for the Control group and the Experimental groups for the Total Execute percentage. ◆ Control group, ■ Experimental groups.

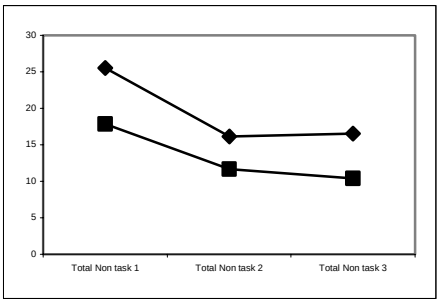

Figure 4.3: Chart showing the differences between phases for the Control group and the Experimental groups for the Total Non task percentage. ◆ Control group, ■ Experimental groups.

4.2.2 Relation between Task acts and text quality

Of course, we are not just interested in differences in writing strategies as such: it would also be very interesting to see whether differences in these strategies can be related to the quality of the final product – the shared argumentative text. Table 4.9 shows the correlations between the Task act percentages in all phases and the five text quality scores for the Control group and the Experimental groups. Similar tables for the three phases of the collaborative writing process are given in Tables 3.1, 3.2 and 3.3 in Appendix 10.

Table 4.9: Correlations between Task act percentages in all phases and text quality for the Control group and the Experimental conditions.

	Textual structure		Segment argumentation		Overall argumentation		Audience focus		Mean text core	
	C	**E**	**C**	**E**	**C**	**E**	**C**	**E**	**C**	**E**
Plan Advisor		.04		.16		-.09		.25*		.13
Plan turn alternation	-.13	.22**	-.12	-.02	-.03	.18**	-.04	.17*	-.08	.20**
Plan coordination	-.12	-.04	-.21	.02	-.11	.07	-.19	-.04	-.19	.00
Plan Diagram		-.10		-.02		.10		.07		.03
Plan Diagram layout		-.09		.01		.07		.03		.00
Plan external source	-.19	-.10	.09	-.07	-.18	-.01	.10	.10	-.05	-.02
Plan goals	.17	.02	.14	-.08	.08	-.15*	.15	.00	.15	-.06
Plan knowledge	-.07	.15*	.06	.12	-.01	.21**	-.10	.04	-.04	.18**
Plan layout	-.02	.15*	-.01	-.13	-.08	-.12	-.09	-.16*	-.07	-.09
Plan notes	-.05	-.03	-.22	-.07	-.05	-.01	-.04	-.18*	-.10	-.10
Plan Outline		-.17		.03		-.08		-.06		-.11
Plan Outline layout		-.29**		-.01		-.04		.11		-.09
Plan revision	-.10	.06	.32**	-.06	.13	-.09	.21	-.12	.19	-.09
Plan revision Diagram		.02		-.01		.05		.02		.06
Plan revision Outline		-.10		.08		-.15		.02		-.06
Plan source	-.27*	-.12	-.22*	-.05	-.19	.04	-.11	-.08	-.23*	-.07
Plan text	-.03	-.05	.25*	-.06	-.04	-.18**	.17	-.06	.10	-.13
Total percentage Plan	-.33**	.05	-.02	-.14*	-.19	.15*	-.02	-.09	-.15	-.01
Execute Advisor										
Execute word count	-.04	-.02	-.16	.19**	-.04	-.09	-.15	.10	-.12	.05
Execute Diagram		-.23**		-.01		-.06		.00		-.11
Execute Diagram layout		-.12		-.08		.06		.21**		.02
Execute external source	.06	.03	.02	.08	-.06	-.02	.08	.04	.02	.04
Execute goals	.04	.11	.08	.01	.18	.07	.50**	-.03	.28*	.05
Execute knowledge	.29*	.18**	.15	.12	.23*	.23**	.45**	.14*	.35**	.23**
Execute notes	-.25*	-.18**	-.24*	-.08	-.05	-.09	-.10	-.11	-.16	-.17*
Execute Outline		-.32**		.00		-.07		.01		-.14
Execute Outline layout		-.09		-.02		-.06		-.08		-.09
Execute revision	.22	.05	.18	-.15*	.14	-.22**	.12	-.05	.19	-.13
Execute revision Diagram		.03		.02		-.07		-.06		-.03
Execute revision Outline		.06		.18*		.06		-.01		.08
Execute source	.02	.04	-.08	.14*	.00	-.06	-.05	.05	-.03	.06
Execute text	.25*	-.14	.21	-.05	.20	-.18**	.34**	.10	.31**	-.09
Total percentage Execute	.37**	-.03	.20	.08	.28*	-.10	.45**	.16*	.40**	.04
Non task program	-.08	.02	-.12	-.07	-.11	.00	-.27*	-.19**	-.19	-.08
Non task social	-.20	-.05	-.25*	.13	-.20	-.09	-.53**	.05	-.37**	.00
Total percentage Non task	-.20	-.03	-.26*	.09	-.21	-.08	-.54**	-.04	-.39**	-.03

** $p < .01$, * $p < .05$.

The total Planning percentage shows a negative tendency for the Control group, while the Experimental group gives a mixed picture here: the chat on planning in the Experimental conditions correlates negatively with Segment argumentation,

but positively with Overall argumentation. Within planning, the general tendency of the relation between the Control group discussion and text quality is negative, with a few interesting exceptions. Planning the goals shows a slight positive tendency for the Control group, though the correlations for the Experimental conditions do not show a clear direction. For Planning revision the tendency for the Control group is also positive, except for the correlation with Textual structure. Planning the text gives a positive tendency for the Control group, but an negative one for the Experimental conditions. The correlations for Plan turn alternation show similar tendencies: the Control group gives a slight negative tendency, whereas the Experimental group correlates positively with all text quality scores except Segment argumentation. These positive correlations can be easily explained: the Experimental groups had to take turns in both writing and using the Diagram and/or Outline, and this made the logistics of the task more complicated, necessitating more elaborate negotiation of task division through the chat. Although there is no clear tendency between planning knowledge and text quality for the Control group, the Experimental group shows a clear positive tendency, with positive correlations for Textual structure, Overall argumentation, and Mean text score.

The total Executing percentage gives a strong positive tendency for the Control group. When we look at the sub measures for the discussion of executing the task, this positive tendency is clearly present throughout, except for Execute word count and Execute notes. The latter also shows a clear negative tendency for the Experimental groups. As we saw in Section 4.1.2 of this chapter, counting words frequently is related slightly negatively to text quality. Apparently, the same is true for talking about the number of words. Just as for Planning, both Execute text and Execute goals show positive tendencies for the Control group. Execute revision gives a positive tendency for the Control group, but a negative tendency for the Experimental conditions. The executing chat about the planning tools both show negative tendencies for text quality, and – quite surprisingly – for Textual structure in particular. Both for the Control group and the Experimental group there is a strong positive relation between Execute knowledge and text quality.

As we predicted, the relation between Non task chat and text quality is negative throughout the groups, sub measures of Non task chat, and text quality measures, although the relation is the most clear for the Control group.

When we look at the three different phases of the writing process, we find that planning knowledge during the first phase is beneficial for text quality in the Control group, but planning knowledge during the final phase has a negative relation with text quality. The Experimental group, on the other hand, seems to benefit from Plan knowledge throughout the process. Plan revision, though negative during the first and last phase for the Experimental groups, shows a positive tendency for text quality in the second phase. Finally, the negative tendency for the total Plan percentage in the Control group seems to stem mainly from the final phase, which makes sense as planning is usually done before executing, and the amount of planning should normally decrease towards the completion of a task.

Execute notes, though clearly negative in the entire process for the Control group, shows a clear positive tendency for the first phase. The positive tendency for Execute revision in the Control group is not found in the first phase: of course, there is still very little to revise at that stage, and thus revision is quite futile. Although we did not find a clear relation between Execute revision Diagram and text quality in the process as a whole, the second phase gives a clear positive tendency: updating the Diagram during this phase is positively related to Overall argumentation, Audience focus, and Mean text score. Surprisingly, the Experimental groups show a positive tendency for Execute text in the first phase, and a negative one for the third phase: we expected quite the opposite.

4.3 Conclusion

The main research question for this chapter suggests that we expect students to develop different types of knowledge construction activities during the first phase versus the second and third phases. Two clear indicators of knowledge construction in the Task acts are Plan knowledge and Execute knowledge. Both are most frequent during the first phase, and their relative percentage decreases steadily for both the Control group and the Experimental groups as the writing process advances. The increase in the chat discussion of Execute Task acts versus the decrease of Non task and – to some extent – of Plan Task acts suggest that the activities are indeed different in the different phases.

The analyses of the use of the tools within TC3 show that most students follow a logical pattern of activities: first they read the sources and the help on the task, the program, and the planning tools, and towards the end of the process they make less

use of the private notes, and more of the planning tools – which are thus used for online planning rather than preplanning, as we would expect from these relative novice writers. In general, the results of the tool use analyses match the results for the Task act analyses, and we can conclude that the students do indeed develop different types of knowledge construction activities during the first phase versus the second and third phases.

CHAPTER 5 SUPPORT OF ORGANIZATION AND LINEARIZATION

5.1 Organizing ideas with the Diagram tool

The Diagram tool was designed to give insight into the relationship between conceptualization in three phases of the writing process and the quality of the collaborative writing process, as measured by the final product – an argumentative text. This key problem was translated into the following research questions mentioned below[2]. In general, we expected to find that better use of the Diagram tool – more and better structured argumentative units and relations – will lead to higher text quality.

- What is the relation between use of the Diagram and text quality and what differences can be observed between different Diagram conditions?
- What is the correspondence of the diagram with the text, and what is the relation of this correspondence with text quality?
- What is the relation between the (argumentative) contents of the diagram and the argumentation of the text?
- What is the relation between the number and origin of arguments in the diagram and text quality?

The first question in this part of our research concerns the relation between the use of the Diagram and the quality of the final text. As we saw in the previous section, there were very few significant differences in text quality between the different conditions. For the Diagram conditions, we found that the Diagram group and the Diagram-Outline-Advisor condition score slightly higher than the Diagram-Advisor condition on textual structure. This difference in text quality might be explained with the distinction made by Suthers and Hundhausen (2001), who state that argumentation 'within' the representation (the diagram) as expected by the researchers, and argumentation 'from' the representation, often found in practice, are two different types of argumentation. In this view, participants show a strong inclination to use a diagram as a communication medium, as a full report of the argumentation, and as a medium "for expressing formal models - in favor of their role in stimulating and guiding collaborative learning discourse". It seems that the

[2] The analyses of the diagrams were done by Paulien Honkoop as part of her Master's thesis.

participants in the Diagram-Advisor condition used the Diagram as a full report of their argumentation, which means they did not let the Diagram guide them in developing an argumentative structure. The categories present in the Diagram are no guideline to them, nor do the participants discuss the contents of the diagram: they merely describe the contents of the ideas generated in their discussion, or even of their text. Thus, the Diagram only 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. If a diagram is used to report on the contents of the text, it can still have a structuring function during the revision of the text.

The next question on the relation between diagram and text is concerned with the content of both the diagram and the text, and more specifically with the correspondence of arguments. Table 5.1 shows the descriptive statistics for the correspondence of arguments between text and Diagram. On the whole, the correspondence of arguments in the text and the Diagram is not optimal: on average four arguments correspond, while three arguments are found only in the Diagram and six are unique to the text. As a first conclusion we can state that the Diagram does not seem to have been used optimally, neither as a summary of the text, nor by stimulating clarification of arguments in the text.

Table 5.1: Descriptive statistics for correspondence of arguments.

	Mean	**SD**	**Minimum**	**Maximum**
Arguments only in diagram	3.03	1.76	0	9
Arguments only in text	6.25	3.39	0	19
Similar arguments text-diagram	4.30	2.30	0	10

Table 5.2 shows the Pearson correlations between the occurrence of arguments in diagram or text or both diagram and text on the one hand, and the quality of the final text on the other hand. There is no relation for Textual structure, which means that the difference between conditions in Textual structure score found in the multiple comparison analysis (see Chapter 3) is not accounted for by the difference in arguments between diagram and text. This is surprising, as we expected that the Diagram would contribute to a better overview of the text for the participants.

Table 5.2: Pearson correlations between correspondence of arguments and text quality.

	Textual structure	Segment argumentation	Overall argumentation	Audience focus	Mean score
Arguments only in diagram	.04	.01	-.25**	-.03	-.06
Arguments only in text	-.03	-.01	-.12	.00	-.06
Similar arguments text-diagram	-.08	-.05	-.03	-.04	-.06

** p < .01, * p < .05; N = 73 dyads.

Like Textual structure, Segment argumentation does not show any significant relations with any of the correspondence measures either. This might be explained from the fact that these measures are not logically related to the specific presentation of each argument within the paragraphs. However, we did expect to find that the segment argumentation would improve with effective use of the Diagram tool, because the Diagram offers labels such as support and refutation that help participants formulate an argument more precisely. We did find a significant negative correlation (-.25) between overall argumentation and the number of arguments present in the diagram but not in the text. This is not surprising: if participants incorporate arguments in the diagram that they do not go on to include in the text, the argumentation of the text as a whole is likely to be incomplete. There are no correlations for any of the other text quality scores.

Table 5.3: Descriptive statistics of diagram elements.

	Mean	SD	Minimum	Maximum
Total nr. of elements	23.68	9.14	0	49
Total no. of objects	11.97	4.70	0	26
Claim	.99	.35	0	2
Conclusion	.77	.56	0	3
Contra	2.59	1.61	0	10
Info	.97	1.65	0	10
Pro	2.74	1.55	0	7
Refutation	2.12	1.61	0	8
Support	1.79	1.53	0	7
Arrow	9.58	5.26	0	23
Line	2.14	3.69	0	16

Another question on the diagram-text relation concerns the structure of the diagram. We looked at the type of units used in the diagrams and their frequencies, and the means, standard deviations, and extremes are shown in Table 5.3. As can be seen from the table, the Diagrams contain some 24 elements, about half of them text boxes and the other half arrows and lines between them. In most Diagrams only one claim was present with almost three supporting arguments (Pro), 2.5 counterarguments (Contra), almost two supporting elements, and two refutations.

Table 5.4 shows the Pearson correlations between the diagram structure and the five text quality measures. Both counterarguments and informative elements were found to correlate negatively with Textual structure. In addition, the informative elements correlate negatively with Overall argumentation and with the Mean text score. These negative correlations for the Information stem from the manner of assessment of the text. We looked for a clear introduction, body, and conclusion, and we expected to find neutral information in the introduction, but not so much in the rest of the text, as it does not contribute to the argumentative structure. A text with a disproportionately long introduction, an informative body, or lots of new information in the conclusion would get a lower score than a text with more concise neutral informative parts. Neutral information does not form a part of the argumentative reasoning, but functions to clarify the topic to the reader. Too much information leaves no room for argumentation, leading to a more informative, and less argumentative text. Excess use of informative elements thus leads to unclearly structured diagrams that are confusing and thus do not help the participants write clear texts.

Table 5.4: Pearson correlations between diagram structure and text quality.

	Textual structure	**Segment argumentation**	**Overall argumentation**	**Audience focus**	**Mean score**
Total nr. of elements	-.14	.02	-.04	-.13	-.11
Total no. of objects	-.11	.07	-.02	-.09	-.06
Claim	.10	.03	-.07	.08	.06
Conclusion	-.01	-.05	-.08	-.11	-.08
Contra	-.27**	-.13	-.03	-.04	-.16
Info	-.20*	-.01	-.17*	-.13	-.17*
Pro	-.02	.05	-.01	-.14	-.04
Refutation	.01	.13	.12	.09	.12
Support	.16	.21*	.07	-.04	.10
Arrow	-.13	-.00	-.02	-.09	-.08
Line	-.02	-.04	-.04	-.07	-.07

** $p < .01$, * $p < .05$.

The positive correlation between the presence of Support items and Segment argumentation suggests that the Support elements encourage participants to support their arguments, something they might have paid less attention to in other conditions. An important factor in determining the Segment argumentation score is the presence of supporting elements. The use of Refutation and Support of arguments both show an overall positive tendency, and this might indicate that more advanced and detailed argumentative structure and content enhance the writing process and thus the final product.

Using sources for argumentation

The final question on the Diagram-text relation is concerned with the origin of the arguments in the text and the Diagram: whether these were present in the given sources, or generated by the participants themselves. The descriptive statistics for the origin of arguments are shown in Table 5.5. It shows that most of the arguments the participants use are taken from the given information sources. Only about five arguments out of 38 – about 1 in 8 – cannot be traced back to these sources and are thus considered self-generated.

Table 5.5: Descriptive statistics for number of arguments by source in the text and the Diagram.

	Mean	**SD**	**Minimum**	**Maximum**
Total arguments	37.52	14.96	11	82
Self-generated arguments	4.58	3.13	0	13
Arguments from sources	32.95	14.25	10	78

N = 73 dyads.

The results of the correlation analysis are shown in Table 5.6. We found a negative relation between arguments from the sources and the Textual structure, the Overall argumentation and the Mean scores, but no significant correlations for the Self-generated arguments. The Total number of arguments also correlates negatively with Mean text score. The Arguments from sources and the Total number of arguments both show negative tendencies, whereas the Self-generated arguments show a slight positive tendency for the text scores. It seems that a well written argumentative text contains a limited number of properly supported arguments. The more arguments there are, the more a text might resemble an enumeration, or the less thoroughly each argument can be supported, or the more likely the participant is to lose the overview of the text, leading to a less transparent textual structure and overall argumentation. We did not find any significant correlations for Segment argumentation or Audience focus.

Table 5.6: Pearson correlations between number of arguments by source and text quality.

	Textual structure	**Segment argumentation**	**Overall argumentation**	**Audience focus**	**Mean score**
Total no. of arguments	-.16	-.09	-.14	-.14	-.18*
Self-generated arguments	.02	.05	.16	.05	.09
Arguments from sources	-.17*	-.10	-.18*	-.16	-.21*

** $p < .01$, * $p < .05$.

The absence of significant correlations for Self-generated arguments stems from the fact that the participants rarely used arguments that could not be traced back to the given sources, as can be seen in Table 5.5. This might be a pity, as the positive tendency for Self-generated arguments suggests that wholly original contributions add to the quality of the text. In an earlier study we also found a positive relation between the number of self-generated arguments and the quality of the argumentative text (Andriessen, Erkens & Overeem, 1996).

5.2 Linearizing content with the Outline tool

Four out of six experimental conditions contained the Outline tool: the Outline, Diagram-Outline, Diagram-Outline-Advisor and Outline-Advisor conditions. The results of the Outline analysis are described in this section, and the results are further compared to text quality and tool use, and between conditions. The first section deals with the structural complexity of the outlines, and in the second section the content of the outlines is discussed. The third section deals with the comparison of the structure and content of the outlines to the final texts.

5.2.1 Structural complexity

Two dimensions of structural complexity were distinguished: formal structure and argumentative structure. The formal structure of an outline is formed by the number of hierarchical levels, the number of organizational items per paragraph, and the number of subordinate items per paragraph. The results for these two measures are shown in Table 5.7.

Table 5.7: Descriptive statistics for the structure measures of the outlines.

	Mean	SD	Minimum	Maximum
1 - formal structure	4.89	2.45	.00	12.67
1.1 - no. of hierarchy levels	2.03	.74	.00	4.00
1.2 - no. of organizational items per paragraph	1.30	.79	.00	4.00
1.3 - no. of sub items per paragraph	1.56	1.42	.00	5.67
2 - argumentative structure	1.40	.74	.00	3.50
2.1 - no. of argumentative lines per paragraph	.84	.54	.00	2.67
2.2 - variation in argumentative types	.55	.27	.00	1.11

N = 63 dyads.

The formal structure shows that on average, all paragraphs consisted of the main level plus one (sub measure 1.1), and had approximately 1.5 subordinate lines (sub

measure 1.3). Although we found some extremes, the average outline was neither simple – just one level – nor too complex – with a large number of levels and items for each paragraph. The argumentative structure is reflected in only about half of the lines: the mean number of explicitly argumentative lines is .84, against 1.56 for the number of sub items per paragraph. On the other hand, when the additional organizational items are taken into account, the difference is much smaller: 1.30 to 1.56. On the whole, it seems that the participants did try to explicitly indicate the structure of their text in the Outline.

Multiple comparisons (two-way ANOVA; $p < .05$; Bonferroni differences) showed several differences between the four conditions for formal and argumentative structure. The number of organizational paragraphs is significantly lower in the Diagram-Outline condition than in the Diagram-Outline-Advisor (mean difference -.585) and Outline-Advisor (mean difference -.532) conditions. In addition to the organizational elements, there is also an increase in the variation in argumentative elements when the Advisor is added to the Outline condition (mean difference .210). This suggests that the presence of the Advisor, and the extra oral instruction that accompanies it, encourage the participants to use more formal organizational and argumentative terminology in their outlines, instead of just content items. Also, it encourages them to use richer argumentation by including terms like refutation and counterargument.

At the same time, Table 5.8 shows that the number of organizational items is positively correlated with textual structure. This suggests that the Advisor has an indirect positive influence on textual structure. The higher variation in argumentative types in the outlines seems to be reflected in the score for segment argumentation. However, these correlations are not supported by the ANOVAs for text quality: they do not show any significant differences for the Outline conditions. Also, the variation in argumentative types correlates negatively with audience focus. It seems likely that the Advisor has an indirect effect on segment argumentation through the Outline tool.

Table 5.8: Correlations of outline structure measures with text quality.

	Textual structure	Segment argumentation	Overall argumentation	Audience focus	Mean text score
1. Formal structure	.05	.05	.02	-.07	.02
1.1 No. of hierarchy levels	.07	.05	.02	-.11	.02
1.2 No. of organizational items per paragraph	.20*	.08	-.04	-.04	.07
1.3 No. of sub items per paragraph	-.07	.01	.04	-.04	-.02
2. Argumentative structure	.05	.18*	-.10	-.16	-.03
2.1 No. of argumentative lines per paragraph	.09	.13	-.06	-.10	.00
2.2 Variation in argumentative types	-.03	.25**	-.14	-.23**	-.09

** $p < .01$; * $p < .05$.

Although not supported by differences between Outline conditions, the correlation between segment argumentation and total argumentative structure of the Outline is also positive, and this strengthens the assumption that there is a positive influence of the Outline tool on this aspect of text quality. On the other hand, there is a correlation between variation in argumentative types and total argumentative structure in the outlines and the total score on the Wild Cat Test (.20 and .18, resp.; $p < .05$). Thus, the positive relation between argumentative structure in the outlines and segment argumentation in the text might be caused partly by differences in skill between students.

In our correlational study of tool use we already found that the use of the Outline throughout the task relates positively to argumentation in the final text at paragraph level, although we also found a weak negative relation between clicking into the Outline and overall argumentation of the shared text. Table 5.9 shows the correlations for tool use. The table reveals a clear positive relation between use of the Outline-Advisor (*to Outline tips*) and structure complexity. This strengthens the idea that the Advisor changes the writing process.

Table 5.9: Correlations between structure complexity and Outline use percentage.

	To Outline tips	Open Outline	To Outline
1. Formal structure	.22*	.34**	.29**
1.1 No. of hierarchy levels	.11	.32**	.32**
1.2 No. of organizational items per paragraph	.30**	.20*	.16
1.3 No. of sub items per paragraph	.16	.31**	.24*
2. Argumentative structure	.27**	.07	.20*
2.1 No. of argumentative lines per paragraph	.27**	.10	.16
2.2 Variation in argumentative types	.21*	-.02	.23*

** $p < .01$; * $p < .05$.

The formal outline structure is positively correlated throughout its sub measures with opening the Outline window with the Outline button on the TC3 toolbar. This is also the only outline variable that correlates with opening the Outline tool. Clicking into the Outline that is already open on the participant's desktop also gives positive correlations for formal structure as well as for argumentative structure.

5.2.2 Content complexity

In addition to the structure of the outline and its relation to the structure of the text, we were also interested in finding out what differences there were in type of content in the outline and how these different types might contribute to text quality. The measures used relate to formal content, and comprehensiveness of content. The descriptive data are shown in Table 5.10, and the correlations with text quality in Table 5.11.

Table 5.10: Descriptive statistics for the content measures of the outlines.

	Mean	SD	Minimum	Maximum
3 - formal content: abstract/concrete/mixed	-.07	.44	-.78	1.00
proportion of abstract items	.32	.25	.00	1.00
proportion of mixed items	.26	.32	.00	1.00
proportion of concrete items	.39	.30	.00	.89
4 - comprehensiveness of content: phrase complexity	.55	.44	-.47	1.00
proportion of key word items	.35	.23	.00	1.00
proportion of clause items	.41	.22	.00	.83
proportion of sentence items	.20	.20	.00	.70
proportion of paragraph items	.01	.04	.00	.27

Table 5.11: Correlations of outline content measures with text quality.

	Textual structure	**Segment argumentation**	**Overall argumentation**	**Audience focus**	**Mean text score**
3. Formal content: abstract/concrete/mixed	.13	.14	-.05	-.06	.05
proportion of abstract items	.08	.13	-.03	-.15	.00
proportion of mixed items	.07	.06	-.15	.08	.01
proportion of concrete items	-.13	-.10	.05	-.04	-.07
4. Comprehensiveness of content: phrase complexity	.16	-.10	-.11	-.10	-.03
proportion of key word items	.18*	-.09	-.03	-.04	.04
proportion of clause items	-.04	.04	-.17	-.12	-.12
proportion of sentence items	-.13	.15	.01	.03	-.01
proportion of paragraph items	-.19*	.06	.06	.10	.00

** $p < .01$; * $p < .05$.

At -.07, the mean for formal content is nicely in between the possible extremes -1 and +1. This means that the average outline is neither abstract nor concrete in nature, but rather a mixture of abstract and concrete lines. This measure does not show any correlations with text quality. There are a few differences between conditions that seem to confirm the findings for the Advisor in the section on outline structure. The Outline-Advisor group used significantly more abstract items than the Diagram-Outline and the Outline groups (mean differences -.254 and -.203; $p < .05$). The overall formal content also shows a significant difference between the Diagram-Outline and Outline-Advisor conditions (mean difference of -.427). In addition, the Outline group used more mixed items than the Diagram-Outline group (mean difference -.196). These results suggest that adding the Diagram might lead to fewer mixed items, possibly because of the time divided between Diagram and Outline: the more time can be spent on the Outline, the more precise the content of the lines will be – stating both content (concrete) and function (abstract) of the paragraph. To check this assumption, the data for formal content were compared to the tool use percentages, as shown in Table 5.12. The positive correlations with the Outline tips means that those dyads who used the Advisor more, used more purely abstract items in their outlines. On the other hand, those dyads who frequently activated the Outline tool used significantly fewer abstract items, so their outlines were more concrete and mixed.

Table 5.12: Correlations between content complexity and Outline use percentage throughout the task.

	To Outline tips	Open Outline	To Outline
3. Formal content: abstract/concrete/mixed	.24*	-.14	-.18*
proportion of abstract items	.28**	-.05	-.17
proportion of mixed items	-.06	-.04	.14
proportion of concrete items	-.11	.16	.13
4. Comprehensiveness of content: phrase complexity	.16	.14	-.08
proportion of key word items	.10	.02	-.23*
proportion of clause items	.09	.17	.26**
proportion of sentence items	-.09	-.09	.19*
proportion of paragraph items	-.14	.01	-.02

** $p < .01$; * $p < .05$.

The second outline content measure relates to the comprehensiveness, or elaborateness, of the content of the lines. The mean proportions in Table 5.10 indicate that the emphasis was on the shorter types, and only very few full paragraphs were found. This is encouraging, as Table 5.11 shows that the correlations between use of paragraph items and textual structure is negative, whereas there is a positive relation between use of key words in the Outline and the structure of the text. When we look at the differences between conditions, the Bonferroni analyses again show some interesting differences for the Diagram-Outline condition (see Table 5.13). This condition has the most elaborate average line type in the outlines. Although we found no significantly lower text scores for this group, these results suggest, in combination with the correlations for comprehensiveness, that less complex lines are related to better textual structure. When we look at tool use percentage in Table 5.12, we find that the use of key words correlates negatively with clicking to the Outline tool, while clauses and sentences correlate positively. This is not surprising, as writing longer items is likely to take more time than writing key words, and using a tool more often generally implies using it longer.

Table 5.13: Comparison of conditions on comprehensiveness of content. Mean differences (Bonferroni).

		DOA	O	OA
Total	DO	.356	.269	.473
Clauses	DO	.186	.157	.146
Sentences	DO			-.180

5.2.3 Correspondence to the final text

The last three measures were meant to determine the correspondence of the outline structure and content to the text and vice versa. The descriptive data are shown in Table 5.14. No less than 89% of the outline items were ordered correctly in the text (measure 5). An average of 87% of all outline items are present in the text, and 79% of all text paragraphs are included in the outlines. However, this does not mean that students worked from the outlines, as they were required to make sure that the final version of their Outline matched the final version of their text. Some participants therefore chose to make their outline after finishing the text, instead of using it as a planning tool. There were no significant differences between conditions for the three correspondence measures.

Table 5.14: Descriptive statistics for the correspondence measures of the outlines.

	Mean	SD	Minimum	Maximum
5 - structure: proportion of items in correct order and place	.89	.23	.00	1.00
6 - content: proportion of outline item presence in text	.87	.25	.00	1.00
7 - content: proportion of presence of text paragraphs in outline	.79	.25	.00	1.00

Table 5.15: Correlations of outline correspondence measures with text quality.

	Textual structure	Segment argumentation	Overall argumentation	Audience focus	Mean text score
5. structure: order and place of items	.04	.19*	-.15	-.15	-.05
6. content: item presence in text	.11	.23**	-.05	-.17	.02
7. content: presence of text paragraphs in outline	-.09	.26**	-.09	-.01	.00

** $p < .01$; * $p < .05$.

As Table 5.15 shows, all three measures correlate positively with segment argumentation. The content correspondence measures also correlate positively with the linearization measure in the Wild Cat Test (both at .20; $p < .05$), which suggests that participants who are good at linearizing contents are also good at keeping an overview of their plan compared to their final product. This is supported by the positive correlations for *to Outline* percentage shown in Table 5.16.

Table 5.16: Correlations between structure and content correspondence and Outline use percentage throughout the task.

	To Outline tips	Open Outline	To Outline
5. structure: order and place of items	-.05	.15	.29**
6. content: item presence in text	.07	-.04	.24**
7. content: presence of text paragraphs in outline	.00	.01	.20*

** p < .01; * p < .05.

If most students did indeed use the Outline after writing the text, as we suggested before, we would expect to find a negative relation between use of the Outline tool during the last phase of the writing process and text quality. However, for segment argumentation, the text quality measure that seems to be influenced by the content and structure of the outline, we did not find any significant correlations with Outline or Outline-Advisor use. Also, paired samples T-tests between the phases showed that the Outline tool is used less frequently as the writing process advances. On the other hand, the correlations in Table 5.17 show that there is a positive relation between correspondence and Outline tool use during the third phase of the writing process. This could be the result of checking and updating the text with the Outline as well as vice versa.

Table 5.17: Correlations between structure and content correspondence and Outline use percentage in different phases of the writing process.

		Outline open	To Outline
5. structure: order and place of items	1st phase	.15	.16
	2nd phase	.03	.06
	3rd phase	.20*	.36**
6. content: item presence in text	1st phase	.08	.09
	2nd phase	-.19*	.01
	3rd phase	.15	.38**
7. content: presence of text paragraphs in outline	1st phase	.08	.22*
	2nd phase	-.12	-.04
	3rd phase	.14	.26**

5.3 Conclusion

The analyses for 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 text reveal a painful discrepancy between the two: only about a third of the arguments are found both in the text and the Diagram. When we look at the types of elements used in the Diagrams, we notice that using Support and Refutation has a slight positive relation with the final written product,

whereas the use of counterarguments and plain information shows the opposite. Finally, 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.

On the whole, we found that there is a positive effect of the Advisor on the structure of the outlines, and possibly also a positive influence of the argumentative structure of the outline on Segment argumentation. At the same time, there is a slight positive effect of use of the Advisor on the conciseness of the content of the outline, and a more concise outline is in turn positively related to Textual structure. Finally, there is a strong relation between Segment argumentation and outline-text correspondence. These results taken together lead to the conclusion that there is a positive influence of the availability and proper use of the Outline tool and its Advisor on Textual structure and on argumentation at paragraph level in the argumentative text.

CHAPTER 6 THE PLANNING PROCESS AND COORDINATION

6.1 The communicative function of the dialogue

6.1.1 Structural characteristics of the dialogue

This section[3] 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 6.1 shows the distribution for the five communicative functions for the Control group and for each experimental condition.

Table 6.1: Distribution of communicative function in the dialogue in percentages.

	Total M	**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	21.26	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	**C SD**	**D SD**	**DA SD**	**DO SD**	**DOA SD**	**O SD**	**OA 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
Total number of contributions	192.40	162.37	126.82	207.14	220.87	164.16	151.61	263.67

The distribution for all groups together shows that Informatives occur most frequently (38%), followed by Responsives (24%). Argumentatives make out an encouraging 10% of the communicative functions, and Imperatives are the least frequent with 8%.

[3] The analyses of the Dialogue Acts in Section 6.1 were done by Floor Scheltens as part of her Master's thesis

Table 6.2 shows the significant differences between conditions. Compared to the other conditions, the Control group uses 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 6.2: Mean differences of communicative functions between conditions (Bonferroni). Values are row label – column label. Only significant differences are shown.

Argumentatives	**Experimental**	**D**	**DA**	**DO**	**DOA**	**O**	**OA**
C	1.12	1.75	1.53			1.72	
Imperatives							
C					2.68		
D					4.34	2.78	2.78
DA					4.38	2.82	2.82
DO					3.06		
Informatives							
C					-4.71		
DA					-4.34		
OA					-6.28		
Responsives							
C							-3.58
D							-5.30
DA							-4.19
DOA							-6.72

Table 6.3 shows the mean percentages of the specific Dialogue acts. The corresponding standard deviations can be found in Table 1.1 in Appendix 11. 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 (4%). 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 (10%), followed by proposals (6%) and open questions (5%). Urging the partner to take action or fulfill a task is the more frequent Imperative with 5%, 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 (4%). Finally, within Responsives the most frequent Dialogue acts are Confirmation (13%) and plain replies (Reply Statement 8%).

Table 6.3: Means of Dialogue act percentages.

	Total	C	D	DA	DO	DOA	O	OA
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	.98
Contra	3.86	3.39	4.17	4.32	3.50	3.65	4.49	3.93
Disjunctive	.69	.59	.78	.77	.80	.63	.70	.52
Reason	1.66	1.59	1.80	1.91	1.66	1.26	1.60	1.61
Then	.80	.81	.74	.88	.73	.81	.82	.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	.60	.56	.80	.40	.56	.52	.79	.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	.55	.93	.35	.41	.37	.32	.48	.55
Evaluation Neutral	.35	.41	.36	.31	.46	.20	.21	.38
Evaluation Positive	2.84	2.92	2.52	2.99	2.66	2.38	2.63	3.93
Performative	.97	.94	1.02	1.11	.76	.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	.67	1.51	.21	.48	.52	.23	.27	.29
Statement Social	1.27	1.72	1.17	.79	1.39	1.11	.93	1.44
Responsives								
Acceptation	1.39	1.79	1.32	1.37	1.20	.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	.14	.16	.22	.15	.13	.12	.05	.11
Reply Confirm	3.03	2.09	3.96	3.37	3.31	3.62	3.01	2.98
Reply Deny	.61	.41	.81	.67	.65	.64	.69	.67
Reply Performative	.04	.04	.11	.06	.01	.01	.03	.03
Reply Statement	3.78	3.30	4.46	3.77	4.29	3.68	3.61	3.77

Transitions between Dialogue acts

Figures 6.1 to 6.7 show the MEPA transition diagrams for each condition. 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, shown in Figure 6.1, differs from the experimental conditions: 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 stimulate structuring of the dialogue.

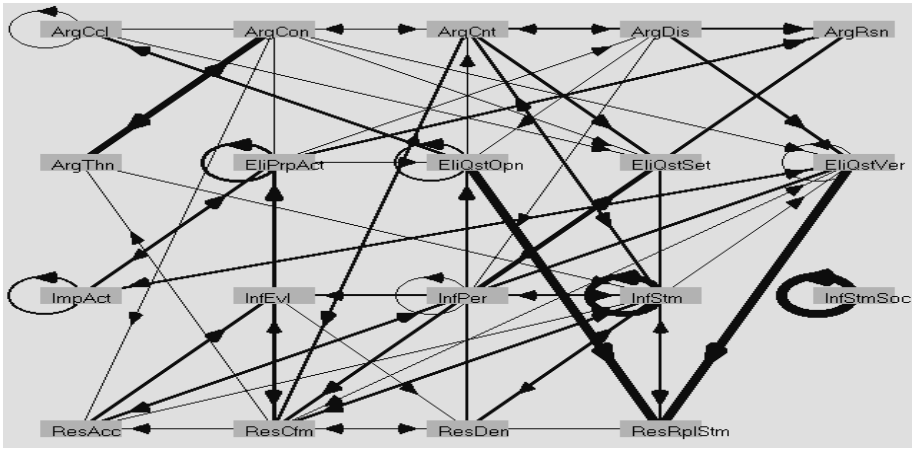

Figure 6.1: Transition diagram for the Control group.

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. Only the Diagram-Advisor condition – shown in Figure 6.5 – 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.

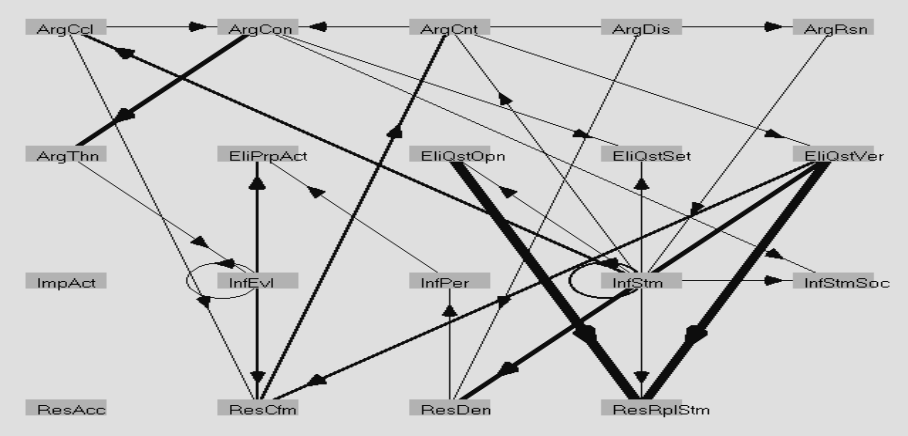

Figure 6.2: Transition diagram for the Diagram condition.

When we look at the transition between 'if'-argumentatives (ArgCon) and 'then'-argumentatives (ArgThn), we find that this transition is not significant for the Outline and Outline-Advisor conditions (Figures 6.3 and 6.6), 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.

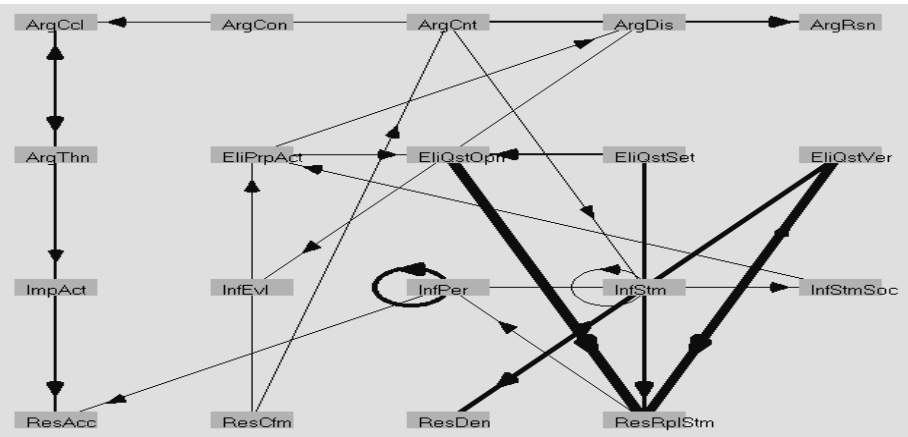

Figure 6.3: Transition diagram for the Outline condition.

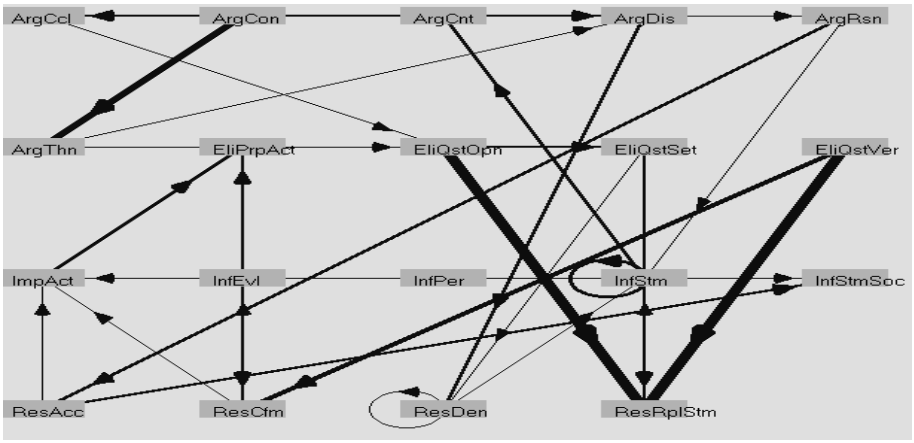

Figure 6.4: Transition diagram for the diagram-Outline condition.

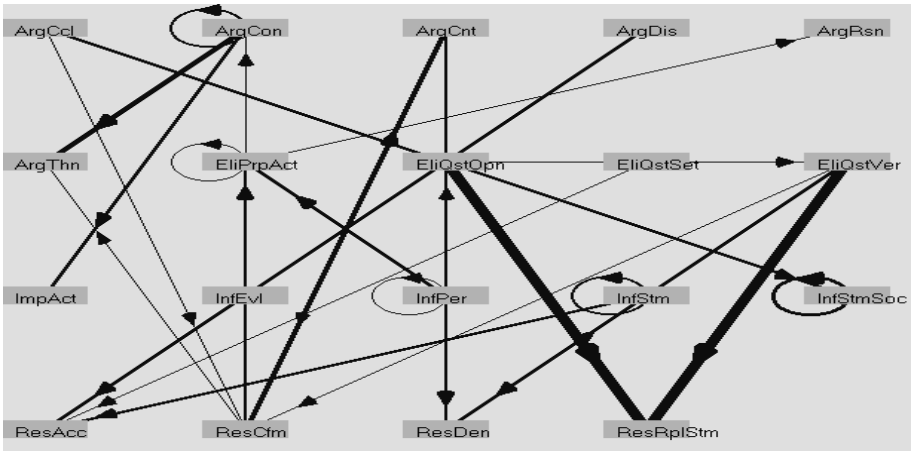

Figure 6.5: Transition diagram for the Diagram-Advisor condition.

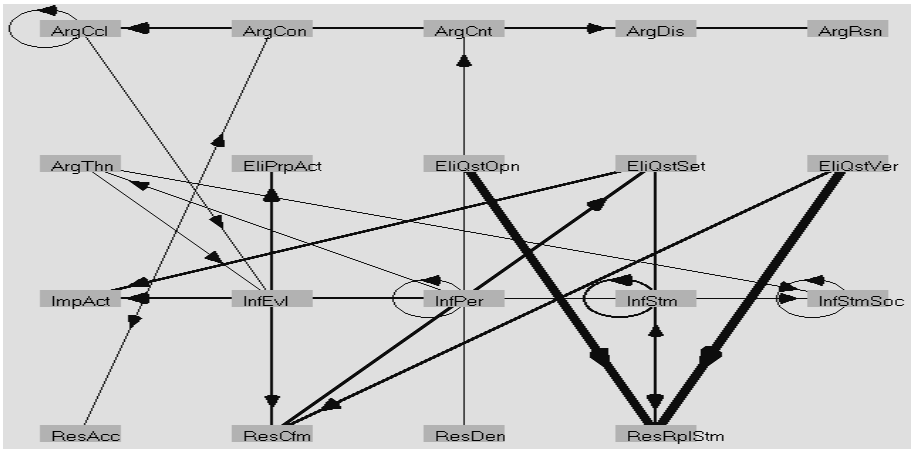

Figure 6.6: Transition diagram for the Outline-Advisor condition.

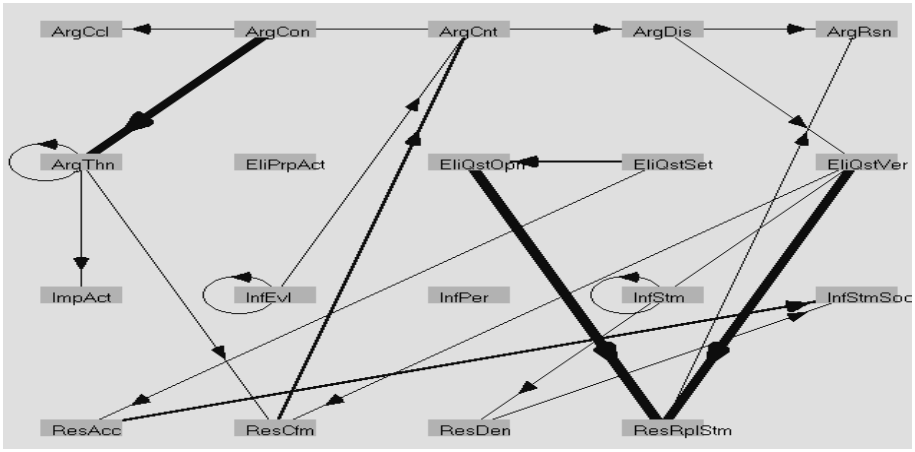

Figure 6.7: Transition diagram for the Diagram-Outline-Advisor condition.

High and low text quality

Figure 6.8 shows the transition diagram for the Control group dyads with a mean text score above 7.0, and Figure 6.9 shows the transition diagram for the Control group dyads scoring below 5.5 on mean text quality. On the whole, the transition diagrams of the Control group were the least structured of all conditions. In groups with unstructured transition diagrams we would expect the largest and most readily visible differences between high and low scoring dyads. Compared to high scoring dyads, low scoring pairs show more different transitions, and more autocorrelations. The high scoring partners show a more clearly structured dialogue, with recurring functional patterns, such as condition-consequence (ArgCon-ArgThn). The two transition diagrams show little correspondence,

although responding to questions with a statement is significant in both groups, and so is the autocorrelation for informative statements. The difference between high and low scoring dyads is not just visible in the level of structuredness, but also in the types of dialogue patterns.

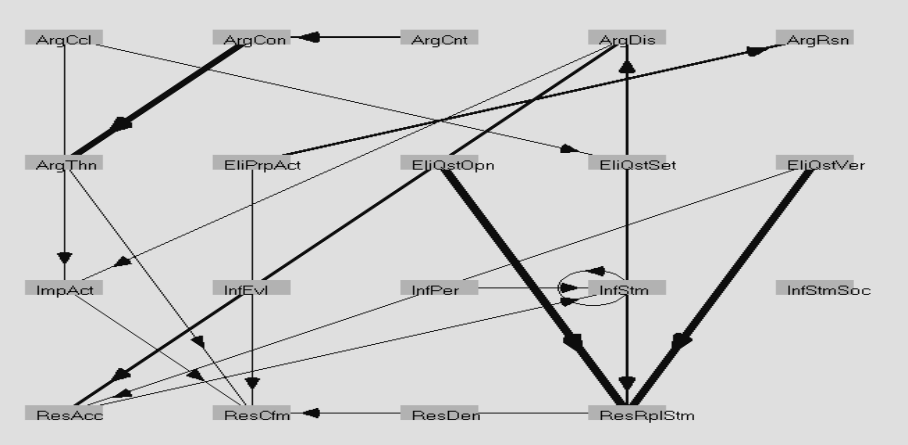

Figure 6.8: Transition diagram for dyads scoring > 7.0 on text quality.

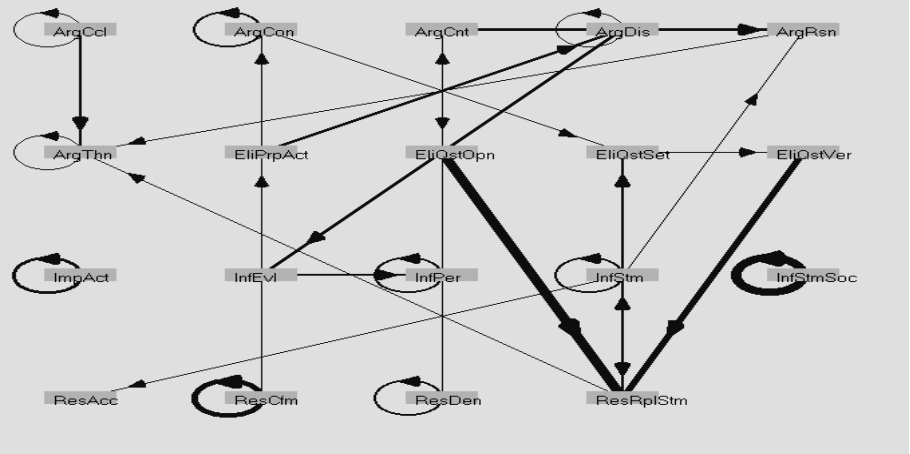

Figure 6.9: Transition diagram for dyads scoring < 5.5 on text quality.

When we compare the transition patterns in the dialogues of the different conditions, we find that the Control group shows very different patterns from the experimental conditions. The dialogue patterns of the Control group are much less clearly structured. This might imply that planning tools stimulate a structured dialogue. In addition, the Control group shows a lot more autocorrelations than the other conditions: the group displays more sequences of utterances with the same

communicative function. Within the experimental conditions, the Diagram-Advisor is the odd one out: its dialogue patterns are less structured, and it shows a relatively high occurrence of autocorrelations. The Diagram-Advisor condition has a negative effect on the quality of the textual structure. Possibly, an unstructured dialogue influences the structure of the text. All conditions have significant numbers of question-statement sequences, and we found the same significant pattern for high and low scoring groups. Apart from this, there are very few similarities between high and low scoring dyads. Low performance dyads show less structured dialogue patterns than high scoring dyads, and they also display a lot more autocorrelations, which means that, more than high scoring pairs, the low scoring dyads show repetitive patterns of Dialogue acts in their chats.

Relation of dialogue structure with text quality

Four out of five measures for dialogue structure show some significant correlations with the quality of the final text (Table 6.4). The Elicitatives correlate positively with most of the text scores, while the Informatives are predominantly negatively correlated. This suggests that asking questions and making proposals leads to a productive argumentative writing process, whereas exchanging neutral information brings about the opposite. This assumption is supported by the more detailed analyses of the subtypes of Dialogue acts: these show that the main contributors to the negative correlations for Informatives are the nonsense statements and the social talk. The Argumentatives and Imperatives each correlate positively with only one text quality measure. The Responsives do not correlate with text quality at all.

Table 6.4: Correlations between communicative functions and text scores.

	Textual structure	**Segment argumentation**	**Overall argumentation**	**Audience focus**	**Mean text score**
Argumentatives	-.01	-.01	.13*	.05	.06
Elicitatives	.00	.17**	.12*	.21**	.18**
Imperatives	.14*	.01	.01	-.09	.02
Informatives	-.06	-.10	-.24**	-.13*	-.19**
Responsives	-.03	-.02	.10	.01	.02

* $p < .01$; ** $p < .05$.

We tested three models for the effects of dialogue structure and condition on the quality of the final text. Figure 6.10 shows the first model, stating that the experimental condition influences the structure of the task related chat as measured by communicative function.

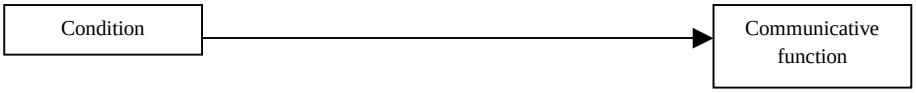

Figure 6.10: Effect of condition on dialogue structure.

Table 6.5 shows the directions of the effects of condition on the communicative function we found in comparison with the Control group. The table is based on regression analyses with the communicative function measures as the dependent variable and the experimental conditions as the independent variable (dummy variable). In the table only the direction of the significant regression weights are shown by means of a + or a -. The condition does not have an effect on the use of Elicitatives. The Diagram-Outline condition does not influence any of the communicative functions in either direction. The Diagram, Outline, and Diagram-Advisor conditions all have a positive effect on the number of Argumentatives. This suggests that a moderate availability of planning tools has a positive influence on the number of arguments in the chat, and that the effect disappears when too many planning tools are present.

Table 6.5: Relation between experimental condition and communicative function.

	Argumentative	**Elicitative**	**Imperative**	**Informative**	**Responsive**
D	+		-	-	+
O	+			-	
DO					
DA	+		-		
OA					-
DOA			+	-	+

The Diagram and Diagram-Advisor conditions both negatively influence the number of Imperatives in the chat, whereas the Diagram-Outline-Advisor condition has a positive effect on the number of Imperatives. Perhaps the larger amount of work for the participants in the latter condition decreased the opportunities for extensive deliberation, leading to more direct behavior by using more Imperatives. The Diagram and the Diagram-Outline-Advisor conditions both show positive effects on the number of Responsives, whereas the Outline-Advisor condition influences the number of Responsives negatively. Although there is an effect of condition on communicative function, Table 6.5 shows no clear direction for this relation. Still, the conditions might have a clear positive or negative effect on text quality through the communicative functions.

Figure 6.11 shows the second theoretical model, where the dialogue structure as measured by communicative function influences the quality of the final written product.

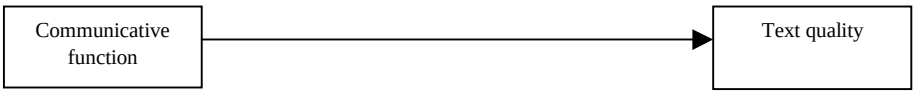

Figure 6.11: Effect of dialogue activity on text quality.

Table 6.6 shows the directions of influence for the model in Figure 6.11, based on significant regression weights in analyses with the communicative function measures as independent and the text quality measures as dependent variables. A relatively high number of elicitatives positively influences the segment argumentation and the audience focus of the final text, but the experimental condition does not affect the elicitatives. Thus, the positive effect of the Elicitatives was not caused by the presence or absence of the planning tools. Informatives negatively influence the mean score and the argumentation in the text as a whole. In turn, the Diagram-Outline-Advisor condition has a negative effect on the Informatives, which leads us to suspect that this condition positively influences overall argumentation through the Informatives. Imperatives positively influence textual structure, possibly because the use of Imperatives reduces ambiguity in the dialogue between the collaborating partners, leading to reduced ambiguity of the textual structure. At the same time, the Diagram and Diagram-Advisor conditions negatively influence the use of Imperatives. We expect that these conditions negatively influence textual structure through the Imperatives. On the other hand, the Diagram-Outline-Advisor condition positively influences the use of Imperatives, and thus might positively influence text quality through this communicative function.

Table 6.6: Relation between communicative function and text quality.

	Textual structure	**Segment argumentation**	**Overall argumentation**	**Audience focus**	**Mean score**
Argumentative					
Elicitative		+		+	
Imperative	+				
Informative			-		-
Responsive					

To check the possibility that condition and communicative function also affect text quality independently of each other, we tested the model presented in Figure 6.12.

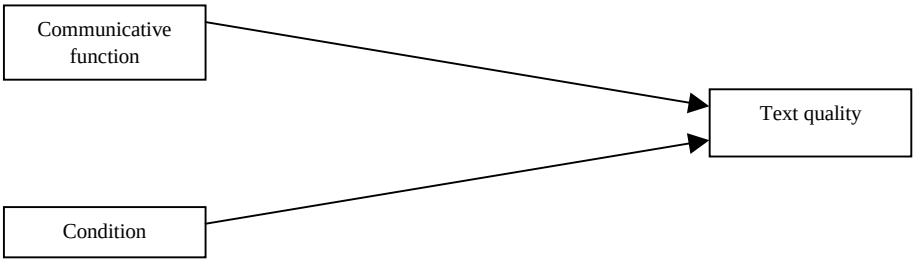

Figure 6.12: Effect of dialogue structure and condition on text quality.

Table 6.7 shows the directions of the effects of condition and communicative function on text quality. In these regression analyses all communicative function measures and all conditions were entered in the regression with the quality measures as dependent variables. Independent of the dialogue activity, the Diagram-Advisor condition negatively influences textual structure, whereas the Diagram-Outline-Advisor condition has a positive effect. Adding a third planning aid – the Outline – seems to enhance the structural quality of the final text (either directly or through some unidentified factor).

Table 6.7: Relation of communicative function and experimental condition with text quality.

	Textual structure	**Segment argumentation**	**Overall argumentation**	**Audience focus**	**Mean score**
Argumentative					
Elicitative		+		+	
Imperative					
Informative			-		-
Responsive					
D					
O					
DO					
DA	-	-			
OA					
DOA	+				

Elicitatives have a positive effect on audience focus and on quality of the argumentation at segment level. Just as in the second model in Figure 6.11 Informatives have a negative effect on the argumentative quality of the text as a whole. In addition, a lower percentage of Informatives in the chat goes together with higher overall text quality. Collaboration with an argumentatively and structurally good result requires little informing, but frequent argumentation, asking accurate questions, responding to the partner, and use of imperatives.

None of the three models presented here needs to be rejected on the basis of the regression analyses. Combining the three models results in the model shown in Figure 6.13. This model hypothesizes that the experimental condition affects text quality through the dialogue, but also directly, independent of the communicative function of the chat.

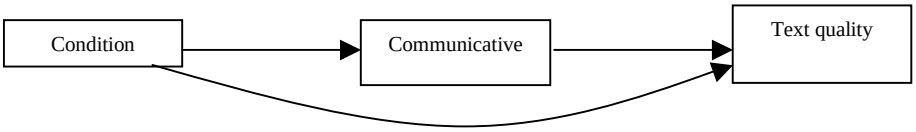

Figure 6.13: Effect of condition and dialogue structure on text quality.

6.1.2 Asymmetry of contribution

This section contains a description of the results for the analyses of symmetry of contribution. The used measure of asymmetry is the absolute difference between the contribution percentages of the partners. A higher value indicates more unequal contributions to the task related chat. When the difference is 0%, there is full symmetry; asymmetry of 100% means that only one partner contributed to the corresponding communicative function. We used five measures of asymmetry, one for each main type of communicative function, and a total measure of asymmetry. Table 6.8 shows the means and standard deviations for asymmetry in the different conditions.

Table 6.8: Asymmetry of contribution in percentages.

	All	**C**	**D**	**DA**	**DO**	**DOA**	**O**	**OA**
	M	**M**	**M**	**M**	**M**	**M**	**M**	**M**
Asymmetry of argumentatives	24.40	25.70	13.68	23.42	26.30	27.54	28.52	24.79
Asymmetry of elicitatives	20.26	17.65	15.62	19.92	27.72	16.93	18.07	28.83
Asymmetry of imperatives	27.61	25.31	28.64	32.14	26.00	22.93	32.95	22.80
Asymmetry of informatives	17.26	18.45	11.29	17.51	16.31	15.30	21.16	19.21
Asymmetry of responsives	16.61	17.40	9.05	16.96	18.93	16.44	21.62	11.87
Asymmetry of total	11.24	12.56	7.86	10.96	9.60	10.11	12.87	14.32
N (dyads)	145	39	17	26	23	11	18	11
	All	**C**	**D**	**DA**	**DO**	**DOA**	**O**	**OA**
	SD	**SD**	**SD**	**SD**	**SD**	**SD**	**SD**	**SD**
Asymmetry of argumentatives	18.40	16.99	11.74	19.91	15.52	25.68	19.91	18.51
Asymmetry of elicitatives	16.26	13.70	12.78	14.09	21.77	14.02	18.24	12.97
Asymmetry of imperatives	18.34	17.52	18.94	22.29	11.62	19.50	20.02	14.91
Asymmetry of informatives	11.41	13.37	7.99	8.15	9.42	15.64	9.09	14.23
Asymmetry of responsives	13.47	13.80	9.49	12.88	12.87	12.05	16.21	11.91
Asymmetry of total	10.30	11.89	6.73	7.99	10.42	13.04	9.23	11.12

The average asymmetry for all conditions and communicative functions taken together is 11%. This means that in each dyad one participant contributes about 11% more utterances to the task related chat than his/her partner. The difference increases when we look at the five types of communicative function separately. The mean difference between the partners in the use of Imperatives is about 28%, while the differences for Argumentatives, Elicitatives, Informatives, and Responsives are 24%, 20%, 17%, and 17%, respectively. Bonferroni comparisons of the conditions give the significant differences shown in Table 6.9. This confirms the observation that the contributions are more equally distributed in the Diagram condition than in any of the other groups, especially in comparison to the Control group, and the Outline and Diagram-Outline conditions, with regard to Argumentatives, Informatives and Responsives.

Table 6.9: Mean differences of symmetry in Dialogue acts between conditions (Bonferroni). Values are row label – column label. Only significant differences are shown.

Argumentatives	**C**	**D**	**DA**	**DO**	**DOA**	**O**	**OA**
D	12.02			12.62		14.84	
Elicitatives							
C				10.07			
D				12.09			
Informatives							
C		-7.16					
O		-9.86					
Responsives							
C		-8.34					
DO		-9.88					
O		-12.57					

Symmetry of contribution and text quality

Table 6.10 shows the correlations between the measures of asymmetry and the measures of text quality. The total asymmetry correlates negatively with all text measures except Segment argumentation. This confirms the hypothesis that equality of contribution is positively related to text quality for all communicative functions. The asymmetries of Argumentatives, Imperatives, Informatives, and Responsives all correlate negatively with textual structure: equal use of these communicative functions is positively related to the formal structure of the text. Only the asymmetry of Elicitatives does not correlate with any of the text quality measures. This is striking, as we would expect a Responsive to follow an Elicitative, thus leading to similar correlations for both of their measures of asymmetry. It seems that symmetry of communication, especially in terms of argumentation, providing information, and responding to the partner, is an important factor in delivering a good collaborative product.

Table 6.10: Pearson correlations between asymmetry of contribution and text quality.

Asymmetry measure	**Textual structure**	**Segment argumentation**	**Overall argumentation**	**Audience focus**	**Mean text score**
Total	-.18**	-.08	-.16**	-.17**	-.22**
Argumentatives	-.17**	-.13*	-.10	-.22**	-.22**
Elicitatives	-.07	-.03	.00	-.05	-.06
Imperatives	-.21**	-.08	-.01	-.11	-.14*
Informatives	-.14*	-.16**	-.13*	-.10	-.19**
Responsives	-.12*	-.09	-.14*	-.13*	-.15*

** $p < .01$; * $p < .05$.

Regression analyses were performed to test the effects of experimental condition on symmetry of contribution, the effect of asymmetry of contribution on text

quality, and the effect of both condition and asymmetry of contribution on text quality. The model for the first effect is shown in Figure 6.14.

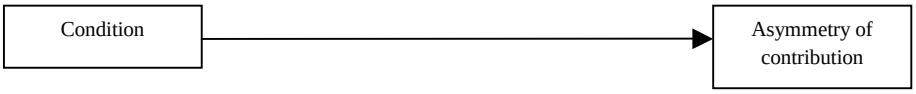

Figure 6.14: Effect of experimental condition on asymmetry of contribution in the dialogue.

The above suggests that the Diagram condition shows a higher symmetry of contribution to the discussion than the other conditions. Table 6.11 shows the directions of the effects for the different conditions on the five communicative functions and their total.

The Diagram condition shows a positive effect on equality of contribution for Argumentatives, Informatives and Responsives. The Outline condition negatively affects the symmetry of contribution of Imperatives and Informatives. The Diagram-Outline condition and the Outline-Advisor condition both contribute negatively to the symmetry of Elicitatives, and the Diagram-Advisor condition reduces the symmetry of Imperatives. When the Diagram is the only available planning tool, the task oriented chat is more symmetrical.

Table 6.11: Relation between condition and symmetry of contribution.

	Total	**Argumentatives**	**Elicitatives**	**Imperatives**	**Informatives**	**Responsives**
D	-	-			-	-
O				+		
DO			+			
DA				+		
OA			+			-
DOA						

On the whole, the Diagram condition shows a lower level of asymmetry than the Control group. This implies that the presence of the Diagram tool enhances symmetry of contribution. The Outline, Diagram-Outline, Outline-Advisor, and Diagram-Advisor conditions generally show higher levels of asymmetry than the Control group. This indicates that these conditions have a negative influence on equality of contribution.

A multiple regression analysis was performed to test the effect of asymmetry of contribution on text quality, visualized in the model in Figure 6.15. Table 6.12 shows the direction of the effects found for this model. Equal contribution of Argumentatives, Imperatives, Informatives, and Responsives has some positive effect on text quality. Equal contribution of Elicitatives does not influence the final text, and there are no negative influences.

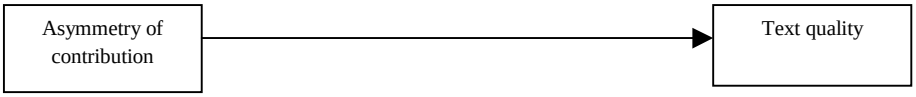

Figure 6.15: Effect of asymmetry of contribution in the dialogue on text quality.

Table 6.12: Effect of asymmetry of contribution on text quality

Asymmetry measures	**Textual structure**	**Segment argumentation**	**Overall argumentation**	**Audience focus**	**Mean text score**
Argumentatives	+			+	+
Elicitatives					
Imperatives	+				+
Informatives		+			
Responsives			+		

Combining the results from Table 6.11 and Table 6.12 leads to the following conclusions. Through equal contribution, the Diagram condition has a positive effect on all five text quality scores. The Outline condition and the Diagram-Advisor condition through the asymmetry of Imperatives have a negative influence on Textual structure and Mean text score. The Outline-Advisor condition has a positive effect on Overall argumentation through asymmetry of Responsives.

To check whether the condition and the symmetry of contribution also influence text quality independent of each other, the model in Figure 6.16 was tested. Table 6.13 shows the directions of effects for this model from the regression analyses.

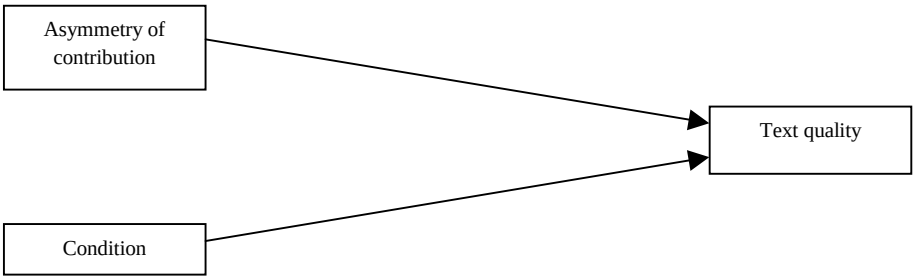

Figure 6.16: Effects of asymmetry of contribution and condition on text quality.

Table 6.13: Effects of asymmetry of contribution and condition on text quality.

	Textual structure	**Segment argumentation**	**Overall argumentation**	**Audience focus**	**Mean text score**
Argumentatives	+			+	+
Elicitatives					
Imperatives	+				+
Informatives		+			
Responsives			+		
D					
O					
DO					
DA	-				
OA		+			
DOA	+				

With the influence of the asymmetry measures kept constant, the Diagram-Outline-Advisor condition does influence Textual structure positively. The Outline-Advisor condition has a positive influence on Segment argumentation, but the Diagram-Advisor condition influences Textual structure negatively. The Diagram, Outline and Diagram-Outline conditions did not have any significant effects on text quality compared to the Control group. Both asymmetry of Argumentatives and of Imperatives have a positive influence on Textual structure as well as on the Mean text score. Asymmetry of Argumentatives also influences Audience focus positively, and equal contribution to Informatives has a positive effect on Segment argumentation. Equal contributions to Responsives result in better Overall argumentation.

The regression analyses lead to the conclusion that the models discussed above do not need to be rejected. Combining the models gives the model shown in Figure 6.17. The condition influences text quality both through asymmetry of contribution and directly. The Diagram condition influences the text positively through equality of contribution, whereas the conditions with both the Outline tool and the Advisor have this positive effect independently of asymmetry of contribution.

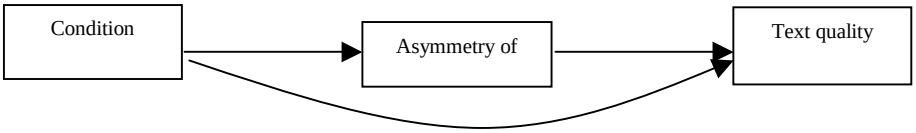

Figure 6.17: Effect of symmetry of contribution and dialogue structure on text quality.

6.1.3 Checking, focusing, and argumentation

In this section we will discuss the three coordination processes found in the collaborative dialogues – focusing, checking and argumentation – and their relation to the conditions and to text quality. First we will present some examples of chats of students collaborating in writing an argumentative text and we will discuss the episodes of coordinated action we looked for in our analyses. The central question here is how collaborating students manage to coordinate and adjust their actions to the processes of shared knowledge construction and problem solving that occur between them. As discussed before we will focus on Focusing (maintaining the same topic of discourse or task strategy), Checking (guarding consistency in shared knowledge construction) and Argumentation (negotiating and coming to agreement about inferences and conclusions). In these coordination processes, there is interaction between the students on task related strategies, cooperative intentions and communication processes.

Focusing episodes

The main goal for the students is to achieve a mutual agreement on the steps to be taken and the inferences to be made (that is, as long as they choose to stay in the collaborative situation). Our impression is that most students are relatively opportunistic (or economically thinking) in this respect. In most cases they will be convinced easily by legitimate arguments of the partner and will not hold on to a lost case (in the ‘studiehuis’ curriculum students are quite experienced in collaborative learning situations). So when focus divergence is noticed, this can be explicitly brought forward and repaired.

Table 6.14: Chat episode with focus divergence (male & female).

Line	Time	Sq	Actor	Dialogue
767	2:07:23	1	0	have you got a title yet?
770	2:08:17	1	1	I'll think it up. ~
768	2:07:31	1	0	okay
771	2:08:17	2	1	if you are going to type now, I'll write down titles via the chat thingie ~
772	2:08:17	3	1	and you only have to say something if you see one you like and not if it's stupid (you can just continue typing then)
773	2:08:32	1	1	okay
774	2:08:51	1	1	from organa to banana
775	2:09:06	1	1	the organ of aunt megan
777	2:09:28	1	1	with an organ in my hand I will walk through the entire land
778	2:09:30	1	0	David!!!!!
779	2:09:35	1	1	Christy
780	2:09:44	1	0	at least I'm working...
782	2:09:50	1	1	I feel sorry for you
784	2:10:29	1	0	what percentage of those things came back?
786	2:10:35	1	1	I'll go look
789	2:10:51	1	0	or do you know numbers?
792	2:12:24	1	1	of the 12 million Dutch receivers (above 18) scant 4.5 million returned the form filled out. ~
793	2:12:24	2	1	of those, 2 million made their organs available after their death. ~
794	2:12:24	3	1	the rest filled out no ~
795	2:12:24	4	1	or left the choice to family or surviving relatives
796	2:13:28	1	1	for all Dutch people who did not fill out the form the family will decide . ~
797	2:13:28	2	1	old codicils will also remain valid
800	2:13:49	1	0	we've got 128 words already.
803	2:13:57	1	1	yes

Table 6.14 shows a short episode of a female and male student working on a text on organ donation after about two hours. The columns show from left to right: Line – protocol line number; Time from start; Sq – sequence number for split utterances; Actor – the student number; and Dialogue – the actual chat text. Before the episode starts the girl (Christy, student 0 in the protocol) is writing a conclusion in the text they both agreed upon before based on one of her sources. Meanwhile, the boy (David, student 1) has to think of a suitable title. Many dyads divide tasks during short periods, mainly during two stages of the task: whilst reading the sources and while actually writing, as is the case here. The writing student, however, expects the partner to watch the paragraph under construcion being developed, so that synchronous collaboration can restart immediately after completing the new draft of the paragraph. On being asked about the title (line 767), the boy gives a humorous series of alternatives (lines 774-777), but is quickly called to order by the girl: "at least I'm working ..." (lines 778-782). So when asked about some facts (the number of codicil forms returned) the boy is

serious again and even gives more information than was asked for. Note that the girl does not explicitly acknowledge the information. She does not need to, as the boy can see that she uses the information in the text. She then remarks satisfied that they have already written 128 words in the shared text.

Checking episodes

Earlier in the same protocol a nice example of a Checking episode occurred where the meanings of several concepts were checked, resulting in taking a position. The episode is shown in Table 6.15 and takes place at the beginning of the session after about 13 to 18 minutes).

The boy refers to the assignment in which it is stated that to start writing an argumentative text and to determine a position on a controversial issue, it is a good idea to begin with a brainstorm session (line 120). The girl accepts this implicit proposal (line 125) and they take turns in writing down concepts: ‘organ donor’ (line 127), ‘transplantation’ (line 133) and ‘failing organs’ (line 134). The girl finds the last concept unclear: ‘Do you mean rejection?’ (line 135) and the boy has to go on explaining his remark until she explicitly accepts (lines 136-140). The meaning of the next concept ‘brain dead’ by the girl is in his turn questioned by the boy (‘What does that have to do with anything’) and the girl has to elaborate (lines 143-144). After an intermezzo about the lollipops they are eating (lines 148-155), they continue their brainstorm session: ‘codicil’, ‘donor registration’ (lines 156 & 157). The codicil concept gives them the opportunity to check each other’s position on the subject (lines 158-163), and they agree on the position they will defend in the text (line 164), and on the breadth of the codicil concept covered, that is every organ, including eyes (lines 165-171). By checking each other’s viewpoints and the meaning given to different concepts, they build a shared frame of reference and construct a collective conceptual landscape (Andriessen, Erkens & Peters, 2002).

Table 6.15: Chat episode with Checking (male & female).

Line	Time	Sq	Actor	Dialogue
120	0:13:12	1	1	in the intro it says that it's a good idea first to brainstorm
121	0:13:22	1	0	yes, I know
123	0:13:23	1	1	so ..
124	0:13:28	1	0	hey, sorry
125	0:13:35	1	0	okay
126	0:13:41	1	0	where are we going to start?
127	0:13:50	1	1	organ donor
128	0:14:13	1	1	(now you have to add a word that has to do with this, guy)
132	0:14:21	1	0	lemme think
133	0:14:29	1	0	transplantation
134	0:14:40	1	1	failing organs
135	0:14:50	1	0	(do you mean rejection?)
136	0:14:52	1	1	no
137	0:14:56	1	0	then what?
138	0:15:24	1	1	that an organ is needed ~
139	0:15:24	2	1	because the original organ failed (your own organ)
140	0:15:31	1	0	ok, that's clear.
142	0:15:39	1	0	brain dead
143	0:16:10	1	1	what does that have to do with anything
144	0:16:33	1	0	but it does! if y're braindead, a doctor can use your organs, not before
148	0:16:51	1	1	I have a lollipop
149	0:16:55	1	0	from ilan?
150	0:16:59	1	1	what do you think..
151	0:17:05	1	0	how could I know!
152	0:17:16	1	1	vanilla chocolate
153	0:17:25	1	0	I had strawberry
154	0:17:27	1	1	yuck
155	0:17:29	1	0	yum!
156	0:17:37	1	1	codicil
157	0:17:46	1	0	donorregistration
158	0:17:49	1	1	you?
159	0:17:54	1	0	what me?
160	0:18:00	1	1	codicil?
161	0:18:04	1	0	yes, you?
163	0:18:06	1	1	yes
164	0:18:14	1	0	so we are for organ donation
165	0:18:22	1	1	are you giving away everything?
166	0:18:25	1	0	yes, you?
168	0:18:33	1	1	your eyes too?
169	0:18:37	1	0	yes, my eyes too
171	0:18:45	1	0	you won't be able to tell later on you know

Argumentation episodes

Table 6.16 presents an example of argumentative negotiation. The fragment is taken from a dialogue between two girls (Esther & Nicole) and takes place early in the session after both students have first read their sources. The whole episode takes 15 minutes and concerns the position the girls will take. Student 1 proposes

to write on anonymous organ donation (lines 1500 & 1505). Student 0 reminds her that anonymity reflects an opinion, goes on to ask her partner's opinion on the subject (line 1508) and gives her own opinion ('I'm ok with it'). Student 1 sets a restriction – not to rapists (line 1511) – which starts a discussion on pros and cons of anonymity in organ donation.

Student 0 defends her position on anonymity by giving a line of argumentation consisting of an example and reason in favor of anonymity (rapists and racists are human beings, line 1515) and a counterargument (difficult choice, line 1517). Student 0 accepts the last argument (line 1518) but indicates her restriction on the codicil (line 1519). Student 0 counters this argument by referring to the separate administrative systems ('won't be marked as criminal', line 1520) and with the general conclusion that it is better not to know these things (line 1521). Student 1 elaborates her position by stating that in that case she would not donate at all (line 1523). Student 0 gives a new argument in favor (line 1527). In her reaction, student 1 refers to 'marc dutrout' (a Belgian serial killer and child molester). Student 0 tries again (lines 1531-1533) and once more rhetorically (why is it that you so badly want to know, line 1538), but student 1 sticks to her principles (line 1540). A final joking argument (line 1545) does not convince her either. This results in an awkward situation: the partners cannot come to an agreement. Remarkable in this respect is the question student 1 poses at the end of this episode (line 1546), which is the same question the episode started with. Eventually, they ended up writing a well-composed paper defending good regulation of organ donation with several arguments and counterarguments for the position. However, the question of anonymity or the possibility of restriction in future receivers of organs is not raised in the argumentative text at all. Although it is clear from the dialogue that they can understand each other's opinions and are able to ground these in a common frame of reference, the impossibility to come to an agreement prevented them from coordinating activities on this topic.

Table 6.16: Chat episode with Argumentation (pair 413, female & female).

Line	Time	Sq	Actor	Dialogue
1500	0:12:20	1	1	I propose to write a paper about:
1505	0:14:09	1	1	that a donor is not allowed to choose whom he gives his organs to and that a receiver doesn't get to find out whose organ he gets
1506	0:15:02	1	0	in the intro it says that you should write an argumentative paper with your opinion about organ donation ~
1507	0:15:02	2	0	not knowing about the receiver and donor is a part of that
1508	0:15:39	1	0	anyway, what do you think about organ donation ~
1509	0:15:39	2	0	I'm ok with it
1510	0:16:12	1	1	I think it's fine too ~
1511	0:16:12	2	1	but I wouldn't want to donate my organs to a rapist or someone like that..
1512	0:18:08	1	0	well, that's why people don't know who they're donating to ~
1513	0:18:08	2	0	I saw in a tv series once that a black boy was going to donate his organs to a racist white guy. ~
1514	0:18:08	3	0	the mother found out but let the donation go through anyway ~
1515	0:18:08	4	0	because she didn't want to see that person as a racist but just as a human being who would die without the donation
1516	0:18:59	1	0	do you feel that people ought to know, then , ~
1517	0:18:59	2	0	I think that if you do that you only make the choice more difficult
1518	0:19:32	1	1	okay ~
1519	0:19:32	2	1	but I do think that you should be able to indicate that your organ isn't donated to a criminal
1520	0:20:36	1	0	on a waiting list a person like that won't be marked as a criminal but rather as ill that person is just plain ill and needs help. ~
1521	0:20:36	2	0	but it's a good thing that you don't know that sort of stuff
1522	0:21:49	1	1	no ~
1523	0:21:49	2	1	but for me this would be a reason not to donate my organs ~
1524	0:21:49	3	1	because they could end up in a criminal
1525	0:23:01	1	0	well, I think this is nonsense criminals are people too, and
1527	0:23:01	2	0	because they maybe made 1 (ok, maybe more at times) mistake their chance at life should be lessened
1528	0:23:22	1	1	oh so you would want to give your heart to marc dutrout
1530	0:24:08	1	1	and I think exactly that is totally wrong
1531	0:24:14	1	0	you don't know who'll get your heart, ~
1532	0:24:14	2	0	you register as donor and who will get your heart you will never find out.
1533	0:24:14	3	0	for sure not because you die after that
1536	0:24:25	1	1	see above
1538	0:24:50	1	0	why is it that you so badly want to know who will get your heart
1539	0:25:24	1	1	I don't want to know who ~
1540	0:25:24	2	1	but I don't want my organ to go to a criminal, that's my point
1541	0:25:32	1	0	well, I see a criminal like that, in that case, just like a regular human being
1543	0:25:53	1	1	well, I -don't- like a regular human being
1545	0:26:42	1	0	shanna says that you could only let him suffer longer that way (joke...)
1546	0:27:03	1	1	ok, what are we going to do this paper about

We will now go on to discuss the statistical analyses for the three main coordination processes Focusing, Checking and Argumentation. The relative frequencies of these specific coordination processes are derived by a number of indicative Dialogue acts (see section 2.5.5 in Chapter 2: Method). Table 6.17

shows the mean percentages and standard deviations for each of the processes for all conditions.

Table 6.17: Means and standard deviations of coordination processes.

	All		C		D		DA	
	M	**SD**	**M**	**SD**	**M**	**SD**	**M**	**SD**
Focusing	18.34	4.23	18.24	3.76	17.45	3.06	16.58	2.63
Checking	26.59	4.64	28.11	4.61	26.22	4.35	26.60	4.78
Argumentation	9.80	2.82	8.98	2.89	10.74	2.98	10.51	2.77
N	145		39		17		26	
	DO		**DOA**		**O**		**OA**	
	M	**SD**	**M**	**SD**	**M**	**SD**	**M**	**SD**
Focusing	18.44	3.53	20.71	6.64	19.96	6.08	19.04	3.19
Checking	25.52	3.79	28.25	5.63	25.45	4.73	24.15	3.33
Argumentation	9.72	2.39	9.03	2.89	10.70	2.51	9.04	2.71
N	23		11		18		11	

Over 25 % of the task related chat consists of Checking activities, 18 % is spent on Focusing, and 10 % on Argumentation. This means that some 55 % of the task related chat is devoted to coordinating the collaborative discussion. There were few significant differences between the conditions. The conditions with the Outline tool focus more frequently than the other groups. The difference is particularly clear when we look at the Diagram-Advisor condition, which focuses significantly less frequently than the Diagram-Outline-Advisor (4.13) and Outline (3.37, both Bonferroni) conditions.

Table 6.18 shows the correlations of the coordination processes with text quality. Focusing correlates positively with Textual structure, Overall argumentation and the Mean text score. Together with the positive tendencies of the other two text scores, this suggests that there is an overall positive relation between Focusing and text quality. Checking, on the other hand, does not show any significant correlations with text quality. Argumentation correlates positively with Overall argumentation, but its correlations do not show a clear positive or negative tendency. As expected, there is a positive link between argumentation processes in the dialogue and the argumentative structure of the text. On the whole, the correlations found are small.

Table 6.18: Correlations between coordination processes and text quality.

	Textual structure	**Segment argumentation**	**Overall argumentation**	**Audience focus**	**Mean text score**
Focusing	.14*	.09	.12*	.02	.12*
Checking	-.01	.02	.04	.03	.02
Argumentation	-.01	-.01	.13*	.05	.07

* $p < .05$.

To check whether particular conditions have an effect on the coordination processes, the model in Figure 6.18 was tested. The directions of the effect of condition hypothesized in Figure 6.18 are shown in Table 6.19. The Diagram condition has a positive effect on argumentation in the chat, and so do the Outline and Diagram-Advisor conditions. The Outline condition also influences focusing positively, but it has a negative effect on checking, as do the diagram-Outline and Outline-Advisor conditions. The Diagram-Advisor condition influences focusing negatively, whereas the Diagram-Outline-Advisor condition has a positive effect on focusing. The overall picture suggests that conditions with fewer tools have a positive effect on the use of argumentation, whereas the amount of checking is influenced negatively by the availability of the Outline tool.

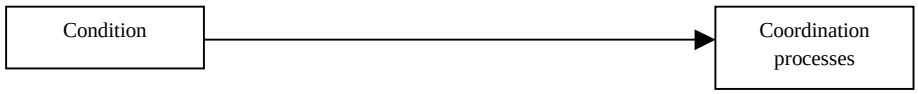

Figure 6.18: Effect of condition on coordination processes.

Table 6.19: Effects of conditions on coordination processes.

	Focusing	**Checking**	**Argumentation**
D			+
DA	-		+
DO		-	
DOA	+		
O	+	-	+
OA		-	

To check the effect of coordination processes on text quality, we tested the model shown in Figure 6.19. Table 6.20 shows the directions of effects for this model. Focusing has a positive effect on overall text quality, and on Textual structure and Overall argumentation in particular. All three coordination processes have positive effects on Overall argumentation. These results support our expectation that the specific processes of coordinating the communication of content facilitates the collaboration between students and thus influences the resulting argumentative text.

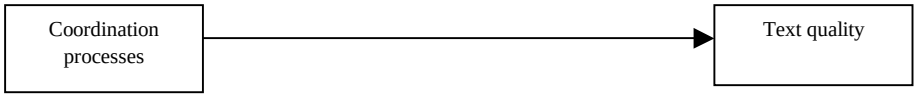

Figure 6.19: Effect of coordination processes on text quality.

Table 6.20: Effects of coordination processes on text quality.

	Textual structure	**Segment argumentation**	**Overall argumentation**	**Audience focus**	**Mean text score**
Focusing	+		+		+
Checking			+		
Argumentation			+		

To check whether the condition and the coordination processes also influence text quality independently of each other, we tested the model shown in Figure 6.20. The directions of the effects are shown in Table 6.21. In this model, focusing still has a positive effect on Overall argumentation and on the Mean text score, but not on Textual structure. As in the previous model, all three coordination process positively influence the Overall argumentation in the final text. There is very little effect of condition on text quality in this model. The Diagram and Diagram-Advisor conditions both influence the Overall argumentation positively, and the Diagram-Advisor condition has a negative influence on Textual structure and Segment argumentation. The Diagram-Outline-Advisor condition has a positive influence on Textual structure. None of the factors contribute to the Audience score.

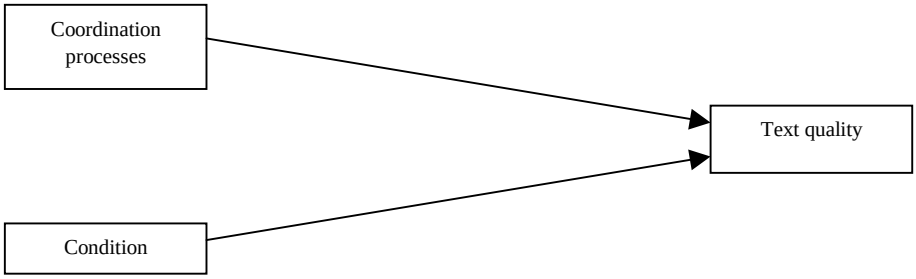

Figure 6.20: Effects of condition and coordination processes on text quality.

Table 6.21: Effects of condition and coordination processes on text quality.

	Textual structure	Segment argumentation	Overall argumentation	Audience focus	Mean text score
Focusing			+		+
Checking			+		
Argumentation			+		
D			+		
DA	-	-	+		
DO					
DOA	+				
O					
OA					

On the whole, none of the models discussed above need to be rejected on the basis of the multiple regression analyses. Combining them gives the model shown in Figure 6.21. Condition has an effect on Textual structure and Overall argumentation of the final text, both independent of and in combination with the coordination processes Focusing, Checking, and Argumentation. The Diagram-Outline-Advisor and the Diagram condition both have a positive effect on text quality through the coordination processes, whereas the Outline-Advisor condition has a negative effect. Independently of the coordination processes, the Diagram-Advisor condition has a negative effect on text quality. The combination of the Advisor with one other planning tool negatively influences the text, but combining the Advisor with both planning tools has a positive effect on text quality.

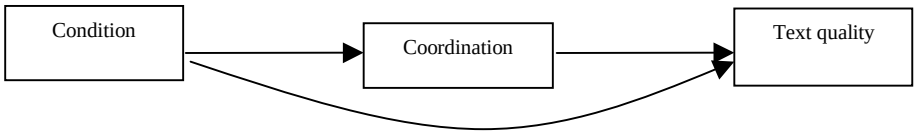

Figure 6.21: Effect of condition and coordination processes on text quality.

6.2 A closer look at argumentation

In the previous section, one of the specific coordination processes in the chat discussion was argumentation. In this section[4], we will take a closer look at the argumentation in the chat, widening our view to five different types of argumentative episodes, based on their content. The purpose is to determine further the relation between argumentation in the collaborative discussion and the

[4] The analyses for this section were done by Tobi Boas, Chris Phielix, Nicolette van der Meijden and Jan-Willem Schoonhoven as part of their second-year research class in Educational Science.

quality of the final written product. First, three of the argumentation types are illustrated.

An example of argumentation is shown in Table 6.22: a brief episode by two male students working on a text on organ donation after about three hours. From left to right the columns show: Line – protocol line number; Time from start; Sq – sequence number for split utterances; Actor – students number; Argumentation episode code; and Dialogue – the actual chat text. The fragment starts with one of the boys asking whether they already have a title for the text and after his partner's response asking whether they should use a nicer one (lines 509-512). When his partner asks for clarification, and after a joking answer, student 0 explains that he thinks the title should have a reference to the position they take (line 521). The argumentation episode stops there and they do not come to an agreement. Other topics are discussed in the chat; the fragment includes an argumentation episode about task division (lines 544-549), and the problem of the title re-enters the discussion initiated by student 1 after 5 minutes, starting a new argumentation episode. They agree on the fact that the current title ('organ transplantation') is not good enough and next they discuss the alternatives (adding 'donation' proposed by student 1 versus adding 'waiting lists' proposed by student 0). Student 1 gives the counterargument that the waiting lists are only a part of the problem (line 566). In the end, after a pause of almost one minute, student 0 accepts the title proposed by student 1 by typing 'all right' and the episode ends (line 568). Although the argumentation is not very elaborated or deep, the fragment shows that, for collaboration to proceed in a coordinated and mutually acceptable way, arguments need to be resolved. The issue of the title had not been solved yet and was bound to be reopened. Furthermore, the fragment shows that students do not argue for the sake of arguing. As soon as an acceptable agreement can be reached, it is taken. In general, argumentation episodes are rather short, 5 to 10 utterances on average.

Table 6.22: Argumentation episodes in chat protocol of two students (both male).

Line	Time	Sq.	Actor	Arg. Epi.	Dialogue
509	2:57:34	2	1	MetaBeg	by the way, do we have a title?
511	2:57:35	2	0		organ donation
512	2:57:47	1	1		that's what I meant shall we think of a nicer one
513	2:57:57	1	0		how very creative
514	2:58:06	1	0		what do you have in mind
515	2:58:32	1	1		mmmmmmm...let me think
516	2:58:48	1	1		how about....everything you ever wanted to know about organs!!! ~
518	2:59:53	1	0		we are not writing an essay on organs, ~
519	2:59:56	1	1		just kidding.
520	2:59:56	2	1		I know I meant it as a joke.
521	3:00:30	1	0		ha ha
522	3:00:30	2	0	Stop	but we give our opinion so that should be clear in the title
544	3:05:00	2	0	CoopBeg	if you give the text to me, then I will read through it first, ~
545	3:05:12	1	1		lets do it!!!!!!!
546	3:05:12	2	1		ok, ~
547	3:06:04	1	0		you can ~
548	3:06:04	2	0		read whilst I type and ~
549	3:06:04	3	0	Stop	if you have any comments send them to me
557	3:06:06	1	1	MetaBeg	are you going to do the title as well or do we have one already
558	3:06:25	6	0		just organ transplantation~
559	3:06:25	7	1		we should also something in the title with donation and transpl.
560	3:06:50	1	0		if we can't think of anything else in the mean time
562	3:07:20	1	0		or with waiting lists ~
563	3:07:27	1	0		because our conclusion is that it must be shortened, ~
564	3:07:27	2	0		or maybe with donor register
566	3:08:01	1	1		waiting list is only a part of it I think
567	3:08:23	1	1		what do you think of donating and transplanting organs
568	3:08:54	1	0	Stop	all right

Another example of the necessity of coming to agreement is illustrated in an episode of a chat protocol of two other students (male and female) in Table 6.23. It is an example of arguing about subject matter, in this case the position the students want to defend. Although the students are not obliged to write a text about a position that reflects their own opinion, students in general try to come to agreement about a common position to defend, but also about other content matter, such as general opinion, the interpretation of factual information and controversial issues. Although arguments about subject matter are not very frequent in the protocols they seem to be very important. The collaboration seems to develop more disorganized if the students do not reach a mutual commitment on matters of content. In almost all protocols the question "What position should we take?" is the first question asked after reading the information sources. Sometimes this problem is resolved pragmatically ("Let's be pro it seems easier"). In other cases,

as in the example in Table 6.23, a more or less elaborated argumentative debate develops early on in the protocol. In the fragment student 1 (the girl) opposes the idea of being an organ donor herself ("...how egotistically it may sound") (lines 122-123). When student 0 (male) gives his point of view (line 124) the girl, after joking, makes a reopening in stating that her ideas might change later (line 129), that is, in a few years time when she has become more social. After an interruption with a cooperation argument, and after rereading some sources on donor registration, student 1 reconsiders her position and changes her point of view ("so I'm in favor") (line 158).

Table 6.23: Subject matter argumentation in chat protocol of two students (male & female).

Line	Time	Sq.	Actor	Arg. Epi.	Dialogue
119	00:24:12	2	1	SubjBeg	if you haven't completed a donor registration form, ~
121	00:24:35	1	1		your family can decide for you.
122	00:24:35	2	1		I never want to lose my organs, ~
123	00:25:05	1	1		no matter how egotistically it may sound
124	00:24:59	1	0		I always think, you don't notice a thing when you're dead.
127	00:25:41	1	0		so as far as I'm concerned they can have my organs.
125	00:25:05	2	1		you're too good for this world
128	00:25:41	2	0		maybe I am
129	00:25:51	1	1		maybe, I will think differently in a few years time
130	00:25:51	2	1		i would be more social
136	00:26:38	2	0		that is not important when you're dead, ~
138	00:27:02	1	0		whether you are social or not.
137	00:26:53	1	1		nice statement
139	00:27:02	2	0	Stop	I know
142	00:27:29	1	1	CoopBeg	I am reading more sources, ~
143	00:27:40	1	1		what are you doing at the moment
144	00:27:40	2	1		kick ass
145	00:27:43	1	0	Stop	I'm reading the sources as well
146	00:27:59	1	1	SubjBeg	in source 1 it says that you need to register, ~
147	00:28:24	1	0		and I wonder whether the text should perhaps be on whether or not to register donors.
149	00:28:47	1	1		you read the first bit
152	00:28:50	2	0		already done so, ~
155	00:29:04	2	1		you're right. ~
156	00:29:19	1	0		that's how they are trying to prevent losing possible donors.
158	00:29:28	1	1	Stop	so I'm in favor

The protocols show that arguing is a crucial process for managing coordination, and that it is necessary in order to come to agreement. Arguments are found in the protocols about every aspect of the task and about many non-task matters as well. So, an interesting question might be whether arguments occur the most on the

topics that we think matter. What topics do the students argue on? For this last question we have collapsed the dialogues into five major topics of argumentation: subject matter, meta-cognitive, cooperation, technical and other/social.

We expected to find that higher percentages of argumentation on Content, Coordination, or Metacognitive strategy would relate positively to text quality, in particular to the argumentative quality of the text in the form of the Segment argumentation and Overall argumentation scores. On the other hand, we expected to find that argumentation on Technical aspects and Miscellaneous topics would relate negatively to text quality, as these are not related to the topic or execution of the writing task itself.

Table 6.24 shows the means, standard deviations and extremes for each of the five types of argumentation episode. It is immediately obvious that argumentation on Metacognitive strategies is the most frequent with more than 50% of the argumentation episodes. This means that the participants spent a lot of their argumentation deciding on *how* to write their text. Coordination is the next most frequent category with about 20%, and it is followed relatively closely by Miscellaneous topics (14%). The Content and Technical aspect are discussed least of all, with only 8% and 7%, respectively. This means that the participants had very little discussion on the content of their text. Of course, the students cannot go on quarreling about the content throughout the task, as this would seriously inhibit the productive writing process.

Table 6.24: Descriptive statistics for types of argumentation episodes.

Topic of the argumentation	Mean	SD	Minimum	Maximum
Content	8.40	5.84	.00	23.10
Coordination	19.30	13.59	.00	43.80
Metacognitive strategy	50.90	13.12	26.30	71.4
Technical aspects	6.50	3.95	.00	13.60
Miscellaneous topic	13.80	13.04	.00	52.60

Means are % of the total number of argumentation episodes.

Table 6.25 shows the correlations between the argument episode types and the text quality measures. The table shows that the argumentation for Coordination is not significantly related to text quality, although there is a clear positive tendency. The non task-related argumentation (Miscellaneous), however, does have a clear relation with text quality: it correlates negatively with Segment argumentation, Overall argumentation, and the Mean text score, and shows an overall negative tendency. Argumentation on Metacognitive strategies are also quite clearly related

to text quality, but in a positive direction. The same goes for discussion on the content and main position of the shared text. Like the Miscellaneous argumentation, argumentation on computer related issues is negatively related with text quality. In short, the results come up to all our expectations.

Table 6.25: Pearson correlations between argument episode types and text quality.

	Textual structure	**Segment argumentation**	**Overall argumentation**	**Audience focus**	**Mean text score**
Content	.21	.16	.16	.37*	.27
Coordination	.12	.08	.28	.08	.19
Metacognitive strategy	.11	.47**	.12	-.02	.18
Technical aspects	-.24	-.34*	-.21	-.04	-.23
Miscellaneous topic	-.14	-.52**	-.36*	-.33	-.42*

** $p < .01$; * $p < .05$.

6.3 Conclusion

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 stimulate a more structured dialogue. The same can be observed when comparing high scoring and low scoring dyads. Several regression models were analyzed in this chapter, none of which had to be rejected. This leads us to conclude that the experimental condition has a direct effect on text quality, but also through the communicative function, and through asymmetry of contribution by the partners, and through coordination processes checking, focusing and argumentation. In addition, we 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.

CHAPTER 7 STUDENT EVALUATIONS

In this Chapter we will give a summary of the outcomes of the evaluation questionnaires. The questions that were suitable for statistic analyses were divided into three main topics: questions about the writing assignment, questions about the computer program, and questions about working collaboratively. All questions were rated by the students on a 3-point scale. The descriptive statistics for the first set of questions are shown in Table 7.1.

Table 7.1: Evaluation of the assignment.

	Total			**C**			**D**			**DA**		
	M	**SD**	**N**	**M**	**SD**	**N**	**M**	**SD**	**N**	**M**	**SD**	**N**
Difficulty of writing assignment	2.06	.59	266	1.83	.62	77	2.25	.44	32	2.22	.53	40
Difficulty of sources	1.75	.63	256	1.65	.63	74	1.91	.39	32	1.81	.70	37
Computer supported collaboration	2.16	.87	267	2.50	.73	78	1.56	.84	32	1.75	.84	40
	DO			**DOA**			**O**			**OA**		
	M	**SD**	**N**	**M**	**SD**	**N**	**M**	**SD**	**N**	**M**	**SD**	**N**
Difficulty of writing assignment	2.11	.54	44	2.14	.56	22	2.17	.71	35	1.81	.40	16
Difficulty of sources	1.76	.62	41	1.50	.67	22	1.79	.64	34	2.06	.68	16
Computer supported collaboration	2.48	.76	44	1.82	.73	22	2.29	.86	35	2.00	.89	16

On average, the writing assignment itself was rated as relatively difficult, although the Control group found it more easy, and even showed significant differences with some of the experimental conditions (see Appendix 12). The sources were rated as relatively easy by all groups except the Outline-Advisor condition, but there were no significant differences for this question. In general, the participants were quite positive about the collaborative and computer supported setup of the assignment. However, there are quite a few significant differences between the conditions. The Control group and the Diagram-Outline condition were the most positive, while the Diagram group, and to some extent the Diagram-Advisor condition, were the least positive about the type of task (see Appendix 12).

Table 7.2 shows the means, standard deviations and numbers of respondents for the questions concerned with the computer program. On the whole, the participants were quite positive about the basic TC3 environment. When we look at the differences between conditions, what strikes us is the negative tendency for the Diagram condition for the program features, most obviously for logging on and off (see Appendix 12). This is not surprising, as both classes in the Diagram group encountered a lot of technical difficulties with their school servers. The Control group found the buttons of the basic program not very clear, in contrast with the

experimental conditions (see Appendix 12). Although most participants claimed they understood the function of the Diagram and/or Outline tool, the planning tools were not viewed very positively by any of the experimental groups.

Table 7.2: Evaluation of the computer program.

	Total			**C**			**D**			**DA**		
	M	**N**	**SD**	**M**	**N**	**SD**	**M**	**N**	**SD**	**M**	**N**	**SD**
Logging on/off	2.50	267	.74	2.59	78	.73	1.53	32	.72	2.48	40	.68
Private notes window	2.47	260	.67	2.52	75	.68	2.34	32	.60	2.40	40	.74
Information window	2.52	264	.65	2.65	78	.58	2.10	31	.65	2.48	40	.75
Chat window	2.77	266	.49	2.86	78	.39	2.44	32	.56	2.72	40	.55
Shared text window	2.37	266	.71	2.36	78	.74	2.06	32	.80	2.35	40	.80
Traffic light	2.22	266	.77	2.23	78	.68	1.71	31	.74	2.15	40	.86
Clarity of buttons	2.56	266	.83	1.77	78	.98	2.87	32	.49	2.75	40	.67
Use of buttons	2.59	265	.81	1.83	77	.99	3.00	32	.00	2.80	40	.61
Diagram window	1.67	137	.76				1.32	31	.60	1.73	40	.82
Oral instruction for Diagram	1.98	61	.87							1.85	39	.87
Advisor window Diagram	1.68	50	.82							1.73	37	.80
Clarity of Diagram function	2.39	61	.71							2.35	40	.74
Outline window	1.91	112	.77									
Oral instruction for Outline	1.71	38	.84									
Advisor window Outline	1.60	25	.82									
Clarity of Outline function	2.47	36	.65									
	DO			**DOA**			**O**			**OA**		
	M	**N**	**SD**	**M**	**N**	**SD**	**M**	**N**	**SD**	**M**	**N**	**SD**
Logging on/off	2.52	44	.66	2.68	22	.65	2.89	35	.32	2.94	16	.25
Private notes window	2.50	44	.63	2.65	20	.59	2.41	34	.70	2.47	15	.64
Information window	2.68	44	.52	2.62	21	.67	2.35	34	.69	2.56	16	.51
Chat window	2.93	43	.26	2.68	22	.65	2.86	35	.36	2.69	16	.70
Shared text window	2.44	43	.59	2.68	22	.48	2.37	35	.65	2.44	16	.63
Traffic light	2.34	44	.71	2.50	22	.74	2.37	35	.84	2.31	16	.70
Clarity of buttons	2.86	43	.52	2.91	22	.43	3.00	35	.00	3.00	16	.00
Use of buttons	2.86	43	.52	2.91	22	.43	2.94	35	.34	3.00	16	.00
Diagram window	1.80	44	.73	1.82	22	.80						
Oral instruction for Diagram				2.23	22	.81						
Advisor window Diagram				1.54	13	.88						
Clarity of Diagram function				2.48	21	.68						
Outline window	1.80	40	.79	2.18	22	.80	1.71	34	.68	2.25	16	.68
Oral instruction for Outline				1.64	22	.90				1.81	16	.75
Advisor window Outline				1.58	12	.79				1.62	13	.87
Clarity of Outline function				2.48	21	.68				2.47	15	.64

Table 7.3 shows the descriptive statistics for the set of questions on working collaboratively. In general, the collaborative work on the planning tools was viewed as quite positive by all experimental groups. Even though the Diagram condition was quite negative about the program, the participants were still positive about the collaboration with their partner, though not as positive as the other groups (see Appendix 12). The Diagram group was also less enthusiastic about

turn taking in the Diagram window. At the same time, the participants in the Control group were more positive about turn taking in writing than their colleagues in the experimental groups.

Table 7.3: Evaluation of working collaboratively.

	Total			**C**			**D**			**DA**		
	M	**N**	**SD**	**M**	**N**	**SD**	**M**	**N**	**SD**	**M**	**N**	**SD**
Collaboration	2.70	266	.52	2.82	78	.42	2.44	32	.67	2.59	39	.64
Turn-taking for writing	2.38	266	.69	2.63	78	.58	2.06	32	.67	2.23	39	.84
Equality of contribution to text	2.67	226	.65	2.79	38	.58	2.56	32	.76	2.62	39	.63
Turn-taking for Diagram	1.99	136	.81				1.41	32	.61	2.23	39	.81
Equality of contribution to Diagram	2.27	135	.86				2.20	30	.85	2.35	40	.86
Turn-taking for Outline	2.15	110	.76									
Equality of contribution to Outline	2.32	111	.81									
Deliberation in chat	2.78	267	.47	2.83	78	.41	2.66	32	.55	2.65	40	.58
	DO			**DOA**			**O**			**OA**		
	M	**N**	**SD**	**M**	**N**	**SD**	**M**	**N**	**SD**	**M**	**N**	**SD**
Collaboration	2.70	44	.46	2.55	22	.60	2.74	35	.44	2.94	16	.25
Turn-taking for writing	2.25	44	.61	2.41	22	.73	2.46	35	.66	2.31	16	.70
Equality of contribution to text	2.59	44	.76	2.86	22	.35	2.80	35	.53	2.44	16	.73
Turn-taking for Diagram	2.14	44	.73	2.10	21	.83						
Equality of contribution to Diagram	2.35	43	.87	2.09	22	.87						
Turn-taking for Outline	2.05	40	.78	2.45	22	.74	2.06	32	.76	2.19	16	.66
Equality of contribution to Outline	2.35	40	.86	2.14	22	.83	2.58	33	.66	2.00	16	.82
Deliberation in chat	2.80	44	.41	2.73	22	.63	2.83	35	.38	2.94	16	.25

In general, the students were quite positive about several aspects of TC3, the writing task, and the process of collaborative learning.

CHAPTER 8 DISCUSSION

The experiments in the COSAR project were conducted in the natural environment of the classroom. The ecological validity will thus be relatively high. As it was not possible technically to assign participants to conditions randomly within schools, entire classes were assigned to single conditions. To prevent group effects each condition – except the Diagram-Outline-Advisor condition – was assigned to at least two different classes from different schools. The effects found for the Diagram-Outline-Advisor condition might thus be attributable to the school rather than to the condition.

The final sample consisted of 145 dyads, with numbers per condition varying from 11 to 39. The large number of participants increased the reliability of the research. In addition, we tried to keep a close control of intervention: the participants were not allowed to use information sources other than those given in TC3, and during sessions the students could only communicate through the program. They were not allowed to take the assignment home. In spite of all this, the experimental situation could not be controlled completely. For instance, partners could communicate, or read extra information in between sessions. In addition, some participants had more prior knowledge of the topics than others. For example, in one of the Control groups a classmate had had a heart transplant. The input of the teachers was also difficult to control, even though they were asked not to interfere and not to answer content-related questions. In some schools the teachers were hardly ever present, whereas in other groups they were very eager to help their students.

The effects of the Diagram-Advisor condition on text quality are very different from the effects of the other three Diagram conditions. Comparing the text quality scores of the three groups within the Diagram-Advisor condition showed that two groups had significantly different scores on audience focus and segment argumentation. However, in a comparison of all groups these differences were not found again. In one group in the Diagram-Advisor condition the teachers had redefined our assignment. Instead of writing an argumentative text working from a position to a conclusion (‘betoog’), the participants were told to write a discussion, working from a question towards a position (‘beschouwing’). Initially, the participants had difficulty fitting their ideas into the constraints of the diagram, as there was no ‘question’ box, and only a ‘position’ box. This might explain the differences found within the Diagram-Advisor condition.

All groups spent approximately the same amount of time on the assignments. However, due to scheduling differences the number of sessions used to complete the assignment ranged from two to six. Obviously, the different sessions were spread out over several days up to three weeks. We expected that more sessions further apart would lead to lower text quality, as the participants in these groups had to trace where they had left off each time they got back to work. However, comparison of the groups did not show clear differences that might be attributable to session effects.

The validity of the measure of asymmetry of contribution in the Dialogue acts might be questioned: does it really measure equality of contribution? The measure uses the difference in percentages of chat messages between partners. The measure cannot take into account the possibility that a participant repeats the same arguments, echoes the partner's utterances, uses false arguments, and so on. Each of these chat messages is counted as such in the measure of asymmetry. Also, some people are good at formulating in brief sentences, whereas others are more long-winded. The measure cannot take into account the exact content of the utterances.

We worked form the assumption that certain Dialogue acts are indicators of the coordination processes focusing, checking and argumentation. However, these processes do not exist in separate utterances, but rather in episodes. A better way to examine coordination processes would therefore be to use the Task act episodes and link these to aspects of the coordination process. For example, an episode beginning with a verifying question would then be a checking episode. Unfortunately, this was not an option, as the Task act coding had to be scheduled for a later time during the research project.

In coding the chat protocols, multifunctionality of utterances was not taken into account, as each utterance received a single coding. This was partly compensated by splitting sentences into messages with different communicative functions. The result is that only very short utterances might still have multiple functions in the Dialogue act or Task act coding systems. For example, a proposal can be a question at the same time: "Shall we continue working on the Outline?". To remove coding problems, a hierarchical order of communicative functions and task acts was determined on the basis of the value of the information contained in the utterances.

An automatic coding filter was used for coding the communicative functions in the dialogue. The coding system was based on the VOS system (Erkens, 1997), and went through several stages of development by different researchers under the supervision of the project leader. The system was tried and tested, and changes were made on the basis of these tests. The result was a filter that coded approximately 80 % of the utterances on the basis of discourse markers, and filled the rest as informative statements. The 20 % that was not coded by the filter was checked manually, thus reducing reliability, but increasing validity. An educational scientist and a linguist (both closely involved in the project) agreed with more than 98 % of the filter coding independently of each other.

Unfortunately, technical difficulties with the TC3 program occurred in most schools. Some problems could be solved fairly easily and hardly hindered the writing process, but in some cases the problems might have had a negative effect on the participants in completing the assignment properly. However, the final results for text quality do not indicate this.

CHAPTER 9 CONCLUSION

The objective of the research project 'Computer Support for Collaborative and Argumentative Writing' (the COSAR project) was to study the relation between the collaborative process and support of the planning process in argumentative writing. Subject of our investigations were students in the 'studiehuis' – a recent innovation in the Dutch secondary school curriculum. Groupware was developed – called TC3: Text Composer, Computer supported & Collaborative – that allows collaborative writing by pairs of students, with or without support by specially designed planning tools for organization and linearization (the Diagram and the Outline). The three main research questions were concerned with support of organization and linearization through the Diagram and Outline, constructive planning and coordination processes, and differences in constructive activities in different phases of the writing process.

9.1 Answering the research questions

We found that constructive activities are indeed different in different phases of the collaborative writing process, both in terms of the use of the software and in terms of the task discussion. Although planning activities occur throughout the process, discussing knowledge and actually writing the text seem to be the two most influential factors for text quality.

It seems that the participants in the Diagram-Advisor condition used the Diagram as a full report of their argumentation, which means they do not let the Diagram guide them in developing an argumentative structure. The categories present in the Diagram are no guideline to them, nor do the participants discuss the contents of the diagram: they merely describe the contents of the ideas generated in their discussion, or even of their text. Thus, the Diagram only 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. If a diagram is used to report on the contents of the text, it can still have a structuring function during the revision of the text. On the whole, however, we found little or no evidence for a positive effect of the Diagram condition on coordination or on text quality. 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. The Outline tool, on the other hand, 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 general idea that preplanning is rare in novice writers was not confirmed by our analysis of three different phases of the writing process: planning activities are found throughout the collaboration of our participants, and planning is even the most frequent activity in the first phase. Support of online planning through the Outline can be helpful, as noted above, even during the final stage of writing. On the whole, most activities that we would logically expect to find in a certain phase do indeed occur in those places most frequently, and if found in unexpected places often have a negative effect on the final product, for example checking the program manual and reading the sources.

In general, the data confirm our ideas that coordination is necessary on all aspects of the task, both in activities and in the dialogue, and that the collaboration needs to be adapted to the phase of the writing process.

9.2 Implications and future developments

The COSAR project has produced a number of articles, conference presentations and book chapters. A list of all these written products can be found in the Bibliography. In addition, the program TC3 and its planning tools were developed, and we have had several requests from teachers to use TC3 in their own classroom as a teaching tool. Unfortunately, the administrator's interface is not WYSIWYG, and requires some instruction. Nevertheless, one school so far was provided with the full TC3 software and has successfully implemented and used the program. In addition to the collaborative software, the COSAR project has also resulted in a new, expanded version of MEPA, the protocol analysis software we used to analyze the chat and activity protocols.

At our own department, adapted versions of TC3 are currently being used in the 'Twins' PhD project (Munneke), in the PhD project 'Computer-supported History and Argumentative Text writing' (CHAT; Van Drie, Van Boxtel & Van der Linden, in press), in the European SCALE project ('Internet-based intelligent tool to Support Collaborative Argumentation-based LEarning in secondary schools', Van Amelsvoort), and in the EC-cole PhD project (Van der Puil).

The PRO-ICT project (NWO project number 411 211 11), in some respects a follow-up of the COSAR project, is in full progress at the time of writing. The PRO-ICT project deals with cognitive and metacognitive support of collaborative learning, and uses a new, more flexible program based on TC3. It is a virtual research environment that allows for both synchronous and asynchronous communication, is suitable for larger project groups of students working together on a complex research task, and has monitoring facilities for the teacher. In addition, the text editor supports graphical representations, the program contains a forum facility for discussion between different project groups, and the different tools are presented in separate windows so the screen layout can be determined fully by the user.

REFERENCES

Alamargot, D. (1997). *Processus de récupération et d'organisation dans l'activité de rédaction de texte: effet de l'acquisition des connaissances référentielles.* Unpublished PhD thesis, University of Poitiers, Poitiers.

Alamargot, D., & Chanquoy, L. (2001). *Through the models of writing, with commentaries by Ronald T. Kellogg & John R. Hayes* (Vol. 9). Dordrecht: Kluwer Academic Publishers.

Andriessen, J. E. B., Coirier, P., Roos, L., Passerault, J.-M., & Bert-Erboul, A. (1996). Thematic and structural planning in constrained argumentative text production. In H. Van den Bergh & G. Rijlaarsdam & M. Couzijn (Eds.), *Theories, Models and Methodology in writing research* (pp. 237-251). Amsterdam: Amsterdam University Press.

Andriessen, J.E.B., Erkens, G., Overeem, E., & Jaspers, J. G. M. (1996, September). *Using complex information in argumentation for collaborative text production.* Paper presented at the Conference on Using Complex Information Systems (UCIS ’96), University of Poitiers, France.

Antaki, C. (1994). Explaining and Arguing: The social organization of accounts. London: Sage Publications.

Baker, M. (1999). Argumentation and constructive interaction. In J. E. B. Andriessen & P. Coirier (Eds.), *Foundations of Argumentative Text Processing* (pp. 179-203). Amsterdam: Amsterdam University Press.

Baltzer, J. E. (1989). *Hardop denken bij het schrijven van een opstel: Een onderzoek naar de effecten van twee steldidactieken op schrijfprocessen van 15-jarigen, met behulp van een analyse-schema voor hardop-denkprotocollen*. Amsterdam: Stichting Centrum voor Onderwijsonderzoek.

Bannon, L. J. (1995). Issues in computer supported collaborative learning. In C. O'Malley (Ed.), *Computer Supported Collaborative Learning* (Vol. 128, pp. 267-283). Berlin, Heidelberg: Springer.

Bereiter, C., & Scardamalia, M. (1987). *The psychology of written composition*. Hillsdale, NJ: Lawrence Erlbaum.

Bolhuis, S. M., & Kluvers, C. (1996). *Op weg naar zelfstandig lerende leerlingen* (Research report). Nijmegen, The Netherlands: Catholic University of Nijmegen.

Breetvelt, I. (1991). *Schrijfproces en tekstkwaliteit: Een onderzoek naar het verband tussen schrijfproces en tekstkwaliteit bij leerlingen in het voortgezet onderwijs*. Amsterdam: Stichting Centrum voor Onderwijsonderzoek.

Brown, J., Collins, A., & Duguid, P. (1989). Situated cognition and the culture of learning, *Educational Researcher*, *18*, 32 – 41.

Chanquoy, L. (1996, October 23-25). *Connectives and argumentative text: a developmental study*. Paper presented at the First International Workshop on Argumentative Text Processing, Barcelona, Spain.

Clark, H. H., & Brennan, S. E. (1991). Grounding in communication. In L. B. Resnick & J. M. Levine & S. D. Teasley (Eds.), *Perspectives on socially shared cognition* (pp. 127-150). Washington: American Psychological Association.

Clark, H. H., & Schaeffer, E. F. (1987). Collaborating on contributions to conversation. *Language and Cognitive Processes, 2*(1), 19-41.

Cohen, E.G. (1994). Restructuring the Classroom: Conditions for Productive Small Groups. *Review of Educational Research, 64*, 1, 1-35.

Coirier, P., Andriessen, J. E. B., & Chanquoy, L. (1999). From planning to translating: The specificity of argumentative writing. In J. E. B. Andriessen & P. Coirier (Eds.), *Foundations of argumentative text processing* (pp. 1-29). Amsterdam: Amsterdam University Press.

Coirier, P., Favart, M., & Broggio, E. (submitted). *From conceptual organization to linguistic formulation: developmental aspects*. Poitiers: CNRS-University of Poitiers.

Di Eugenio, B., Jordan, P. W., Thomason, R. H., & Moore, J. D. (2000). The agreement process: an empirical investigation of human-human computer mediated collaborative dialogs. *International Journal of Human-Computer Studies, 53*, 1071-1075.

Dillenbourg, P. (1999). Introduction: what do you mean by "Collaborative Learning"? In P. Dillenbourg (Ed.), *Collaborative Learning: Cognitive and Computational Aspects* (pp. 1-19). New York: Pergamon.

Doise, W., & Mugny, G. (1984). *The social development of the intellect*. New York: Pergamon.

Erkens, G. (in preparation). Dynamics of coordination in collaboration. In J. L. Van der Linden & P. Renshaw (Eds.), *Dialogic learning: Shifting perspectives to learning, instruction and teaching*. Dordrecht: Kluwer Academic Publishers.

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: Modeling cooperative dialogues for the design of intelligent educational systems]. Utrecht, The Netherlands: Utrecht University.

Erkens, G., Andriessen, J. E. B., & Peters, N. (2002). Interaction and performance in computer supported collaborative tasks. In H. Van Oostendorp (Ed.), *Cognition in a Digital World*. Hillsdale, NJ: Lawrence Erlbaum.

Galbraith, D. (1999). Writing as a knowledge-constituting process. In M. Torrance & D. Galbraith (Eds.), *Knowing what to write: conceptual processes in text production* (pp. 139-159). Amsterdam: Amsterdam University Press.

Gere, A., & Stevens, R. S. (1989). The language of writing groups: How oral response shapes revision. In S. W. Freedman (Ed.), *The acquisition of written language: Response and revision* (pp. 85-105). Norwood, NJ: Ablex.

Giroud, A. (1999). Studying argumentative text processing through collaborative writing. In J. E. B. Andriessen & P. Coirier (Eds.), *Foundations of argumentative text processing* (pp. 149-178). Amsterdam: Amsterdam University Press.

Hayes, J. R., & Flower, L. S. (1980). Identifying the organization of writing processes. In L. W. Gregg & E. R. Steinberg (Eds.), *Cognitive processes in writing: An interdisciplinary approach* (pp. 3-30). Hillsdale, NJ: Lawrence Erlbaum.

Hayes, J. R., & Nash, J. G. (1996). On the nature of planning in writing. In C. M. Levy & S. Ransdell (Eds.), *The Science of Writing* (pp. 29-55). Mahwah, NJ: Lawrence Erlbaum.

Henri, F. (1995). Distance learning and computer mediated communication: Interactive, quasi-interactive or monologue? In C. O'Malley (Ed.), *Computer supported collaborative learning* (Vol. 128, pp. 145-165). Berlin: Springer.

Katz, S. (1995, October). *Identifying the support needed in computer supported collaborative learning systems.* Paper presented at the CSCL 95, First International Conference on Computer Support for Collaborative Learning, Bloomington, Indiana.

Kozma, R. B. (1991). The impact of computer-based tools and embedded prompts on writing processes and products of novice and advanced college writers. *Cognition and Instruction, 8*(1), 1-27.

McCutchen, D. (1987). Children's discourse skill: form and modality requirements of schooled writing. *Discourse Processes, 10*(3), 267-286.

Oostdam, R. J. (1991). *Argumentatie in de peiling: Een aanbod- en prestatiepeiling van argumentatievaardigheden in het voortgezet onderwijs.* Unpublished PhD thesis, University of Amsterdam, Amsterdam.

Rijlaarsdam, G., & Van den Bergh, H. (1996). The dynamics of composing - An agenda for research into an interactive compensatory model of writing: Many questions, some answers. In C. M. Levy & S. Ransdell (Eds.), *The Science of Writing* (pp. 107-125). Mahwah, NJ: Lawrence Erlbaum.

Roschelle, J., & Teasley, S. D. (1995). Construction of shared knowledge in collaborative problem solving. In C. O'Malley (Ed.), *Computer-supported collaborative learning* (pp. 69-97). New York: Springer.

Salomon, G. (1993). On the nature of pedagogic computer tools: The case of the writing partner. In S. P. Lajoie & S. J. Derry (Eds.), *Computers as Cognitive Tools* (pp. 289-317). Hillsdale, NJ: Lawrence Erlbaum.

Scardamalia, M., & Bereiter, C. (1985). The development of dialectical processes in composition. In D. Olson & N. Torrance & A. Hildyard (Eds.), *Literacy, language and learning: the nature and consequences of reading and writing* (pp. 307-329). New York: Cambridge University Press.

Scardamalia, M., Bereiter, C., & Lamon, M. (1994). The CSILE project: trying to bring the classroom into world 3. In K. McGilly (Ed.), *Classroom lessons: Integrating cognitive theory and classroom practice* (pp. 201-229). Cambridge, MA: MIT Press.

Schellens, P. J., & Verhoeven, G. (1994). *Argument en tegenargument: Analyse en beoordeling van betogende teksten* [Argument and counterargument: Analysis and assessment of argumentative texts]. Groningen, The Netherlands: Martinus Nijhoff.

Schiffrin, D. (1987). *Discourse markers*. Cambridge: Cambridge University Press.

Schriver, K. A. (1988). *Teaching writers how to plan: which planning heuristics work best?* Paper presented at the Meeting of the American Educational Research Association, St. Louis, MO, USA.

Stein, N. L., Bernas, R. S., & Calicchia, D. (1997). Conflict talk: Understanding and resolving arguments. In T. Giron (Ed.), *Conversation: cognitive, communicative and social perspectives* (Vol. 34). Amsterdam: John Benjamins.

Suthers, D., & Hundhausen, D. (2001). *Learning by constructing collaborative representations: An empirical comparison of three alternatives.* Paper presented at the FIrst European Conference on Computer Supported Collaborative Learning (CSCL), Maastricht, The Netherlands

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.

Teasley, S. D., & Roschelle, J. (1993). Constructing a joint problem space: The computer as a tool for sharing knowledge. In S. P. Lajoie & S. J. Derry (Eds.), *Computers as cognitive tools* (pp. 229-257). Hillsdale, NJ: Lawrence Erlbaum.

The Cognition and Technology Group at Vanderbilt. (1994). From visual word problems to learning communities: Changing conceptions of cognitive research. In K. McGilly (Ed.), *Classroom lessons: Integrating cognitive theory and classroom practice* (pp. 157-201). Cambridge, MA: MIT Press.

Torrance, M., Thomas, G. V., & Robinson, E. J. (1996). Finding something to write about: strategic and automatic processes in idea generation. In C. M. Levy & S. Ransdell (Eds.), *The Science of Writing* (pp. 189-205). Mahwah, NJ: Lawrence Erlbaum.

Van Boxtel, C. (2000). *Collaborative Concept Learning.* Unpublished PhD thesis, Utrecht University, Utrecht, The Netherlands.

Van Cuilenburg, J. J., Kleinnijenhuis, J., & De Ridder, J. A. (1988). *Tekst en betoog: Naar een computergestuurde inhoudsanalyse van betogende teksten* [Text and argumentation: Towards computer supported content analysis of argumentative texts]. Muiderberg, The Netherlands: Coutinho.

Van der Linden, J. L., Erkens, G., Schmidt, H., & Renshaw, P. (2000). Collaborative learning. In P. R. J. Simons & J. L. Van der Linden & T. Duffy (Eds.), *New learning* (pp. 33-48). Dordrecht, The Netherlands: Kluwer Academic.

Van der Pool, E. (1995). *Writing as a conceptual process: a text-analytical study of developmental aspects.* Unpublished PhD thesis, Catholic University Brabant, Tilburg, The Netherlands.

Van Drie, Van Drie, J., van Boxtel, C., & Van der Linden, J.L. (in press). Learning history in an interactive learning environment. In A.M. O'Donnell, J.L. Van der Linden & C.E. Hmelo (Eds): *Collaborative learning, reasoning and technology*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Veerman, A. L., & Andriessen, J. E. B. (1997, September). *Argumentative text production in electronic environments.* Paper presented at the Writing the Future Conference, Brighton, UK.

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

Wampold, B. E. (1992). The Intensive Examination of Social Interaction. In T. R. K. J. R. Levin (Ed.), *Single-Case Research Design and Analysis: New directions for Psychology and Education* (pp. 93-133). Hillsdale, NJ: Lawrence Erlbaum Associates

Wesdorp, H. (1983). *Schrijven in het voortgezet onderwijs [Writing in secondary education]*. Harlingen, The Netherlands: Flevodruk.

BIBLIOGRAPHY

Articles

Erkens, G. (in preparation). Dynamics of coordination in collaboration. In J. L. Van der Linden & P. Renshaw (Eds.), *Dialogic Learning: Shifting Perspectives to Learning, Instruction and Teaching*. Dordrecht: Kluwer Academic Publishers.

Erkens, G., Jaspers, J.G.M., & (Tabachneck-)Schijf, H. (2000). Computerondersteuning bij zelfstandig en coöperatief leren. In K. Stokking, G. Erkens, B. Versloot & L. van Wessum (Eds.), *Van Onderwijs naar Leren* (pp. 129 - 142). Amersfoort: Garant.

Erkens, G., Kanselaar, G., Jaspers, J. G. M., & Schijf, H. J. M. (2001). Computer ondersteund samenwerkend leren. In A. Wald & J. L. Van der Linden (Eds.), *Leren in Perspectief* (pp. 85-97). Leuven, Apeldoorn: Garant.

Erkens, G., Prangsma, M. E., & Jaspers, J. G. M. (2002, in preparation). Planning and coordinating activities in collaborative learning. In A. O'Donnell, J. L. Van der Linden & C. E. Hmelo (Eds.), *Collaborative Learning, Reasoning, and Technology*. Hillsdale, NJ: Lawrence Erlbaum.

Erkens, G., Prangsma, M. E., Jaspers, J. G. M., & Schijf, H. J. M. (2002). Argumentation in Collaborative Writing: Relationships between Task Discussion and the Written Argumentative Product. Paper submitted for publication.

Erkens, G., Theil, J. W., Kanselaar, G., Prangsma, M. E., & Jaspers, J. G. M. (2002). " W8 ff :-) ": 'chattalk' in een coöperatieve leeromgeving [“W8 :-) ”: ‘chattalk’ in a collaborative learning environment]. Paper submitted for publication.

Kanselaar, G., Erkens, G., Jaspers, J. G. M., & Schijf, H. J. M. (2001). Computer supported collaborative learning (essay review). *Teaching and Teacher Education*, 17, 123 - 129.

Software

Jaspers, J. G. M., & Erkens, G. (2000). TC3. Text Composer, Computer supported & Collaborative (Version 1.4) [Computer software]. Utrecht, The Netherlands: Utrecht University.

Erkens, G. (2001, October). MEPA. Multiple Episode Protocol Analysis (Version 4.7) [Computer software]. Utrecht, The Netherlands: Utrecht University.

Paper presentations

Erkens, G., Jaspers, J. G. M., & Prangsma, M. E. (2001, August). *Coordinated activity in collaborative learning*. Paper presented at the Symposium Social interaction in the community of learners: Cases cross age-levels, at the Annual EARLI Meeting, Fribourg, Switzerland.

Erkens, G., Jaspers, J. G. M., Prangsma, M. E., & Kanselaar, G. (2001, March). *Het COSAR-project*. Poster presentation presented at the PROO-conference "Wat heeft onderwijsonderzoek het onderwijs te bieden?" [What can educational research offer education?] Jaarbeurs, Utrecht.

Erkens, G., Jaspers, J. G. M., Schijf, H. J. M., & Prangsma, M. E. (2001, March). *Computer-supported collaboration in argumentative writing*. Paper presented at the Proceedings of the First European Conference on Computer-Supported Collaborative Learning, Maastricht University, The Netherlands.

Erkens, G., Kanselaar, G., Prangsma, M. E., & Jaspers, J. G. M. (2002, January). *Using Tools and Resources in Computer Supported Collaborative Writing.* Paper presented at the Conference on Computer Support for Collaborative Learning, Boulder, CO.

Erkens, G., Schijf, H. J. M., Jaspers, J. G. M., & Van Berlo, J. (2000, September). *How does computer-supported collaboration influence argumentative writing?* Paper presented at the Workshop on Argumentative Text Production, EARLI SIG Writing Conference, Verona, Italy.

Erkens, G., Schijf, H. J. M., Jaspers, J. G. M., & Van Berlo, J. (2000, June). *Dynamics of coordination in collaboration*. Paper presented at the Symposium The Dynamics of Social Interaction in Learning and Instruction: theoretical, methodological and educational challenges. Earli SIG Social Interaction in learning and Instruction, Oulu, Finland.

Jaspers, J. G. M., Erkens, G., & Kanselaar, G. (2001, August). COSAR: *Collaborative writing of argumentative texts*. Paper presented at the ICALT 2001; IEEE International Conference on Advanced Learning Technologies, University of Wisconsin, Madison.

Prangsma, M. E., Erkens, G., Jaspers, J. G. M., & Kanselaar, G. (2002). *Computer support of linearization in collaborative argumentative writing.* Paper submitted for publication.

Schijf, H. J. M., Erkens, G., & Kanselaar, G. (1999, April). *Developing a computer-supported, collaborative learning environment for argumentative writing*. Paper presented at the 11th International Writing and Computers Conference, Cambridge, UK.

Schijf, H. J. M., Erkens, G., Jaspers, J. G. M., & Kanselaar, G. (1999, September). *Developing a computer-supported, collaborative learning environment for argumentative writing*. Paper presented at the Workshop Computers In Psychology 7 (CIP7).

Master's theses and student reports

Boas, T.J.E., Van der Meijden, P.E., Phielix, C., & Schoonhoven, J.-W. (2001). *Moet ik argumenteren om beter te presteren?* [Do I need argumentation for better results?]. Unpublished student report, Utrecht University.

Honkoop, P. (2002). *Diagramgebruik bij samenwerkend schrijven* [Use of diagrams in collaborative writing].Unpublished Master's thesis, Utrecht University.

Scheltens, F. (2001). *Coördinatie in coöperatieve dialogen: De relatie tussen coördinatie in de samenwerkingsdialoog en het schrijfproduct bij coöperatief schrijven* [Coordination in cooperation dialogues: The relation between coordination in the collaboration dialogue and the written product]. Unpublished Master's thesis, Utrecht University.

Theil, J.W. (2002, in preparation). *Chattalk en sociale communicatie in een coöperatieve leeromgeving* [Chat talk and social communication in a collaborative learning environment]. Unpublished Master's thesis, Utrecht University.

Van Berlo, J. (2000). *Computerondersteund samenwerken bij argumentatieve tekstproductie in het studiehuis: De relatie tussen het schrijfproces en het schrijfproduct bij coöperatief schrijven [Computer assisted collaboration in argumentative text production in the 'study house': The relation between the writing process and the written product in collaborative writing]*. Unpublished Master's thesis, Utrecht University.

APPENDIX 1: ADVISOR MODULES

Diagram

The Diagram helps you to generate and order your position, arguments, refutations, support and conclusions. The Diagram shows:

- the elements of your text in text boxes
- key words or brief, concise sentences
- the connections between the elements through arrows and lines

Before writing you jointly set up a diagram by:

- reading sources and brainstorming
- adding positions
- adding and connecting arguments to positions
- adding and connecting support and refutations to arguments
- adding and connecting conclusions

While writing you are supposed to use the Diagram by:

- consulting it
- updating it: extend and organize
    - add text boxes
    - connect with arrows
    - add text to text boxes

After writing, don’t forget to:

- check your Diagram
    - does each text box have at least one arrow?
    - does the Diagram correspond to the text?
- update your Diagram

Outline

The Outline helps you to decide on the order of your position, arguments, refutations, support and conclusions. The Outline shows:

- the content of your text
- in headings and subheadings of paragraphs (numbered automatically)
- key words or brief, concise sentences
- the order of the elements in the text

Before writing you jointly set up an outline by:

- reading sources
- adding parts
- organizing parts:
    - updating the order by changing lines with the up and down arrows
    - adding subheadings by changing the heading level with the right and left arrows

While writing you are supposed to use the Outline by:

- consulting it
- updating it: extend and organize

After writing, don't forget to:

- check your Outline
    - does the content correspond to the text?
    - does the order correspond to the text?
- update your Outline

APPENDIX 2: ASSESSMENT OF PRE-TESTS

The Wild Cat Test

Instructions and task

On the next page you will find a list with 11 characteristics of the wild cat. Please, read through them attentively. The assignment is to write a well constructed text on the wild cat. The text must contain all 11 characteristics, without adding or deleting any characteristics. You are free to change the order of the characteristics, and to replace them by synonyms. You can change the form of the sentences, add conjunctions or other words, and connect as many characteristics as you like into longer sentences.

So, you have to write a text on the wild cat using all 11 characteristics. The text must be as straightforward and clear as possible.

The 11 characteristics:

1	The wild cat has a gray coat.	De wilde kat heeft een grijze vacht.
2	The wild cat is accused of eating rabbits and chicken.	De wilde kat wordt ervan beschuldigd konijnen en kippen te eten.
3	The wild cat is mostly active during the night.	De wilde kat is vooral ’s nachts actief.
4	The wild cat is slightly bigger than the domestic cat.	De wilde kat is wat groter dan de huiskat.
5	The wild cat lives in the forest.	De wilde kat leeft in de bossen.
6	The wild cat is not a pest for human beings.	De wilde kat is geen plaag voor mensen.
7	The wild cat is an animal species that should be protected.	De wilde kat is een diersoort die beschermd moet worden.
8	The wild cat has a foxtail.	De wilde kat heeft een pluimstaart.
9	The wild cat only eats small rodents.	De wilde kat eet alleen kleine knaagdieren.
10	The wild cat is not very well known.	De wilde kat is niet erg bekend.
11	The wild cat is a shy animal.	De wilde kat is een schuw dier.

Assessment

Indicator 1: Exactness of characteristics

Total score exactness of characteristics	11 – (# missing) – (# changed) – (# added)
# of characteristics present	Characteristics that are literally present or paraphrased. Equals the maximum of 11 minus the number of changed and missing characteristics.
# of changed characteristics	Examples: Addition: The wild cat is a very shy animal. Substitution: The wild cat mainly eats small rodents. Changed meaning: The wild cat is a protected animal.
# of missing characteristics	Characteristics that were omitted completely.
# of added characteristics	New characteristics of the wild cat that do not occur on the list of 11 given characteristics. For example: The wild cat is on the verge of extinction.

Indicator 2: Linearization

The total score for linearization is the # of correctly clustered clusters plus the # of properly placed clusters.

Clustering
The characteristics should be logically clustered as follows. The order within the cluster is not important, but the cluster should be complete, there must be no interfering characteristics. Each proper cluster is worth 1 point.

Appearance (1, 4, 8)	The wild cat has a gray coat. The wild cat is slightly bigger than the domestic cat. The wild cat has a foxtail.
Way of life (3, 5, 11)	The wild cat is mostly active during the night. The wild cat lives in the forest. The wild cat is a shy animal.
Argumentative syllogism (2, 6, 9)	The wild cat is accused of eating rabbits and chicken The wild cat is not a pest for human beings The wild cat only eats small rodents.

Location

Each full cluster placed in the correct location is worth 1 point. Clusters are allowed to overlap, as long as all the sentences are in the right range of locations. Added characteristics are not counted here, but changed ones are. Missing characteristics mean 0 points for that cluster.

	Characteristic	Location
Opening sentence: The wild cat is not very well known.	10	1
C1: Appearance	1,4,8	1-5
C2: Way of life	3,5,11	4-8
C3: Argumentative syllogism	2,6,9	7-11
Conclusion: The wild cat is an animal species that should be protected.	7	11

Indicator 3: Linguistic aspects – Anaphoric use of pronouns

The total score for anaphor complexity is the total number of anaphora divided by the total number of different types of anaphora. The total number of anaphoric pronouns includes all personal, possessive, demonstrative, and relative pronouns, nouns, and ellipsis, insofar they refer to or substitute the wild cat. The following types of anaphora are distinguished here:

Ellipsis	-
Nouns	the animal; this cat
Personal pronoun	he; it; she; they
Possessive pronoun	his; its; her; their
Demonstrative pronoun	that; this; these; those
Relative Pronoun	who; whose; that

Indicator 4: Linguistic aspects – Sentence complexity

The total score for sentence complexity is the total number of sentence constituents divided by the total number of sentences.

Indicator 5: Semantic properties

The total score for semantic properties is the score for correctness of the syllogistic cluster plus the score for correctness of the familiarity cluster.

Correctness of the syllogistic cluster (2, 6, 9)
Example of correct construction: Although the wild cat is accused of eating rabbits and chicken, he only eats small rodents, so he is not a pest for human beings.

1 point	The three parts are connected correctly as shown in the example above, using conjunctions to reflect the appropriate meaning.
0.5 points	Two of the three parts are connected appropriately.
0 points	None of the three parts are connected correctly to reflect the appropriate meaning.

Correctness of the familiarity cluster (10; 1, 3, 4, 5, 11)
This cluster was added to the original system of Coirier. Example of correct construction: The wild cat is not very well known, because it is a shy animal, it is mostly active during the night, and it lives in the forest. Moreover, it is slightly bigger than the domestic cat and it is gray (which makes it rather inconspicuous).

1 point	3, 4 or 5 of these characteristics are connected with the 'well known' characteristic to reflect the appropriate meaning.
0.5 points	1 or 2 of these characteristics are connected with the 'well known' characteristic to reflect the appropriate meaning.
0 points	None of these characteristics are connected with the 'well known' characteristic to reflect the appropriate meaning.

Total score Wild Cat Test

For each of the five indicators, the means of the entire original sample were determined (N Control group = 76; N total sample = 427). Next, the participants scoring below or equal to the mean received a norm score of 0, and the participants scoring above the mean received a norm score of 1. The five norm scores are then added up for each participant to form the total score for the Wild Cat Test.

Underline Arguments Test

Instructions

In this exercise you will find 20 short texts. For each text, you are asked to underline the statement that functions both as an argument and as a position. Here are two examples:

> You need to put on an extra 10p stamp, because a letter must be stamped sufficiently. Otherwise it won't arrive on time.
>
> The position that you need to put on an extra stamp is supported with the argument that a letter needs proper stamping. At the same time, this is another position, supported by the argument that it won't arrive on time.
>
> Fish are certainly in pain when they are caught, because they struggle fiercely. Fishing is therefore a horrible hobby.
>
> The position that fishing is a horrible hobby is supported with the argument that fish are certainly in pain when they are caught. At the same time, this is another position, supported by the argument that they struggle fiercely.

If you have underlined the wrong statement, please correct by crossing out the wrong statement and underlining the correct one.

Task

1. Volgens mij moet je nodig eens wat aan lichaamsbeweging gaan doen, want je bent veel te zwaar. Je kunt beter geen brommer kopen.
2. Hij is al twee keer failliet gegaan, dus mijns inziens moeten we met hem geen zaken doen. We moeten de directie vertellen dat we geen gebruik zullen maken van zijn diensten.
3. Ik ga altijd met de tent op vakantie. Het hotelpubliek staat me niet aan. Je moet altijd zo gemaakt gezellig doen.
4. Het aardgastransport is voor de overheid duurder geworden, dus de aardgasprijzen worden dit jaar vast flink verhoogd. Onze stookkosten zullen dit jaar wel behoorlijk tegenvallen.
5. Saskia moeten we niet mee laten gaan met het schoolreisje naar Parijs. Ik vind dat zulke schoolreisjes moeten worden afgeschaft. Voor veel ouders zijn de kosten te hoog.
6. Hij is onbeschoft. Laatst begon hij in een restaurant zomaar de ober uit te schelden. Ik wil hem niet op jouw verjaarsfeestje hebben.
7. Het wordt voor jou tijd eens wat aan sport te gaan doen. Toen je laatst de tram moest halen was je helemaal buiten adem, dus volgens mij heb je helemaal geen conditie.
8. Moderne mensen willen graag een telefoontoestel dat niet uit de toon valt bij de inrichting, want zij schenken veel aandacht aan hun interieur. De PTT moet meer modern vormgegeven toestellen op de markt brengen.
9. Volgens mij werkt de televisie verslavend. Bij sommige mensen staat hij aan ongeacht of er televisie gekeken wordt. Het is beter geen televisie te kopen.
10. Neem maar geen bananen mee van de markt. Bananen bevatten tamelijk veel zetmeel. Ik vind dat bananen niet in een fruitsalade verwerkt moeten worden.
11. Je moet niet intekenen op die dure encyclopedie. Je zoekt er hooguit twee keer per jaar iets in op. Eigenlijk is het een overbodige aanschaf.
12. We moeten een afspraak maken om samen naar de stad te gaan. Je loopt de laatste tijd steeds in dezelfde kleren, dus je hebt dringend iets nieuws nodig.
13. Ik vind dat we een nieuwe computer moeten kopen. De harde schijf in deze computer is veel te klein. Ik krijg steeds de melding ‘disk full’.
14. Ik denk dat we Herman maar op drumles moeten doen. Hij is erg ritmisch, want hij zit voortdurend met zijn vingers op tafel te tikken.
15. U heeft zich niet aan de gemaakte afspraken gehouden. U heeft uw auto nog niet van de weg gehaald. We hebben hem gisteren nog rijdend gesignaleerd.

16. Je zit de hele tijd te gapen. Je bent vast hartstikke moe. Je kunt maar beter vroeg naar bed gaan vanavond.
17. Nee, je krijgt echt geen hond, ze moeten wel drie keer per dag hun behoefte doen dus je moet er veel te vaak mee naar buiten.
18. Alex kan beter vanavond naar huis rijden. Hij rijdt veel beter dan jij. Hij heeft in één keer zijn rijbewijs gehaald.
19. Volgens mij moet je met de baby naar de dokter gaan. Ze heeft vast en zeker koorts. Haar voorhoofd heeft zeker nog nooit zo warm gevoeld.
20. Carpoolen is goedkoper voor de automobilist. Je hebt lagere afschrijvingskosten voor je auto. Dus mensen moeten overgaan tot carpoolen.

APPENDIX 3: INTEGRATED ACTIVITY PROTOCOLS (IAPS)

Table 1: Fragment from collaborative protocol in Diagram-Outline-Advisor condition.

Line	Time	Sequence	Actor	Activity	Content
515	1:12:12	1	1	chat	we're doing fine
516	1:12:19	1	0	to-chat	
517	1:12:21	1	0	chat	what now?
518	1:12:25	1	0	to-text	
519	1:12:27	1	1	to-src9	
520	1:12:27	2	1	to-src9	
521	1:12:41	1	1	to-src11	
522	1:12:42	1	1	to-src11	
523	1:12:50	1	0	to-chat	
524	1:12:51	1	1	to-chat	
525	1:13:00	1	1	chat	do you have sources there that support our position?
526	1:13:02	1	0	chat	no
527	1:13:04	1	0	to-src1	
528	1:13:06	1	0	to-src6	
529	1:13:07	1	1	chat	huh?
530	1:13:12	1	1	chat	no supporting arguments?/
531	1:13:22	1	0	chat	nothing just about brain death etc.
532	1:13:25	1	0	to-src6	
533	1:13:36	1	0	to-chat	
534	1:13:38	1	1	chat	I found something
535	1:13:43	1	0	chat	Hey I found this
536	1:13:45	1	0	to-notes	
537	1:13:51	1	0	to-text	
538	1:14:02	1	0	to-chat	
539	1:14:07	1	0	chat	The existence.....
540	1:14:21	1	0	to-text	
541	1:14:58	1	1	diag-open	
542	1:15:13	1	0	to-chat	
543	1:15:17	1	1	diag-close	
544	1:15:20	1	0	chat	John is this right or not
545	1:15:28	1	0	to-text	
546	1:15:50	1	1	chat	is ok
547	1:16:26	1	1	turn-ask	
548	1:16:27	1	0	to-chat	
549	1:16:31	1	0	chat	wait what is it
550	1:16:35	1	1	chat	let me have a go
551	1:16:40	1	1	chat	you can go quickly after
552	1:16:41	1	0	turn-give	

Table 2: Fragment from collaborative protocol in Diagram-Outline-Advisor condition.

Line	Time	Sequence	Actor	Activity	Content
1342	3:14:57	1	1	to-text	
1343	3:15:00	1	0	outl-close	
1344	3:15:18	1	0	wordcount	757
1345	3:16:15	1	0	chat	connection between paragraphs
1346	3:16:59	1	0	wordcount	755
1347	3:17:12	1	0	outl-open	
1348	3:17:15	1	0	outl-close	
1349	3:17:25	1	0	search-mark	
1350	3:17:32	1	0	turn-ask	
1351	3:17:34	1	0	turn-ask	
1352	3:17:41	1	0	wordcount	755
1353	3:18:08	1	0	chat	Hey those statistics who are they from?
1354	3:18:11	1	1	to-src1	
1355	3:18:11	2	1	to-src3	
1356	3:18:14	1	1	to-src3	
1357	3:18:20	1	0	chat	write down who they are from
1358	3:18:26	1	1	to-text	
1359	3:18:27	1	0	outl-open	
1360	3:18:30	1	0	outl-close	
1361	3:19:23	1	0	chat	ok
1362	3:19:25	1	0	wordcount	759
1363	3:19:30	1	0	wordcount	759
1364	3:20:24	1	0	chat	John do you have a source that says the chances of success are slim
1365	3:20:34	1	0	chat	replace it with that medical journal
1366	3:21:10	1	0	wordcount	757
1367	3:21:33	1	0	wordcount	757
1368	3:21:52	1	1	diag-open	
1369	3:22:08	1	1	to-chat	
1370	3:22:09	1	0	wordcount	757
1371	3:22:37	1	1	chat	I do have a source that says organs age and anyway slim the chances of success
1372	3:22:47	1	1	chat	but that way we refute our position
1373	3:22:52	1	1	chat	that source is tough
1374	3:23:01	1	0	chat	So??
1375	3:23:04	1	1	chat	leave it like it is
1376	3:23:14	1	1	turn-give	

APPENDIX 4: CODING AND ASSESSMENT OF ARGUMENTATIVE TEXTS

Coding

Table 1.1: Coding categories for argumentative texts in Control group.

Category	Description
Part of argument	Generic term for argumentative utterances: claim, conclusion, solution, support, refutation, put-in-perspective.
Claim	Utterance that provokes difference of opinion, or poses a problem; "I assert that …".
Part of argument & claim	Utterance that functions both as a [PA] and as a position for the next paragraph, and indicates a subordinate structure.
Conclusion	An inference (*so*), a summary (*in short*), a consequence (*therefore*), or a paraphrase of the claim.
Solution	A solution to a problem mentioned earlier.
Support	Reason or fact used to underpin a claim, even if the reason or fact is not true, reliable, or even genuinely grounding.
Put in perspective	A partial refutation of a claim (*unless, if*).
Refutation	A full invalidation of a claim.
Organizer	Often the first or final utterance in a paragraph, with little or no content, and used to introduce a new topic or paragraph.
Information	Utterance that does not contribute directly to the argumentation.
Elaboration	An argument that is a continuation of an argument mentioned earlier; often an example.

Table 1.2: Coding categories for argumentative texts in the experimental groups.

Category	Description
Position	The overall position or main claim of the text. If the same position was mentioned more than once, only the first occurrence was coded as such, and the rest was coded as arguments, to avoid scoring of multiple positions in the assessment.
Argument pro	Argument supporting the position.
Argument contra	Argument refuting the position, or a contra position.
Support	Support for an argument pro or contra.
Refutation	Refutation of an argument pro or contra.
Conclusion	The final conclusion, a solution, or a conclusion to an argument.
Information	All phrases that do not contribute to the argumentation.
Organizer	A phrase with little or no content: rhetorical, announcing or indicating the structure or function of the next or the preceding section.
Title	The title at the top of the text.

Assessment Sheet

Category	Questions	Points	Totals
A. Structure of the text	Is there a title? If no, 0. If yes, 1.	0 – 1	Title …
	Is there an introduction? If no, 0. If yes, 1, does it: attract the attention of the reader? state the topic? contain counterarguments or different points of view? state the writers' position?	0 – 1 0 – 1 0 – 1 0 – 1 0 – 1	Intro …
	Is there a body? If no, 0. If yes, 1, does it: state the writers' position? contain supporting arguments? contain refutations of counterarguments? suggest solutions?	0 – 1 0 – 1 0 – 1 0 – 1 0 – 1	Body …
	Is there a conclusion? If no, 0. If yes, 1, does it: state a conclusion or suggest a solution? summarize the main arguments? indicate possible consequences or give recommendations?	0 – 1 0 – 1 0 – 1 0 – 1	Concl …
			A: …
B. Audience focus (presentation, level of formality, ability to empathize)	Is the division into paragraphs correct, appropriate, and sufficient? Have connecting sentences been used correctly, appropriately, and sufficiently? Is the tone of voice appropriate for the intended audience (formal)? Do the writers succeed in showing their enthusiasm and their commitment to the topic?	 0 – 1 – 2 0 – 1 – 2 0 – 1 – 2 0 – 1 – 2	B: …
C. Quality of argumentation at segment level	What is the number of main segments? Determine the quality of argumentation for each main segment (0 - 8 points). a. _____ i. _____ b. _____ j. _____ c. _____ k. _____ d. _____ l. _____ e. _____ m. _____ f. _____ n. _____ g. _____ o. _____ h. _____ p. _____	# of segments: … # of points: … 	C: (# of points divided by # of segments) …
D. Quality of argumentation of the text as a whole	Determine the quality of argumentation for the text as a whole (0 - 6 points).	0 – 1 – 2 – 3 – 4 – 5 – 6	D: …
Total score	A + B + C + D =		…

Conversion to a 10-point scale:
(10 x A) divided by the maximum score of 15
(10 x B) divided by the maximum score of 8
(10 x C) divided by the maximum score of 8
(10 x D) divided by the maximum score of 6
(10 x total score) divided by the total maximum score of 37

Assessment instructions

Ad. C & D: Quality of argumentation in the segments and of the text as a whole. Segments and total text were assessed separately, because the argumentation within a segment can be perfect whilst at the same time the line of argumentation is not maintained throughout the text. The opposite occurs as well: the argumentation for the text as a whole is not properly developed in the segments. All MEPA text codings except *information* and *organization* are considered part of the argumentative structure. The number of qualifying lines is divided by the total number of lines in the segment.

C. Quality of argumentation within the segments ONLY.
C.1. The presence of an argumentative structure.

0 points	Less than 30%
1 point	30 to 70%
2 points	> 70%

Exceptions: The first segment is an exception to this rule if it functions as an introduction. For the introductory segment we assessed whether the information was to the point.

C.2. Supporting the segment claim or argument.

0 points	There is no claim or argument, or the segment does not contain any supports or refutations.
1 point	The segment claim or argument is supported/refuted through arguments or facts that clearly relate to the topic of the claim, but do not sufficiently support or refute it, for example because they have the wrong level of specificity or because they only support/refute part of the claim/argument. If it contains no supports or refutations, the last segment receives a maximum of 1 point, but only if it functions as a conclusion and introduces new arguments supporting or refuting the conclusion or the position mentioned in the final segment.

2 points	The segment claim/argument is supported/refuted by appropriate and sufficient arguments/facts. A general claim like "Through improved science and surgical techniques the chances of success for transplants are increasing." needs to be supported by facts that prove it, or at least refer to a source containing such proof. An example of a single successful transplant is not sufficient evidence. A specific claim like "The life of John Smith did not improve after his transplant." needs to be supported by facts about that transplant.

Exceptions: The first segment is an exception to this rule if it functions as an introduction. 0 points: random copying of source information; 1 point: a good summary of general information or a logical order of information; 2 points: informative introduction with well summarized information in a logical order.

C.3. The conclusion of the segment.

0 points	There is no conclusion, or the conclusion does not refer to the segment topic.
1 point	The conclusion does refer to the segment topic, but it is not appropriate, for example because it has the wrong level of specificity.
2 points	The grounds for the conclusion lie within the segment. It is clear what the conclusion is based on – usually a claim/argument and support/refutations. The level of specificity of the conclusion is appropriate.

C.4. Structure of the segment.

0 points	The structure cannot be derived. The segment is a collection of loose sentences.
1 point	The segment is slightly disorganized, but the structure can still be derived, OR it is an informative segment with good structure.
2 points	The segment is well composed, well structured. It clearly contains a beginning, a middle and an end.

D. Quality of argumentation in the text as a whole.

An argumentative text consists of the main elements introduction, position, support, refutation, and conclusion. The main position for organ donation was generally "We are in favor of/against organ donation", or "We are in favor of/against donor codicils", and for cloning the position was "We are in favor of/against cloning". Here, we assessed the strength of the arguments and conclusions in supporting the position.

D.1. Support of the main position.

0 points	EITHER there is no position, OR the rest of the text does not support or refute the position, OR there are two or more fully informative segments while the rest of the text is worth 1 point.
1 point	EITHER the position is supported/refuted by arguments or facts that are clearly related to the topic of the position, but do not have sufficient strength, for example by lacking proper specificity or by only covering part of the position OR there are two or more fully informative segments while the rest of the text is worth 2 points.
2 points	The main position (sometimes found in the last segment) is supported/refuted by fitting and sufficient arguments and conclusions AND there is no more than one informative segment (not counting the introduction).

D.2. Conclusion of the text.

0 points	EITHER there is no conclusion OR the conclusion does not relate to the topic of the text.
1 point	The conclusion does relate to the topic, but it is not have the proper specificity (either too specific or too general).
2 points	The basis for the conclusion is found within the text: it is clear what the conclusion is drawn from. The specificity of the conclusion is about right.

D.3. Structure of the text.

0 points	The structure of the text is very obscure. The text is a collection of segments that do not seem to be related to the main position.
1 point	The text is a bit untidy, but the structure can still be derived.
2 points	The text is well structured. Introduction, body and conclusion are clearly present.

APPENDIX 5: OUTLINE ASSESSMENT

Complexity of the outline

1. Formal structure	Total of 1.1 + 1.2 + 1.3
1.1. Number of hierarchy levels	# of levels as indicated by the numbering (depth – 1/2/3/4/5; e.g., 1.1.1.1.1)
1.2. Number of organizational items per paragraph	1.2.1 divided by 1.3.2
1.2.1. Total number of lines with explicit organizational items	Organizational items being: title / introduction / body / closing / transition / question / topic / problem definition / general explanation of problem / acknowledgement of sources / argumentative part / anecdote / example / information / facts / paragraph # / summary AND position / argument pro / argument contra / support / refutation / conclusion
1.3. Number of sub items per paragraph	1.3.1 divided by 1.3.2
1.3.1. Total number of sub items	# of numbered lines in outline below first level
1.3.2. Total number of paragraphs	# of paragraphs; equals number of lines at first hierarchical level

A paragraph is defined as a set of hierarchically linked lines. For example, the following outline consists of three paragraphs and six lines. There are three sub items. This gives the following scores: (1) = **5**; (1.1) = **2**; (1.2) = (1.2.1)/(1.3.2) = 6/3 = **2**; (1.3) = (1.3.1)/(1.3.2) = 3/3 = **1**.

- introduction
    - position
- body
    - argument 1
    - argument 2
- conclusion

Items with more than one organizational item were counted as one item, and items with organizational and content items were counted as an organizational item as well.

2. Argumentative structure	Total of 2.1 + 2.2
2.1. Number of argumentative lines per paragraph	2.1.1 divided by 1.3.2
2.1.1. Total number of lines with explicit argumentative items	Sum of values under 2.1.2 = # position + # argument pro + # argument contra + # support + # refutation + # conclusion
2.1.2. Number of argumentative items per type	gives six scores: # position / # argument pro / # argument contra / # support / # refutation / # conclusion
2.2. Variation in argumentative types	2.2.1 divided by six
2.2.1. Total number of different types of argumentative items	(1/2/3/4/5/6)

Only explicitly argumentative items were counted here, so content items that obviously function argumentatively, but are not explicitly labeled as such were not counted. Some dyads put multiple information or argumentative units in one line, against the instructions. For measure 2.1, the argumentative lines were counted, instead of the number of argumentative items. The outline example under measure 1 above would give the following scores: (2) = **1.33**; (2.1) = (2.1.1)/(1.3.2) = 3/3 = **1**; (2.1.2) = 1 position / 2 argument pro / 0 argument contra / 0 support / 0 refutation / 0 conclusion; (2.2) = (2.2.1)/6 = 2/6 = **.33**.

3. Formal content: abstract/mixed/concrete	Total three scores is score for each type divided by number of items abstract, meta level: no content, only argumentative and/or organizational indicators – see list under 1.2.1 concrete, content level: lines without indication of abstract function, only content mixed: abstract and concrete combined in one item

This measure gives three separate scores adding up to a total of 1.0. The example under measure 1 above gives the following scores: (3) = 6 abstract / 0 concrete / 0 mixed = 1.0 / 0 / 0. The total score was calculated by converting the scores to a scale of abstractness from –1 to +1 through the following formula:

Total = 1 x (abstract) – 1 x (concrete)

4. Comprehensiveness: phrase complexity	Total four scores is score for each type divided by number of items
4.1. key words (0/1)	loose words, noun phrases (adj. + noun)
4.2. clauses (0/1)	sentence fragments lacking a subject or a finite verb; dependent clauses; prepositional phrases; etc.
4.3. sentences (0/1)	full sentences with SVO
4.4. paragraphs (0/1)	multiple sentences or phrases, as shown by punctuation

This measure gives four separate scores adding up to a total of 1.0. The example under measure 1 above gives the following scores: (4) = 6 key words / 0 clauses / 0 sentences / 0 paragraphs = 1.0 / 0 / 0 / 0. The total score was calculated by converting the scores to a scale of abstractness from –1 to +1 through the following formula:

Total = 1 x (abstract) – 1 x (concrete)

An item was scored as the highest ranking/most complex phrase type it contained. For example, if an item consists of a keyword followed by a full sentence, the item was scored as a sentence.

Correspondence to the argumentative text

5. Order correspondence	Are all the items from the outline ordered in the same way in the text?

To check the order of items in the outline with the order of the contents in the text, the outline and text were put next to each other and arrows were drawn from the outline to the corresponding items in the text. Crossing arrows indicate discrepancies. When two arrows cross, both items are in the wrong place. If an arrow crosses more than one other arrow, the crossing arrow is in the wrong place, but the others are not. The total score is the total number of properly placed outline items divided by the total number of lines in the outline.

6. Item correspondence	Are all items from the outline present in the text?

The score is the (total number of lines in the outline minus the number of outline lines missing from the text) divided by the total number of lines in the outline.

7. Paragraph correspondence	Are all paragraphs from the text as assigned by the experimenters present in the outline?

The score is the (total number of paragraphs in the text minus the number of paragraphs missing from the outline) divided by the total number of paragraphs in the text.

APPENDIX 6: EVALUATION QUESTIONNAIRE

In addition to answering the questions, students were asked to state the date, their name, login number, school name, and class. They were asked to tick only one box per question. We kept a record of double-ticked questions and coded the highest and the lowest of two ticked options alternately. Most questions were followed by an open question "And why?".

Questions about the writing assignment	**Possible answers**
1. Did you find the writing assignment easy or difficult?	Difficult; average; easy
2. What did you find easy or difficult?	(Open)
3. Did you find the sources easy or difficult?	Difficult; average; easy
4. What was it like to perform a writing task in this way (together through the computer)? And why?	Annoying; average; nice
Questions about the computer program	**Possible answers**
5. How did logging on and off go? And why?	Bad; reasonable; good
6. What did you think of the notes window? And why?	Bad; reasonable; good
7. What did you think of the information window? And why?	Bad; reasonable; good
8. What did you think of the chat windows? And why?	Bad; reasonable; good
9. What did you think of the shared text window? And why?	Bad; reasonable; good
10. What did you think of the Diagram window? And why?	Bad; reasonable; good
11. We gave extra instruction on the use of the Diagram window. Did this help you? And why?	No; a little; yes
12. In the information window we gave extra tips on the use of the Diagram window. Did this help you? And why?	No; a little; yes
13. Has the function of the Diagram become sufficiently clear through the instruction and/or the tips? And why?	Bad; reasonable; good
14. What did you think of the Outline window? And why?	Bad; reasonable; good
15. We gave extra instruction on the use of the Outline window. Did this help you? And why?	No; a little; yes
16. In the information window we gave extra tips on the use of the Outline window. Did this help you? And why?	No; a little; yes
17. Has the function of the Outline become sufficiently clear through the instruction and/or the tips? And why?	Bad; reasonable; good
18. What did you think of the traffic light? And why?	Bad; reasonable; good
19. Was the use of the screen buttons clear? And why?	Yes; no
20. Were the screen buttons easy to use? And why?	Yes; no

Questions about working collaboratively	**Possible answers**
21. How did the collaboration go? And why?	Bad; reasonable; good
22. How did turn taking in writing go? And why?	Bad; reasonable; good
23. Did you contribute equally to writing the text? And why?	Yes; reasonable; no
24. How did turn taking in the Diagram go? And why?	Bad; reasonable; good
25. Did you contribute equally to the Diagram? And why?	Yes; reasonable; no
26. How did turn taking in the Outline go? And why?	Bad; reasonable; good
27. Did you contribute equally to the Outline? And why?	Yes; reasonable; no
28. How did the deliberation go between you and your partner? And why?	Bad; reasonable; good
Miscellaneous questions	
29. Did you spend time on the task away from class? For example, did you discuss it with others, or did you search on the internet or in the library? (You may tick several boxes.)	No; yes
30. If yes, what did you do and how much time did you spend?	Discussion; searching; other, namely: …
31. What advantages does this way (working together through the computer) of performing school or learning tasks have according to you?	(Open)
32. What disadvantages does this way (working together through the computer) of performing school or learning tasks have according to you?	(Open)
33. Do you have any ideas for improving the TC3 environment?	(Open)

APPENDIX 7: TOOL USE RESULTS

Section 1: Descriptive statistics of tool use percentages

Means and standard deviations for percentages of tool use in the three phases of the writing process, for the entire sample (N = 139 dyads). Means are in percentages of the phase total.

Table 1.1: Means of percentages all conditions 1st phase.

	Total	C	D	DA	DO	DOA	O	OA
To chat	12.32	14.24	11.15	10.48	12.77	12.29	11.54	13.14
Chat	43.71	56.39	40.98	37.38	43.67	37.76	39.69	37.45
To source	18.24	11.41	21.76	18.80	19.16	17.64	22.97	22.93
Mark source	5.58	8.87	3.23	4.85	3.85	.91	8.87	3.98
To notes	4.19	.32	3.92	4.12	6.34	3.02	5.15	11.49
To text	2.65	3.16	2.48	2.02	2.96	2.55	2.47	2.57
To assignment	3.26	4.86	3.08	1.69	3.80	3.85	2.25	2.44
To manual	.71	.41	.80	.69	.85	.42	1.41	.43
To Diagram tips	1.08			1.12		.99		
To Outline tips	1.15					.93		1.37
Word count	.33	.31	.33	.41	.44	.18	.27	.28
Stop	.27	.04	.38	.51	.15	.31	.38	.24
Diagram open	3.22		3.16	4.47	.92	5.12		
To Diagram	1.62		1.51	2.44	.49	2.17		
Diagram close	2.47		2.69	3.63	.83	2.80		
Diagram activities within Diagram	4.72		4.53	7.40	1.01	6.41		
Diagram delete link	.11		.04	.22	.00	.17		
Diagram delete object	.51		.53	.78	.09	.72		
Diagram new link	1.06		.86	1.62	.16	1.91		
Diagram new object	1.44		1.53	2.25	.36	1.65		
Diagram update object	1.60		1.57	2.53	.40	1.96		
Outline open	1.54				1.19	.93	2.32	1.60
To Outline	.57				.39	.76	.61	.69
Outline close	1.44				1.18	.96	2.09	1.38
Total no. of acts	247.78	141.73	275.47	348.77	247.43	292.00	268.33	207.27

Table 1.2: Standard deviations of percentages all conditions 1st phase.

	Total	**C**	**D**	**DA**	**DO**	**DOA**	**O**	**OA**
To chat	3.51	4.49	1.93	2.65	2.86	1.54	3.27	3.88
Chat	15.03	17.32	13.14	12.45	10.81	11.45	11.52	8.75
To source	8.48	6.83	9.47	6.72	8.10	8.75	6.66	5.48
Mark source	7.54	11.55	4.15	5.37	3.22	2.11	7.53	4.82
To notes	4.51	1.21	2.37	2.99	4.99	2.18	3.44	6.24
To text	2.03	2.72	2.37	1.54	1.50	2.09	1.50	1.38
To assignment	3.06	4.35	2.46	1.54	2.87	2.98	1.32	1.59
To manual	.96	.58	.78	.76	1.42	.55	1.19	.56
To Diagram tips	.89			.93		.79		
To Outline tips	.89					.96		.77
Word count	.57	.65	.54	.67	.62	.38	.39	.38
Stop	.41	.16	.44	.43	.22	.51	.39	.58
Diagram open	4.21		2.69	4.20	1.41	7.19		
To Diagram	1.68		1.68	1.63	.95	1.70		
Diagram close	3.00		2.44	3.91	1.30	2.44		
Diagram activities within Diagram	5.53		5.83	5.99	2.00	4.80		
Diagram delete link	.29		.09	.43	.00	.30		
Diagram delete object	.97		.88	1.31	.20	.92		
Diagram new link	1.63		1.49	1.90	.51	1.74		
Diagram new object	1.75		1.93	1.99	.66	1.29		
Diagram update object	1.85		2.05	1.88	.92	1.62		
Outline open	1.83				1.96	1.15	2.08	1.25
To Outline	.73				.68	1.03	.63	.62
Outline close	1.73				1.93	1.20	1.95	.98
Total no. of acts	150.90	80.60	112.40	190.85	123.79	236.87	86.73	64.06

Table 1.3: Means of percentages all conditions 2nd phase.

	Total	**C**	**D**	**DA**	**DO**	**DOA**	**O**	**OA**
To chat	13.87	16.96	12.77	11.61	12.86	13.36	13.63	14.66
Chat	38.65	46.01	33.95	33.31	40.29	32.53	38.70	38.99
To source	14.00	10.51	18.21	13.51	14.92	15.34	15.77	12.89
Mark source	1.52	2.76	.74	.59	.97	.83	1.99	2.28
To notes	4.74	1.68	7.28	5.33	5.28	6.32	4.35	6.50
To text	9.95	14.21	8.97	5.88	8.75	8.95	10.36	11.18
To assignment	.94	1.13	1.05	.82	.91	.51	.90	1.00
To manual	.33	.37	.33	.17	.54	.19	.29	.34
To Diagram tips	.24			.22		.29		
To Outline tips	.33					.14		.53
Word count	4.37	5.94	5.05	4.00	2.95	3.08	3.97	4.43
Stop	.59	.43	.72	.91	.29	.51	.48	.95
Diagram open	3.68		3.20	5.95	1.84	2.86		
To Diagram	1.71		.89	2.81	1.19	1.43		
Diagram close	3.35		2.99	5.50	1.57	2.52		
Diagram activities within Diagram	6.22		3.84	9.37	4.47	6.12		
Diagram delete link	.19		.08	.30	.10	.26		
Diagram delete object	.72		.68	.79	.72	.60		
Diagram new link	1.73		.54	3.31	.85	1.65		
Diagram new object	1.60		1.26	1.86	1.52	1.67		
Diagram update object	1.99		1.27	3.12	1.28	1.95		
Outline open	2.56				1.40	2.14	4.17	2.81
To Outline	.97				.55	1.03	1.48	.97
Outline close	2.31				1.20	1.84	3.91	2.49
Total no. of acts	358.47	319.91	236.41	390.85	462.83	438.36	317.50	355.27

Table 1.4: Standard deviations of percentages all conditions 2nd phase.

	Total	**C**	**D**	**DA**	**DO**	**DOA**	**O**	**OA**
To chat	3.14	2.90	2.53	2.52	2.13	1.63	2.73	2.35
Chat	11.37	11.28	11.60	11.13	9.07	8.55	9.76	8.38
To source	6.99	4.57	9.03	7.47	6.98	4.35	7.77	3.76
Mark source	2.71	4.37	1.23	.87	1.32	.84	2.65	2.70
To notes	4.13	2.76	5.24	3.47	3.59	3.33	3.08	5.16
To text	4.41	3.95	4.12	2.67	2.79	3.40	3.19	3.60
To assignment	.98	1.17	1.01	.95	.95	.74	.86	.72
To manual	.46	.57	.51	.29	.52	.29	.38	.28
To Diagram tips	.34			.34		.33		
To Outline tips	.38					.17		.44
Word count	3.55	3.84	4.07	2.98	1.99	1.55	3.99	4.47
Stop	.89	1.58	.62	.53	.16	.48	.17	.67
Diagram open	3.51		3.39	4.19	1.57	1.63		
To Diagram	1.76		1.55	1.89	1.36	1.15		
Diagram close	3.24		3.08	3.89	1.21	1.50		
Diagram activities within Diagram	6.10		5.81	6.36	5.24	4.48		
Diagram delete link	.34		.21	.41	.21	.41		
Diagram delete object	.90		1.03	.91	.94	.48		
Diagram new link	2.21		.99	2.65	1.20	1.94		
Diagram new object	1.61		2.24	1.34	1.55	.99		
Diagram update object	2.09		2.21	2.03	1.72	1.66		
Outline open	2.47				1.20	2.18	3.16	2.10
To Outline	1.16				.71	1.47	1.24	1.18
Outline close	2.35				1.04	1.93	3.05	1.99
Total no. of acts	157.34	151.90	101.73	162.18	134.57	156.23	112.53	160.79

Table 1.5: Means of percentages all conditions 3rd phase.

	Total	**C**	**D**	**DA**	**DO**	**DOA**	**O**	**OA**
To chat	13.24	17.07	11.68	11.85	11.06	12.16	12.52	14.22
Chat	44.86	52.51	37.93	43.02	42.15	44.26	41.85	48.11
To source	5.92	3.47	6.13	6.80	4.70	5.08	10.12	7.43
Mark source	.18	.17	.05	.10	.13	.05	.62	.16
To notes	2.58	.90	3.48	3.23	2.18	2.50	3.52	4.07
To text	10.57	16.69	7.61	7.98	6.89	8.91	11.13	11.36
To assignment	.59	.47	1.08	.49	.44	.31	.56	1.11
To manual	.18	.13	.26	.19	.21	.02	.23	.16
To Diagram tips	.11			.13		.06		
To Outline tips	.20					.10		.30
Word count	6.87	8.04	6.72	6.92	4.44	6.90	9.03	5.02
Stop	.83	.54	1.22	1.35	.48	.73	.54	1.12
Diagram open	4.52		4.51	4.05	5.85	2.90		
To Diagram	2.48		2.96	2.53	2.54	1.48		
Diagram close	2.89		3.22	3.53	2.11	2.51		
Diagram activities within Diagram	9.88		13.14	7.82	11.53	6.27		
Diagram delete link	.45		.67	.23	.61	.28		
Diagram delete object	1.17		1.67	.58	1.76	.55		
Diagram new link	3.33		4.23	3.27	3.30	2.14		
Diagram new object	1.96		2.83	1.14	2.50	1.41		
Diagram update object	2.97		3.73	2.59	3.36	1.88		
Outline open	2.95				2.25	2.53	4.12	2.90
To Outline	1.57				1.36	1.45	2.02	1.36
Outline close	2.46				1.67	1.77	3.76	2.68
Total no. of acts	349.23	360.18	238.41	336.77	443.91	416.18	284.94	357.36

Table 1.6: Standard deviations of percentages all conditions 3rd phase.

	Total	**C**	**D**	**DA**	**DO**	**DOA**	**O**	**OA**
To chat	3.38	2.47	2.47	2.27	2.23	2.30	2.89	3.37
Chat	11.54	8.09	12.49	9.23	11.52	8.30	12.17	13.67
To source	5.38	2.60	7.45	5.24	3.59	2.46	7.03	5.83
Mark source	.85	.70	.22	.30	.41	.12	2.03	.29
To notes	3.03	1.05	4.44	3.03	2.81	1.86	3.41	3.01
To text	4.62	3.16	2.66	2.92	2.59	2.31	2.15	3.03
To assignment	.78	.58	1.09	.64	.55	.57	.51	1.38
To manual	.35	.26	.43	.41	.36	.07	.42	.27
To Diagram tips	.21			.23		.14		
To Outline tips	.34					.21		.42
Word count	4.08	3.85	5.36	2.94	2.91	2.81	4.97	3.36
Stop	.68	.38	.72	.89	.32	.47	.54	.55
Diagram open	7.59		2.60	2.51	13.39	1.56		
To Diagram	1.61		1.50	1.57	1.80	1.00		
Diagram close	2.00		1.80	2.39	1.65	1.26		
Diagram activities within Diagram	7.36		9.25	5.91	7.27	3.86		
Diagram delete link	.78		1.20	.38	.80	.40		
Diagram delete object	2.14		2.02	.68	3.27	.64		
Diagram new link	2.81		3.94	2.68	2.27	1.33		
Diagram new object	1.97		2.46	1.29	2.00	1.49		
Diagram update object	2.19		2.66	2.14	1.94	1.27		
Outline open	2.32				1.29	1.82	2.63	3.17
To Outline	1.09				.90	.74	1.40	.99
Outline close	2.15				.89	.87	2.43	3.23
Total no. of acts	143.73	123.19	82.96	129.72	181.74	107.79	92.84	157.99

Section 2: T-test differences of Control group vs. Experimental group

Table 2.1: Independent samples T-test for Control group vs. Experimental group. Mean differences of tool use percentages. Values are Control group minus Experimental group.

	All phases	**1st phase**	**2nd phase**	**3rd phase**
To chat	4.33	2.51	4.05	5.02
Chat	11.97	16.63	9.66	10.04
To source	-5.94	-8.96	-4.57	-3.22
Mark source	1.00	4.31	1.62	
To notes	-3.54	-5.07	-4.01	-2.20
To text	6.67	.67	5.58	8.03
To assignment		2.09		
To manual	-.15	-.40		
Word count	2.25		2.05	1.53
Stop	-.25	-.30		-.37
Total no. of acts	-175.26	-139.07	-50.57	

Table 2.2: Independent samples T-test for Control group vs. Experimental group. Mean differences of tool use duration.

	1st phase	**2nd phase**	**3rd phase**
In chat			
To chat			-1.42
Chat			
In source	9.01	16.94	10.79
To source	9.99	11.60	12.28
Mark source			
To notes	-21.77		-15.80
To text			-7.94
In instruction		18.92	
To assignment		12.40	
To manual		8.77	3.99
Mean duration per activity		2.85	

Section 3: Differences for tool use percentages in all phases

Tables 3.1 to 3.14: Mean differences of tool use percentages for all phases between conditions (Bonferroni). Values are row label – column label. Only significant differences are shown.

Table 3.1: To chat

	C	**D**	**DA**	**DO**	**DOA**	**O**	**OA**
C	-	4.70	5.35	4.38	4.05	3.75	2.51
D	-4.70	-					-2.19
DA	-5.35		-			-1.60	-2.84
DO	-4.38			-			-1.87
DOA	-4.05				-		
O	-3.75		1.60			-	
OA	-2.51	2.19	2.84	1.87			-

Table 3.2: Chat

	C	**D**	**DA**	**DO**	**DOA**	**O**	**OA**
C	-	13.79	13.85	9.92	13.69	11.00	8.86
D	-13.79	-					
DA	-13.85		-				
DO	-9.92			-			
DOA	-13.69				-		
O	-11.00					-	
OA	-8.86						-

Table 3.3: To source

	C	**D**	**DA**	**DO**	**DOA**	**O**	**OA**
C	-	-7.70	-5.00	-4.52	-4.82	-8.32	-5.63
D	7.70	-		3.18			
DA	5.00		-			-3.33	
DO	4.52	-3.18		-		-3.80	
DOA	4.82				-		
O	8.32		3.33	3.80		-	
OA	5.63						-

Table 3.4: Mark source

	C	D	DA	DO	DOA	O	OA
C	-			1.57	2.29		
D		-				-2.29	
DA			-			-1.93	
DO	-1.57			-		-2.45	
DOA	-2.29				-	-3.17	
O		2.29	1.93	2.45	3.17	-	
OA							-

Table 3.5: To notes

	C	D	DA	DO	DOA	O	OA
C	-	-3.78	-3.17	-3.19	-2.76	-3.32	-5.88
D	3.78	-					-2.10
DA	3.17		-				-2.72
DO	3.19			-			-2.70
DOA	2.76				-		-3.13
O	3.32					-	-2.57
OA	5.88	2.10	2.72	2.70	3.13	2.57	-

Table 3.6: To text

	C	D	DA	DO	DOA	O	OA
C	-	7.37	8.27	6.76	6.42	5.18	4.29
D	-7.37	-				-2.19	-3.08
DA	-8.27		-	-1.51	-1.85	-3.09	-3.98
DO	-6.76		1.51	-		-1.59	-2.47
DOA	-6.42		1.85		-		-2.13
O	-5.18	2.19	3.09	1.59		-	
OA	-4.29	3.08	3.98	2.47	2.13		-

Table 3.7: To assignment

	C	D	DA	DO	DOA	O	OA
C	-						
D		-	.84				
DA		-.84	-				
DO				-			
DOA					-		
O						-	
OA							-

Table 3.8: To manual

	C	D	DA	DO	DOA	O	OA
C	-	-.24				-.37	
D	.24	-			.32		
DA			-			-.33	
DO				-	.28		
DOA		-.32		-.28	-	-.45	
O	.37		.33		.45	-	.33
OA						-.33	-

Table 3.9: Word count

	C	D	DA	DO	DOA	O	OA
C	-	2.16	2.12	3.10	2.04	1.52	2.36
D	-2.16	-					
DA	-2.12		-				
DO	-3.10			-		-1.57	
DOA	-2.04				-		
O	-1.52			1.57		-	
OA	-2.36						-

Table 3.10: Stop

	C	D	DA	DO	DOA	O	OA
C	-	-.33	-.53				-.40
D	.33	-		.37			
DA	.53		-	.57	.29	.46	
DO		-.37	-.57	-			-.44
DOA			-.29		-		
O			-.46			-	-.34
OA	.40			.44		.34	-

Table 3.11: To Diagram

	D	DA	DO	DOA
D	-	-.80		
DA	.80	-	1.15	.94
DO		-1.15	-	
DOA		-.94		-

Table 3.12: Diagram activities within Diagram

	D	DA	DO	DOA
D	-			
DA		-	2.38	
DO		-2.38	-	
DOA				-

Table 3.13: Outline open

	DO	DOA	O	OA
DO	-		-2.00	
DOA		-	-1.74	
O	2.00	1.74	-	1.04
OA			-1.04	-

Table 3.14: Total no. of acts

	C	D	DA	DO	DOA	O	OA
C	-		-254.57	-339.41	-324.73		
D		-	-326.09	-410.93	-396.25		
DA	254.57	326.09	-			205.61	
DO	339.41	410.93		-		290.45	241.32
DOA	324.73	396.25			-	275.77	
O			-205.61	-290.45	-275.77	-	
OA				-241.32			-

Section 4: Differences for tool use percentages in the 1st phase

Tables 4.1 to 4.13: Mean differences of tool use percentages for the first phase between conditions (Bonferroni). Values are row label – column label. Only significant differences are shown.

Table 4.1: To chat

	C	D	DA	DO	DOA	O	OA
C	-	3.09	3.76			2.70	
D	-3.09	-					
DA	-3.76		-	-2.30			-2.66
DO	-2.70		2.30	-			
DOA					-		
O						-	
OA			2.66				-

Table 4.2: Chat

	C	D	DA	DO	DOA	O	OA
C	-	15.41	19.02	12.73	18.63	16.71	18.95
D	-15.41	-					
DA	-19.02		-				
DO	-12.73			-			
DOA	-18.63				-		
O	-16.71					-	
OA	-18.95						-

Table 4.3: To source

	C	**D**	**DA**	**DO**	**DOA**	**O**	**OA**
C	-	-10.35	-7.39	-7.75	-6.23	-11.57	-11.53
D	10.35	-					
DA	7.39		-				
DO	7.75			-			
DOA	6.23				-		
O	11.57					-	
OA	11.53						-

Table 4.4: Mark source

	C	**D**	**DA**	**DO**	**DOA**	**O**	**OA**
C	-	5.64		5.02	7.95		
D	-5.64	-				-5.64	
DA			-				
DO	-5.02			-		-5.02	
DOA	-7.95				-	-7.96	
O		5.64		5.02	7.96	-	
OA							-

Table 4.5: To notes

	C	**D**	**DA**	**DO**	**DOA**	**O**	**OA**
C	-	-3.60	-3.80	-6.01	-2.69	-4.82	-11.17
D	3.60	-		-2.41			-7.57
DA	3.80		-	-2.22			-7.37
DO	6.01	2.41	2.22	-	3.32		-5.16
DOA	2.69			-3.32	-		-8.48
O	4.82					-	-6.34
OA	11.17	7.57	7.37	5.16	8.48	6.34	-

Table 4.6: To assignment

	C	**D**	**DA**	**DO**	**DOA**	**O**	**OA**
C	-		3.18			2.61	2.42
D		-					
DA	-3.18		-	-2.12			
DO			2.12	-			
DOA					-		
O	-2.61					-	
OA	-2.42						-

Table 4.7: To manual

	C	D	DA	DO	DOA	O	OA
C	-					-1.00	
D		-					
DA			-			-.72	
DO				-			
DOA					-	-.98	
O	1.00		.72		.98	-	.98
OA						-.98	-

Table 4.8: Stop

	C	D	DA	DO	DOA	O	OA
C	-	-.34	-.47			-.34	
D	.34	-					
DA	.47		-	.36			
DO			-.36	-			
DOA					-		
O	.34					-	
OA							-

Table 4.9: Diagram open

	D	DA	DO	DOA
D	-			
DA		-	3.55	
DO		-3.55	-	-4.20
DOA			4.20	-

Table 4.10: To Diagram

	D	DA	DO	DOA
D	-	-.93	1.01	
DA	.93	-	1.95	
DO	-1.01	-1.95	-	-1.67
DOA			1.67	-

Table 4.11: Diagram activities within Diagram

	D	DA	DO	DOA
D	-		3.52	
DA		-	6.39	
DO	-3.52	-6.39	-	-5.40
DOA			5.40	-

Table 4.12: Outline open

	DO	DOA	O	OA
DO	-		-1.13	
DOA		-	-1.39	
O	1.13	1.39	-	
OA				-

Table 4.13: Total no. of acts

	C	D	DA	DO	DOA	O	OA
C	-	-133.74	-207.04	-105.71	-150.27	-126.61	
D	133.74	-					
DA	207.04		-	101.33			141.50
DO	105.71		-101.33	-			
DOA	150.27				-		
O	126.61					-	
OA			-141.50				-

Section 5: Differences for tool use percentages in the 2nd phase

Tables 5.1 to 5.14: Mean differences of tool use percentages for the second phase between conditions (Bonferroni). Values are row label – column label. Only significant differences are shown.

Table 5.1: To chat

	C	D	DA	DO	DOA	O	OA
C	-	4.19	5.35	4.10	3.60	3.33	2.30
D	-4.19	-					
DA	-5.35		-			-2.02	-3.05
DO	-4.10			-			
DOA	-3.60				-		
O	-3.33		2.02			-	
OA	-2.30		3.05				-

Table 5.2: Chat

	C	D	DA	DO	DOA	O	OA
C	-	12.06	12.69		13.48	7.31	
D	-12.06	-					
DA	-12.69		-	-6.98			
DO			6.98	-			
DOA	-13.48				-		
O	-7.31					-	
OA							-

Table 5.3: To source

	C	D	DA	DO	DOA	O	OA
C	-	-7.69		-4.41		-5.26	
D	7.69	-	4.70				
DA		-4.70	-				
DO	4.41			-			
DOA					-		
O	5.26					-	
OA							-

Table 5.4: Mark source

	C	D	DA	DO	DOA	O	OA
C	-	2.02	2.17	1.79			
D	-2.02	-					
DA	-2.17		-				
DO	-1.79			-			
DOA					-		
O						-	
OA							-

Table 5.5: To notes

	C	D	DA	DO	DOA	O	OA
C	-	-5.60	-3.65	-3.60	-4.64	-2.67	-4.82
D	5.60	-				2.93	
DA	3.65		-				
DO	3.60			-			
DOA	4.64				-		
O	2.67	-2.93				-	
OA	4.82						-

Table 5.6: To text

	C	D	DA	DO	DOA	O	OA
C	-	5.24	8.33	5.46	5.26	3.85	3.03
D	-5.24	-	3.10				
DA	-8.33	-3.10	-	-2.87	-3.08	-4.48	-5.30
DO	-5.46		2.87	-			
DOA	-5.26		3.08		-		
O	-3.85		4.48			-	
OA	-3.03		5.30				-

Table 5.7: To manual

	C	D	DA	DO	DOA	O	OA
C	-						
D		-					
DA			-	-.37			
DO			.37	-			
DOA					-		
O						-	
OA							-

Table 5.8: Word count

	C	D	DA	DO	DOA	O	OA
C	-			2.98	2.86		
D		-					
DA			-				
DO	-2.98			-			
DOA	-2.86				-		
O						-	
OA							-

Table 5.9: Stop

	C	D	DA	DO	DOA	O	OA
C	-						
D		-					
DA			-	.62			
DO			-.62	-			
DOA					-		
O						-	
OA							-

Table 5.10: Diagram open

	D	DA	DO	DOA
D	-	-2.75		
DA	2.75	-	4.11	3.09
DO		-4.11	-	
DOA		-3.09		-

Table 5.11: To Diagram

	D	DA	DO	DOA
D	-	-1.92		
DA	1.92	-	1.62	1.38
DO		-1.62	-	
DOA		-1.38		-

Table 5.12: Diagram activities within Diagram

	D	DA	DO	DOA
D	-	-5.54		
DA	5.54	-	4.91	
DO		-4.91	-	
DOA				-

Table 5.13: Outline open

	DO	DOA	O	OA
DO	-		-2.77	
DOA		-	-2.03	
O	2.77	2.03	-	
OA				-

Table 5.14: Total no. of acts

	C	D	DA	DO	DOA	O	OA
C	-			-142.92	-118.45		
D		-	-154.43	-226.41	-201.95		
DA		154.43	-				
DO	142.92	226.41		-		145.33	
DOA	118.45	201.95			-	120.86	
O				-145.33	-120.86	-	
OA							-

Section 6: Differences for tool use percentages in the 3rd phase

Tables 6.1 to 6.12: Mean differences of tool use percentages for the third phase between conditions (Bonferroni). Values are row label – column label. Only significant differences are shown.

Table 6.1: To chat

	C	D	DA	DO	DOA	O	OA
C	-	5.38	5.22	6.00	4.91	4.55	2.85
D	-5.38	-					-2.53
DA	-5.22		-				-2.37
DO	-6.00			-			-3.15
DOA	-4.91				-		
O	-4.55					-	
OA	-2.85	2.53	2.37	3.15			-

Table 6.2: Chat

	C	**D**	**DA**	**DO**	**DOA**	**O**	**OA**
C	-	14.59	9.49	10.36	8.25	10.66	
D	-14.59	-					-10.19
DA	-9.49		-				
DO	-10.36			-			
DOA	-8.25				-		
O	-10.66					-	
OA		10.19					-

Table 6.3: To source

	C	**D**	**DA**	**DO**	**DOA**	**O**	**OA**
C	-		-3.34			-6.65	-3.96
D		-				-3.99	
DA	3.34		-				
DO				-		-5.42	
DOA					-	-5.04	
O	6.65	3.99		5.42	5.04	-	
OA	3.96						-

Table 6.4: To notes

	C	**D**	**DA**	**DO**	**DOA**	**O**	**OA**
C	-	-2.58	-2.32			-2.62	-3.16
D	2.58	-					
DA	2.32		-				
DO				-			
DOA					-		
O	2.62					-	
OA	3.16						-

Table 6.5: To text

	C	**D**	**DA**	**DO**	**DOA**	**O**	**OA**
C	-	9.08	8.71	9.80	7.78	5.57	5.33
D	-9.08	-				-3.52	-3.75
DA	-8.71		-			-3.15	-3.38
DO	-9.80			-		-4.24	-4.47
DOA	-7.78				-		
O	-5.57	3.52	3.15	4.24		-	
OA	-5.33	3.75	3.38	4.47			-

Table 6.6: To assignment

	C	D	DA	DO	DOA	O	OA
C	-	-.61					-.64
D	.61	-	.60	.64	.77		
DA		-.60	-				-.62
DO		-.64		-			-.66
DOA		-.77			-		-.80
O						-	
OA	.64		.62	.66	.80		-

Table 6.7: Word count

	C	D	DA	DO	DOA	O	OA
C	-			3.61			3.02
D		-					
DA			-	2.48			
DO	-3.61		-2.48	-		-4.60	
DOA					-		
O				4.60		-	4.01
OA	-3.02					-4.01	-

Table 6.8: Stop

	C	D	DA	DO	DOA	O	OA
C	-	-.68	-.81				-.58
D	.68	-		.74		.68	
DA	.81		-	.88	.62	.82	
DO		-.74	-.88	-			-.65
DOA			-.62		-		
O		-.68	-.82			-	-.59
OA	.58			.65		.59	-

Table 6.9: To Diagram

	D	DA	DO	DOA
D	-			1.47
DA		-		
DO			-	
DOA	-1.47			-

Table 6.10: Diagram activities within Diagram

	D	DA	DO	DOA
D	-	5.32		6.87
DA	-5.32	-		
DO			-	5.26
DOA	-6.87		-5.26	-

Table 6.11: Outline open

	DO	DOA	O	OA
DO	-		-1.86	
DOA		-		
O	1.86		-	
OA				-

Table 6.12: Total no. of acts

	C	D	DA	DO	DOA	O	OA
C	-	121.77		-83.73			
D	-121.77	-	-98.36	-205.50	-177.77		-118.95
DA		98.36	-	-107.14			
DO	83.73	205.50	107.14	-		158.97	
DOA		177.77			-	131.24	
O				-158.97	-131.24	-	
OA		118.95					-

Section 7: Differences for activity duration in the 1st phase

Tables 7.1 to 7.6: Mean differences of tool use duration for the first phase between conditions (Bonferroni). Values are row label – column label. Only significant differences are shown.

Table 7.1: In chat duration

	C	D	DA	DO	DOA	O	OA
C	-						
D		-			8.64	5.66	
DA			-				
DO				-			
DOA		-8.64			-		
O		-5.66				-	
OA							-

Table 7.2: In source duration

	C	D	DA	DO	DOA	O	OA
C	-				15.57	12.39	
D		-					
DA			-				
DO				-			
DOA	-15.57				-		
O	-12.39					-	
OA							-

Table 7.3: To notes duration

	C	D	DA	DO	DOA	O	OA
C	-			-39.22			
D		-					
DA			-	-24.13			
DO	39.22		24.13	-			
DOA					-		
O						-	
OA							-

Table 7.4: In Diagram duration

	D	DA	DO	DOA
D	-		28.11	
DA		-	31.91	
DO	-28.11	-31.91	-	
DOA				-

Table 7.5: In Outline duration

	DO	DOA	O	OA
DO	-		-8.13	-14.82
DOA		-		
O	8.13		-	
OA	14.82			-

Table 7.6: Mean duration of activities

	C	D	DA	DO	DOA	O	OA
C	-				3.72		
D		-			4.53	3.14	
DA			-				
DO				-	3.77		
DOA	-3.72	-4.53		-3.77	-		-4.54
O		-3.14				-	
OA					4.54		-

Section 8: Differences for activity duration in the 2nd phase

Tables 8.1 to 8.7: Mean differences of tool use percentages for the second phase between conditions (Bonferroni). Values are row label – column label. Only significant differences are shown.

Table 8.1: In chat duration

	C	D	DA	DO	DOA	O	OA
C	-	-5.40			6.60		
D	5.40	-	8.68	7.78	12.00	9.29	7.47
DA		-8.68	-				
DO		-7.78		-			
DOA	-6.60	-12.00			-		
O		-9.29				-	
OA		-7.47					-

Table 8.2: In source duration

	C	D	DA	DO	DOA	O	OA
C	-	17.93	16.30	17.06		19.65	17.19
D	-17.93	-					
DA	-16.30		-				
DO	-17.06			-			
DOA					-		
O	-19.65					-	
OA	-17.19						-

Table 8.3: In instruction duration

	C	D	DA	DO	DOA	O	OA
C	-		23.05				
D		-					
DA	-23.05		-				
DO				-			
DOA					-		
O						-	
OA							-

Table 8.4: To assignment duration

	C	D	DA	DO	DOA	O	OA
C	-	13.90	14.23		18.09	12.90	
D	-13.90	-					
DA	-14.23		-				
DO				-			
DOA	-18.09				-		
O	-12.90					-	
OA							-

Table 8.5: In Diagram duration

	D	DA	DO	DOA
D	-	-39.75		-22.85
DA	39.75	-	26.25	
DO		-26.25	-	
DOA	22.85			-

Table 8.6: In Outline duration

	DO	DOA	O	OA
DO	-		-10.09	
DOA		-		
O	10.09		-	
OA				-

Table 8.7: Mean duration of activities

	C	D	DA	DO	DOA	O	OA
C	-		4.61	3.36	5.02	3.78	
D		-	6.24	4.99	6.65	5.41	
DA	-4.61	-6.24	-				
DO	-3.36	-4.99		-			
DOA	-5.02	-6.65			-		
O	-3.78	-5.41				-	
OA							-

Section 9: Differences for activity duration in the 3rd phase

Tables 9.1 to 9.8: Mean differences of tool use percentages for the third phase between conditions (Bonferroni). Values are row label – column label. Only significant differences are shown.

Table 9.1: In chat duration

	C	**D**	**DA**	**DO**	**DOA**	**O**	**OA**
C	-	-5.25					
D	5.25	-	4.55		7.87	4.74	
DA		-4.55	-				
DO				-	5.09		
DOA		-7.87		-5.09	-		
O		-4.74				-	
OA							-

Table 9.2: In source duration

	C	**D**	**DA**	**DO**	**DOA**	**O**	**OA**
C	-	12.92	11.18		14.56	13.99	
D	-12.92	-					
DA	-11.18		-				
DO				-			
DOA	-14.56				-		
O	-13.99					-	
OA							-

Table 9.3: To notes duration

	C	**D**	**DA**	**DO**	**DOA**	**O**	**OA**
C	-			-33.03	-18.63		-19.60
D		-		-26.66			
DA			-	-24.11			
DO	33.03	26.66	24.11	-		24.43	
DOA	18.63				-		
O				-24.43		-	
OA	19.60						-

Table 9.4: To text duration

	C	D	DA	DO	DOA	O	OA
C	-	-11.02	-10.84	-8.13			
D	11.02	-					
DA	10.84		-				
DO	8.13			-			
DOA					-		
O						-	
OA							-

Table 9.5: In instruction duration

	C	D	DA	DO	DOA	O	OA
C	-						
D		-	13.70			14.47	
DA		13.70	-				
DO				-			
DOA					-		
O		14.47				-	
OA							-

Table 9.6: To assignment duration

	C	D	DA	DO	DOA	O	OA
C	-	-12.75					
D	12.75	-	13.35		14.95	12.01	
DA		-13.35	-				
DO				-			
DOA		-14.95			-		
O		-12.01				-	
OA							-

Table 9.7: In Diagram duration

	D	DA	DO	DOA
D	-			34.38
DA		-		
DO			-	
DOA	-34.38			-

Table 9.8: Mean duration of activities

	C	D	DA	DO	DOA	O	OA
C	-						
D		-			4.70		
DA			-				
DO				-			
DOA		-4.70			-		
O						-	
OA							-

Section 10: T-test differences between the phases

Table 10.1: Mean differences of percentages between phases for the entire sample.

	1st phase – 2nd phase	**2nd phase – 3rd phase**
To chat	-1.55	.63
Chat	5.06	-6.21
To source	4.25	8.07
Mark source	4.06	1.34
To notes		2.16
To text	-7.31	-.62
To assignment	2.33	.35
To manual	.38	.15
To Diagram tips	.84	.13
To Outline tips	.81	.13
Word count	-4.04	-2.50
Stop	-.32	-.24
To Diagram		-.77
Diagram close	-.88	
Diagram activities within Diagram	-1.50	-3.66
Diagram delete link	-.08	-.26
Diagram delete object		-.45
Diagram new link	-.67	-1.60
Diagram update object		-.98
Outline open	-1.03	
To Outline	-.40	-.59
Outline close	-.87	
Total no. of acts	-110.70	

$p < .05$; only significant differences are shown.

Section 11: Correlations between activity duration and text quality

Table 11.1: Duration correlations 1st phase all conditions.

	Textual structure	Segment argumentation	Overall argumentation	Audience focus	Mean text score
In chat	.01	.07	.10	.10	.09
To chat	.01	.09	.08	.10	.09
Chat	.02	.03	.09	.08	.07
In source	.06	.11	.07	.10	.10
To source	.00	.14*	.05	.21**	.12
Mark source	.08	.04	.06	-.04	.05
To notes	-.01	.03	-.01	-.09	-.02
To text	-.06	.07	-.04	.10	.03
In instruction	.12	.07	.13*	.04	.11
To assignment	.17*	.05	.16*	.09	.15*
To manual	.01	.09	-.01	.00	.03
To Diagram tips	-.20	-.14	.05	-.25*	-.21
To Outline tips	.14	.45**	.43**	.14	.42**
In Diagram	-.03	-.06	.01	.10	.03
Diagram open	-.08	-.12	-.14	-.08	-.12
To Diagram	.02	.10	.15	.19*	.17*
Diagram delete link	.16	-.15	-.06	-.05	-.03
Diagram delete object	.02	.00	.09	.04	.03
Diagram new link	.00	-.14	-.11	.03	-.03
Diagram new object	-.12	-.02	.05	.14	.01
Diagram update object	.01	.04	.10	.15	.10
In Outline	-.15	.07	-.26**	.11	-.09
Outline open	-.08	-.13	-.28**	-.01	-.17
To Outline	-.16	.28**	-.10	.20*	.06
Mean duration per activity	-.02	.11	.08	.12*	.09

** $p < .01$; * $p < .05$.

Table 11.2: Duration correlations 2nd phase all conditions.

	Textual structure	Segment argumentation	Overall argumentation	Audience focus	Mean text score
In chat	.04	.05	.04	.10	.05
To chat	.04	.01	.03	.08	.05
Chat	.02	.06	.03	.09	.04
In source	.13*	.11	-.01	.00	.07
To source	.14*	.13*	.00	.20**	.13*
Mark source	.08	.06	-.01	-.09	.02
To notes	-.10	.04	.00	.14*	.05
To text	.11	-.02	.04	.05	.08
In instruction	-.02	.17**	.01	-.01	.04
To assignment	-.04	.14*	-.02	-.05	-.01
To manual	-.02	.17**	.05	.03	.08
To Diagram tips	.03	.04	-.04	-.10	-.03
To Outline tips	.11	-.05	-.06	.13	.08
In Diagram	-.20*	-.11	-.09	-.07	-.15
Diagram open	-.05	.01	-.05	-.01	-.04
To Diagram	-.31**	-.31**	-.18*	-.15	-.32**
Diagram delete link	.18*	-.01	.01	.02	.10
Diagram delete object	-.08	.03	.02	.24**	.08
Diagram new link	-.11	.01	-.09	-.08	-.08
Diagram new object	-.20*	-.10	-.03	-.18*	-.18*
Diagram update object	-.26**	-.23*	-.16	-.22*	-.28**
In Outline	-.18*	.07	-.15	-.08	-.13
Outline open	-.08	.06	-.09	-.02	-.06
To Outline	-.18	.06	-.13	-.09	-.13
Mean duration per activity	.10	.19**	.12	.16*	.17*

** p < .01; * p <.05.

Table 11.3: Duration correlations 3rd phase all conditions.

	Textual structure	Segment argumentation	Overall argumentation	Audience focus	Mean text score
In chat	-.10	-.01	.02	-.02	-.04
To chat	-.11	-.07	-.02	-.09	-.09
Chat	-.06	.05	.05	.05	.02
In source	-.15*	-.03	-.18**	.00	-.14*
To source	-.19**	-.10	-.19**	-.10	-.22**
Mark source	.00	.06	-.04	.13*	.06
To notes	.02	-.03	.07	.03	.04
To text	-.12*	.10	.18**	.08	.09
In instruction	.08	.05	-.01	-.01	.05
To assignment	.06	.02	.00	-.01	.04
To manual	.07	.06	-.01	-.01	.04
To Diagram tips	-.02	.00	-.19	.05	-.02
To Outline tips	-.07	.16	.38*	.27	.31*
In Diagram	-.08	.02	.01	-.11	-.06
Diagram open	-.07	.21*	.22*	.12	.14
To Diagram	-.01	.14	.18*	.00	.09
Diagram delete link	-.03	-.30**	-.23*	-.16	-.23**
Diagram delete object	-.05	.06	-.01	.08	.03
Diagram new link	.05	-.07	-.06	-.20*	-.09
Diagram new object	-.03	.06	.03	-.04	.00
Diagram update object	-.19*	-.14	-.17*	-.17*	-.21*
In Outline	.00	.07	-.04	-.19*	-.07
Outline open	-.08	.06	.05	.01	.00
To Outline	.04	.04	-.08	-.23*	-.09
Mean duration per activity	-.07	.09	.04	.02	.02

** $p < .01$; * $p < .05$.

APPENDIX 8: TWO EXAMPLES OF ARGUMENTATIVE TEXTS (IN DUTCH)

Example 1

Dyad: 746	Textual structure	6.67
Condition: Diagram	Segment argumentation	5.97
Topic: Cloning	Overall argumentation	8.33
	Audience focus	8.75
	Mean text score	7.24

Steeds vaker zijn er discussies over het voor of tegen van klonen van dieren en mensen. Er zijn veel mensen die klonen veel te ver vinden gaan, omdat ze het geknoei met de natuur vinden. Wij zijn van mening dat het klonen van dieren mogelijk moet kunnen zijn, maar hele mensen klonen vinden we toch te ver gaan. Het klonen van organen moet daarentegen volgens ons wel weer mogelijk zijn.

Als men mensen kloont om organen over te zetten heeft dat alleen maar voordelen.
Omdat weinig mensen tegenwoordig niet zoveel meer nadenken over het donorcodicil hebben de ziekenhuizen steeds minder transplantatieorganen beschikbaar. Maar als we beginnen met het klonen van mensen voor de organen zullen de ellenlange wachtlijsten voor bijv. een hart of een nier gewoon verdwijnen en zal er een stuk minder leed op de wereld zijn. Wel moeten er geen volledige mensen gekloond worden voor dit doel (geen hoofd mee klonen), omdat het een kind al in de baarmoeder een bewustzijn heeft gecreerd en op die manier krijg je dus moord.
Ook is de kans dat het donororgaan door het natuurlijke afweersysteem van de mens wordt afgestoten. Tegenwoordig moeten mensen die een orgaan ontvangen allemaal medicijnen slikken, zodat het afweersysteem het donororgaan niet zo snel afstoot. Deze medicijnen zorgen er ook voor dat het afweersysteem andere dingen minder snel aanvalt, zodat de patient sneller ziek word. Dat slikken van medicijnen hoeft niet als je het orgaan van een kloon krijgt, omdat het precies dezelfde DNA-code heeft, en daardoor is het veel gezonder voor de patient.

Ook het klonen van dieren geeft vele voordelen. Zo heb je dan de mogelijkheid om dieren te klonen die uitermate productief zijn (koeien die meer melk geven dan andere koeien, enz.). Hiermee kan de produktie van voedsel een goed eind opgekrikt worden, zodat er op een kleiner stuk land meer voedsel verbouwd kan worden (gezien in aantal vlees en melk) en helpt dit weer tegen het ruimtetekort in vele landen. Ook kunnen deze dieren worden geexporteerd naar Afrika of andere derde wereld landen om daar dan het voedsel probleem op te lossen.

Natuurlijk zijn er ook mensen die het hier allemaal niet mee eens zijn en die komen vervolgens met tegenargumenten. Een zo'n argument kan zijn dat er dictators, zoals bijvoorbeeld Hitler, en mensen die heel erg bewonderd worden, zoals Albert Einstein, gekopieerd zullen gaan worden. Maar deze mensen beseffen waarschijnlijk niet dat alleen het lichaam van deze mensen zal worden gekopieerd en niet de persoonlijkheid. Dus er hoeft niet gevreesd te worden voor een tweede Hitler. Een voorbeeld hiervan is het boek "The boys of Brazil".

Een ander tegenargument kan zijn dat een oorspronkelijk, uniek exemplaar een grotere waarde heeft voor mensen dan een kloon. Maar de kloon heeft ook zijn eigen meerwaarde. Ten eerste voor de wetenschappers, die de kloon natuurlijk willen onderzoeken en daarmee in de publiciteit willen komen. Maar er is ook een meerwaarde voor de gekloonde. Die kan de wens hebben voor een bepaalde onsterfelijkheid of kan een kinderwens hebben. Ten slotte kan een cellijn ontwikkeld worden, die kan dienen als weefsel-leverancier.

Verder brengen de mensen die tegen het klonen zijn naar voren dat de genetische diversiteit van de mensen en dieren die gekloond worden, wordt verminderd.
Als antwoord daarop hebben wij, dat bij mensen dit niet gebeurd, omdat de klonen alleen voor organen worden gebruikt en niet om andere mensen erbij te maken. Bij dieren ligt het iets anders, omdat deze wel blijven leven, moet er op gelet worden dat er wel met de niet gekloonde gefokt moet blijven worden zodat er toch een diversiteit aan genen komt tussen de populatie dieren.

Het tegenargument dat nog opgeworpen kan worden door de tegenstanders van het klonen, is dat de klonen van mensen die worden gebruikt worden voor organen- transplantaties een bewustzijn hebben die dan word vernietigd wordt (moord).
Dit argument kan heel snel weerlegd worden door het feit dat de klonen zonder hoofd gekweekt kunnen worden (dit word in bron 1 uitgelegd door het feit dat het gelukt is kikkerembryo's zonder hoofd te kweken) zodat er geen hersenen zijn en dus ook geen bewustzijn.

Dus na al deze voordelen en nadelen met elkaar te vergeleken hebben, kunnen we tot de conclusie komen dat de voordelen van het klonen sterker zijn dan de nadelen. De nadelen die vaak door mensen als tegenargument gebruikt worden, kunnen meestal gelijk weer verworpen worden. Wij zijn dus van mening dat het klonen van dieren en organen alleen maar zal bijdragen aan een verbetering van de maatschappij.

Example 2

Dyad: 1340	Textual structure	7.33
Condition: Diagram-Outline-Advisor	Segment argumentation	6.25
Topic: Organ donation	Overall argumentation	6.67
	Audience focus	6.25
	Mean text score	6.76

Orgaandonatie een zaak van levensbelang.

Orgaandonaties zijn steeds harder nodig. Er is erg veel vraag naar organen en de wachtlijsten worden steed langer. Onze medemens die deze organen hard nodig heeft moet worden geholpen en dat kan maar op één manier: er moeten meer donoren komen. Om hier iets aan te doen kunt U met een wetsvoorstel komen om orgaan- en weefseldonatie stimuleren. Dit is namelijk heel hard nodig, het hele "donatie systeem" is hard toe aan hervorming. Een goed plan, is bijvoorbeeld geld geven aan mensen die een "JA" op hun donorcodicil invullen en hiermee dus hun organen ter beschikking stellen na hun dood. Orgaandonaties zijn dus een goede zaak en moeten worden gestimuleerd.

Met een orgaan van een donateur die hij zelf niet meer nodig heeft kan een mensenleven gered worden. De organen die gedoneerd worden kunnen nog voor andere mensen worden gebruikt en moeten dus niet nutteloos begraven worden, want wat heeft een mens nou nog aan zijn of haar organen na diegene zijn dood, ze kunnen beter deze organen afstaan zodat er mensen levens mee worden gered. Maar dit moet eerst doordringen tot de mensen voordat ze daadwerkelijk toestemming geven om na hun dood organen af te staan.

Tuurlijk zullen er altijd nog gevallen blijven waarbij het slechter gaat na de transplantatie dan voor de transplantatie, zoals in het geval van dhr. S. Goossens die een nieuwe nier heeft gekregen maar na een tijdje kwamen er allemaal ongewenste bijwerkingen. De kans op mislukking of afstoting is altijd aanwezig, maar als men dit afweegt tegen de keren dat een transplantatie wel lukt dan zult U zien dat deze veel meer voorkomen. Door medische onwikkelingen is er namelijk een steeds toenemende overlevingskans dit wordt tevens bevestigd door een onderzoek van de stichting donorvoorlichting die hebben bewezen dat de kans op overleven van de mens na het eerste jaar 60-70% is bij hart/long transplantatie, 70-75% bij een dubbele long transplantatie, 75-80% bij een enkele long transplantatie, 75% bij een nier/pancreas transplantatie, 85% bij een long transplantatie, 90-92% bij een hart transplantatie en maar liefst 95-98% bij een niertransplantatie.

Wat nog een nadeel is dat een transplantatie erg veel geld kost, maar wat maakt dat nog uit als er levens mee worden gered en dan zijn de kosten nog niet eens voor de donateur maar voor degene die het orgaan krijgt.

Er worden in sommige landen zonder toestemming organen van overledenen mensen weggehaald, dit is natuurlijk zeer kwade zaak maar de nood naar organen is blijkbaar zo hoog dat ze zelfs al illegaal organen worden weggehaald en verhandeld. Volgens Ana Beatriz Magno da Silva in haar krant "Correio Brasiliense" worden er illegaal organen van Braziliaanse straatkinderen verhandeld en deze gaan meestal naar rijke Europese en Amerikaanse kinderen. In de derde wereld is er wel handel in organen maar is dit puur voor het geld , deze handel is ook gewoon legaal. Mensen geven bijvoorbeeld een niet en daar krijgen ze in verhouding een veelvoud van hun jaarsalaris voorterug, volgens dokter Rafael Matesanz.

Het hele "donatie systeem" moet worden gereorganiseerd zodat we in ieder geval de mensen die wel donor zijn ook als donor kunnen gebruiken, want volgens de stichting donorvoorlichting komt het met enige regelmaat voor dat de nabestaanden toestemming weigeren omdat zij niet op de hoogte zijn van wat de overledene gewild zou hebben. Ook bestaat de mogelijkheid dat de donatieprocedure in een ziekenhuis, om praktische redenen niet gestart kan worden of omdat de arts niet aan de mogelijkheid van donatie heeft gedacht dit is natuurlijk ook nog voor verbetering vatbaar.

Kortom orgaan- en weefseldonatie is een goede zaak en moet absoluut niet aan zijn eigen lot worden overgelaten want dan is het niet haalbaar, maar moet daarentegen wel worden gestimuleerd door allerlei subsidies en acties. Het moet verkomen worden dat organen gewoon verloren gaan doordat de implantatie niet op tijd kan worden georganiseerd. Donateuren van organen zijn nu dus harder nodig dan ooit en er moeten zeker meer reclame acties komen om deze donatie te stimuleren.

APPENDIX 9: CODING THE DIAGRAMS

Table 1 gives an example of an if-then coding filter. The second column indicates whether the line contains input or output for the if-then rule; basically, 0 means 'if', and 999 means 'then'. If a protocol line contains the text in the last column, it receives the coding of the next 'then' line in the third column. The coding is based on a categorization of all given sources on the topic and the arguments that can be found in them.

Table 1: Example of a coding filter for the Diagram analyses.

Line number	If-then	Coding (then)	Pointer (if)
1	0		[Kk]lonen*
2	0		[Gg]enen*
3	999	1A	
4	0		[Mm]edi*
5	0		[Pp]roduct*
6	0		[Oo]ntwikkel*
7	999	1B	
8	0		[Mm]edicijn*
9	0		[Mm]aken*
10	999	1B	
11	0		[Mm]edicijn*
12	0		ontwikkel*
13	999	1B	
14	0		[Vv]ooruit*
15	0		[Mm]edi*
16	999	1C	
17	0		[Mm]edisc*
18	0		[Mm]ogelijk*
19	999	1C	
20	0		[Mm]ensen*
21	0		[Rr]edden*
22	999	5H	
23	0		[Ll]even*
24	0		[Rr]edden*
25	999	3E	
26	0		[Ll]evensreddend*
27	0		[Mm]edicijnen*
28	999	3E	
29	0		[Vv]erbeter*
30	0		[Vv]oortplantingstechn*
31	999	3E	
32	0		[Kk]lonen*
33	0		[Oo]rganen*
34	999	1M	
35	0		[Dd]onororga*
36	999	1M	

APPENDIX 10: TASK ACT RESULTS

Section 1: Descriptive statistics for Task act percentages

Table 1.1: Descriptive statistics Task act percentages in all phases for the separate experimental conditions.

	D M	D SD	DA M	DA SD	DO M	DO SD	DOA M	DOA SD	O M	O SD	OA M	OA SD
Plan advisor			.00	.00			.00	.00			.04	.12
Plan turn alternation	5.19	1.94	4.19	2.37	4.29	2.90	4.51	3.43	6.02	2.83	4.24	3.26
Plan coordination	14.24	5.34	13.36	4.20	13.21	4.44	13.96	5.00	13.35	4.76	13.28	3.91
Plan Diagram	7.56	2.77	9.24	4.28	6.51	2.82	6.19	2.61				
Plan Diagram layout	.29	.57	.39	.51	.36	.63	.10	.22				
Plan experimenter	.42	.78	.09	.19	.20	.29	.15	.26	.35	.63	.31	.53
Plan external source	.81	1.30	1.72	1.34	.98	.82	.81	.97	1.10	1.01	1.62	1.88
Plan goals	1.44	.89	1.36	1.28	1.52	1.11	1.80	1.44	1.89	1.28	1.83	1.25
Plan knowledge	9.51	4.96	9.10	4.54	8.70	4.29	6.69	4.12	6.69	4.14	6.16	2.88
Plan layout	.31	.94	.37	.75	.24	.38	.97	.53	.64	.78	.38	.40
Plan notes	1.42	1.17	1.33	1.40	1.29	1.04	1.24	.58	1.70	1.74	1.52	.91
Plan Outline					3.30	1.55	3.05	1.51	5.80	2.74	3.68	2.04
Plan Outline layout					.05	.14	.00	.00	.04	.13	.00	.00
Plan revision	3.47	2.17	3.14	2.33	2.88	1.27	6.07	2.95	4.18	1.95	3.64	1.67
Plan revision Diagram	.68	1.00	1.31	1.38	.44	.56	.53	.55				
Plan revision Outline					.19	.43	.32	.40	.37	.70	.42	.47
Plan source	7.13	3.13	7.22	2.55	7.55	2.76	5.72	1.27	6.79	2.91	7.18	2.59
Plan text	18.05	4.41	17.92	4.13	17.79	4.14	18.75	4.36	21.83	3.56	19.26	5.27
Total percentage Plan	70.11	6.07	70.64	4.66	69.32	4.46	70.73	5.71	70.39	6.51	63.24	8.43
Execute advisor			.00	.00			.00	.00			.00	.00
Execute word count	1.52	1.48	1.16	1.28	1.45	.78	1.32	.86	1.99	1.22	3.13	3.15
Execute Diagram	.81	1.16	1.79	2.18	1.41	1.48	1.32	1.19				
Execute Diagram layout	.00	.00	.05	.15	.04	.14	.00	.00				
Execute external source	.15	.41	.37	.71	.14	.29	.25	.44	.34	.50	.29	.33
Execute goals	.82	1.18	1.06	1.01	1.07	.96	1.28	1.28	1.16	.90	1.17	.87
Execute knowledge	5.90	3.20	4.71	2.72	4.93	2.67	4.46	3.60	4.93	2.98	3.76	2.47
Execute notes	.10	.40	.03	.14	.00	.00	.02	.07	.00	.00	.00	.00
Execute Outline					.40	1.07	.50	.63	.56	.81	.55	1.21
Execute Outline layout					.01	.06	.00	.00	.00	.00	.00	.00
Execute revision	1.28	1.28	1.02	1.06	1.21	1.01	2.72	2.65	2.04	1.67	2.07	1.47
Execute revision Diagram	.33	.62	.29	.51	.40	.60	.26	.30				
Execute revision Outline					.07	.20	.12	.26	.17	.52	.00	.00
Execute source	2.18	1.52	1.72	1.14	2.71	1.74	2.22	1.41	2.30	1.83	1.69	1.71
Execute text	3.94	3.44	5.02	3.27	4.77	2.25	5.20	2.72	5.56	2.32	5.77	5.09
Total percentage Execute	17.03	4.72	17.23	5.00	18.61	4.81	19.66	5.63	19.05	4.34	18.43	7.00
Non task program	5.49	2.94	4.04	2.29	3.74	1.75	2.75	2.15	3.33	2.02	2.78	2.34
Non task social	6.91	3.88	8.01	4.00	8.12	4.12	6.71	3.54	6.89	4.15	15.24	7.70
Total percentage Non task	12.40	3.67	12.04	4.79	11.87	4.41	9.46	4.22	10.21	4.74	18.02	8.62

Table 1.2: Descriptive statistics Task act percentages in the 1st phase for the Control group and the Experimental conditions and mean differences on independent samples T-tests.

	Control group		Experimental group		T-test
	Mean	**SD**	**Mean**	**SD**	**Mean differences**
Plan advisor			.05	.35	
Plan turn alternation	3.14	3.68	2.97	3.02	
Plan coordination	8.42	7.45	13.32	6.13	-4.90**
Plan Diagram			5.63	6.06	
Plan Diagram layout			.19	.71	
Plan external source	.86	1.71	1.48	2.59	
Plan goals	1.90	4.27	1.13	1.91	.77*
Plan knowledge	7.00	7.71	13.56	8.53	-6.56**
Plan layout	.04	.24	.09	.57	
Plan notes	2.31	4.32	1.91	2.62	
Plan Outline			2.94	4.50	
Plan Outline layout			.00	.00	
Plan revision	.62	1.66	.76	2.22	
Plan revision Diagram			.51	1.42	
Plan revision Outline			.24	1.00	
Plan source	18.13	11.23	15.41	6.93	2.72*
Plan text	9.02	8.66	9.57	5.60	
Total percentage Plan	51.44	16.57	66.71	8.85	-15.28**
Execute word count	.44	1.20	.13	.49	.31**
Execute Diagram			1.25	2.47	
Execute Diagram layout			.05	.28	
Execute external source	1.18	2.56	.16	.66	1.02**
Execute goals	1.35	3.06	.12	.62	1.23**
Execute knowledge	8.57	9.51	8.62	7.08	
Execute notes	.11	.49	.01	.13	.09*
Execute Outline			.30	1.05	
Execute Outline layout			.00	.00	
Execute revision	.55	1.42	.15	.58	.41**
Execute revision Diagram			.14	.57	
Execute revision Outline			.00	.00	
Execute source	7.03	7.58	3.69	4.27	3.35**
Execute text	3.80	5.85	1.35	2.37	2.46**
Total percentage Execute	23.04	14.27	15.45	9.31	7.59**
Non task program	2.88	3.24	5.29	4.41	-2.41**
Non task social	22.36	20.54	12.15	9.34	10.21**
Total percentage Non task	25.53	19.77	17.84	10.86	7.69**

** $p < .01$, * $p < .05$. Only significant differences are shown.

Table 1.3: Descriptive statistics Task act percentages in the 2nd phase for the Control group and the Experimental conditions and mean differences on independent samples T-tests.

	Control group		Experimental group		T-test
	Mean	SD	Mean	SD	Mean differences
Plan advisor			.00	.00	
Plan turn alternation	7.02	4.57	5.60	4.84	1.42*
Plan coordination	9.09	5.01	13.75	6.35	-4.66**
Plan Diagram			5.82	6.72	
Plan Diagram layout			.26	.83	
Plan external source	.80	1.93	1.17	1.61	
Plan goals	2.23	2.10	1.53	1.83	.70**
Plan knowledge	2.40	2.39	7.09	5.47	-4.69**
Plan layout	1.62	1.83	.35	.76	1.27**
Plan notes	2.11	2.51	1.84	2.45	
Plan Outline			2.85	3.34	
Plan Outline layout			.07	.33	
Plan revision	2.46	2.15	3.69	3.54	-1.23**
Plan revision Diagram			.53	1.22	
Plan revision Outline			.08	.30	
Plan source	6.06	3.78	6.71	5.15	
Plan text	14.75	6.13	22.07	7.34	-7.32**
Total percentage Plan	48.53	9.19	70.38	9.75	-21.85**
Execute advisor			.00	.00	
Execute word count	3.35	2.97	1.88	2.77	1.47**
Execute Diagram			.68	1.73	
Execute Diagram layout			.02	.13	
Execute external source	1.03	1.87	.25	.85	.77**
Execute goals	1.52	1.67	.37	1.04	1.15**
Execute knowledge	5.27	5.10	4.58	5.08	
Execute notes	.73	1.26	.00	.00	.73**
Execute Outline			.29	.94	
Execute Outline layout			.00	.00	
Execute revision	6.93	5.65	1.83	2.87	5.10**
Execute revision Diagram			.06	.33	
Execute revision Outline			.09	.46	
Execute source	6.31	4.77	2.27	2.47	4.04**
Execute text	10.18	5.89	5.97	4.91	4.20**
Total percentage Execute	35.32	12.04	17.93	8.85	17.38**
Non task program	3.18	2.45	3.81	3.58	
Non task social	11.96	8.05	7.63	6.06	4.33**
Total percentage Non task	16.15	8.81	11.68	7.48	4.47**

** $p < .01$, * $p < .05$. Only significant differences are shown.

Table 1.4: Descriptive statistics Task act percentages in the 3rd phase for the Control group and the Experimental conditions and mean differences on independent samples T-tests.

	Control group		Experimental group		T-test
	Mean	SD	Mean	SD	Mean differences
Plan advisor			.00	.00	
Plan turn alternation	6.69	4.07	5.35	4.23	1.34*
Plan coordination	9.96	4.01	13.57	6.72	-3.61**
Plan Diagram			8.91	5.52	
Plan Diagram layout			.48	1.09	
Plan external source	.40	.74	1.04	1.57	-.64**
Plan goals	1.54	1.35	2.01	2.21	
Plan knowledge	1.04	1.56	5.08	4.55	-4.05**
Plan layout	2.49	2.85	.64	1.15	1.85**
Plan notes	.81	1.72	.76	1.18	
Plan Outline			5.28	4.23	
Plan Outline layout			.02	.13	
Plan revision	5.23	3.09	5.43	4.03	
Plan revision Diagram			1.17	1.92	
Plan revision Outline			.44	.98	
Plan source	1.72	2.27	2.46	2.70	-.74*
Plan text	12.48	4.94	22.54	6.82	-10.07**
Total percentage Plan	42.37	7.90	69.96	7.75	-27.59**
Execute advisor			.00	.00	
Execute word count	5.35	4.10	2.44	2.53	2.91**
Execute Diagram			1.57	2.74	
Execute Diagram layout			.02	.18	
Execute external source	.75	1.13	.29	.66	.46**
Execute goals	4.26	2.86	2.40	2.54	1.86**
Execute knowledge	2.87	2.28	2.54	2.94	
Execute notes	.21	.61	.04	.32	.17**
Execute Outline			.62	1.90	
Execute Outline layout			.01	.08	
Execute revision	14.96	7.96	2.36	2.81	12.60**
Execute revision Diagram			.60	1.18	
Execute revision Outline			.17	.66	
Execute source	1.68	1.95	1.02	1.67	.66**
Execute text	11.15	5.44	6.51	6.41	4.64**
Total percentage Execute	41.22	9.71	19.65	7.61	21.57**
Non task program	2.98	2.59	2.97	2.85	
Non task social	11.84	6.88	7.25	5.58	4.59**
Total percentage Non task	16.55	8.37	10.38	6.30	6.18**

** $p < .01$, * $p < .05$. Only significant differences are shown.

Section 2: Descriptive statistics of Task act frequencies

Table 2.1: Task acts frequencies in all phases for the Control group, the Experimental conditions, and the Total sample.

	Control group		Experimental condition		T-test
	Mean	SD	Mean	SD	Mean differences
Plan alternate turn	46.02	42.45	24.82	17.12	21.20**
Plan coordination	69.32	48.69	72.60	40.51	-3.28
Plan external source	4.09	5.30	6.99	9.25	-2.90**
Plan goals	12.56	9.52	8.65	8.21	3.91**
Plan knowledge	12.44	12.82	2.43	4.06	10.01**
Plan layout	12.44	12.82	2.11	3.53	10.32**
Plan notes	10.67	11.72	7.18	6.33	3.49**
Plan revision	21.62	13.76	20.52	15.97	1.10
Plan source	40.71	32.91	37.16	18.07	3.55
Plan text	87.23	46.59	101.79	50.27	-14.57*
Total frequency Plan	321.57	180.71	324.09	134.68	-2.52
Execute word count	27.71	25.93	8.63	7.64	19.08**
Execute external source	6.51	7.34	1.48	2.80	5.02**
Execute goals	17.46	13.55	5.34	4.94	12.12**
Execute knowledge	33.98	27.89	25.13	16.70	8.84**
Execute notes	3.13	4.70	.07	.41	3.06**
Execute revision	68.23	49.16	9.05	10.66	59.19**
Execute source	35.11	31.14	11.57	9.19	23.54**
Execute text	64.44	36.49	28.41	22.97	36.03**
Total frequency Execute	256.56	131.27	89.68	44.32	166.88**
Non task program	25.09	25.99	20.69	15.58	4.40
Non task social	98.57	95.52	45.78	33.80	52.79**
Total frequency Non task	123.66	117.46	66.47	41.18	57.19**

N Control group = 39 dyads; N Experimental conditions = 106 dyads; N Total = 145 dyads. ** $p < .01$, * $p < .05$. Only significant differences are shown.

Section 3: Correlations of Task act percentages with text quality

Table 3.1: Correlations between Task act percentages in the 1st phase and text quality for the Control group and the Experimental conditions.

	Textual structure		Segment argumentation		Overall argumentation		Audience focus		Mean text core	
	C	**E**	**C**	**E**	**C**	**E**	**C**	**E**	**C**	**E**
Plan advisor		.04		.16		-.09		.25*		.13
Plan turn alternation	-.11	.22**	-.23*	.12	-.18	.22**	-.01	.23**	-.16	.28**
Plan coordination	-.26*	-.06	-.27*	.04	-.20	.02	-.04	-.18*	-.21	-.08
Plan Diagram		-.04		-.01		.08		.11		.04
Plan Diagram layout		.12		-.10		.07		.04		.06
Plan external source	.14	.05	-.09	.02	-.06	.13	-.08	.12	-.05	.13
Plan goals	.05	.00	.06	-.04	.20	.03	.32**	-.04	.22	-.02
Plan knowledge	.24*	.14*	.34**	.09	.20	.19**	.06	.13	.23*	.19**
Plan layout	.18	.15*	-.05	-.04	-.05	-.03	.10	.05	.04	.05
Plan notes	.00	-.03	-.18	-.07	-.04	-.18**	-.10	-.15*	-.10	-.15*
Plan Outline		-.21*		.15		-.16		.09		-.05
Plan Outline layout										
Plan revision	-.17	-.12	-.08	-.18**	.05	-.10	.27*	-.16*	.07	-.20**
Plan revision Diagram		.20*		.10		.13		.08		.17*
Plan revision Outline		-.17		-.02		-.14		-.06		-.15
Plan source	-.16	-.04	-.11	-.11	-.08	.01	-.05	-.10	-.11	-.07
Plan text	.15	.02	.32**	-.14*	.13	-.23**	.35**	-.17*	.29**	-.18**
Total percentage Plan	-.05	.08	.03	-.09	.02	.12	.24*	-.08	.09	.02
Execute advisor										
Execute word count	-.11	.10	-.07	-.10	-.09	-.03	.11	.14*	-.03	.05
Execute Diagram		-.18*		-.02		-.05		.07		-.07
Execute Diagram layout		-.08		-.03		.08		.14		.04
Execute external source	.25*	.08	.23**	.03	-.02	-.03	.14	.06	.15	.05
Execute goals	.05	.13	.05	-.08	.25*	-.05	.11	.08	.16	.04
Execute knowledge	.16	-.02	.12	.05	.12	.07	.22	.16*	.19	.09
Execute notes	-.02	.02	.16	-.16*	.19	-.14*	.10	.01	.15	-.08
Execute Outline		-.21*		-.03		-.14		.05		-.12
Execute Outline layout										
Execute revision	-.36**	-.01	-.14	.01	-.18	-.04	.00	-.01	-.18	-.02
Execute revision Diagram		.08		.09		.04		.06		.08
Execute revision Outline										
Execute source	.09	-.07	.08	.07	.19	-.14*	.06	.08	.14	-.01
Execute text	.26*	.15*	.29*	.07	.20	.07	.33**	.15*	.32**	.17*
Total percentage Execute	.27*	-.04	.28*	.07	.30**	-.01	.37**	.23**	.38**	.09
Non task program	-.19	-.05	-.18	-.04	-.12	-.08	-.10	-.16*	-.17	-.13
Non task social	-.12	-.02	-.19	.04	-.20	-.07	-.43**	-.09	-.31**	-.06
Total percentage Non task	-.15	-.03	-.23*	.02	-.23*	-.09	-.47**	-.13	-.35**	-.10

** $p < .01$, * $p < .05$.

Table 3.2: Correlations between Task act percentages in the 2nd phase and text quality for the Control group and the Experimental conditions.

	Textual structure		Segment argumentation		Overall argumentation		Audience focus		Mean text core	
	C	**E**	**C**	**E**	**C**	**E**	**C**	**E**	**C**	**E**
Plan advisor										
Plan turn alternation	-.03	.05	-.05	-.04	.01	.15*	.05	.09	.00	.09
Plan coordination	-.11	-.08	-.18	-.12	-.15	-.01	-.13	-.04	-.18	-.09
Plan Diagram		-.13		-.11		.03		.11		-.01
Plan Diagram layout		-.11		-.01		.09		-.01		-.02
Plan external source	-.25*	-.12	-.04	-.17*	-.27*	-.09	-.01	-.03	-.17	-.14*
Plan goals	.08	.08	.19	-.09	.18	-.07	.18	.02	.20	-.02
Plan knowledge	-.12	.05	-.08	.06	.04	.09	-.18	-.09	-.10	.04
Plan layout	-.07	.11	.14	-.10	-.01	-.25**	-.03	-.14*	.01	-.12
Plan notes	.02	.04	-.20	-.03	.01	.09	.03	-.15*	-.03	-.02
Plan Outline		-.10		-.04		.01		.07		-.02
Plan Outline layout		-.28**		.01		-.08		.08		-.10
Plan revision	-.01	.01	.30**	.15*	.18	.15*	.26*	.00	.24*	.08
Plan revision Diagram		.00		-.05		.10		.11		.08
Plan revision Outline		-.09		-.08		-.07		-.15		-.14
Plan source	-.13	-.12	-.07	-.04	-.02	-.05	-.09	-.13	-.08	-.12
Plan text	.01	.00	.25*	.07	.08	-.07	.19	-.02	.16	-.04
Total percentage Plan	-.20	-.13	.07	-.15*	.00	.11	.10	-.10	.01	-.10
Execute advisor										
Execute word count	.01	.06	.09	.23**	.07	-.08	.12	.10	.10	.10
Execute Diagram		-.07		.01		.06		.06		.04
Execute Diagram layout		-.06		-.07		-.16*		-.07		-.12
Execute external source	.05	-.11	-.12	-.08	-.11	-.03	-.05	.03	-.08	-.06
Execute goals	.24*	.14*	-.05	.03	.17	.10	.22	.03	.19	.11
Execute knowledge	.10	.23**	-.02	.11	.12	.14*	.32**	.06	.18	.20**
Execute notes	-.12		-.18		-.08		-.05		-.12	
Execute Outline		-.13		-.02		-.02		-.11		-.11
Execute Outline layout										
Execute revision	.17	-.05	.01	-.08	-.01	-.16*	-.08	-.03	.01	-.13
Execute revision Diagram		.03		.02		.17*		.26**		.19*
Execute revision Outline		.12		.12		.08		-.07		.08
Execute source	.04	.02	-.01	.07	-.02	-.09	-.06	-.01	-.02	.00
Execute text	.27*	-.10	.06	-.03	.06	-.12	.31**	.13	.21	-.04
Total percentage Execute	.30**	.07	-.00	.10	.09	-.06	.28*	.13	.20	.09
Non task program	-.14	.14*	.01	.00	-.04	.03	-.15	-.07	-.10	.05
Non task social	-.21	.02	-.13	.11	-.14	-.10	-.50**	.03	-.31**	.02
Total percentage Non task	-.21	.08	-.07	.08	-.12	-.08	-.48**	-.02	-.28*	.03

** $p < .01$, * $p < .05$.

Table 3.3: Correlations between Task act percentages in the 3rd phase and text quality for the Control group and the Experimental conditions.

	Textual structure		Segment argumentation		Overall argumentation		Audience focus		Mean text core	
	C	**E**	**C**	**E**	**C**	**E**	**C**	**E**	**C**	**E**
Plan advisor										
Plan turn alternation	-.15	.15*	-.09	-.05	-.03	.10	.04	.09	-.05	.12
Plan coordination	.00	-.02	-.09	.02	-.01	.10	-.09	.05	-.06	.06
Plan Diagram		.05		.16		.09		-.05		.09
Plan Diagram layout		-.03		.09		.04		.05		.04
Plan external source	-.17	-.17*	.33**	-.04	.12	-.14*	.27**	.04	.19	-.11
Plan goals	.12	.03	-.01	.00	-.05	-.05	-.06	.10	-.02	.04
Plan knowledge	-.21	.19**	.03	.09	-.15	.11	-.14	-.01	-.15	.13
Plan layout	.06	.06	-.13	-.09	-.11	-.01	-.08	-.13	-.09	-.06
Plan notes	-.12	-.02	.09	-.12	-.05	.07	.12	.07	.02	.01
Plan Outline		-.02		.06		.04		-.04		.01
Plan Outline layout		-.06		-.03		.08		.09		.03
Plan revision	-.13	.00	.14	-.13	-.11	-.16*	.06	-.09	-.01	-.14*
Plan revision Diagram		-.07		-.01		-.09		-.11		-.06
Plan revision Outline		.01		.18*		-.06		.13		.09
Plan source	-.29*	.00	-.05	.09	-.15	.02	-.18	.06	-.20	.05
Plan text	-.04	-.07	.11	-.10	-.13	-.02	.00	.00	-.03	-.06
Total percentage Plan	-.28*	.14*	.02	-.08	-.26*	.15*	-.07	.02	-.18	.10
Execute advisor										
Execute word count	.00	-.07	-.25*	.06	-.07	-.02	-.24*	.03	-.18	-.02
Execute Diagram		-.16		.04		-.04		-.06		-.08
Execute Diagram layout		-.06		.08		.11		.29**		.14
Execute external source	-.17	.13	.14	.19**	.04	.02	.06	.03	.03	.12
Execute goals	-.13	.10	.07	.04	-.05	.06	.44**	-.02	.13	.06
Execute knowledge	.04	.15*	.38**	.07	.26*	.11	.55**	.02	.41**	.12
Execute notes	-.13	-.19**	-.09	-.07	.05	-.08	-.02	-.11	-.03	-.16*
Execute Outline		-.24**		.05		.02		.11		-.01
Execute Outline layout		-.09		-.02		-.06		-.08		-.09
Execute revision	.27*	.06	.22	-.13	.20	-.15*	.23*	-.05	.27*	-.10
Execute revision Diagram		.02		.02		-.09		-.10		-.05
Execute revision Outline		.03		.16		.04		.03		.07
Execute source	-.20	.11	-.17	.16*	-.17	.16*	.06	.08	-.13	.17*
Execute text	.07	-.19**	.01	-.05	.13	-.18**	.12	-.05	.11	-.17*
Total percentage Execute	.16	-.11	.17	.03	.23*	-.12	.43**	-.04	.32**	-.11
Non task program	.09	.05	-.13	-.03	-.05	.08	-.23*	-.07	-.12	.02
Non task social	-.01	-.07	-.20	.08	-.04	-.10	-.42**	.09	-.22*	.00
Total percentage Non task	.07	-.03	-.20	.06	-.01	-.04	-.41**	.04	-.19	.01

** $p < .01$, * $p < .05$.

APPENDIX 11: DIALOGUE ACT RESULTS

Table 1.1: Standard deviations of Dialogue act percentages.

	Total SD	C SD	D SD	DA SD	DO SD	DOA SD	O SD	OA SD
Argumentatives								
Conclusion	.92	.83	1.12	.76	.85	.67	1.17	.96
Conditional	.73	.75	.89	.67	.68	.59	.83	.47
Contra	1.47	1.37	1.33	1.76	1.17	1.48	1.31	1.44
Disjunctive	.54	.53	.56	.65	.57	.39	.43	.37
Reason	.94	1.03	1.20	.79	.82	1.00	.74	.95
Then	.56	.59	.80	.56	.36	.55	.34	.74
Elicitatives								
Proposal Action	1.96	2.32	1.84	2.03	1.41	1.42	2.15	1.42
Question	3.64	4.69	3.65	2.91	3.29	1.40	3.15	3.87
Question Open	1.66	1.61	2.07	1.33	1.43	1.74	2.01	1.14
Question Set	.51	.49	.60	.30	.45	.30	.75	.38
Question Verify	2.91	3.58	2.49	2.67	2.68	1.18	2.55	3.33
Imperatives								
Action	2.38	2.24	2.03	1.41	1.71	2.28	3.84	1.85
Focus	1.86	1.59	1.27	1.54	1.42	3.20	2.22	1.70
Informatives								
Evaluation	1.62	1.76	1.32	1.43	1.28	1.21	1.78	1.66
Evaluation Negative	.57	.70	.38	.46	.36	.34	.48	.59
Evaluation Neutral	.35	.35	.39	.26	.46	.14	.26	.39
Evaluation Positive	1.41	1.68	1.02	1.35	1.10	1.17	1.44	1.26
Performative	.77	.67	.59	.88	.61	.69	1.04	.81
Statement	4.90	5.17	4.19	5.58	5.02	3.09	4.73	3.84
Statement Action	1.77	1.74	1.98	1.41	1.76	1.78	1.84	1.92
Statement Nonsense	2.08	3.85	.31	.46	.41	.22	.34	.26
Statement Social	1.16	1.39	.89	.74	1.35	.71	.87	1.19
Task	5.31	5.62	4.23	5.46	5.24	2.79	5.98	4.57
Responsives								
Acceptation	1.11	1.33	1.11	1.10	.92	.68	1.00	.73
Confirmation	4.22	3.89	3.30	5.09	2.59	4.69	4.47	3.90
Deny	.85	.89	.82	.79	.74	.98	.56	.89
Reply	2.27	2.00	2.27	2.22	1.89	1.78	1.76	1.58
Reply Accept	.20	.20	.27	.17	.18	.17	.11	.22
Reply Confirm	1.46	1.05	1.91	1.62	1.17	1.46	1.13	.77
Reply Deny	.44	.37	.46	.41	.53	.38	.47	.39
Reply Performative	.11	.12	.18	.11	.07	.04	.08	.09
Reply Statement	1.30	1.26	1.51	1.31	1.20	1.08	1.27	.76

APPENDIX 12: STUDENT EVALUATION RESULTS

Bonferroni mean differences. Only significant differences are shown ($p < .05$).

Table 1: Difficulty of writing assignment.

	C	D	DA	DO	DOA	O	OA
C	-	-.42	-.37				
D	.42	-					
DA	.37		-				
DO				-			
DOA					-		
O						-	
OA							-

Table 2: Computer supported writing.

	C	D	DA	DO	DOA	O	OA
C	-	.94	.77		.68		
D	-.94	-		-.91		-.72	
DA	-.77		-	-.75		-.56	
DO		.91	.75	-	.66		
DOA	-.68			-.66	-		
O		.72	.56			-	
OA							-

Table 3: Logging on and off.

	C	D	DA	DO	DOA	O	OA
C	-	1.06					
D	-1.06	-	-.95	-.99	-1.15	-1.35	-1.41
DA		.95	-				
DO		.99		-			
DOA		1.15			-		
O		1.35				-	
OA		1.41					-

Table 4: Information window.

	C	D	DA	DO	DOA	O	OA
C	-	.56					
D	-.56	-		-.59			
DA			-				
DO		.59		-			
DOA					-		
O						-	
OA							-

Table 5: Chat window.

	C	**D**	**DA**	**DO**	**DOA**	**O**	**OA**
C	-	.42					
D	-.42	-		-.49		-.42	
DA			-				
DO		.49		-			
DOA					-		
O		.42				-	
OA							-

Table 6: Shared text window.

	C	**D**	**DA**	**DO**	**DOA**	**O**	**OA**
C	-						
D		-			-.62		
DA			-				
DO				-			
DOA		.62			-		
O						-	
OA							-

Table 7: Traffic light.

	C	**D**	**DA**	**DO**	**DOA**	**O**	**OA**
C	-	.52					
D	-.52	-		-.63	-.79	-.66	
DA			-				
DO		.63		-			
DOA		.79			-		
O		.66				-	
OA							-

Table 8: Clarity of buttons.

	C	**D**	**DA**	**DO**	**DOA**	**O**	**OA**
C	-	-1.11	-1.00	-1.09	-1.14	-1.23	-1.23
D	1.11	-					
DA	1.00		-				
DO	1.09			-			
DOA	1.14				-		
O	1.23					-	
OA	1.23						-

Table 9: Use of buttons.

	C	**D**	**DA**	**DO**	**DOA**	**O**	**OA**
C	-	-1.17	-.99	-1.03	-1.08	-1.11	-1.17
D	1.17	-					
DA	.99		-				
DO	1.03			-			
DOA	1.08				-		
O	1.11					-	
OA	1.17						-

Table 10: Collaboration.

	C	**D**	**DA**	**DO**	**DOA**	**O**	**OA**
C	-	.38					
D	-.38	-					-.50
DA			-				
DO				-			
DOA					-		
O						-	
OA		.50					-

Table 11: Turn-taking in writing.

	C	**D**	**DA**	**DO**	**DOA**	**O**	**OA**
C	-	.57					
D	-.57	-					
DA			-				
DO				-			
DOA					-		
O						-	
OA							-

Table 12: Diagram window.

	D	**DA**	**DO**	**DOA**
D	-		-.47	
DA		-		
DO	.47		-	
DOA				-

Table 13: Turn-taking Diagram.

	D	**DA**	**DO**	**DOA**
D	-	-.82	-.73	-.69
DA	.82	-		
DO	.73		-	
DOA	.69			-