Using Visualizations to Support Collaboration and Coordination during Computer-Supported Collaborative Learning

Jeroen Janssen

Leden beoordelingscommissie:
Prof. dr. P. R. J. Simons
Prof. dr. W. Herrlitz
Prof. dr. T. de Jong
Prof. dr. M. Volman
Dr. C. Hmelo-Silver

NWO
Netherlands Organisation for Scientific Research
The research reported in this dissertation was supported by the Netherlands Organisation for Scientific Research (project no. 411-02-121).

ico
This research was carried out in the context of the Dutch Interuniversity Centre for Educational Research.

© 2008, Jeroen Janssen
ISBN-978-90-393-4761-4

Printed by: Print Partners Ipskamp

Cover design: Gabry Karsdorp en Jeroen Janssen

Using visualizations to support collaboration and coordination during computer-supported collaborative learning

Het gebruik van visualisaties ter ondersteuning van samenwerking en coördinatie tijdens computer-ondersteund samenwerkend leren

(met een samenvatting in het Nederlands)

Proefschrift

ter verkrijging van de graad van doctor aan de Universiteit Utrecht op
gezag van de rector magnificus, prof. dr. J.C. Stoof,
ingevolge het besluit van het college voor promoties
in het openbaar te verdedigen op

vrijdag 14 maart des middags te 2.30 uur

door

Jeroen Jacobus Hubertus Maria Janssen
geboren op 7 juni 1977 te Venlo

Promotoren:	Prof. dr. P. A. Kirschner
	Prof. dr. G. Kanselaar
Co-promotor:	Dr. G. Erkens

Dit proefschrift werd (mede) mogelijk gemaakt met financiële steun van NWO.

Voorwoord

"*I used to get mad at my school, the teachers that taught me weren't cool*" – Lennon & McCartney

Ooit begon ik aan de studie Onderwijskunde met het idee dat ons onderwijs soms zoveel beter, mooier, leuker en ook cooler kan zijn. De laatste vier jaar heb ik kunnen werken aan een onderzoek waarin we op onze eigen manier hebben geprobeerd hieraan een klein steentje bij te dragen. Of dat gelukt is? Dat mag de lezer zelf beoordelen.

Natuurlijk wil ik verschillende mensen bedanken. Zonder deze mensen zou dit onderzoek niet mogelijk zijn geweest. Dat geldt in de eerste plaats voor de docenten en leerlingen van het Bonifatius College in Utrecht en het Rijnlands Lyceum in Wassenaar die ons iedere keer met veel enthousiasme binnen lieten in hun lessen. Bedankt Cilla, Dirk, Frans, Hans, Helma, Mehmet, Paul en Rik!

Drie mensen zijn altijd heel nauw betrokken geweest bij dit project. Hierdoor heb ik ook nooit het gevoel gehad dat ik het allemaal alleen heb gedaan. Mede dankzij Gijsbert, Jos en Marcel heb ik deze klus kunnen klaren. Wat jammer dat de wekelijkse CRoCiCL-overleggen voorbij zijn! Marcel, bedankt voor al je programmeerkunsten! Doordat je zo zorgvuldig en precies was, liepen we tijdens de dataverzameling nooit tegen *fatal errors* aan. Je was bovendien een geweldige kamergenoot met een interessante voorkeur voor muziek! Gelukkig waart je geest hier nog rond in de vorm van je iTunes collectie en de OverlegAgenda Wiki (het actiepuntje notulen hebben we maar verwijderd). En Jos, je hebt je baby VCRI geleidelijk aan uit handen moeten geven aan Marcel, maar nu heb je hem als volgroeid volwassene weer terug gekregen. In de toekomst zullen we zeker nog samenwerken aan nieuwe ontwikkelingen. Bedankt ook voor al de tripjes naar de scholen om servers en dergelijke te installeren. Gijsbert, ik heb al vaker gezegd dat ik vond dat ik het getroffen had met mijn dagelijks begeleider en hier zeg ik het nog een laatste keer. Betrokken, eerlijk, gezellig, grappig, overtuigend, slim, vasthoudend, vindingrijk en vriendelijk zijn allemaal woorden die op jou van toepassing zijn en ze geven aan waarom je zo'n goede begeleider was. Gelukkig blijven we in de toekomst samenwerken en komen al deze eigenschappen weer goed van pas!

Natuurlijk moet ik ook Gellof en Paul, mijn promotoren, bedanken voor hun inzet, hulp en advies. Jullie waren op de achtergrond aanwezig maar lazen geduldig alle stukken en voorzagen die van hun commentaar. Van jullie beiden heb ik veel geleerd. Gellof, bedankt voor je vertrouwen in mij en ons team. Paul, je nam het roer over in een roerige tijd. Ondertussen is de rust wel weer weergekeerd en wordt er hard aan de toekomst van ons onderzoekscentrum gewerkt. Ik zie er naar uit om hieraan mee te kunnen werken.

Ook zonder leuke collega's red je het niet. Daarom verdienen ook zij een bedankje. Allereerst natuurlijk collega's en oud-collega's van de 'wandelploeg':

Voorwoord

Agaath, Bert, Chiel, Chris, Crina, Dortie, Harmen, Hendrien, Jannet, Jantine, Joris, Liesbeth, Lisette, Maaike, Maarten, Marije, Patrick, Sandy, Tim, Wouter en Yvette. En verder natuurlijk alle andere collega's van de afdeling Onderwijskunde.

Janny en Dick, bedankt voor het de gezelligheid en warmte jullie mij en Gabry de afgelopen jaren gegeven hebben. Pap, mam en Peter, jullie hebben ook meegekeken en meegeleefd bij het tot stand komen van dit boekje. Het is af! Bedankt voor alles!

En Gabry, van alle mensen die dit boek ooit zullen lezen weet ik dat jij het meest nieuwsgierig bent naar de inhoud ervan en dan vooral naar de tekst op deze twee pagina's. Voor jou zijn de laatste woorden die ik schrijf aan dit proefschrift. Al bijna 7 jaar zijn we om-de-beurt kapitein op ons eigen schip. Bedankt voor je vertrouwen, je liefde, de schop onder mijn kont wanneer nodig en alle mooie en leuke momenten. Je maakt nog steeds het verschil!

Jeroen Janssen
Utercht, januari 2008

Contents

[1] This chapter contains an English summary.

1. General Introduction

The Computer-Supported Collaborative Learning (CSCL) research community is a thriving one with its own journal (International Journal of CSCL), a conference every two years (CSCL conference), and numerous journal articles and theses written and published each year by scientists from a broad range of disciplines (e.g., psychology, computer science, ethnography, educational sciences, human-computer interaction). In spite of all the research that has been carried out and all the CSCL environments that have been developed over the last 20 or so years, educators and researchers still encounter problems when they let students collaborate using computer technologies. This thesis is about some of these problems and how they might be addressed. This chapter serves as an introduction to the studies reported on in this thesis. First, we describe several commonly observed problems during CSCL. Next, we explain how visualizations aid in addressing these problems. We conclude by describing the main research question of this thesis and by giving an overview of the chapters of this thesis.

1.1 CSCL: What Can and Does Go Wrong…

Although CSCL has been identified as a promising educational approach, the research on the effectiveness of CSCL and the processes that take place during CSCL demonstrate that the collaboration in these environments is not always free of problems. A great deal of these problems are not unique to CSCL; they have also been documented in 'traditional' face-to-face (FTF) collaborative learning (cf., O'Donnell & O'Kelly, 1994; Salomon & Globerson, 1989). These include conflicts between group members (e.g., Hobman, Bordia, Irmer, & Chang, 2002), free riding behavior and unequal participation (e.g., Fjermestad, 2004), and discussions between group members that lack depth and high-quality reasoning and argumentation (Munneke, Andriessen, Kanselaar, & Kirschner, 2007). On the other hand, in CSCL environments some of the problems may be less severe. For example, it has been shown that in CSCL environments communicative contributions are more equally distributed among group members compared to FTF collaboration (e.g., Dubrovsky, Kiesler, & Sethna, 1991; Fjermestad, 2004; Straus, 1997; Warschauer, 1996). Some problems however, may be unique to CSCL environments or may be exacerbated in these environments, such as difficulties coordinating one's actions with other group members' actions (e.g., Baker, Greenberg, & Gutwin, 2001; Ellis, Gibbs, & Rein, 1992; Erkens, Jaspers, Prangsma, & Kanselaar, 2005) and confusion while communicating electronically at a distance (e.g., Fuks, Pimentel, & Lucena, 2006).

The remainder of this section describes several problems that may arise during CSCL. This description is not meant to be exhaustive. Rather, it is meant to describe problems that we think can be addressed by implementing visualization tools in CSCL environments. This is elaborated upon in the next section. Furthermore, the problems described below are often not unique to CSCL; they are inherent to the collaborative process itself.

1.1.1 Lack of Awareness

Rafael: What do you think of the Debate now?
Casey: Fine.
Rafael: And what did you do Casey?
Rafael: Nothing was added to the Debate…?
Rafael: What are you doing now in the Debate, Case? Don't put anything in there cos you'll mess up the order.
Casey: Oops. Too late!
Casey: I've added source 3 to propaganda as a new argument. By accident!!
Rafael: I see. Grrrrr :-|
Rafael: Do you know what Lara is doing?
Casey: She's supposed to be working on the Martyrs position…
Rafael: I get the feeling she's letting us do most of the work.
Casey: So do I!

The fragment above comes from two male secondary education students working in a CSCL environment. They are members of a 3-person group, and they are currently working on a representation of a historical debate in an argumentative diagram. This fragment highlights one of the most common problems that group members encounter during online collaboration. Rafael is uncertain about what Casey is doing in the Debate-tool. He warns Casey about messing up the order in the tool, but he is too late. Additionally, they wonder about their other group member, Lara. They have no idea what she's doing, if she is even online, and whether she is doing what she is supposed to be doing. On top of that, they are afraid she is letting them do the lion's share of the work, but they do not know that for sure. In other words, Rafael and Casey lack *awareness* information (Dourish & Bellotti, 1992).

The issue of awareness has received considerable attention in the area of computer-supported cooperative work (CSCW, Schmidt, 2002). This has led to a multitude of definitions of awareness and to the identification of a large number of different forms of awareness such as passive awareness (Dourish & Bellotti, 1992), workspace awareness (Gutwin & Greenberg, 2002), social awareness (Bødker & Christiansen, 2006; Prasolova-Forland, 2002), conversational awareness (Mendoza-Chapa, Romero-Salcedo, & Oktaba, 2000), group awareness (Mendoza-Chapa et al., 2000; Romero-Salcedo et al., 2004), and history awareness (Kreijns, 2004; Kreijns & Kirschner, 2001). Although there are differences between these forms of awareness and their definitions, their main commonality is their focus on information, or rather, the lack thereof in CSCL environments. In such environments it is often difficult to obtain needed information on what the other is doing, whether he or she is available for communication and interaction, what the others know about the task at hand, what group members will do next, and so on (Gutwin, Stark, & Greenberg, 1995). This has to do with the fact that CSCL environments offer only a small fraction of the perceptual information that is usually available during FTF collaboration (Gutwin & Greenberg). For instance, when Lara, Casey, and Rafael had been sitting around a table creating a representation of the historical debate with colored markers, it would have been much easier for Rafael to know what Casey and Lara were doing. This information would be easily available by looking at Casey and Lara. In a FTF situation, Rafael would not have to spend so much time inquiring what Casey and Lara were doing.

Thus, awareness is a problem of perception and information (Romero-Salcedo et al., 2004). Why would this be problematic for group members working in a CSCL environment? Consider the chat-fragment above. Would Rafael also be asking Casey what he was doing to the same degree if they were collaborating FTF? Probably not. Because Rafael and Casey lack information about their group members' activities, their collaboration is far from smooth. Note for instance Rafael's irritation after Casey's mistake. Awareness information can reduce group members' efforts to coordinate their action, can increase their efficiency, and reduce the chance of errors (Gutwin & Greenberg, 2004).

When group members collaborate, they often switch between episodes of individual work and collaborative discussion (Barron, 2003). These switches from individual to collaborative phases are usually triggered by the need to ask group members for help, to decide on the next step in the problem, to correct mistakes made by other group members, and so on. But to be able to decide when it is necessary to start a collaborative episode, group members need to be aware of what the others are doing. Without this information, they may miss out on collaboration opportunities or may unnecessarily interrupt other group members (Gutwin & Greenberg, 2004). Furthermore, awareness information is crucial for mutual performance monitoring. When working on a group task, responsibilities are often shared among group members. However, group members need to monitor whether other group members are doing their tasks correctly. To be able to do so, awareness information is needed (e.g., Who is doing what? Is performance of group members on a sufficient level?). Finally, awareness information can be used to articulate ideas and notions and to facilitate conversation. In the fragment above, students refer to parts of the environment (e.g., "Debate", "source 3", "Martyrs position"). This is called deictic referencing (Mühlpfordt & Stahl, 2007; Stahl et al., 2006; Suthers & Hundhausen, 2003). They use this strategy because it helps to facilitate the other's interpretation of the message. Rafael refers to the information he gets from the Debate-tool ('Nothing was added') to express his concerns. Without this information, communication would have been more difficult. In sum, awareness information is an important factor in facilitating online group work.

During collaboration, group members have to engage in different types of activities. These activities often have to do with the execution of the task, while others have to with the regulation of the task. But group members also need to regulate the social aspect of collaboration. For instance, they need to coordinate their collaboration: Who is available for discussion and communication? Who needs help? Is the collaboration going fine or should changes be made? This means that group members need awareness information of these social aspects of collaboration as well (Kreijns, 2004). A common problem in collaboration is the free rider effect: one student lets the other group members do most of the work (Salomon & Globerson, 1989). This is obviously not in the best interest of the group and therefore needs to be avoided. But it is often very difficult to determine whether free riding behavior is occurring. Rafael and Casey think that Lara might be taking a free ride, but without the proper information they cannot be certain. Thus, while working in a CSCL environment, group members not only require awareness information about task-related, but about social aspects of the collaboration as well.

1.1.2 Communication Problems

Research has demonstrated that it is difficult for group members to communicate effectively and efficiently during CSCL. For example, Fjermestad (2004) found that in many studies, communication was more difficult and cumbersome in CSCL environments as compared to face-to-face (FTF) conditions, concluding that “it is still easier to communicate verbally than through the computer.” (p. 250) Some researchers have argued that the communication problems found during CSCL may be due to the medium itself. More precisely, traditional CSCL media, such as e-mail or chat, are seen as media that are low in media richness (Daft & Lengel, 1986; Dennis, Kinney, & Hung, 1999). Media richness is defined as a medium’s ability to facilitate communication and the establishment of shared meaning. Factors such as the immediacy of feedback or the ability of the medium to transmit multiple cues (e.g., facial expressions, gestures, or intonation of voice) influence its richness. As media richness decreases, group members will have more difficulties conveying their opinions and ideas (i.e., lack of nuance), will have more difficulties determining the meaning of group member’s messages, and will experience more difficulties knowing whether their arguments are understood or accepted by group members (i.e., lack of "body language", Adrianson & Hjelmquist, 1999; Kiesler, Siegel, & McGuire, 1984; Thompson & Coovert, 2003). CSCL environments that are low in media richness use primarily discussion boards or chat as a means to communicate. This means communication difficulties between group members are bound to surface from time to time.

Furthermore, when working on group tasks, group members often are required to work on complex problems without demonstrably correct answers and which require them to resolve differing viewpoints. The communication channels usually used during CSCL (i.e., discussion boards or chat), may not be suited to such challenging tasks involving a high degree of uncertainty and equivocality. The low media richness of these environments may constrain collaboration in such a way that it does not transmit the type of communication that group members need to solve their task successfully (i.e., formulating arguments and engaging in critical discussion) leading to communication problems and decreased task performance (Adrianson & Hjelmquist, 1991; Hollingshead & McGrath, 1995; Mennecke, Valacich, & Wheeler, 2000).

This "problem" however is not as simple as it may appear. The user's experience with the medium may strongly affect how one uses a medium and what is possible within that medium (Adrianson & Hjelmquist, 1999; Carlson & Zmud, 1999; Munneke, Andriessen, Kirschner, & Kanselaar, 2007). When users have had more experience with communication in a CSCL environment, they may be more adept at overcoming the communication problems that often arise in these environments (i.e., using emoticons, using special language or abbreviations). Furthermore, it should be noted that in recent years a lot of technological progress has been made to overcome the possible leanness of CSCL environments. Video conferencing possibilities can for example be incorporated in these environments. By including auditory and visual channels in CSCL environments, the richness of these environments can be increased.

1.1.3 Coordination Problems

The potential of collaborative learning is thought to reside in the social interaction between group members (Cohen, 1994; Lou, Abrami, & d'Apollonia, 2001). Social interaction may elicit processes that are beneficial to group members' understanding of the subject matter and the construction of new knowledge (i.e., giving and receiving explanations, conducting critical discussions of the subject matter, giving arguments and counterarguments). However, as stated before, efficient, effective, and meaningful interaction should not be taken for granted (Kreijns, Kirschner, & Jochems, 2003).

One of the reasons why good collaboration is often so difficult has to do with the different types of activities that group members need to engage in while working on a group task. During collaboration, they need to exchange ideas and opinions, ask questions, produce arguments and counterarguments, and generally work towards producing a group product (Dennis & Valacich, 1999). This has been called the *production function* of groups (McGrath, 1991) or the *content space of group work* (Barron, 2003; Cole & Nast-Cole, 1992). On the other hand, collaboration involves a social-relational aspect as well. group members have to perform social and communicative activities that establish group well-being, create a sound social space and a common frame of reference (Clark & Brennan, 1991; Kreijns et al., 2003; Massey, Montoya-Weiss, & Hung, 2003; Thompson & Fine, 1999). In this respect, McGrath refers to the *member-support* and *well-being functions* of groups. Similarly, Barron (2003) referred to the *relational space* that needs to be created and maintained by group members during collaboration. Cole and Nast-Cole (1992) refer to these activities as *maintenance activities* and state that "effective groups have learned to integrate maintenance activities into the very fabric of the group's work, and to view maintenance activities as the route to excellence." (p. 49)

Because collaboration involves so many different activities that need to be performed by group members, the need for *coordination* arises (Erkens, 2004; Erkens, Prangsma, & Jaspers, 2006; McGrath, 1991). This need for coordination is created by the interdependence between group members during collaboration (Malone & Crowston, 1992). Groups require the mutual input of all group members to be successful (Damon & Phelps, 1989; Johnson & Johnson, 1999). Coordination, be it FTF or distributed, is in itself is a difficult task. Without coordination, group members are likely to engage in conflicting or repetitive activities, disrupting the collaborative process (Malone & Crowston, 1992). It can also lead to irritation among group members (see for example the fragment at the beginning of this chapter) which may undermine the group's social space. Successful groups have been found to coordinate their activities better than unsuccessful groups (Barron, 2003). Coordination is thus an important activity during collaboration, one that "can be viewed as an activity in itself, as a necessary overhead when several parties are performing a task." (Ellis et al., 1992, p. 25)

But what is coordination exactly? Malone and Crowston (1992) note many different definitions of coordination have been proposed, but that coordination becomes visible only when coordination problems arise: good coordination is often invisible. In spite of this, Erkens (2004) and Erkens et al. (2006) distinguish three types of activities that play a role during the coordination of the collaborative process, namely (a) activation of knowledge and skills, (b) grounding, and (c) negotiation and coming to agreement. These processes are explained below and summarized in Table 1.1.

Activation of knowledge and skills involves initiating communication and collaboration episodes. Because group members can benefit from the skills and knowledge of their group members during collaboration, knowledge and information exchange are important processes: unshared knowledge needs to be externalized. This can take the form of verbalizations or externalization of knowledge (Kirschner, Beers, Boshuizen, & Gijselaers, in press; Teasley, 1995), elaborate explanations (Webb & Mastergeorge, 2003; Webb, Troper, & Fall, 1995) and so on. Furthermore, it is important that all group members engage in these activities to an equal extent (Damon & Phelps, 1989). Equal participation and symmetry in contributions between participants is important for successful collaboration and student learning (Cohen, 1994). When group members participate equally, each group member has the opportunity to contribute to the group process, to engage in reciprocal knowledge construction, and to put his or her skills into action (Barron, 2000). This ensures learning opportunities for all group members.

Grounding is another important activity that group members have to perform in order to coordinate their collaboration. Grounding is a necessary process for group members to establish mutual understanding and a common frame of reference (Clark & Brennan, 1991; Erkens et al., 2006; Van der Pol, 2007). To communicate and collaborate effectively, group members need to ensure they understand each other. Grounding can be seen as the strategies employed to create this understanding. One such strategy is *tuning*, defined as adapting to the collaboration partner(s). In a collaborative situation different perspectives and interpretations may for example exist between group members as a result of differing experiences and knowledge bases. If one "tunes" to the other group member, these differences are taken into account during conversation and thus misunderstandings may be prevented. In other words, the listeners develop a framework in which they interpret the externalizations of others. This does not mean that they agree, but that they understand the origin of the statement. One way of tuning is to match one's linguistic style to that of the partner (i.e., adapting one's word use to that of the communication partner, see Niederhoffer & Pennebaker, 2002).

Another important strategy is to ensure *joint attention* when needed. Joint attention exists when group members respond appropriately or engagedly to the proposals of a group member (Barron, 2000). Appropriate or engaged responses are, for example, acceptations or starting a constructive discussion. Inappropriate responses are ignorance of the other or outright rejections. Successful groups display higher levels of these engaged responses compared to unsuccessful groups (Barron, 2003). In this respect, Erkens et al. (2005) and Veerman (2003) refer to *focusing*. This pertains to the way group members try to maintain a shared discourse topic. This can be done by asking questions, asking for attention, and repairing focus divergence. Barron (2003) for example, noted that successful groups' contributions were more often in keeping with the previous discussion. Successful groups were better at maintaining a shared focus during their discussions.

A final strategy that group members may use during the grounding process is *checking*. When group members engage in discussion it is important they check whether the information that is being exchanged fits with the common frame of reference that has been created thus far. Furthermore, it is necessary to check whether group members understand the information or proposals being communicated. Group members can check this by asking verification questions or by indicating agreement or disagreement (Krol, Janssen, Veenman, & Van der Linden, 2004; Van der Linden, Erkens, & Nieuwenhuysen, 1995). Kirschner et al. (in press) emphasize the reciprocal relationship between externalization of information by one group member and the processing thereof by other group members. When knowledge or information is externalized, group members have to verify whether their understanding of the information matches the other's understanding of the information (e.g., by asking verification question or by giving more elaborate examples). This can lead to an extensive process of negotiation of meaning. Activation of knowledge through the externalization of information is not enough. Rather, through the process of negotiation of meaning group members have to come to a form of shared understanding about the information that is being discussed.

The previous coordination strategy, grounding, referred to understanding others. However, understanding the other is not equal to agreeing with the other. Thus, the final coordination strategy distinguished by Erkens (2004) and Erkens et al. (2006) is *negotiation and coming to agreement*. Group members need for instance to negotiate and come to agreement about solutions or strategies.

This implies that explanation and argumentation are important processes. Group members need to give elaborate explanations for their remarks, proposals, and solutions. Furthermore, they should engage in a process of constructive argumentation. During collaboration there should for example be room for challenging ideas and offering counter arguments (Wegerif, Mercer, & Dawes, 1999). Group members engage in argumentation when they confront each others' ideas, opinions, or beliefs and exchange, criticize, and complete the underlying ideas, reasons, and motives (Andriessen, Erkens, Van de Laak, Peters, & Coirier, 2003). Finally, group members need to reach agreement on a shared decision or solution (Di Eugenio, Jordan, Thomason, & Moore, 2000). Although conflicts can be productive from a neo-Piagetian point of view in that they may stimulate explanation and argumentation (cf., De Lisi & Golbeck, 1999; Dimant & Bearison, 1991), they need to be resolved at some point or otherwise they may become detrimental (cf., Erkens; Erkens et al.; Tudge, 1989). In this respect, Kirschner et al. (in press) make a distinction between negotiation of meaning, and negotiation of position. Whereas group members can agree on what something means (i.e., negotiation of meaning), they can disagree about its effects, implications, et cetera (i.e., negotiation of position).

Table 1.1: Summary of coordination processes (adapted from Erkens, 2004; Erkens et al., 2006).

Coordination process	*Example*
1. *Activation of knowledge and skills*	• Initiating communication • Articulating and externalizing knowledge and information • Equality
2. *Grounding*	• Tuning to the other group member(s) • Maintaining joint attention • Focusing • Checking and negotiation of meaning
3. *Negotiation and coming to agreement*	• Giving explanations, engaging in argumentation • Coming to agreement

From the above it becomes clear why coordination of collaborative learning is often problematic. Group members need to carry out many different activities and coordinating these activities requires employing different and often difficult strategies. There is thus a great risk that the coordination of these activities is far from smooth. Considering the awareness and communication problems listed in the previous sections, coordination may be even more difficult during CSCL (Thompson & Coovert, 2003).

1.1.4 Lack of Quality Discussions

The final problem we wish to discuss here concerns the quality of the discussions group members engage in during online collaboration. When group members collaborate, they are often working on complex problems which require the input of all group members (Damon & Phelps, 1989). Ideally, group members engage in discussions that are critical, but also constructive. This means that (1) group members are critical of their own and the other group members' ideas, (2) criticism is accepted, (3) that they offer explanations for their opinions, and (4) all group members participate in the interaction process. Such discussions have been called *exploratory discussions* (Wegerif et al., 1999) or *interactive argumentation* (Chinn & Anderson, 1998; Munneke, Andriessen, Kanselaar et al., 2007) and have been found to enhance learning during group work. However, research has shown that group members rarely give arguments and counter arguments (Kuhn & Udell, 2003), and that they often do not offer explanations for their ideas during CSCL (Van der Meijden & Veenman, 2005).

This absence of critical constructive discussion may be explained in several ways. First, students may not know how to conduct such discussions and may not possess the necessary skills (Weinberger & Fischer, 2006). For example, during discussions it is important to give explanations and evidence for assertions. However, students often do not do this. Furthermore, group members are often quite adept at giving arguments that support their ideas and theories, but they usually fail to search for counter arguments or alternative theories and tend to focus arguments that only confirm their position (Felton & Kuhn, 2001; Hightower & Sayeed, 1995; Kuhn & Udell, 2003). Research by Kuhn (1991) demonstrated that adults also tend to have difficulties with for example supporting their opinions with relevant evidence and considering counterarguments. The question arises then whether it is the case that group members (and adults) *cannot* engage in argumentation or whether they do not *want* to. Kuhn and Udell (2003) showed that students can be taught to employ powerful argumentative discourse strategies (i.e., presenting genuine evidence, considering counterarguments). This seems to suggest that a lack of argumentation skills (at least partly) explains why students often do not argue effectively.

A second reason for the lack of critical online discussion may be that group members find it difficult to conduct constructive debates in a CSCL environment and have difficulties interpreting these discussions. For example, they may not know whether group members agree or disagree with them. This possibly hampers argumentation and discussion (Adrianson & Hjelmquist, 1991).

Finally, groups may possess group norms that stimulate consensus among group members, instead of critical or exploratory discussion. Group norms that stimulate consensus instead of critical discussion can contribute to the low quality of some online discussions (Erkens et al., 2006; Postmes, Spears, & Cihangir, 2001).

In conclusion, the relative absence of critical discussion during CSCL may have different causes. These causes need to be addressed in order to facilitate critical but constructive discussions during CSCL. Different solutions have been proposed to enhance the quality of collaborative discussions. These range from teaching or training group members to carry out such constructive discussions and argumentations (e.g., Saab, Van Joolingen, & Van Hout-Wolters, 2007;

Veenman, Denessen, Van den Akker, & Van der Rijt, 2005; Voss & Means, 1991; Wegerif et al., 1999), to offering computer tools to foster and structure argumentation (e.g., Munneke, Andriessen, Kanselaar et al., 2007; Munneke, Van Amelsvoort, & Andriessen, 2003; Suthers & Hundhausen, 2003; Van Drie, Van Boxtel, Jaspers, & Kanselaar, 2005), or to scripting argumentation (e.g., Lazonder, Wilhelm, & Ootes, 2003; Weinberger, Ertl, Fischer, & Mandl, 2005). In the next section we propose a different solution.

1.2 Possible solution: Visualizations

The previous section described four commonly encountered problems during CSCL. Research on CSCL has identified promising approaches to solve these problems, such as scripting, assigning specific roles to group members (e.g., Schellens, Van Keer, & Valcke, 2005; Strijbos, Martens, Jochems, & Broers, 2004), using external representations to foster argumentation and knowledge construction, and instructing group members in proper collaboration strategies. However, we propose a different solution, namely to visualize aspects of the collaborative process to mitigate the described problems of CSCL. Before we describe the reasons why collaboration may improve when visualizations are implemented in CSCL environments, we will first give several examples of visualizations employed in CSCL environments. Visualizations can be used to make several features of the online collaborative process salient.

1.2.1 Examples of Visualizations in CSCL Environments

Working in CSCL environments on a group task can be quite a challenging experience. Usually group members work on complex or even wicked problems (Van Bruggen, Boshuizen, & Kirschner, 2003), which require them to carry out different activities for both solving the problem and managing the process of collaboration.

Figure 1.1 shows the CSCL environment that is used throughout the studies reported on in this thesis. This environment is called *Virtual Collaborative Research Institute* (VCRI) and has been developed over the last two decades at Utrecht University. The program and its predecessors have been used in several research projects (e.g., Andriessen et al., 2003; Erkens et al., 2005; Munneke, Andriessen, Kanselaar et al., 2007; Van Amelsvoort, 2006; Van Drie, 2005). VCRI is a groupware program designed to support collaborative learning on inquiry tasks and research projects. Students use VCRI to communicate with each other, access information sources, and co-author texts and essays. While working with VCRI, students share several tools, such as a *Sources*-tool which contains information sources that students can use to gather important information. This tool is shown in the bottom two windows of the screenshot shown in Figure 1.1. The left window shows a list of all the sources students can access, while the right window displays a source that has been opened by the user. Furthermore, a *Chat*-tool is available for synchronous communication with group members. This tool is displayed in the upper left window of the screenshot. Also, the VCRI contains a *Cowriter* for shared word processing, which students can use to simultaneously compose their texts or answers. The Cowriter is displayed in the upper right window of the Figure. Furthermore, the VCRI contains several more tools not shown in the screenshot, such as a *Diagrammer* for making external representations of ideas or arguments, and a *Forum* to communicate asynchronously. Other tools not shown in Figure 1.1 include a *Planner* and a *Logbook*.

From this Figure it becomes clear why collaboration in CSCL environment is such a daunting task. Besides conducting a research or inquiry task, students have to stay aware of what their group members are doing (e.g., by asking this in the Chat or by watching what happens in the tools of the program), communicate and discuss with their group members (e.g., about how they are going to approach the task or about conceptual issues which need clarification), and coordinate their activities (e.g., making sure that what student A types in the Cowriter does not conflict with what the other group members typed). While working in a CSCL environment, group members need to collect, interpret, monitor, and respond to large quantities of information. It has been suggested that visualization can support group members during these processes (Keller & Tergan, 2005).

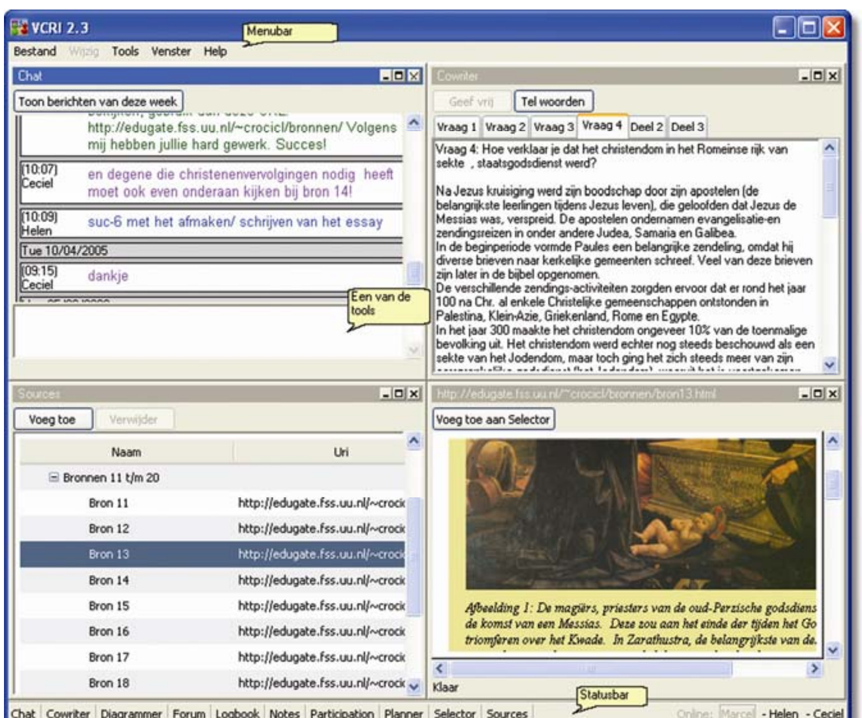


Figure 1.1: Screenshot of the CSCL environment Virtual Collaborative Research Institute (VCRI).

When designing visualizations, designers need to decide *what* information they want to visualize. It might, for example, be useful to visualize task-related aspects (e.g., How many sub problems have been solved by the group?) or social aspects (e.g., How many messages have been sent by each group member?) of the collaborative process. Starting point for the design could be the abovementioned production, member-support, or well-being functions of groups (McGrath, 1991; Zumbach & Reimann, 2003). Designers have to decide which function(s) they want to focus on. Ideally, the user and his or her experiences in the environment are taken as a starting point for this decision (Kirschner, 2002). The decision about what information to visualize should be driven by a motivation *why* it is important to visualize those specific aspects of the collaborative process. After deciding what they want to visualize (i.e., task-related aspects of the collaboration, social aspects, or a combination of both) and why they want to visualize it, designers have to decide *how* they want to visualize it. It is important they design the visualization in such a way that users can perceive and interpret it correctly (Keller & Tergan, 2005). Some information can best be visualized using textual representations (e.g. tables or matrices), while other information may be better visualized using graphical representations (e.g., line graphs or pie charts).

In the following paragraphs we give several examples of visualizations that have been implemented in learning environments. These examples can be placed on a continuum ranging from a focus on *task-related* to a focus on *social* aspects of the collaborative process and a continuum ranging from completely *textual* to completely *graphical*.

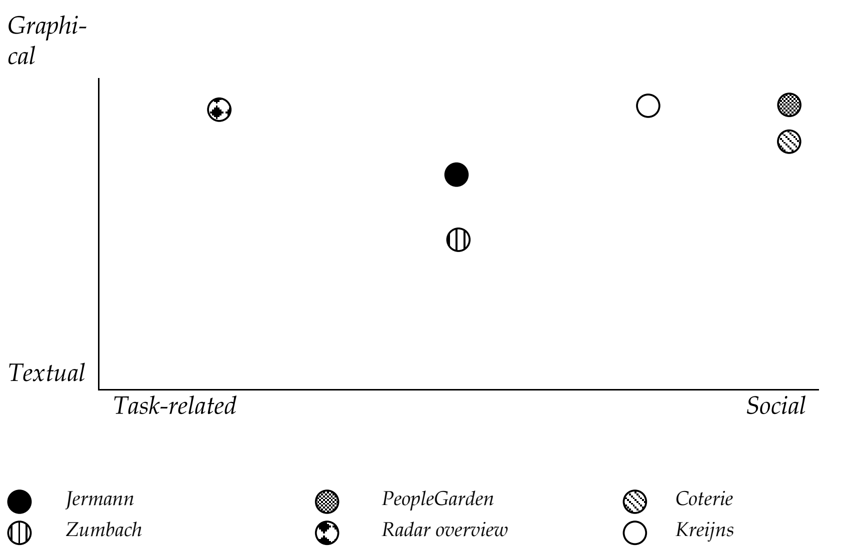


Figure 1.2: Classification of visualizations implemented in learning environments.

The work of Jermann (2004) is a good example of how information about the collaborative process might be visualized. The visualizations shown in

Figure 1.3 and Figure 1.4 have been implemented in a traffic simulation. Dyads collaborate on a task which requires them to tune the lights of a traffic simulation so as to minimize waiting time for car drivers. This requires them to 'talk' (visualized by the number of chat messages sent) and to 'tune' the traffic lights (visualized by the number of times the traffic lights were tuned in the simulation). Both visualizations show the amount of talking and tuning done by the group members. This allows group members to compare themselves with one another. Both examples visualized a social aspect (talking) and a task-related aspect of the collaboration (tuning), and both combine textual with graphical elements (see Figure 1.2).

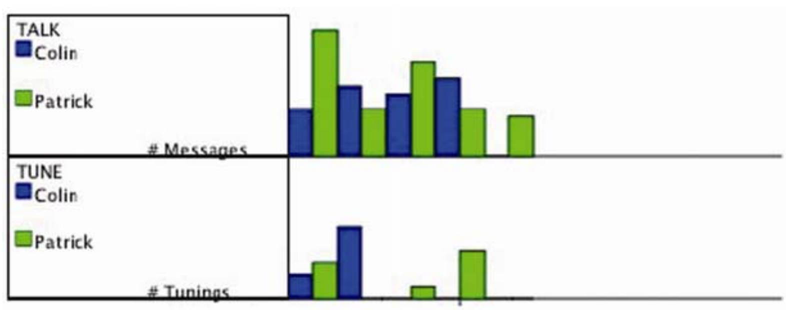


Figure 1.3: Visualization of cognitive and social aspects of the collaborative process. From Jermann (2004).

An important difference between the two visualizations concerns the use of a metaphor to enhance group members' understanding of the visualization (Keller & Tergan, 2005). The second visualization employs a 'dashboard'-metaphor. The needles visualize the proportion of talking and tuning of the group members (Christina and Billy) and the group as a whole (the thickest needle). With the use of color, group members are encouraged to keep a balance between talking and tuning; their needles should stay in the green area on the right side of the dashboard meter. This kind of value judgment is not given in the visualization displayed in Figure 1.3.

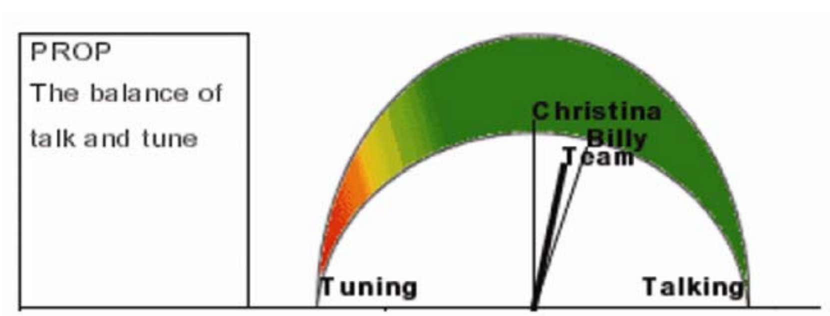


Figure 1.4: Visualization of cognitive and social aspects of the collaborative process. From Jermann (2004).

The PeopleGarden visualization (Figure 1.5) uses a different metaphor to describe participants' activity on a message board, namely that of gardens and flowers (Donath, 2002). PeopleGarden is an example of a visualization which focuses completely on the social aspect of collaboration (participation in group discussion) and which uses only graphical elements to represent this information (see Figure 1.2). Each message board is visualized as a garden containing flowers, whereas each participant is represented by a flower. The length of the stems of the flowers indicates the time participants have been active in the discussion, while the number of petals of their flower indicates the number of messages they have posted. This gives an idea how 'healthy' the garden is. Ideally, the garden should have many flowers with stems of different lengths and a large number of petals. In contrast to the visualizations developed by Jermann (2004), the PeopleGarden does not reveal who the active participants and lurkers are because it does not display users' names. In other words, this does not allow for a direct comparison between participants. Social comparison may form a motivational incentive for group members to put extra effort into the collaboration. This will be discussed in more detail in section 1.2.2.

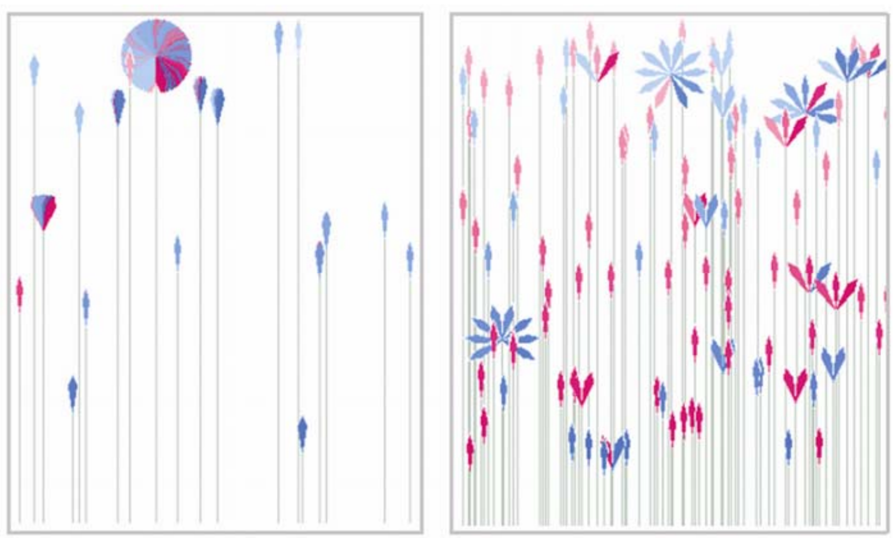


Figure 1.5: Two PeopleGarden screenshots. From Donath (2002).

Figure 1.6 shows two screenshots of a visualization called Coterie (Spiegel, 2001)[2] where users can also discern the most active participants in an IRC chat conversation. Participants are visualized by colored ovals. Coterie uses color effects and movements to indicate active participation: chatters whose ovals are colored brighter and whose ovals bounce more up and down are participating

[2] A Quicktime movie of Coterie can be downloaded at http://alumni.media.mit.edu/~spiegel/thesis/coterie.mov

more actively than chatters whose ovals have faded colors and remain motionless. Furthermore, because multiple conversations are often taking place in IRC channels, Coterie tries sort participants' chat messages into conversation threads. To do so, it analyzes the content of the chat messages. In Figure 1.6 several threads are visible, highlighting the multiple conversational topics of the IRC channel. Chatters whose ovals remain in one thread are chatters who stay in one conversation, whereas chatters whose ovals bounce between threads are chatters who contribute to multiple conversations. It is assumed that this makes it easier to follow the discussions taking place in the IRC channel (e.g., by examining the different threads in the channel), who the initiators of new discussions are (e.g., their ovals will move to new threads), who the prime contributors are (e.g., their ovals are colored brightly and bounce up and down a lot), and conversational cohesion (e.g., in chat rooms with cohesive discussions participants tend to stay in one thread; Donath, 2002). Coterie is another example of a visualization that focuses on the social aspect of collaboration and uses mostly graphical elements to visualize discussion participation although it also contains textual elements (e.g., chat messages).

Figure 1.6: Two Coterie screenshots. From Spiegel (2001).

Another example of how visualizations can be embedded in CSCL environments comes from the work of Zumbach and colleagues (Zumbach, Hillers, & Reimann, 2004; Zumbach & Reimann, 2003). The CSCL environment tracks group members' activity in the environment and feeds this back to the users in a pie chart (see Figure 1.7). Furthermore, at regular intervals group members are asked to rate their motivation. These data are then aggregated and visualized in a line graph, showing the development of each group member's motivation over time. Group members can use this information to track whether someone, for example, is dominating the collaboration, or if the motivation of a group member is dropping below a critical value. Furthermore, the environment gives feedback about the way group members have approached the problem. This feedback is given in the form of problem-solving protocols. It is assumed that this feedback will enable group members to identify successful and unsuccessful strategies, helping them to adapt their problem-solving process if this is necessary. Thus, the environment developed by Zumbach and colleagues visualizes social as well as task-related aspects of the collaborative process and combines textual and graphical elements to visualize this information (see Figure 1.2).

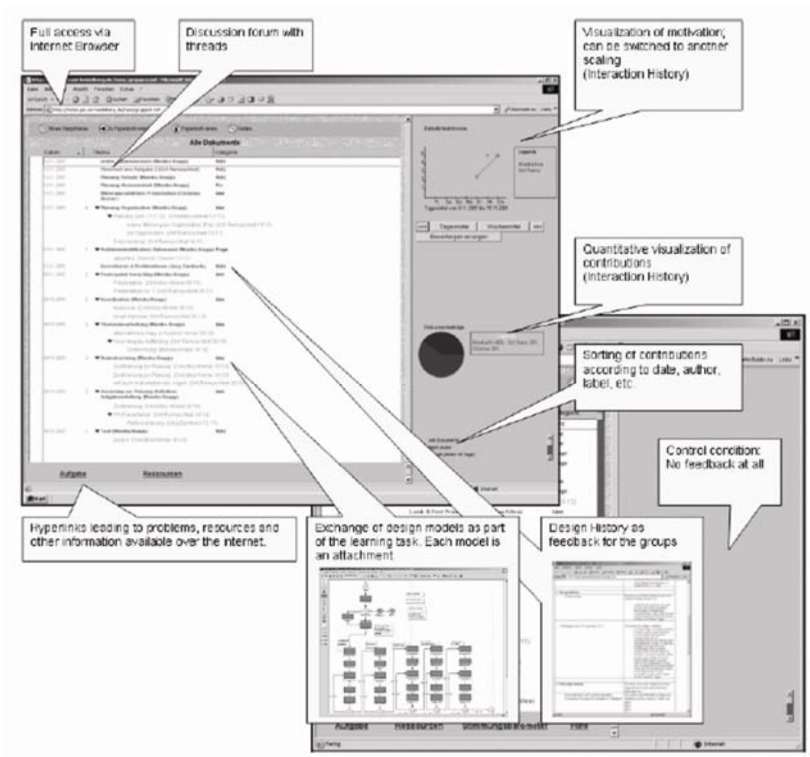


Figure 1.7: Screenshot of a CSCL environment with visualizations of group members' motivation (line graph) and participation. From Zumbach and Reimann (2003).

Gutwin and Greenberg (1999) have developed a so-called radar overview for their groupware concept map editor (Figure 1.8). Because concept maps can become too big to fit on a user's computer screen, it is often difficult to see what other group members are doing and what objects they are working on (i.e., they are working on objects not visible on the screen). The radar overview addresses this problem. It visualizes a small version of the entire concept map on top of the user's detailed view of the concept map. Using the radar overview, users can easily see who is currently working on the concept map, what they are doing, and on which part of the concept map they are working. The radar overview is an example of how task-related aspects of the collaboration (i.e., Who is working on which part of the concept map?) can be visualized using mostly graphical elements.

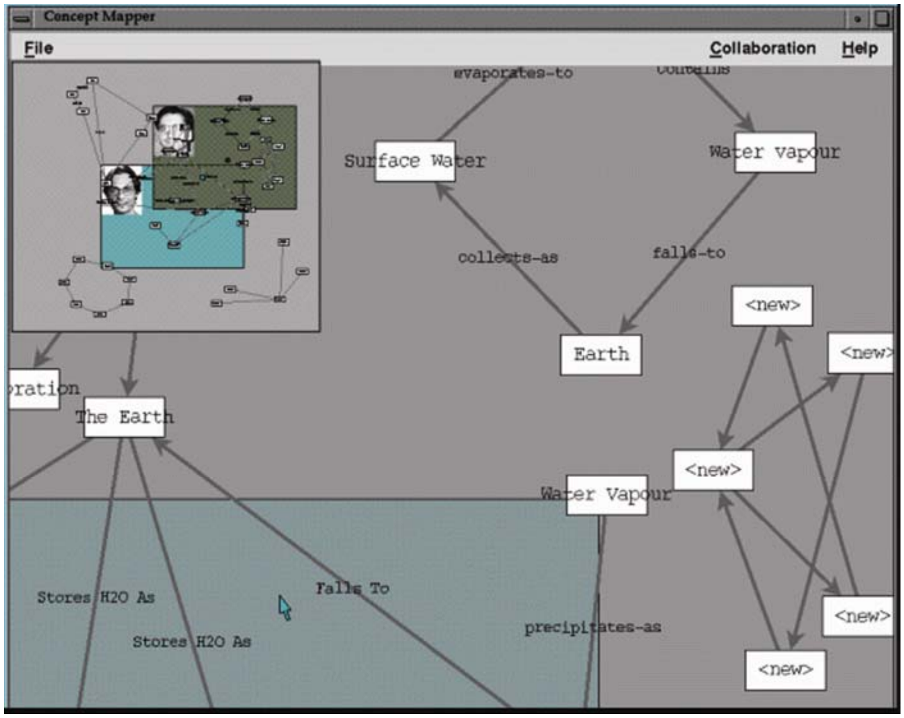


Figure 1.8: Radar overview for a concept map editor. From Gutwin and Greenberg (1999).

Finally, Kreijns (2004) describes developed a group awareness widget that visualizes several aspects of online collaboration (see Figure 1.9). These mostly concern the social aspect of collaboration (e.g., participation in a discussion forum, participation in social chat space), but also address task-related aspects (e.g., number of times participants access the course web-site). Kreijns' widget not only visualizes group members' current social and task-related activities, but also how these activities have developed over time. Current activities are placed to the left side of the bars displayed in Figure 1.9, while past activities are located near the right side of the bars. Kreijns' awareness widget is an example of how both social and task-related aspects of collaboration may be visualized, using mostly graphical elements (see Figure 1.2).

In this section we gave several examples of visualizations embedded in CSCL environments. These examples give an idea about what aspects of the collaborative process can be visualized (task-related or social) and how these aspects can be visualized (using graphical or textual elements). In the next section we will explain why these types of visualizations can possibly be of use to increase the effectiveness of CSCL environments.

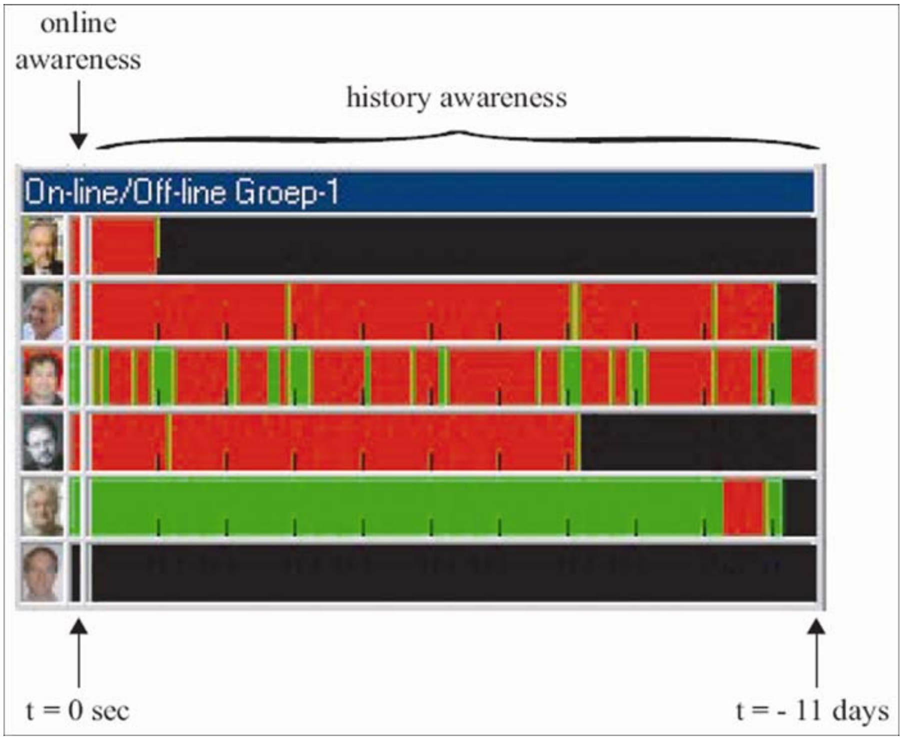


Figure 1.9: Kreijns' group awareness widget. From Kreijns (2004).

1.2.2 Reasons for the Possible Effectiveness of Visualizations

There may be several reasons why visualizations may be of help to increase the effectiveness of CSCL environments, namely (a) making complex information easier to interpret thereby decreasing the cognitive demands placed on group members, (b) giving feedback about import aspects of the collaborative process, (c) raising awareness, (d) facilitating coordination, and (e) providing a motivational incentive. These reasons will be described below.

As stated before, collaborating in CSCL environments is a complex endeavor. Group members have to carry out many different activities, while keeping track of the overwhelming amount of information that is available in the environment (e.g., the chat history detailing all the decisions that were made by the group or the version history of shared documents that are being written). The collection and interpretation of such information is a cognitively demanding task. However, visualizations can *make it easier to collect and interpret this information*, because "it is possible to have a far more complex concept structure represented externally in a visual display than can be held in visual and verbal working memories." (Ware, 2005, p. 29). Because visualizations can display large amounts of information and can facilitate its interpretation, they can decrease the cognitive demands placed on group members working in CSCL environments (Keller & Tergan, 2005; Sweller & Chandler, 1994). Visualizations for example, facilitate computational offloading (Ainsworth, 2006; Cox, 1999; Scaife & Rogers, 1996) since team members need to invest less effort to collect and interpret the information they need to collaborate successfully in a CSCL environment. Most of the examples described in the previous section indeed visualize large quantities of information. The awareness widget described by Kreijns (2004) for example, visualizes not only group members' current activities, but their past activity as well.

Visualization can be used to generate *external feedback* (Butler & Winne, 1995). This feedback provides group members with information they can use to monitor the progress of their collaboration. It allows them to determine whether selected strategies are working as expected, and whether group performance and products are up to standard. The visualizations designed by Jermann (2004) and Zumbach and Reimann (2003) for example, provide group members with feedback on how well they are collaborating (i.e., Are group members participating equally in the collaborative process?). Additionally, such visualizations can be used for group processing such as when group members discuss how well the group is functioning and how collaboration may be improved (Webb & Palincsar, 1996). These discussions may help groups pinpoint, comprehend, and solve collaboration problems (e.g., free riding by some group members) and can thus positively affect collaboration (Yager, Johnson, Johnson, & Snider, 1986). Group processing is also facilitated because visualizations can mediate discussion as is the case when they help group members externalize and articulate their thoughts about collaboration processes by providing them with appropriate information and concepts (Fischer, Bruhn, Gräsel, & Mandl, 2002; Teasley & Roschelle, 1993). After examining the pie chart in Figure 1.7 for example, a group member may feel that someone is free riding, which may stimulate him or her to discuss this with other group members by referring to the visualization.

Previously we identified lack of awareness as an important problem group members encounter while working in CSCL environments as they offer only a small fraction of the information that is readily available during FTF collaboration. Visualizations can enhance the awareness information CSCL environments offer to group members, and could thus *raise group members' awareness of the collaborative process* (Gutwin & Greenberg, 2004; Kirschner, Strijbos, Kreijns, & Beers, 2004). The VCRI for example, tries to increase group members' workspace awareness (Gutwin & Greenberg, 2002) through the Statusbar which can be seen at the bottom of the VCRI window (see Figure 1.1) and which lists all the tools available to students. Each group member has her or his own, distinct color in the VCRI (see for example the colors used in the Chat window in Figure 1.1), and these colors are used in the Statusbar. When a group member is working in a certain tool, the tool's name in the Statusbar will blink in his/her corresponding color (see Figure 1.10). This raises users' awareness of the activities of other group members in the environment. PeopleGarden and Coterie (Figure 1.5 and Figure 1.6) are examples of how social or conversational awareness can be enhanced by visualizations. Coterie visualizes different conversation threads by analyzing the content of chat messages which may raise group members' awareness of the different conversations that are being conducted and who is initiating them Without such a visualization, this would be much more difficult to determine, especially if the number of participants in the discussion is large. Visualizations can also be used to raise group members' (and teachers') awareness about the quality of the discussions being conducted. Recent work by De Groot et al. (2007), Hever et al., (2007), and Dönmez, Rosé, Stegmann, Weinberger, and Fischer (2005) has been aimed at developing discussion-quality indicators and tools for the automatic analysis of online conversations. These tools can be embedded in CSCL environments and the results of the automatic analysis can be fed back to group members in order to raise their awareness about the quality of their discussions. In sum, visualizations can be used to raise group members' awareness about several important aspects of the collaborative process.

Figure 1.10: VCRI's statusbar gives information about group members' activities.

Because visualizations give feedback and thus can raise group members' awareness about important collaborative processes, they can be used to *facilitate the coordination of collaboration*. In the previous section, we described repetitive or conflicting activities as indications of coordination problems. By visualizing group members' activity in the different tools, the VCRI's Statusbar tries to prevent this. Group members can use the information the Statusbar provides to prevent coordination problems (e.g., They notice a group member is working in the Cowriter, and then examine what he/she is doing in order to avoid duplicating the work). The Radar overview developed by Gutwin and Greenberg (Gutwin & Greenberg, 1999) has also been developed to facilitate coordination.

The coordination strategies identified by Erkens (2004; Erkens et al., 2006) can be facilitated by visualizations. Ensuring equal participation for example, can be facilitated by the visualizations developed by Jermann (2004), Spiegel (2001), and Zumbach and Reimann (2003), because they provide information about who the most active participants of the group are. This helps group members to decide whether someone is for example dominating the group, or whether someone is taking a free ride. If this is the case, group members can take steps to address this issue. The Coterie visualization (Figure 1.6) is a good example of how visualizations may help to maintain joint attention and a common focus. Coterie provides information about the discussion threads in a chat room and whether participants' contributions fit within these threads. Coterie therefore helps participants identify whether there are divergences of the conversational focus, and whether they should apply strategies to re-establish joint attention and a common focus. The abovementioned automated analyses of the quality of online discussions (De Groot et al., 2007; Dönmez et al., 2005; Hever et al., 2007), may be used by group members to determine whether their argumentation is of sufficient quality (e.g., Are they merely offering arguments in favor of a position, without considering arguments against or are they engaging in conflicts, without coming to an agreement?). If needed, they can decide whether or not they should try to increase the argumentative quality of their discussions.

Finally, visualizations can influence collaboration through *motivational processes*. Motivational processes have been used to explain why group members put effort into collaboration (Abrami & Chambers, 1996; O'Donnell & O'Kelly, 1994; Slavin, 1996). To counter productivity-loss in groups (e.g., caused by free-riding), one could provide group members with an incentive that enhances their motivation to contribute to collaboration (Shepperd, 1993). When participation of group members is visualized as is done with Zumbach and Reimann's (2003) pie chart (Figure 1.7) or Jermann's (2004) "dashboard", the contributions of each group member to the group are identifiable; establishing a link between a group member and his or her contribution to the collaboration (Jermann). This identifiability may provide motivational incentives for group members to invest extra effort into the collaboration. Visualization of participation in a pie chart for example, can motivate group members to participate more, because they are unable to "hide in the crowd" and they can be evaluated negatively by their group members when they participate insufficiently. This social evaluation can motivate group members to increase their participation (Shepperd, 1993). In addition, they can compare themselves to other group members, subsequently motivating them to set higher standards for themselves and to try to increase their participation (Michinov & Primois, 2005; Wheeler, Suls, & Martin, 2001). This also applies to the visualization of contribution-quality proposed by De Groot et al. (2007), Hever et al., (2007), and Dönmez et al (2005). By visualizing the quality of each contribution to the online discussion, group members can compare themselves with their group members, and may thus be motivated to raise the quality of their contributions.

Up to now, we only discussed possible advantages of visualizing task-related and social aspects of online collaboration. It may also be possible that such visualizations have detrimental effects on online collaboration. For instance, the visualizations described by Jermann (2004) and Zumbach and Reimann (2003) may stimulate group members to contribute to the online discussion just to make their needle stay in the green area of the dashboard or to increase their share of the pie chart (i.e., contributing for the sake of contribution). This could also lead to behavior that is aimed at deliberately manipulating the visualization (i.e., typing nonsense messages to increase one's participation). Furthermore, social comparison does not always produce positive effects (Buunk, Collins, Taylor, Van Yperen, & Dakof, 1990; Wheeler et al., 2001). For example, group members can become demotivated when they compare themselves with other group members who are doing better (e.g., in terms of participation or offering suggestions for solutions and strategies). In other cases, group members may not compare themselves to group members who are 'doing better' but to those who are 'doing worse'. This may cause them to set lower standards for themselves. Finally, adding visualizations to an already complex and challenging learning environment may lead information overload if the visualizations are not properly designed (e.g., their meaning is difficult to interpret).

1.3 Research Question and Overview of this Thesis

In the previous section we outlined several problems that can occur during CSCL (e.g., lack of awareness, communication and coordination problems, and lack of quality discussions). Furthermore, we gave examples of visualizations that have been designed to address some of these problems and we explained why visualizations might be helpful to solve them. However, the systematic study of embedding visualizations in CSCL environments has yet to begin. Except for a small number of studies which addressed the effects of visualizations of the collaborative process (cf., Jermann, 2004; Michinov & Primois, 2005; Zumbach et al., 2004; Zumbach & Reimann, 2003), most research concerns the development and description of such visualizations for CSCL environments (Keller & Tergan, 2005). The central issue in this thesis is therefore how and why visualizations of the collaborative process affect group members' collaboration in CSCL environments. In order to tackle this issue, three different visualizations were developed. The effects of these visualizations are examined in three studies by giving one group of students access to the visualization, while the other group of students is not given access to the visualization. The three visualizations were implemented in the earlier described VCRI environment. All three studies in this thesis worked with 16 to 18-year old participants in the fifth year of pre-university secondary education. Additionally, all studies were carried out in the domain of history since this subject is well suited for the inquiry tasks to be carried out in the VCRI environment. In all three studies we collected data about the collaborative process, students' perception of the collaborative process, and group and/or individual performance. In sum, the central question addressed in this these can be specified as follows:

> *How and why does giving students access to visualizations of the collaborative process affect the collaborative process, students' perceptions of the process, and group or individual performance?*

Study 1 (Chapter 2) describes the effects of the *Participation-tool* (PT), a tool which visualizes group members' relative contributions to the online discussion through the Chat tool of the VCRI environment. The PT is shown in Figure 1.11, and gives an impression of the number of messages sent by each group member and the relative length of these messages. The PT thus gives feedback about group members' participation during the collaboration (e.g., Is there equal participation in our group?) and allows them to compare themselves to other group members. This may raise their awareness about the manner in which they are collaborating and may stimulate group discussions about the collaborative process. The PT can be considered a visualization that focuses on a specific social aspect of collaboration, namely participation. Furthermore, the PT uses mostly graphical elements to visualize participation, but also some textual elements (see Figure 1.14).

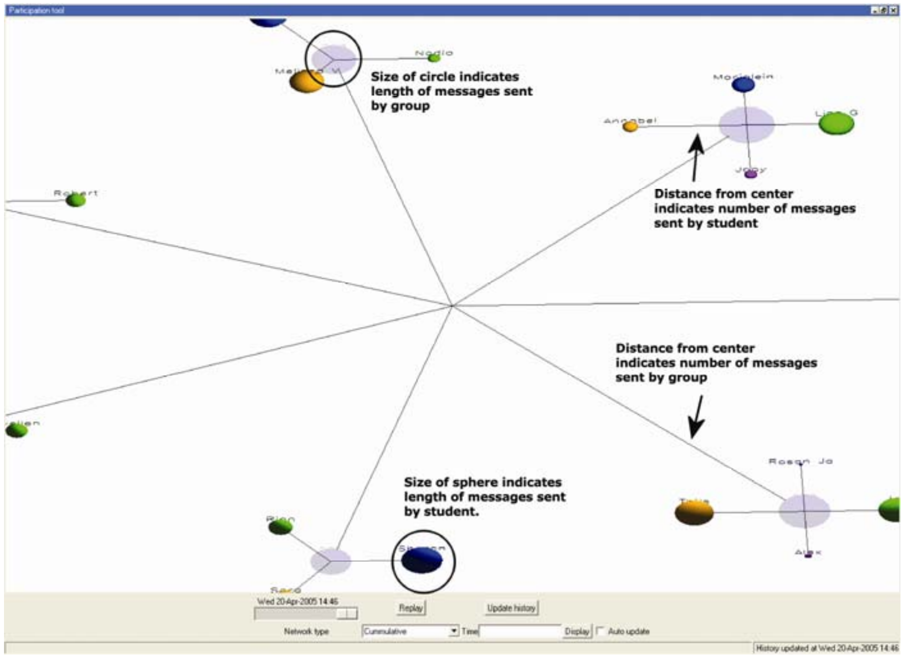


Figure 1.11: Screenshot of the Participation-tool developed for Study 1.

A visualization called *Shared Space* is the focus of Study 2 (Chapter 3). The Shared Space (SS) visualizes the amount of agreement or discussion during group members' chat conversations (see Figure 1.12). To do so, the SS analyzes all messages students type in the Chat tool of the VCRI environment. The SS uses discourse markers (Erkens, 1997; Schiffrin, 1987) to determine the communicative function of each message. Based on this analysis, the SS moves the current chat topic to the left if it contains indications of disagreement (e.g., denials, critical questions, negative evaluations) or to the right if it contains indications of agreement (e.g., confirmations, positive evaluations). The purpose of the SS is to give feedback about the types of discussions group members are conducting (e.g., critical discussions or consensual discussions) and thus to raise their awareness about their conversational strategies. Furthermore, by raising group members' awareness, groups may be stimulated to discuss how well their group is functioning and how the group process may be improved. Because the SS gives feedback about the types of discussions that are taking place in the VCRI environment, it visualizes a social aspect of collaboration. The SS uses graphical elements to visualize this information (see Figure 1.14).

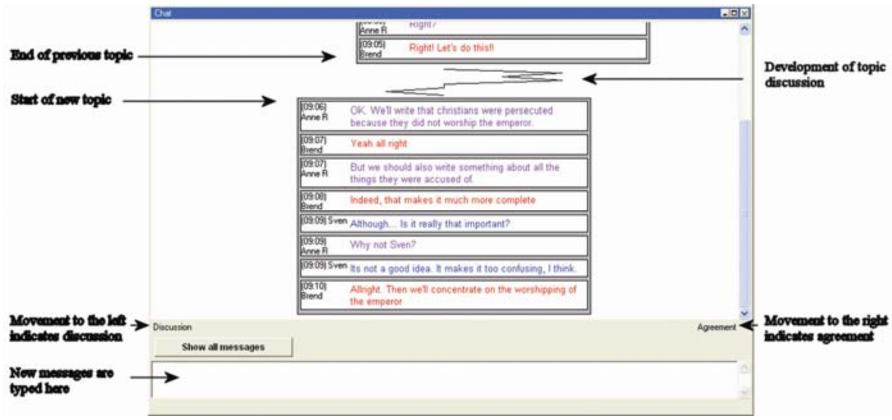


Figure 1.12: Screenshot of the Shared Space developed for Study 2.

Finally, Study 3 (Chapter 4) describes the effects of the *Graphical Debate-tool*. The Graphical Debate-tool uses visualization techniques to offer students representational guidance during the co-construction of an argumentative diagram (see Figure 1.13). The Graphical Debate-tool is an argumentative diagram which is used by group members to make a representation about a debate involving multiple positions, arguments, supporting and refuting information. The augmented Debate-tool visualizes the strength of support of a position through automatic repositioning of the position and accentuation of the complexity of the argumentation. Thus, the Graphical Debate-tool visualizes task-related aspects of the collaboration (i.e., Has one position been given more attention than the other? What is the balance between supporting and refuting information for an argument?) using mostly graphical elements (see Figure 1.14).

This thesis is concluded by a general discussion in Chapter 5. Because the studies reported on in this thesis involve students working in groups, traditional statistical techniques like *t*-tests or ANOVAs were not appropriate to answer some research questions. In these cases multilevel analysis was used instead. The reasons for this choice will be described in Chapter 6. This Chapter will serve as an epilogue to Chapters 2, 3, and 4. Furthermore, this Chapter will highlight the dangers of using inappropriate statistical techniques using data collected for Studies 1, 2, and 3.

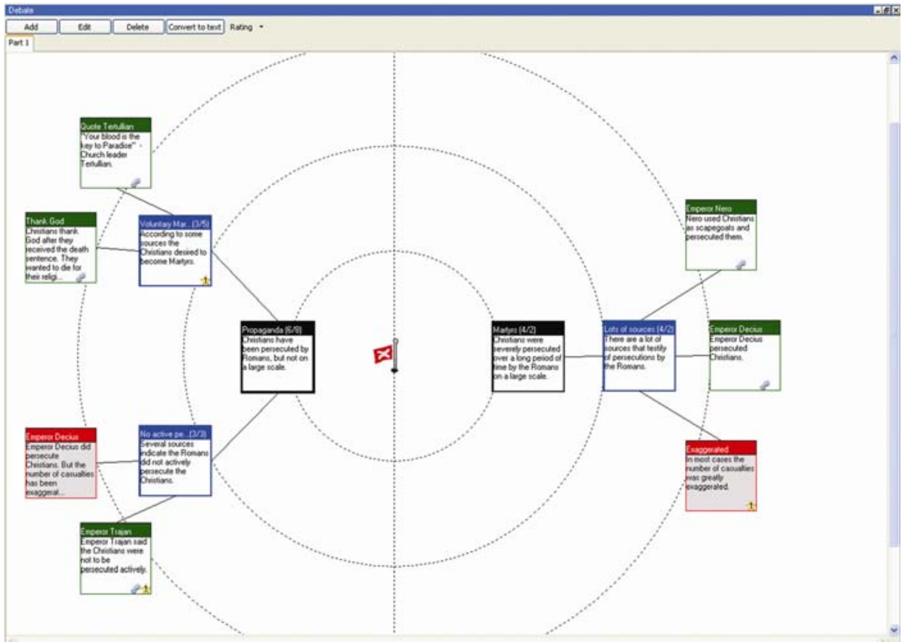


Figure 1.13: Screenshot of the Graphical Debate-tool developed for Study 3.

Figure 1.14: Classification of visualizations implemented in the VCRI environment for this thesis.

2. Visualization of Participation: Does it Contribute to Successful Computer-Supported Collaborative Learning?[1]

This study investigated the effects of visualization of participation during computer-supported collaborative learning (CSCL). It is hypothesized that visualization of participation could contribute to successful CSCL. A CSCL-environment was augmented with the Participation Tool (PT). The PT visualizes how much each group member contributes to his or her group's online communication. Using a posttest-only design with a treatment (N = 52) and a control group (N = 17), it was examined whether students with access to the PT participated more and more equally during collaboration, reported higher awareness of group processes and activities, collaborated differently, and performed better than students without access to the PT. The results show that students used the PT quite intensively. Furthermore, compared to control group students, treatment group students participated more and engaged more in coordination and regulation of social activities during collaboration by sending more statements that addressed the planning of social activities. However, equality of participation, awareness of group processes and quality of the group products was not higher in the treatment condition. Still, the results of this study demonstrate that visualization of participation can contribute to successful CSCL.

Keywords: Cooperative/collaborative learning; Computer-mediated communication; Distributed learning environments; Secondary education; Pedagogical issues

Over the last decades advanced information and communication technologies (ICT) have developed rapidly, which has led to many new computer applications, such as e-mail, chat rooms, video conferencing, simulations, and discussion forums. Many educational designers, policy makers, researchers, and teachers have embraced these applications as potentially useful tools for education. This interest has inspired many comparative studies, examining the effects of using ICT in education. Results of a meta-analysis showed that educational applications of ICT can have moderate positive effects on students' learning (Fletcher-Flinn & Gravatt, 1995).

In addition, new conceptions of learning have emerged. Researchers and theorists have increasingly recognized that learning is not only a cognitive, but also a social, cultural, and interpersonal, constructive process (Salomon & Perkins, 1998). As a result, instructional strategies whereby students work

[1] Janssen, J., Erkens, G., Kanselaar, G., & Jaspers, J. (2007). Visualization of participation: Does it contribute to successful computer-supported collaborative learning? *Computers & Education*, 49, 1037-1065.

together in small groups, also known as *collaborative learning*, are increasingly being used in education, since these instructional strategies seem to fit well with this new conception of learning. Furthermore, the positive effects of collaborative learning have been well documented: it enhances students' cognitive performance (Johnson & Johnson, 1999; Slavin, 1996) and it stimulates students to engage in knowledge construction (Stahl, 2004).

More recently, educational researchers have begun to explore the combination of ICT applications and collaborative learning. As a result, a relatively new field of educational design and educational research has developed. This field, *computer-supported collaborative learning* (CSCL), deals with issues concerning collaboration, learning processes, and the use of *computer-mediated communication* (CMC). The primary aim of CSCL is to provide an environment that supports and enhances collaboration between students, in order to enhance students' learning processes (Kreijns et al., 2003). During CSCL, students work on group tasks, and produce a group product. A CSCL environment usually offers tools that facilitate sharing of information and ideas, and the distribution of expertise among group members (Lipponen, Rahikainen, Lallimo, & Hakkarainen, 2003). When students collaborate in a CSCL environment, they use CMC to communicate with group members. CMC can be either synchronous (e.g., through a chat facility or video conferencing), asynchronous (e.g., through a forum or e-mail), or a combination of both.

Because CSCL combines collaborative learning and the use of ICT, various educational, social, and motivational benefits of CSCL have been suggested and documented by research. First, concerning *educational goals*, students report higher levels of learning in CMC groups, compared to FTF groups (Hertz-Lazarowitz & Bar-Natan, 2002). More importantly, when compared to FTF groups, students in CMC groups deliver more complete reports, make decisions of higher quality, and perform better on tasks that require groups to generate ideas (Fjermestad, 2004). Second, concerning *social goals*, researchers have found that CMC groups, compared to FTF groups, engage in more complex, broader, and cognitively challenging discussions (Benbunan-Fich, Hiltz, & Turoff, 2003), and group members participate more equally (Fjermestad, 2004). Finally, CSCL also seems to affect *motivational outcomes*, since students who collaborated in CMC groups report higher levels of satisfaction (Fjermestad, 2004). Thus, it seems that CSCL can have positive effects for education.

However, many studies have also demonstrated that several things can, and in fact *do* go wrong during CSCL. A number of studies have shown that during CSCL, several communication- and interaction problems can occur between students. These results are in contrast with the studies mentioned above; some results even seem to contradict the results of other studies. For instance, students working in CMC groups sometimes perceive their discussions as more confusing, compared to FTF groups (Thompson & Coovert, 2003). Furthermore, Hobman, Bordia, Irmer, and Chang (2002) found higher levels of personal conflict between students working in CMC groups, compared to FTF groups. As a result, CMC groups need more time to reach consensus and make decisions (Fjermestad, 2004). Moreover, they are less productive, and group cohesiveness is lower (Straus, 1997; Straus & McGrath, 1994). These problems can also influence the results CMC groups attain. For example, compared to FTF groups, CMC groups need more time to complete tasks (Baltes, Dickson, Sherman, Bauer, & LaGanke, 2002; Fjermestad, 2004), perform worse on mixed-motive tasks (Barkhi, Jacob, & Pirkul, 1999), and report lower levels of satisfaction (Baltes et al., 2002). In sum, several problems can occur during CSCL, and therefore its potential may not always be realized.

2.1 Visualization of Participation: A Solution?

In the section above, contradictory results concerning the possible benefits of CSCL were mentioned. Another important contradictory result found in CSCL studies concerns participation levels and equality of participation. Some studies report more equal participation of group members in CMC groups (e.g., Fjermestad, 2004), whereas other studies report dominance of some group members (e.g., Savicki, Kelley, & Ammon, 2002). Furthermore, in some CSCL studies researchers report low participation rates of all participating students (Lipponen et al., 2003; Veldhuis-Diermanse, 2002). It seems CMC groups may suffer from the same debilitating effects that sometimes occur in FTF groups (O'Donnell & O'Kelly, 1994; Salomon & Globerson, 1989), such as social loafing (group members invest less effort in a group, compared to working individually), or the free rider effect (students let other group members do the work).

If CSCL sometimes results in low overall participation rates or unequal participation, this is a cause for concern, since group productivity and student achievement depends on students' participation during collaboration (Cohen, 1994). When students participate equally during collaboration, every group member has the opportunity to contribute to group processes, to participate in knowledge construction, to give or request explanations, and to use and refine his or her skills (Webb, 1995). Given the importance of participation and equal participation, it is therefore important to ensure high levels of participation and equal participation of group members during CSCL.

Various strategies can be used to stimulate high levels of participation and equal participation. For example, a common strategy is to incorporate *positive interdependence* and *individual accountability* in the group task. Positive interdependence exists when all group members realize they have to work together to achieve their common goal. Individual accountability exists when group members are being held responsible for their contribution to the group goal. Positive interdepence and individual accountability may increase motivation to participate and foster social cohesion, and thus may serve to counter the free rider effect (Johnson & Johnson, 1999; O'Donnell & O'Kelly, 1994).

Another way to improve participation in CSCL may be through visualization of participation. Such a technique visualizes how much each group member relatively contributes to group discussion. It can be hypothesized that visualization of participation affects participation through *motivational* and *feedback* processes. Each of these processes will be explained below. For a detailed description of the visualization used in this study, the reader is referred to the method and instrumentation section.

2.1.1 Motivational Processes

Visualization of participation may influence collaboration through motivational processes. Motivational processes have been used to explain *why* students put effort into collaboration (Abrami & Chambers, 1996; O'Donnell & O'Kelly, 1994; Slavin, 1996). To counter productivity loss in groups (e.g., caused by free-riding), a possible solution could be to provide group members with an incentive that enhances their motivation to contribute to collaboration (Shepperd, 1993). When participation of group members is visualized, this makes the contribution of each group member to the group identifiable; establishing a link between a group member and his or her contribution to the collaboration (Jermann, 2004). This identifiability may provide several motivational incentives for group members to invest effort into collaboration. For example, visualization of participation can motivate group members to participate more, because they are unable to "hide in the crowd" and they can be evaluated negatively when they participate insufficiently. This *social evaluation* can motivate students to increase their participation (Shepperd, 1993). In addition, though *social comparison*, that is, through comparing themselves to other group members, students may be motivated to set higher standards and to try to increase their participation (Michinov & Primois, 2005; Wheeler et al., 2001).

2.1.2 Feedback Processes

Additionally, visualization of participation can also be considered a form of *external feedback* (Butler & Winne, 1995), that is, feedback generated by sources other than the student him- or herself (i.e., by teachers, group members, or computer displays). First, external feedback may provide students with information, which they can use to *monitor* their problem solving progress. External feedback allows students to determine whether selected strategies are working as expected, and whether group performance and products are up to standard. Thus, visualization of participation may provide group members with feedback on how well they are collaborating, and whether they have selected an appropriate collaboration strategy (i.e., equal participation of all group members). Similarly, visualization of participation can also be used for *group processing*. Group processing occurs when group members discuss how well their group is functioning and how group processes may be improved (Webb & Palincsar, 1996). These discussions may help groups pinpoint, comprehend, and solve collaboration problems (e.g., free riding by some group members) and may contribute to successful collaborative behavior (Yager et al., 1986). Group processing is also facilitated because visualization of participation may serve a *mediating* purpose, since it may help group members to externalize and articulate their thoughts about collaboration processes by providing them with appropriate information and concepts (Fischer et al., 2002; Teasley & Roschelle, 1993). For example, after examining the visualization, a group member may feel someone is free riding, which may stimulate him or her to discuss this with other group members by referring to the visualization.

Second, the external feedback provided by visualization of participation can also raise students' *awareness* of the group processes and activities taking place. Because visualization of participation shows group members' participation rates, it could raise students' awareness of group processes, and more specifically, of participation. Several researchers have suggested that awareness can play an important role in facilitating CSCL (Dourish & Bellotti, 1992; Gutwin & Greenberg, 2004; Kirschner, Strijbos et al., 2004). When students are collaborating, they have to be aware of the activities of their group members, because it allows them to decide which activities they have to engage in. This enables them to anticipate group members' actions.

2.2 Collaborative Activities during Collaboration

In addition to stimulating participation and equality of participation, visualization of participation may also influence the way students collaborate. For example, as described above, it may stimulate students to engage in group processing. Since one of the aims of this study is to investigate the effects of visualization of participation on the manner in which students collaborate, it is important to describe the different activities students perform during collaboration.

To successfully complete a group task group members have to engage in different types of activities (McGrath, 1991). First, group members have to *perform task-related activities* that are aimed at solving the problem at hand. Group members need to share and discuss task-related information, in order to pool their informational resources, make valuable information available to all group members (Jehn & Shah, 1997), verbalize their ideas and opinions (Van der Linden, Erkens, Schmidt, & Renshaw, 2000), and ask questions that elicit important information (Johnson, Johnson, Roy, & Zaidman, 1985; King, 1994). These task-related activities contribute to a group's production function (McGrath, 1991) and stimulate successful problem solving and individual learning. For example, Henry (1995) instructed group members to share task-relevant information on a judgment task. These groups outperformed groups who did not receive this instruction. Furthermore, in a study by Teasley (1995), dyads were instructed to generate hypotheses. Dyads that verbalized their ideas and opinions produced better hypotheses than dyads that did not verbalize their ideas.

Second, groups also have a member-support and well-being function (McGrath, 1991). Thus, group members have to attend to the social and emotional element of collaboration to successfully complete a group task (Forman & Cazden, 1985; Kumpulainen & Mutanen, 1999; Rourke, Anderson, Garrison, & Archer, 1999). Behaviors such as offering positive comments and praising group members contribute to a sound social space and a positive group atmosphere (Kreijns, 2004), which may increase group members' efforts to complete the group task (Jehn & Shah, 1997; Rourke et al., 1999). On the other hand, behaviors such as swearing or displaying negative emotions may have a negative impact on group cohesion (Johnson et al., 1985). Thus, groups have to *perform social activities* that help to maintain the group. For example, in a study by Jehn and Shah (1997) positive communication (e.g., offering positive comments and motivating group members) was related to performance on certain group tasks.

Table 2.1: Collaborative activities during collaboration.

	Task-related activities	*Social activities*
Performance	• Discussing task information • Sharing task information • Offering task-related opinions • Asking task-related questions	• Maintaining a positive group atmosphere • Disclosing personal information • Indicating understanding or misunderstanding
Coordination / regulation	• Making task-related plans • Discussing task-related strategies • Monitoring of task progress • Evaluation of task progress	• Making plans to collaborate • Discussing collaboration strategies • Monitoring group processes • Evaluating group processes

Third, collaboration also involves *coordination or regulation of task-related activities* (Erkens, 2004; Erkens et al., 2005). Coordination of task-related activities involves performing them in the right order and at the right time, without conflicting with the activities of other students (Gutwin & Greenberg, 2004). During collaboration, group members need to coordinate their activities to determine a common course of action. Therefore, metacognitive activities that regulate task performance, such as making plans and monitoring task progress, are considered important to successful group performance (Artzt & Armour-Thomas, 1997; Van Meter & Stevens, 2000). For instance, in a study on computer-supported collaborative writing, planning activities were related to the quality of written texts (Erkens et al., 2005). Furthermore, Jehn and Shah (1997) demonstrated task monitoring was related to performance on group tasks.

Fourth, similar to task-related activities, collaboration requires *coordination or regulation of social activities* as well (Ellis, 1997; Erkens, 2004; Forman & Cazden, 1985). During collaboration, group members are interdependent, and therefore they have to discuss collaboration strategies, monitor the collaboration process, and evaluate and reflect on the manner in which they collaborated. For instance, studies by Yager et al. (1986) and Johnson et al. (1990) demonstrated the positive influence of group processing. That is, when group members discuss how their group is performing and how collaboration may be improved, group performance is increased.

Table 2.1 depicts the different types of collaborative activities group members have to engage in during collaboration: task-related and social. Furthermore, these activities refer to two levels: a performance level and a coordination or regulation level. Successful collaboration requires that group members attend to both types of activities at both levels.

2.3 Research Questions

This study investigates the effects of visualization of participation during CSCL. An existing CSCL-environment will therefore be enhanced with a new tool that visualizes students' participation during collaboration: the *Participation Tool* (PT, described below). The following research questions will be addressed:

1. How intensively do students use the PT while collaborating online?
2. Do students who have access to the PT participate more and more equally, during collaboration than students who do not have access to the PT?
3. Are students who have access to the PT more aware of group processes and activities during collaboration than students who do not have access to the PT?
4. Do students who have access to the PT engage in different collaborative activities than students who do not have access to the PT?
5. Do groups who have access to the PT perform better on an inquiry group task than groups who do not have access to the PT?

First, it is expected that the PT will contribute to students' participation during CSCL through motivational and feedback processes as described above. Second, it is expected that the PT will help students to become more aware of the group processes and activities taking place during collaboration, since it provides them with feedback about the participation rates of group members. Furthermore, it is expected that the PT will affect the way students collaborate. Because the PT gives students information about the way their group is functioning, it may stimulate students to engage in coordination or regulation of social activities (see Table 2.1). For example, the PT may help group members to monitor group processes, evaluate how their group is collaborating, or help them to make plans for collaboration. Finally, it is expected that, through higher levels of participation, higher awareness and different collaborative activities, the PT will contribute to groups' performance on group tasks, thus increasing the efficacy of collaboration.

2.4 Method and Instrumentation

2.4.1 Design

A post-test-only design with a treatment and a control group was used to answer the research questions. Treatment group students had access to the PT; control group students did not. Three different classes participated in the study. Each class was randomly assigned to either the treatment or the control group. Thus, all students from one class were in the same condition: either treatment or control. Two classes were assigned to the treatment group, and one class was assigned to the control group. The treatment group consisted of 55 students (17 groups), and the control group of 17 students (5 groups).

2.4.2 Participants

Participants were 72 eleventh-grade students (30 male, 42 female) from a secondary school in The Netherlands. Students came from three different classes and were enrolled in the second stage of the pre-university education track. Mean age of the students was 16 years (*SD* = .58, *Min* = 15, *Max* = 18). Three treatment group students were omitted from all analyses because they attended less than half of the lessons. Therefore, the final number of participants was 52 treatment group and 17 control group students.

During the experiment, the participating students collaborated in groups of three or four; students were randomly assigned to a group by the researchers. Therefore, group composition was heterogeneous with respect to ability and gender. In order to eliminate combinations of students who could not get along with each other, the group compositions were verified by their teachers. The initial group compositions were approved.

2.4.3 Task and Materials

CSCL-environment: VCRI

Participating students collaborated in a CSCL-environment called *Virtual Collaborative Research Institute* (VCRI, Jaspers, Broeken, & Erkens, 2004). VCRI is a groupware program designed to support collaborative learning on research projects and inquiry tasks. Every student works at one computer. Figure 2.1 shows a screenshot of the VCRI-program, detailing the most important tools.

The *Chat* tool is used for synchronous communication between group members. The chat history is stored automatically and can be re-read at any time. Students can read the description of the group task and search for relevant historical information using the *Sources* tool. The *Cowriter* is a shared word-processor, which can be used to write a group text. Using the Cowriter, students can work simultaneously on different parts of their texts. The *Statusbar*, in the bottom of the screen, displays who is online, and which tools group members are currently using, and thus serves as a tool to raise students' workspace awareness (Gutwin & Greenberg, 2004). Other tools of the VCRI-program, not shown in Figure 2.1, include for example the *Planner*, which can be used to organize and plan group activities, the *Diagrammer*, which can be used to construct argumentative diagrams, and the *Reflector*, which is used by students to reflect on group processes.

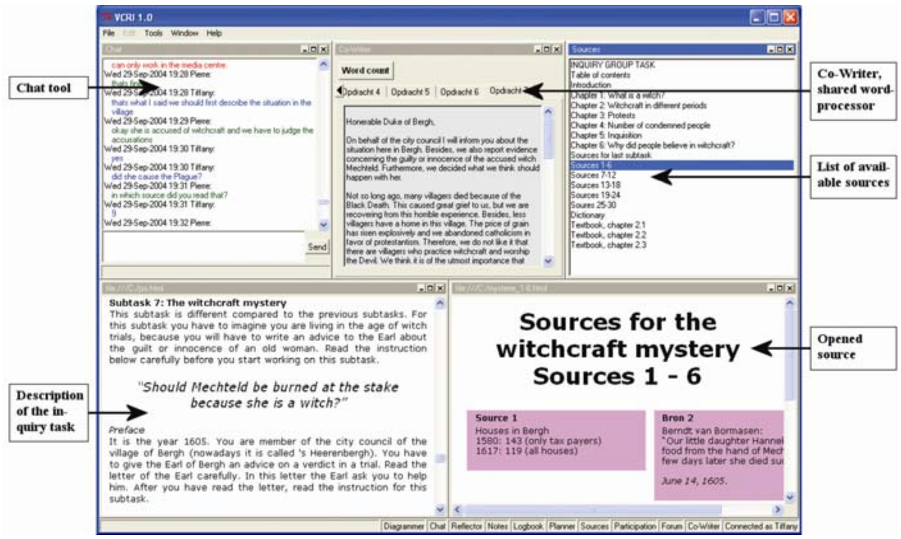


Figure 2.1: Screenshot of the VCRI-program, detailing the most important tools (translated from Dutch).

Inquiry group task

The participating students collaborated on a historical inquiry group task. Inquiry tasks are an important part of the curriculum in the Dutch upper secondary levels. Subject of the task was "Witches and the persecution of witches". This task was developed together with the participating teachers. The task consisted of *seven subtasks* that addressed various aspects of the subject. The introduction of the task stressed the importance of working together as a group on the subtasks, and pointed out that group members were themselves responsible for the successful completion of the task. Students were instructed to use the VCRI program to communicate with group members. Students were told they had eight lessons to complete the inquiry task.

The groups had to use different historical and (more) contemporary *sources* to answer questions about: 1) the definition of a witch, 2) how witches were perceived in different historical periods, 3) protests against the persecution of witches, 4) the number of people that were condemned for witchcraft, 5) the role of the inquisition, and 6) reasons why people believed in witchcraft. Approximately 40 sources from textbooks and the Internet were available to the students through the *Sources* tool.

The final subtask[2] was somewhat different compared to the other subtasks. To complete this subtask, students had to imagine they were 17th century judges. They were asked to write an advice about the guilt or innocence of an old woman. The woman lived in a small village, plagued by the Black Death and famine.

To successfully complete the inquiry group task, all group members had to participate during the group process. *Positive interdependence* and *individual accountability* were incorporated in the group task, in order to ensure collaboration between group members (Johnson & Johnson, 1999). Positive interdependence was realized in different ways. First, because the inquiry task was quite extensive and complex, no group member was likely to solve the task on his or her own. Group members needed to make plans, generate alternatives and solutions, and give or request explanations. Furthermore, students were told they would receive a group grade for their final version of the task and the quality of their collaboration would also partly determine their grade. Finally, for several of the subtasks, group members needed to integrate their perspectives in a jointly written text. Individual accountability was realized in several subtasks. For example, in the final subtask each student was responsible for a specific part of the advice, while the group as a whole was responsible for making the advice into a coherent whole. Thus, although task division between group members could be efficient, the present group task required high levels of collaboration in order to successfully complete the task.

2.4.4 Treatment: Participation Tool

To answer the research questions the VCRI was augmented with a new tool, the *Participation Tool* (PT). The PT visualizes how much each group member contributes to his or her group's *online communication*, through for example the Chat tool shown in Figure 2.1. The PT does not take into account students' activities in other tools, such as the shared word processor.

In the PT, each student is represented by a sphere; group member's spheres are grouped together. For example, Figure 2.2 shows a class of students. The students from this class were assigned to several groups. For instance, the four spheres in the upper right of the Figure represent one group of four students. While students are communicating with each other in the online environment, the PT is continually updated, allowing students to compare their participation rates to those of their group members.

In the PT, the *distance* of a sphere to the group center indicates the *number of messages sent* by the student, compared to the other group members. If a sphere is close to the center, the student has sent more messages compared to a student who is farther away from the center. The *size* of a sphere indicates the *average length of the messages* sent by a student, compared to the other group

[2] This subtask was based on a task developed for the *Active historical thinking* textbook (De Vries, Havekes, Aardema, & Van Rooijen, 2004).

members. If a sphere is smaller, the student has sent shorter messages compared to a student whose sphere is bigger.

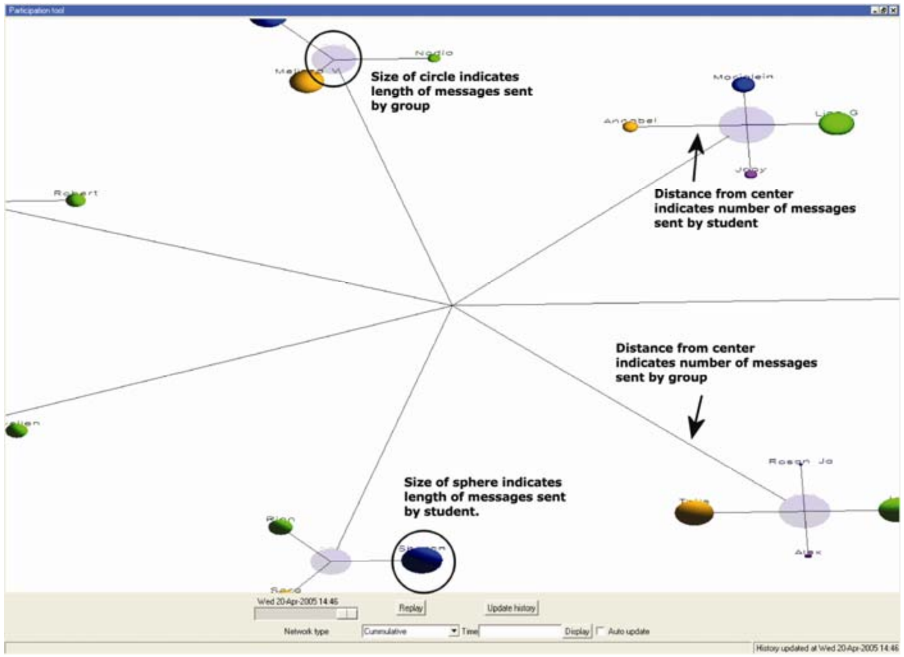


Figure 2.2: Screenshot of the Participation Tool.

Using the PT, students can zoom in, to examine their own group more closely, or zoom out to examine the whole class. In Figure 2.2, the PT is zoomed out, displaying much of the class. This enables students to examine the participation rates of students from other groups. For instance, the *distance of a group* to the center of the whole class indicates the *number of messages* this group has sent, compared to other groups. Thus, a group that has sent many messages is located closer to the center of the class, compared to a group that has sent a few messages. In addition, the *size of the grey circles* in the middle of each group, indicate the *average length of the messages* sent by the groups. If a grey circle is bigger, this group has sent longer messages compared to a group whose grey circle is smaller.

The PT can be opened and examined by students at any time. The visualization can be rotated using the mouse, to examine the visualization from a different perspective. The PT can display students' cumulative participation rates (i.e., total number of messages sent), but can also display a moving average. The moving average displays students' participation rates during the past 20 minutes. After a while, the moving average display is more sensitive to changes than the cumulative display. This is because when students have sent many messages, sending one long message does not influence the cumulative display very much. But in a period of 20 minutes, the number of messages is limited, which means that sending one long message has more impact on the visualization (i.e., the size of the sphere will increase more dramatically). Furthermore, it is important to note that students are *not forced* to use the PT. In order words, it is available and students can use it whenever they want, but they can also choose to ignore or close it whenever they want.

Finally, it should be noted that the PT visualizes the *quantity* of the online communication between group members and the equality of participation between group members. Obviously, the quality of the messages sent by the students is also very important for successful collaboration. The PT does not visualize the quality of the messages sent by the students. Nonetheless, quantity of communication is also important for successful collaboration. For example, when unequal participation exists between group members, this is an indication of free riding behavior. Furthermore, if a group member only types a few messages, he or she cannot be regarded a full-fledged group member, although his or her messages may be of high quality. In short, quality and quantity of collaboration should go hand in hand. Ideally, group members should contribute many, high-quality messages to the online discussion.

2.4.5 Unit of Analysis

To answer the research questions regarding the influence of the PT on participation and collaboration, a decision had to be made regarding the unit of analysis. To determine participation in CSCL settings, researchers have mostly used the number of messages sent (e.g., Adrianson, 2001; Lipponen et al., 2003), or the number of words written (e.g., Savicki et al., 2002; Straus, 1997) as the unit of analysis. However, using the message as a measure of participation can be considered arbitrary in CSCL, since CMC discourse differs considerably from FTF discourse. For example, in synchronous CMC there are fewer conventions about the acceptable length of messages. As a result, some users only send one proposition per chat message, while other users type multiple sentences, combining several propositions (Howell-Richardson & Mellar, 1996). The chat messages sent by the participating students, were therefore segmented into *dialogue acts* (Erkens, 2004). Dialogue acts indicate the communicative function of a chat message (responding, informing, argumentation, commanding, or eliciting). One dialogue act corresponds to one proposition.

The computer program *Multiple Episode Protocol Analysis* (MEPA) was used for the analyses of chat protocols (Erkens, 2003). To segment chat messages, a segmentation filter was used. A filter is a program, which can be specified in the MEPA program for automatic rule based coding or data manipulation. The filter automatically segments message into dialogue acts, using over 300 decision rules. Punctuation marks and connectives (e.g., "and", "but") are used to segment a chat messages into dialogue acts. For example, the chat message "The first answer is okay, but the second is not" is split into two dialogue acts ("The first answer is okay" and "but the second is not"), whereas the message "That's fine" is treated as one dialogue act.

Dialogue acts were also used as the unit of analysis to answer the fourth research question. This was done because chat messages can also refer to multiple collaborative activities. For example, the message "Hi, let's start task 6" contains two communicative functions and can therefore be segmented into "Hi" and "let's start task 6". These two parts also refer to different collaborative activities. The first part is a greeting, whereas the second part refers to the planning of task-related activities.

2.4.6 Use of the Participation Tool

In order to analyze how intensively students used the PT (research question one), all user actions were logged and stored. Based on the log files, two scores were calculated. First, the total number of times a student used the PT (e.g., opening and closing the tool, rotating the view) was calculated. Second, the total time (in minutes) a student displayed the PT on his or her screen was calculated. For example, when a student opened or maximized the PT, and closed or minimized the tool five minutes later, 5 minutes were added to the total time.

2.4.7 Student Participation and Equality of Participation

It is expected that the PT will influence student participation during collaboration. Moreover, it is expected that the PT will lead to more equal participation between group members (research question two).

Measure of student participation. As described above, dialogue acts were used as the unit of analysis to determine participation. A distinction was also made between *long* (>5 words) and *short dialogue acts* (<=5 words). Short dialogue acts are used mainly for back channeling, supporting, and confirming (e.g., "okay", "I agree"), whereas longer dialogue acts are used mainly for transfer of content and regulation of task and group processes. The former can be considered nonsubstantive contributions, since they are less important for the development of the conversation. In contrast, the latter can be considered substantive contributions, which are expected to contribute to whether one is considered a useful participant during conversation (Bonito, 2000).

Measure of equality of participation. To examine the influence of the PT on equality of participation, the Gini coefficient was used. The Gini coefficient is a *group level measure*, which was calculated for all 22 groups involved in this study. The coefficient sums, for each group, the deviation of its group members from equal participation. This sum is divided by the maximum possible value of this deviation (Alker Jr., 1965; Dubrovsky et al., 1991; Warschauer, 1996). Thus, the coefficient ranges between 0 (perfect equality; all students contribute equally to discussion) and 1 (perfect inequality; one student completely dominates discussion).

2.4.8 Awareness of Group Processes and Activities

To measure students' awareness of group processes and activities during collaboration, a questionnaire was administered to the participating students.

Based on the work of other researchers (Gutwin & Greenberg, 2004; Mendoza-Chapa et al., 2000), 14 items were formulated. The items addressed a) awareness of the activities of others in the VCRI, b) awareness of group members' tasks, c) awareness of group members' participation during online collaboration, and d) awareness of conversational processes.

An exploratory factor analysis using principal axis factoring was conducted to identify latent variables underlying the 14 items. Based on the examination of the screeplot and the K1-rule (Hetzel, 1996), two factors were extracted. Using an oblique Promax rotation with a salience level of |.40|, factor one was identified as "Awareness of participation". This factor consisted of four items (e.g., "I knew how much my group members contributed to the discussion") with a Cronbach's alpha of .72. Factor two was identified as "Awareness of group members' tasks". This factor consisted of three items (e.g., "I knew what my group members were working on") with a Cronbach's alpha of .78. In total, the two factors explained 34.30% of the total variance. As can be expected when using oblique rotations, the two factors correlated significantly, r = .62, df = 61, p = .00. Because of this significant correlation, it was examined whether it was possible to extract only one factor instead of two. This was not the case, as the resulting solution could not be interpreted in a meaningful way. It was concluded that the two-factor solution was more meaningful, because the two factors seemed to represent two different types of awareness. Factor scores were subsequently used in analyses of differences between treatment and control groups.

2.4.9 Collaborative Activities

To answer the fourth research question, regarding the influence of the PT on students' collaborative activities, a coding scheme was developed. The aim of this coding scheme was to provide insight into the task- and group-related processes taking place between students while working on the inquiry group task. This section describes the development and interobserver reliability of the coding scheme.

Description of the coding scheme. As described above, and summarized in Table 2.1, different types of activities are necessary to successfully complete a group task. These types of activities are reflected by the four different dimensions of the coding scheme. Each dimension contains two or more coding categories. In total, the scheme consists of 19 categories.

The first dimension referred to *performance of task-related activities*. This dimension contained two categories pertaining to the discussion of relevant task-related information: exchanging and sharing task-related information (*TaskExch)* and asking task-related questions (*TaskQues*). In brackets, the abbreviations of the codes are given. These abbreviations will be used from time to time in the analyses presented below.

The second dimension referred to *regulation and coordination of task-related activities*, encompassing four categories. First, planning (*MTaskPlan*) involved discussion of strategies necessary to complete the task, choice of appropriate strategies, and delegation of task responsibilities. Second, monitoring (*MTaskMoni*) involved exchange of information that could be used to monitor task performance and progress, and assessing the amount of time available. Finally, evaluation involved appraisal and discussion of task performance and progress, which could be positive (*MTaskEvl+*) or negative (*MTaskEvl-*).

Performance of social activities was the third dimension of the coding scheme. This dimension contained five categories. First, greetings (*SociGree*) were included, since they contribute positively to group atmosphere and a feeling of social presence (Rourke et al., 1999). Second, social support remarks (*SociSupp*) referred to comments that contributed positively to group atmosphere, such as exchanging positive comments, displaying positive emotions, and disclosure of personal information. Third, social resistance remarks (*SociResi*) referred to behaviors that contributed negatively to group atmosphere, such as insulting group members and displaying negative emotions. Fourth, shared understanding (*SociUnd+*) referred to confirmations, acceptances, and indications of agreement, which serve to reach and maintain shared understanding during collaboration. Similarly, loss of shared understanding (*SociUnd-*) referred to denials, disagreements, and expressions of incomprehension.

Table 2.2: Collaboration acts (abbreviation) and category Kappa's.

	Task-related activities		*Social activities*	
	Codes	κ	Codes	κ
Perfor-mance	Info exchange (*TaskExch*)	.85	Greetings (*SociGree*)	.89
	Asking questions (*TaskQues*)	.89	Social support (*SociSupp*)	.85
			Social resistance (*SociResi*)	.73
			Mutual understanding (*SociUnd+*)	.92
			Loss of mutual underst. (*SociUnd-*)	.83
Coordi-nation / regulation	Planning (*MTaskPlan*)	.87	Planning (*MSociPlan*)	.86
	Monitoring (*MTaskMoni*)	.81	Monitoring (*MSociMoni*)	.84
	Positive evaluations (*MTaskEvl+*)	.84	Positive evaluations (*MSociEvl+*)	-
	Negative evaluations (*MTaskEvl-*)	1.00	Negative evaluations (*MSociEvl-*)	.88
Other	Neutral technical (*TechNeut*)	1.00	Other / nonsense (*Other*)	.67
	Negative technical (*TechNega*)	-		
	Positive technical (*TechPosi*)	-		

The fourth dimension referred to *regulation and coordination of social activities*. This dimension contained four categories. First, planning (*MSociPlan*) involved discussion of collaboration strategies instead of discussion of task-related strategies, such as helping each other or proposals to work together on certain tasks. Second, monitoring (*MSociMoni*) referred to the exchange of information that could be used to monitor group processes. Finally, evaluation involved appraisal and discussion of group processes and collaboration, which could be positive (*MSociEvl+*) or negative (*MSociEvl-*). These four categories reflect *group processing*; they indicate group members discuss how well their group is performing and how collaboration can be improved. It was expected that the PT would stimulate group members to engage more in these types of activities.

Statements that addressed neutral, negative, or positive technical aspects of the CSCL environment were also included in the coding scheme (codes *TechNeut*, *TechNega*, and *TechPosi*). Although these statements can refer to task-related activities, it was decided to separate them from the other codes, because they are different in nature and focus (i.e., they are aimed more at the discussion of the CSCL environment, instead of discussion of the task). Finally, statements that did not fit into any of the previously mentioned categories were coded as *Other*. These codes mostly referred to nonsense remarks.

Interobserver reliability. Two researchers were involved in the development and refinement of the coding scheme. In order to examine interrater agreement 601 dialogue acts were coded independently by two raters. An overall Cohen's Kappa of .86 was found, a satisfactory result. The category Kappa's (Cicchetti, Lee, Fontana, & Dowds, 1978) are shown in Table 2.2. Note that for the codes *MSociEvl-*, *TechNega*, and *TechPosi* it was impossible to compute a category Kappa, since these codes were not present in the dialogue acts coded by the raters. However, since most other category Kappa's are satisfactory, it can be expected that the category Kappa's for these three codes are also sufficient.

2.4.10 Group Performance Scores

The quality of the texts written by the groups for the subtasks of the inquiry group task was examined in order to answer the final research question regarding the influence of the PT on group performance. For this purpose an assessment form was developed.

Using this assessment form, three quality aspects were assessed for each subtask. *Use of sources* referred to the manner how students incorporated the available sources into their texts. This quality aspect contained two items: completeness of sources used in the written text, and copy-pasting of information from sources to the written texts. *Content and argumentation* referred to the manner in which students formulated their answers and supported their answers with arguments. Since each subtask addressed different aspects of witchcraft, the content and argumentation aspect was formulated differently for each subtask. However, the scoring-scale was the same for each aspect. The amount of items that addressed content and argumentation also differed for each subtask, since some subtasks were more extensive than others. *Text construction and language* referred to the manner how students' written text had an adequate text construction and correct language. This quality aspect contained three items: text construction, text structure, and correctness of language use.

All items of the assessment form were answered on a 3-point scale, with 0 indicating poor quality and 2 good quality. To determine whether students directly copy-pasted information from the sources to their text, the program WCopyFind (http://plagiarism.phys.virginia.edu/Wsoftware.html) was used. This program compares written texts to the available sources and determines how many percent of the written text is copy-pasted directly from the sources. This percentage was used to determine whether the group received 0, 1, or 2 points. Groups that copy-pasted less than 34% of their text from the sources, received 2 points; groups that pasted more than 66% of their text received no points. In total, groups could receive up to 12 points for subtasks one, five and seven, 14 points for subtasks two, four and six, and 18 points for subtask three. Thus, in total 96 could be earned. In the analyses presented below, results will be presented for total points earned and points earned for each subtask.

To check the objectivity of the scoring, two researchers independently scored a number of tasks. Each researcher filled out the assessment form for 8 to 10 groups for four subtasks. For use of sources, content and argumentation, and text construction and language, interrater agreement reached 88.5%, 77.5%, and 75.0% respectively. Furthermore, to examine the internal consistency of the scoring procedure, Cronbach's alpha was calculated and was found to be .81.

2.4.11 Procedure

The participating students worked in small groups on the inquiry group task for a period of four weeks. In the first lesson, the task was introduced to the students by their teachers. During this lesson, the most important features of the CSCL-environment were also explained to the students by the experimenters. After the first lesson, another seven history lessons were devoted to the inquiry group task. During these lessons, the students collaborated on the group task. The teachers were standby to answer task-related questions, while the experimenters were standby to answer technical questions or to solve any technical problems. The students were allowed to work on the inquiry task during free periods. For example, students could work in the media center when they had spare time in their timetable. However, students could only access the CSCL-environment from school, not from their homes. After eight lessons the students were required to hand in their final versions of the group task. These final versions were graded by their teachers. After the last lesson, a questionnaire was administered to the students in order to assess students' awareness of collaborative processes and activities during online collaboration.

2.4.12 Data Analysis

To investigate the effects of the PT on student participation during CSCL, one solution would be to compare the participation rates of students who used the PT to the participation rates of students who did not use the PT, using an independent samples *t* test with *participation* as a dependent variable and *condition* (PT or no PT) as an independent variable. However, it is important to note that students' participation rates are most likely *nonindependent* (Bonito, 2002; Kenny, Mannetti, Pierro, Livi, & Kashy, 2002). According to Kenny et al. (2002), *mutual influence* is the most important source of nonindependence when students collaborate. That is, what one group member says, is influenced by, and influences the contributions of other group members. Therefore, students who are in the same group behave in more or less similar ways. Thus, it can be expected that students who are, for example, in a group with highly active group members, will also be stimulated to participate more; whereas students in groups with low participating group members will participate less.

When nonindependence exists between group members, this violates one of the assumptions of the independent samples *t* test (and other statistical methods, such as ANOVA and regression analysis). Since the independent variable (condition) is a between-groups independent variable (i.e., its value is the same for all group members, but differs across groups), nonindependence makes a *t* test too liberal, thus resulting in an increase of Type I errors (Bonito, 2002; Kenny et al., 2002; Snijders & Bosker, 1999). Therefore, using a *t* test to determine the effect of the PT would not be appropriate.

Presence of nonindependence can be examined by computing the intraclass correlation coefficient for each dependent variable. This correlation represents the dependency between scores of students in the same group. This coefficient can also be tested for significance. However, Kenny et al. (2002) argued that, given the usual small sample sizes in small group research, the correlation coefficient may not be significant, even though it is actually large enough to bias the *t* test. Kenny et al. (2002) therefore propose to assume the data are nonindependent, even though the intraclass correlation coefficient may not be significant. As a result, *multilevel analysis* was used to examine the effects of the PT, since this type of analysis can be used when data have a hierarchically nested structure (e.g., students nested within groups) and nonindependence is present.

Multilevel analysis involved estimating two models: an empty model and a model including predictor variables. For both models, the deviance (a measure of the goodness of fit of the model) was computed. By comparing the deviance of the latter model with the empty model, a decrease in deviance can be calculated. When this decrease in deviance is significant (tested with a χ^2-test), it can be concluded the latter model is a better model. Furthermore, the estimated parameters of the predictors can be tested for significance by dividing the regression coefficient by its standard error. This so-called *t*-ratio has approximately a standard normal distribution (Snijders & Bosker, 1999).

The line of reasoning concerning the nonindependence of students' participation rates can also be extended to the other individual measures in this study (research questions three and four). Therefore, the effects of the PT on students' awareness and the manner in which students collaborated will be examined using multilevel analysis as well.

2.5 Results

2.5.1 Use of the PT

The first research question concerned how intensively students used the PT. On average, treatment group students used the PT 76.04 times (*SD* = 48.03, *Min* = 9, *Max* = 286), and displayed the PT for 64.33 minutes (*SD* = 47.89, *Min* = 0.52, *Max* = 186.80) on their screen. Since the average time a student was online in the VCRI environment was 361.01 minutes (*SD* = 79.65), most students displayed the PT a considerable amount of time (18%) on their screen and used the PT on a regular basis (about once every 5 minutes).

The large standard deviations for use of the PT and display time of the PT show that there were considerable differences between students in the extent to which the PT was actually used. Thus, to examine whether how intensely the PT was used, influenced the dependent measures (e.g., participation, awareness), correlations were calculated between use of the PT and display time of the PT and the dependent measures. Because the total time students were online correlated significantly positively with use of the PT and display time of the PT ($r = .36$, $p = .01$, and $r = .34$, $p = .01$, respectively), we calculated *partial correlations*, which controlled for time online.

2.5.2 Participation and Equality of Participation

Table 2.3 shows the descriptive statistics and effect sizes[3] for differences between treatment and control groups for participation (research question two). Overall, the mean scores of treatment group students were higher compared to control group students.

Table 2.3: Means, standard deviations, and effect sizes for measures of participation for treatment and control groups.

	Treatment group students (*N* = 52)		Control group students (*N* = 17)		
Measure of participation	*M*	*SD*	*M*	*SD*	*ES*
Dialogue acts	301.21	159.86	235.24	75.32	0.46
- Long dialogue acts (> 5 words)	114.08	70.99	72.89	30.39	0.65
- Short dialogue acts (<= 5 words)	187.13	96.69	162.35	60.53	0.28

Before examining the effect of the PT on participation, it was investigated whether there were differences between the two conditions regarding the time students were online. Treatment group students were not longer online (*M* = 370.68 minutes, *SD* = 86.01), compared to control group (*M* = 358.28

[3] Effect sizes (*ES*s) were calculated using Carlson and Schmidt's (1999) formula for a post-test only with control group design.

minutes, *SD* = 68.67) students, *t*(21) = 0.33, *p* = .37. Although these differences were not significant, the total time a student was online was used as an individual level predictor variable. This was done to account for the fact that some group members worked longer in the CSCL-environment than others. For example, some students worked longer because they also worked during free periods, whereas other students worked shorter because they were ill during a lesson.

Table 2.4 shows the results of the multilevel analysis for the three measures of participation. For all measures of participation, time online was a significant predictor. This indicates that students who were online longer, produced more dialogue acts, *t*(68) = 2.90, *p* = .00, more long dialogue acts, *t*(68) = 1.89, *p* = .03, and more short dialogue acts, *t*(68) = 3.15, *p* = .00.

For the total number of dialogue acts, the effect of condition was not significant. Students who used the PT did not produce more dialogue acts compared to students without the PT, *t*(21) = 1.11, *p* = .14. Furthermore, students who had access to the PT did not type more short dialogue acts, compared to students who did not have access to the PT, *t*(21) = 0.68, *p* = .25. However, students who had access to the PT, typed more long dialogue acts, compared to students who did not have access to the PT, *t*(21) = 1.76, *p* = .05.

Table 2.4: Results of multilevel analysis for measures of participation.

	Dialogue acts		Long dialogue acts		Short dialogue acts	
	Coeff.	*SE*	Coeff.	*SE*	Coeff.	*SE*
Intercept	55.74	78.90	30.02	35.78	27.76	48.72
Predictor 1: Minutes online	0.58**	0.20	0.17*	0.09	0.41**	0.13
Predictor 2: Condition	27.34	24.64	19.11*	10.83	8.34	14.94
Variance group level	5563.73	2993.13	990.44	584.81	1952.52	1108.54
Variance student level	1245.26	2569.31	2688.50	554.60	4883.54	1006.67
Total variance explained (%)	13.44		12.57		12.75	
Deviance	865.66		757.52		799.68	
Decrease in deviance compared to empty model (*df* = 2)	8.93*		6.17*		10.08**	

The effect of the PT on equality of participation was examined, using the Gini coefficient. Since this dependent variable is a *group level* variable, a *t* test for independent samples was used to examine differences between treatment and control group. As can be seen in Table 2.5, the Gini coefficients were not far from zero, indicating a, more or less, equal distribution of participation among group members. Note that the number of observations is different in this Table, compared to Table 2.3, because groups are analyzed instead of students. On average, the Gini coefficients are higher for the treatment groups, indicating slightly more inequality of participation. Moderate ESs were found, although the differences did not reach statistical significance.

To examine how the use of the PT was related to students' participation during collaboration, a number of partial correlations (e.g., between times the PT was used and dialogue acts typed) were calculated. One significant partial correlation emerged: the total time the PT was displayed correlated significantly negatively with the Gini coefficient for long dialogue acts ($N = 14$, $r = -.62$, $p = .01$). This indicates that in groups that displayed the PT more on their screen, the number of long dialogue acts was more equally distributed.

Table 2.5: Means and standard deviations for treatment and control groups for equality of participation (group as unit of analysis).

	Treatment groups ($N = 17$)		Control groups ($N = 5$)				
Equality of participation	*M*	*SD*	*M*	*SD*	*t*	*p*	*ES*
Dialogue acts	.17	.09	.09	.04	1.73	.10	0.88
- Long dialogue acts (> 5 words)	.18	.11	.14	.05	0.81	.43	0.41
- Short dialogue acts (<= 5 words)	.17	.09	.12	.05	1.05	.31	0.54

Note. The *group level* measure of equality used, the Gini coefficient, ranges from 0 to 1, with 0 indicating perfect equality and 1 perfect inequality.

2.5.3 Awareness of Group Processes and Activities

The third research question concerned the effect of the PT on students' awareness of the group processes and activities taking place during online collaboration. One treatment and two control group students failed to complete the questionnaire. Table 2.6 shows the findings for students' awareness of group processes and activities. Overall, students indicated medium to high awareness of group processes and activities. Apparently, students were aware of what was happening in the VCRI, and what their group members were doing.

Multilevel analysis was used to examine differences between treatment and control group students by estimating two models. Again, the empty model was estimated first. Second, a model that included condition (PT or no PT) as a predictor was estimated. The results of these analyses are also presented in Table 2.6.

On average, students with access to the PT reported lower awareness of participation, but higher awareness of group members' tasks. Overall, the effect of the PT on the two factors was not significant, $t(21) = -1.22$, $p = .12$, and $t(21) = .89$, $p = .19$, respectively. For the sake of completeness, the effect of the PT was also examined for the seven items that had low pattern and structure coefficients on the two factors. Only for one item, condition was a significant predictor. Students with access to the PT reported they knew better when someone was not working hard, than students without access to the PT, $t(21) = 2.43$, $p = .01$.

Examination of the partial correlations between intensity of PT use and the two awareness measures revealed no significant relationships. Thus, how many times students used the PT, or how long students displayed the PT, did not influence students' awareness of participation and group members' tasks.

Table 2.6: Descriptive statistics indicating treatment and control group students' awareness of group processes and activities, and results of multilevel analyses with condition as predictor.

	Treatment group students (N = 47)		Control group students (N = 15)				
Factor	*M*	*SD*	*M*	*SD*	*B*	*SE*	χ^2
Awareness of participation	-0.09	0.88	0.29	0.96	-0.182	0.149	1.40
Awareness of group members' tasks	0.06	0.83	-0.19	1.15	0.138	0.155	0.77

Note. Factor scores range from negative (= *no awareness*) to positive (= *high awareness*). * $p < .05$.

2.5.4 Collaborative Activities

The fourth research question concerned the effect of the PT on the collaboration processes taking place between students during online collaboration. In Table 2.7, the mean frequencies of collaboration acts per student are presented. For the descriptions of the collaboration acts, the reader is referred to Table 2.2 and the method and instrumentation section above. The numbers in parenthesis indicate how many *percent* of the total number of collaboration acts was devoted to a specific collaboration act. The data in Table 2.7 show most of the collaboration acts to indicate signaling and monitoring mutual understanding (*SociUnd*+, 22%). Furthermore, many collaboration acts involved regulating the completion of the group task, such as making plans (*MTaskPlan*, 19%), or monitoring task progress (*MTaskMoni*, 13%). In order to complete the task, the students exchanged a lot of task related information (*TaskExch*, 9%), but also paid attention to the social aspect of collaboration by sending many social support remarks (*SociSupp*, 7%).

Table 2.7: Mean frequencies and standard deviations of collaboration acts and multilevel analyses of the effects of condition.

	Treatment group students (N = 52)			Control group students (N = 17)			Total (N = 69)			Effect of condition		
	M			*M*			*M*					
	Freq.	(%)	*SD*	Freq.	(%)	*SD*	Freq.	(%)	*SD*	*B*	*SE*	χ^2
Performing task-related activities												
- Info exchange (*TaskExch*)	34.87	(9.76)	45.64	18.06	(7.70)	16.99	30.72	(9.25)	41.03	1.985	5.270	39.62**
- Asking questions (*Taskques*)	7.88	(2.59)	8.67	3.82	(1.62)	3.07	6.88	(2.35)	7.85	1.102	1.046	18.05**
Coordinating task-related activities												
- Planning (*MTaskPlan*)	56.62	(19.35)	33.97	39.41	(16.94)	18.98	52.38	(18.76)	31.72	2.455	2.306	95.75**
- Monitoring (*MTaskMoni*)	37.50	(13.05)	22.18	32.59	(14.51)	9.14	36.29	(13.41)	19.83	-1.256	1.612	75.58**
- Positive evaluations (*MTaskEvl*+)	4.62	(1.50)	3.92	2.59	(1.13)	1.91	4.12	(1.41)	3.63	0.543	0.434	24.43**
- Negative evaluations (*MTaskEvl*-)	5.46	(1.76)	4.92	4.41	(1.73)	3.41	5.20	(1.75)	4.60	-0.130	0.507	33.85**
Performing social activities												
- Greetings (*SociGree*)	11.17	(4.21)	7.39	5.88	(2.55)	3.87	9.87	(3.80)	7.05	1.913*	1.015	21.02**

	Treatment group students (*N* = 52)			Control group students (*N* = 17)			Total (*N* = 69)			Effect of condition		
	M Freq.	(%)	*SD*	*M* Freq.	(%)	*SD*	*M* Freq.	(%)	*SD*	*B*	*SE*	χ^2
- Social support (*SociSupp*)	18.12	(5.88)	14.06	26.06	(10.45)	16.39	20.07	(7.01)	14.94	-6.493**	1.751	45.40**
- Social resistance (*SociResi*)	3.63	(1.28)	3.70	8.71	(3.66)	8.86	4.88	(1.87)	5.79	-2.889**	0.830	14.36**
- Mutual understanding (*SociUnd*+)	67.56	(22.57)	41.35	49.29	(21.20)	22.71	63.06	(22.23)	38.30	2.009	2.994	73.23**
- Loss of mutual understanding (*SociUnd*-)	9.19	(3.37)	4.88	11.00	(4.77)	6.07	9.64	(3.72)	5.21	-1.633**	0.575	32.31**
Coordinating social activities												
- Planning (*MSociPlan*)	6.44	(2.31)	4.65	2.82	(1.15)	2.16	5.55	(2.03)	4.45	1.337*	.543	28.23**
- Monitoring (*MSociMoni*)	17.40	(6.23)	10.94	11.82	(5.16)	5.43	16.03	(5.97)	10.13	1.135	1.567	37.34**
- Positive evaluations (*MSociEvl*+)	0.58	(0.19)	0.85	0.41	(0.17)	0.71	0.54	(0.19)	0.81	0.030	0.110	5.90
- Negative evaluations (*MSociEvl*-)	0.98	(0.37)	1.39	0.35	(0.17)	1.00	0.83	(0.32)	1.33	0.234	0.252	7.17*

	Treatment group students (*N* = 52)			Control group students (*N* = 17)			Total (*N* = 69)			Effect of condition		
	M Freq.	(%)	*SD*	*M* Freq.	(%)	*SD*	*M* Freq.	(%)	*SD*	*B*	*SE*	χ^2
Technical												
- Neutral technical (*TechNeut*)	6.94	(2.66)	5.22	6.82	(2.64)	5.36	6.91	(2.65)	5.22	-0.433	0.778	11.55**
- Negative technical (*TechNega*)	3.92	(1.36)	2.96	4.00	(1.47)	4.56	3.94	(1.39)	3.38	-0.514	0.443	26.24**
- Positive technical (*TechPosi*)	1.10	(0.34)	1.62	0.71	(0.25)	1.05	1.00	(.31)	1.50	0.025	0.231	17.12**
Other	3.19	(1.20)	5.01	6.47	(2.74)	3.86	4.00	(1.58)	4.93	-1.860**	0.660	7.56*

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

Table 2.7 lists the results of the multilevel analyses. Number of dialogue acts typed was a significant predictor for all collaboration acts, except *Other*. For example, the more dialogue acts a student typed, the more questions he or she asked (*TaskQues*).

Table 2.7 also shows that condition was a significant predictor for several types of collaboration acts, independent of the number of dialogue acts typed. First, having access to the PT was related significantly to the number of greetings (*SociGree*) typed by a student, $t(21) = 1.89$, $p = .04$. Students who had access to the PT sent significantly more greetings. Second, the coefficient for *SociSupp* indicates a negative effect of the PT on the number of social support remarks typed by a student, $t(21) = -3.71$, $p = .00$. Thus, students who had access to the PT made significantly less social support remarks. Furthermore, the PT also had a significantly negative effect on the number of social resistance remarks (*SociResi)*, $t(21) = -3.48$, $p = .00$. When students had access to the PT, they typed significantly less social negative remarks. Fourth, the PT influenced the number of messages which signaled loss of mutual understanding (*SociUnd-*), $t(21) = -2.84$, $p = .00$. Students with access to the PT, indicated less incomprehension and disagreement. Fifth, the PT had a positive effect on the number of remarks aimed at planning of group processes *(MSociPlan*), $t(21) = 2.46$, $p = .01$. Thus, students with access to the PT constructed more messages that regulated the planning of group processes. This indicates that students who had access to the PT devoted more of their online discussion to this aspect of group processing. Finally, access to the PT influenced the number of nonsense (*Other*) remarks typed by a student, $t(21) = -2.82$, $p = .01$. Students with access to the PT sent significantly less *Other* remarks. As mentioned earlier, number of dialogue acts was not a significant predictor for *Other* remarks, therefore a model which only included condition as a predictor was also estimated. Condition was also a significant predictor in this model, $t(21) = -2.50$, $p = .01$. Thus, it seems that the PT influenced students to type less nonsense remarks.

To examine how operating the PT influenced students' collaboration, partial correlations (controlling for the time students were online) between use of the PT and collaborative activities were inspected. However, no significant relationships were found.

2.5.5 Group performance scores

The final research question concerned the effects of the PT on groups' performance scores. Table 2.8 shows the results of the comparison between treatment and control groups. Note that since the performance scores were calculated for groups, not individual students, the number of observations is different from, for example, Table 2.8.

Differences were tested using a *t* test for independent samples. On average, treatment groups attained higher total performance scores, but the difference was not significant, and the resulting *ES* was small. Inspection of Table 2.8 shows that on average, treatment groups scored higher compared to control groups on four of the seven subtasks. None of the differences were statistically significant. However, for subtasks two and six, the *ES*s (.80 and .92, respectively) were rather high and in favor of the treatment groups. Finally, inspection of the correlations between use of the PT and group performance scores, revealed no significant relationships.

Table 2.8: Means and standard deviations for treatment and control groups for group performance scores.

Group performance scores	Treatment groups (N = 17)		Control groups (N = 5)				
	M	*SD*	*M*	*SD*	*t*	*p*	*ES*
Total score	*66.47*	*10.66*	*64.80*	*8.70*	*0.32*	*.75*	*0.16*
- Subtask 1	1.37	0.33	1.47	0.18	-0.60	.55	-0.33
- Subtask 2	1.38	0.32	1.14	0.20	1.54	.14	0.80
- Subtask 3	1.27	0.34	1.47	0.27	-1.16	.26	-0.61
- Subtask 4	1.39	0.36	1.49	0.26	-0.53	.60	-0.29
- Subtask 5	1.39	0.30	1.27	0.15	0.89	.38	0.43
- Subtask 6	1.42	0.24	1.20	0.24	1.81	.09	0.92
- Subtask 7	1.51	0.37	1.40	0.45	0.55	.59	0.28

Note. To increase comparability, scores for the subtasks were standardized, with 0 being the minimum amount of points and 2 the maximum.

2.6 Conclusions

In the present study, the effects of visualization of participation during CSCL were examined. A CSCL-environment was augmented with the Participation Tool (PT). The PT visualizes how much each group member contributes to his or her group's online communication. It is assumed that the PT would influence group members' participation, awareness, collaborative activities, and groups' performance scores.

The first research question investigated how intensively students used the PT during online collaboration. Treatment group students used the PT quite intensively, although some students used the PT very little. On average, students displayed the PT on their screen 18% of the total time they were online.

The second research question was whether students who had access to the PT (treatment group students) would participate more, and more equally during collaboration compared to students without access to the PT (control group students). The results show an effect of the PT on the participation of treatment group students. Treatment group students sent more long dialogue acts (messages containing more than five words), compared to control group students. This important since longer messages are mainly used for transfer of information and coordination of activities. These longer, more substantial messages contribute to satisfactory collaboration (Bonito, 2000). Regarding short dialogue acts, and total number of dialogue acts, no differences were found between treatment and control group students. Furthermore, equality of participation was not higher in treatment compared to control groups. However, display time of the PT on students' screens was correlated with equality of participation for long dialogue acts. This indicates that operating the PT has an effect on equality of participation. In conclusion, some positive results were found for the effect of the PT on participation during online collaboration. Having access to the PT and operating the PT seem to partly stimulate participation and equality of participation.

The third research question was whether the PT would influence students' awareness of the group processes and activities taking place during online collaboration. To answer this question, students completed a 14-item questionnaire. From this questionnaire, seven items were grouped into two factors. The results indicate treatment group students reported similar levels of awareness of participation and group members' tasks compared to control group students. Thus, the PT did not stimulate students' awareness of the group processes and activities taking place during collaboration. Furthermore, operating the PT was not correlated with awareness. In sum, the PT did not affect students' awareness of the group processes and activities in their group. It is worth noting however, that on a single questionnaire item that was not included in one of the two extracted factors, a significant difference emerged: treatment group students reported they knew better when a group member was not working hard, compared to control group students. This indicates that the PT may have helped students to assess whether a group member was taking a free ride.

The fourth research question was whether the PT would influence students' collaborative activities during online collaboration. It was expected that the PT would stimulate students to engage in social activities and coordination of social activities. That is, treatment group students were expected to spend more time planning, monitoring, and evaluating their collaboration (group processing). This proved to be partially the case. Treatment group student made more remarks aimed at planning of group processes. Furthermore, treatment group students made fewer remarks that indicated social resistance (e.g., swearing, displaying negative emotions). This is important because such negative behavior undermines collaboration (Jehn & Shah, 1997). Surprisingly, treatment group students also typed fewer social support remarks (e.g., offering positive comments, self-disclosure). Instead, treatment group students typed more greetings. In addition, the treatment group students signaled loss of mutual understanding on fewer occasions. Finally, treatment group students typed less nonsense remarks. Intriguingly, no relationship between use of the PT and collaborative activities was found. This may indicate that even students that did not use the PT intensively, were influenced by the PT. Possibly because they were influenced by group members that *did* use the PT more intensively. Overall, these results indicate that access to PT stimulated students to invest more effort into coordination of social activities.

The final research question was whether the PT would increase the efficacy of collaboration by improving group performance. It was hypothesized that through stimulating participation, equality of participation, awareness of group processes and activities and coordination of social activities, the PT would help groups to perform better on the inquiry group task. The results indicate that this was not the case. Treatment groups did not attain higher performance scores than control groups. Furthermore, partial correlations revealed no relationship between operating the PT and group performance. It is worth noting that for two subtasks moderate to high effects sizes were found in favor of the treatment group. In conclusion, the effects of the PT on group performance were not as profound as expected.

Overall, the results of this study demonstrate the usefulness of visualizing participation during CSCL. It stimulates students to exchange longer, more substantial messages. These results are in line with previous studies by Michinov and Primois (2005) and Zumbach et al. (2004), although these studies used different types of visualizations. In addition, operating the PT influenced equality of participation. In groups that used the PT a lot, the number of long dialogue acts was distributed more equally among group members. Finally, visualization of participation stimulated students to collaborate differently. It helped group members to engage in group processing, but also to decrease off-task and negative behavior (e.g., typing nonsense messages and swearing). This is important, since it demonstrates that the increase in participation is not caused by students sending more nonsense messages in order to manipulate the visualization. Effects of the PT on awareness of group processes and –activities and on group performance were not as strong as expected. Groups that had access to the PT did not report higher awareness and did not earn higher performance scores on the inquiry group task.

2.7 Discussion

In interpreting the results of this study, some possible limitations should be kept in mind. These limitations and alternative interpretations of the results will be discussed hereafter. First, the data were collected on a single school. In addition, the sample used in this study was rather small: 55 treatment and 17 control group students. This may limit the generalizability of the results

Second, students were not forced to use to PT. They could open and close it whenever they wanted. Unsurprisingly, some students used the PT very little, whereas others used it a lot. If the PT would have been on students' screens all the time, this might have produced different, perhaps more positive, effects. However, this study demonstrated that only giving students access to the PT could still be beneficial.

Third, a possible explanation for the mixed results could be that it was difficult for students to perceive differences between the sizes of the spheres in the PT. If students do not perceive differences between each other in terms of participation, they may not change their behavior because they believe they are collaborating efficiently. This study did not investigate whether students actually perceived differences. However, examination of the online discussions revealed that students themselves mentioned and were aware of these differences (e.g., "your sphere is the biggest", "according to the PT, she participated the most"). These examples illustrate that these students were able to perceive differences between each other.

Fourth, the type of group task used in this study may have influenced the results found. This study used an inquiry group task, for which a high level of collaboration and equal participation was necessary to perform well. Positive interdependence and individual accountability were incorporated in the task, thus reducing the chance that free riding would occur (O'Donnell, 2001). As was described above, in both conditions participation was rather equally distributed among group members, indicating there were few free riders. With a different task, one in which free riding is more likely to occur, the effects of the PT may be stronger because free riders are more easily identified. Future research could investigate whether different types of task elicit different effects of the PT.

Fifth, because students were not forced to use the PT, there were considerable differences between students regarding the extent to which the PT was actually used. These differences may have made it more difficult to detect differences between treatment and control group students. However, it was found that operating the PT only had an effect on equality of participation. Thus, it seems that it is more important to be in a group that has access to the PT, than to actually observe and operate the PT. This may seem counterintuitive, but two explanations can account for this finding. Firstly, the measures used may not completely grasp whether students meaningfully used the PT. The total amount of time the PT was displayed on students' screen may not be a good indicator, because students can have the PT on their screen without paying attention to it. Using the PT may not be a good indicator, because even without using the tool, students can still be influenced by it, by simply observing the tool. Secondly, students can be influenced by the PT through their group members. As described above, students with access to the PT typed more long messages and engaged more in coordination of social activities. By observing the increased participation, or different collaborative activities of group members, even students that did not use the PT intensively may have been prompted to change their collaborative behavior.

Finally, group size may have influenced the results (Strijbos, Martens, & Jochems, 2004). This study used groups with three or four members. Bonito (2000) noted that smaller groups minimize participation differences, possibly because in smaller groups the obligation to participate is higher, lack of participation can be noticed more easily, and there is less competition for attention. Therefore, if larger groups had been used in this study, the results might have been different. In larger groups, there is more competition for attention and free riders are more likely to go unnoticed. In these groups, the PT could possibly have a greater impact on equality of participation. Similarly, in larger groups it is more difficult to know what group members are doing and which group members are participating too little. Therefore, under these circumstances the PT could possibly have a greater impact on awareness of group processes and activities. In the future research, it should be examined whether group size influences the effects of visualization of participation.

Overall, the results of this study were quite positive. In this case, access to visualization of participation seemed to stimulate participate more during online collaboration. Furthermore, students also discussed more about manner in which they were collaborating, which may help them to collaborate better. Whether these results can be replicated with other students, other types of groups or using different types of tasks, remains to be seen, although the results seem promising. In our own future research, we will explore the merits of visualization during collaboration further.

3. Visualization of Agreement and Discussion Processes during Computer-Supported Collaborative Learning[1]

This study examined the effects of the *Shared Space* (SS) on students' behaviors in a computer-supported collaborative learning (CSCL) environment. The SS visualizes discussion and agreement during online discussions. It was hypothesized the SS would increase the media richness of the CSCL-environment, would stimulate critical and exploratory group-norms, would lead to more positive perceptions of online collaboration, and would have an impact on students' collaborative activities. In total, 59 students working in 20 groups had access to the SS visualization, while 58 students working in 20 groups did not. The results show that students with access to the SS visualization, a) perceived higher media richness, b) had a more exploratory group-norm perception, c) perceived more positive group behavior, d) perceived their group's task strategies to be more effective, e) engaged in different collaborative activities, and f) performed better on one part of the group task. These results demonstrate the potential benefits of visualizing agreement and discussion during CSCL.

Keywords: Computer-supported collaborative learning; Collaboration; Computer-mediated communication; Visualizations; Secondary education

3.1 Introduction

Teachers and students are increasingly using ICT to facilitate learning of various subjects (Lou et al., 2001). Computer-supported collaborative learning (CSCL) is one application of ICT that has received considerable attention by educational researchers. CSCL aims to provide students with an environment that supports and enhances collaboration, in order to facilitate their learning processes (Kreijns et al., 2003). During CSCL students usually communicate with group members using discussion forums or chat rooms.

Several studies demonstrated CSCL to be an effective tool for education. For example, a meta-analysis by Lou et al. (2001) found that combining small group learning with ICT was more effective than combining individual learning with ICT. Additionally, Cavanaugh (2001) demonstrated the effectiveness of interactive distance education technologies. Thus, the perceived potential of CSCL seems to be, at least partially, supported by research.

---

[1] Janssen, J., Erkens, G., & Kanselaar, G. (2007). Visualization of agreement and discussion processes during computer-supported collaborative learning. *Computers in Human Behavior*, *23*, 1105-1125.

Notwithstanding the positive effects of using CSCL, many studies have also demonstrated possible pitfalls when using CSCL (Kreijns et al., 2003). For example, students communicating through computer-mediated communication (CMC) sometimes perceive their discussions as more confusing (Thompson & Coovert, 2003), demonstrate higher levels of personal conflict (Hobman et al., 2002) or participate in unsustained, low quality discussions (Lipponen et al., 2003). In sum, positive and productive interaction is sometimes lacking during CSCL.

The following sections discuss two problems that may occur during CSCL, namely communication and discussion difficulties, and students' difficulties to conduct critical, yet constructive discussions. Possible explanations for these problems will be described in short. This is followed by a description of how these problems can be addressed.

3.1.1 Communication Difficulties during CSCL

Research has demonstrated that it is difficult for group members to communicate during CSCL (Fjermestad, 2004; Fuks et al., 2006). Some researchers have argued that the communication problems found during CSCL may be due to the medium itself. Traditional text-based CMC systems, such as chat, are seen as media that are low in *media richness* (Daft & Lengel, 1986). Media richness is defined as a medium's ability to facilitate communication and the establishment of shared meaning. Factors such as the ability of the medium to transmit multiple cues (e.g., facial expressions, gestures), and the immediacy of feedback influence its media richness. As media richness decreases, students will have more difficulties conveying their opinions and will have more difficulties determining the meaning of group members' messages.

Furthermore, when working on group tasks students usually work on complex problems without demonstrably correct answers, which require them to resolve differing viewpoints. The type of communication usually used during CSCL, may not be suited to the types of tasks group members work on (Mennecke et al., 2000). The low media richness of CSCL-environments may constrain collaboration in such a way that it does not transmit the type of communication that group members need to solve their task successfully. This may lead to communication problems and decreased task performance.

3.1.2 Lack of Critical but Constructive Discussion during CSCL

When group members collaborate, they are usually working on complex problems, which require the input of all group members. Ideally, group members engage in discussions that are critical, but also constructive. This means that group members are critical of their own and the other group members' ideas, that criticism is accepted, and that they offer explanations for their opinions. These types of discussions have been called exploratory discussions and have been found to enhance learning during group work (Wegerif et al., 1999). However, research has shown that students rarely give arguments and counter arguments (Kuhn & Udell, 2003; Munneke, Andriessen, Kanselaar et al., 2007), nor do they offer explanations for their ideas regularly during CSCL (Van der Meijden & Veenman, 2005).

This absence of critical but constructive discussion may be explained in several ways. First, students may not know how to conduct such discussions and may not posses the necessary skills (Weinberger & Fischer, 2006). Second, as stated above, students may find it difficult to conduct constructive debates in a CSCL-environment and may have difficulties interpreting discussions. For example, they may not know whether group members agree or disagree with them. This possibly hampers argumentation and discussion (Adrianson & Hjelmquist, 1991). Finally, groups may possess group-norms that stimulate consensus among group members, instead of critical or exploratory discussion. Group-norms that stimulate consensus instead of critical discussion can contribute to the low quality of some online discussions (Erkens et al., 2006; Postmes et al., 2001). In conclusion, the relative absence of critical discussion during CSCL may have different causes. These causes need to be addressed in order to facilitate critical but constructive discussions during CSCL.

3.1.3 Addressing Communication Problems Using Visualizations

This section describes how visualizations of online dialogue may help address the previously described communication problems and the relative lack of critical but constructive discussion. Several researchers noted the lack of social cues of CSCL-environments (e.g., Donath, 2002). For users of chat rooms or discussion boards, it is often very difficult to quickly determine who the participants of online discussions are, or what the social norms of the online group are (Lee, Girgensohn, & Zhang, 2004). This lack of awareness may constrain social interaction and lead to lower perceived quality of the social space (Kreijns, Kirschner, Jochems, & Van Buuren, 2004). To address this problem, several researchers have turned to visualization techniques that visualize important social features of the environment. For example, Donath (2002) and Janssen, Erkens, Kanselaar, and Jaspers (2007) developed techniques for visualizing participation during online discussions. It is expected that by using such visualizations, social awareness can be increased, which may in turn lead to more productive interaction.

However, visualizing participation may not be sufficient to overcome communication problems and to stimulate critical discussions, because it only visualizes quantitative aspects of collaboration, not qualitative. Therefore, a visualization called *Shared Space* (SS) was developed. The SS visualizes whether group members are agreeing or disagreeing about a topic during online discussion. This visualization has been implemented in an existing CSCL-environment, called *Virtual Collaborative Research Institute* (VCRI, see Jaspers, Broeken, & Erkens, 2005). More specifically, the SS is an extension of the *Chat tool* of the VCRI-program. Students use this tool to communicate synchronously. The SS analyzes all messages entered in the Chat tool. For a more detailed description of the VCRI-program, the reader is referred to the Tasks and Materials section.

First, the SS discerns discussion topics using time intervals. When students do not type messages for more than 59 seconds, a new topic begins. Figure *3.1* shows a screenshot of VCRI's Chat tool with SS visualization. The screenshot shows the end of one topic, and the beginning of a new one.

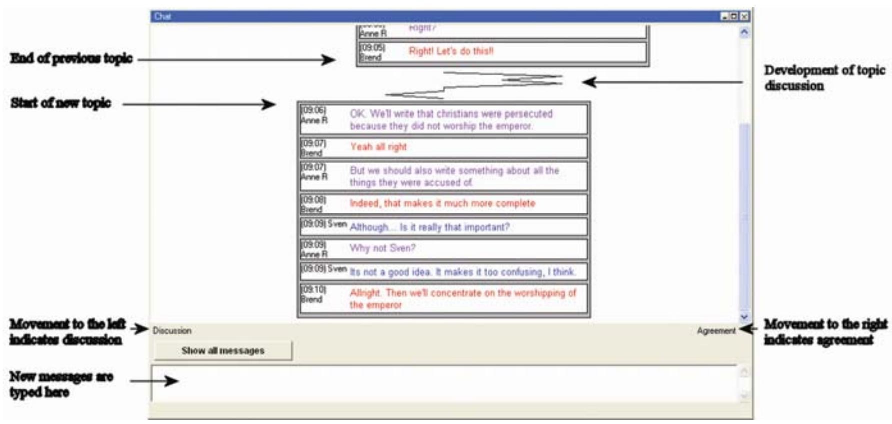


Figure 3.1: Screenshot of the Chat tool with Shared Space visualization.

Second, the SS analyzes the content of each chat message in order to determine whether it indicates discussion or agreement. For this purpose, the SS determines the communicative function of the message. This is done using the *Dialogue Act Coding (DAC) filter* (see Erkens et al., 2005). This filter uses over 1300 rules based on discourse markers to determine the communicative function of a chat message. Discourse markers are characteristic words or phrases signaling the communicative function of a message. In total, five main categories of communicative functions are distinguished: argumentative, responsive, informative, elicitative, and imperative. Each category consists of several subcategories, 29 in total. Of these, confirmations, acknowledgements, and positive evaluations are considered indications of agreement, while denials, verification questions, negative evaluations, and counterarguments are considered indications of discussion or debate. In a prior study (Erkens et al., 2005), the reliability of the DAC filter was tested and found to be acceptable (over 90% of all messages coded correctly).

Finally, after establishing whether the message indicates discussion or agreement, the SS moves the current topic to the left or to the right in small steps. When a message indicates discussion, the SS moves the topic to the left; when it indicates agreement, the SS moves the topic to the right. The lines above the topics visualize the development of the online discussion. For example, in Figure *3.1*, at the beginning of the topic, the SS indicated agreement (the line moves to the right), whereas later on the SS indicated debate (the line goes to the left).

It is hypothesized that the SS visualization will help group members overcome the communication and discussion problems described above for several reasons. First, the SS may increase the media richness of the CSCL-environment. Because the SS visualizes discussion and agreement, it may be easier for students to determine the meaning of messages. Additionally, it may be easier to identify the different views and positions held by group members. Moreover, the SS may help group members to determine whether there is shared understanding about a topic.

Second, the SS provides group members with feedback about the manner in which they are conducting their discussions. For example, when the SS keeps moving to the right, this tells group members they may be engaged in an uncritical discussion. Thus, the feedback provided by the SS visualization may increase students' awareness about their conversational strategies and their group-norms.

Finally, by providing them with feedback and raising their awareness, the SS may help students to engage in group processing. This occurs when group members discuss how well their group is functioning and how group processes may be improved (Yager et al., 1986). During these discussions group members may be stimulated to adopt more critical or exploratory group-norms.

In conclusion, it is expected that SS visualization may address some of the communication problems that occur during CSCL, and may help group members to collaborate and discuss more productively.

3.2 Research Questions

This paper investigates the effects of the SS visualization on online collaboration. The following research questions are addressed: Do students with access to SS visualization, compared to students without access, …

1. … perceive higher media richness when using the Chat tool?
2. … perceive different, more critical or exploratory group-norms?
3. … perceive their online collaboration and communication more positively?
4. … engage in different collaborative activities?
5. … perform better on an inquiry group task?

3.3 Method and Instrumentation

3.3.1 Design

A posttest-only design with a treatment and a control group was used to answer the research questions. Treatment group students had access to the Chat tool with SS visualization, whereas control group students used the same Chat tool but without SS visualization. In order to stimulate productive interaction, students worked in small, three-person groups (Schellens & Valcke, 2006).

However, due to the size of some classes and illness of students, some groups consisted of two or four students. In total, there were 33 groups of three, five groups of two, and two groups of four students. Each group of students was randomly assigned to either the treatment or the control group. The treatment group consisted of 20 groups (59 students; two two-person, 17 three-person, and one four-person group). Similarly, the control group also consisted of 20 groups (58 students; three two-person, 16 three-person, and one four-person group).

3.3.2 Participants

All 117 eleventh-grade (54 male, 63 female) participants came from five different history classes from two secondary schools in The Netherlands. These students were enrolled in the second stage of the pre-university education track. Mean age of the students was 16.17 years (*SD* = .60, *Min* = 15, *Max* = 18). Students were randomly assigned to a group by the researchers. In order to prevent combinations of students who could not get along with each other, their teachers checked the group compositions. As a result, three students were re-assigned.

3.3.3 Task and Materials

CSCL-environment: VCRI

Group members collaborated in a CSCL-environment called VCRI. The VCRI is a groupware program designed to support collaborative learning on research projects and inquiry tasks. Students use the program to collaborate in small groups. Every group member works at one computer. Students use the *Chat* tool to communicate synchronously with group members (see Figure 3.2). To read the description of their group task or to search and read relevant information, students can use the *Sources* tool. This tool lists a number of sources, which can be opened and read from the screen. Group members use the *Cowriter* as a shared word processor. Using the *Cowriter*, group members can work simultaneously on different parts of their texts. To collaboratively construct (argumentative) diagrams, students can use the *Diagrammer*. The VCRI-program contains several other tools not shown in Figure 3.2. For example, the *Planner*, which can be used to develop plans and assign tasks.

An alternative version of the VCRI-program was available for teachers. Using this so-called *Coach*-program, teachers can monitor the online discussions of their students. Teachers can also send messages in order to answer students' questions, or to warn students in case of misbehavior. Furthermore, teachers have access to the texts students are writing in the *Cowriter*. This way, teachers can monitor the progress of their groups.

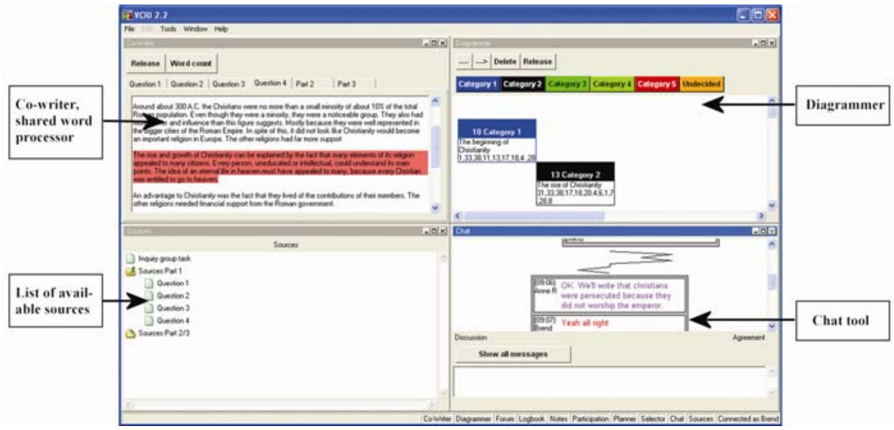


Figure 3.2: Screenshot of the VCRI program.

Inquiry Group Task

Participating students worked together on a historical inquiry group task. Subject of the task was "The first four centuries of Christianity" and consisted of three parts. The introduction of the task stressed the importance of working together as a group to successfully complete the inquiry task. Students were told they had eight lessons to hand in their reports, and they would receive a group grade for their reports.

For the first part of the inquiry task, the groups had to answer four different questions pertaining to the first four centuries of Christianity. To answer these questions, 12 different sources were available to the students. These sources were, for example, fragments from the New Testament, and historical texts from the Roman era. Additionally, students could search the Internet or their textbooks for more information. To complete the second part of the task, the groups had to study 40 different sources about the subject. These sources needed to be categorized into up to five different categories. Furthermore, group members were instructed to construct a diagram of their categorization using the Diagrammer. Finally, students had to write a short text, explaining how and why they categorized the different sources. For the final part of the inquiry task, group members had to collaboratively write an essay of at least 1200 words. The essay had to explain why and how Christianity developed from a small 'cult' into the main religion of the Roman Empire.

Comparable to the task used by Munneke et al. (2007), the task can be characterized as an open-ended task without a standard procedure and no single right answer. Furthermore, the inquiry task was quite complex and comprehensive; therefore, no single group member was likely to solve the task on his or her own. Thus, the participation of all group members was necessary to successfully complete the task.

3.3.4 Procedure

During the lessons each student worked on a separate computer in a computer lab. Students sat as far from their teammates as possible, in order to stimulate them to use to the VCRI-program to communicate with their other group members. Before the first computer lesson, students received information about the task and the group compositions from their teachers. During the lessons, the teachers answered task-related questions, while the experimenters solved technical problems. Students were allowed to work on the inquiry group task during free periods. Thus, they could work on the task in the school's media center when they had spare time in their timetable. After eight lessons, the groups handed in their final versions of the group task.

After the last lesson, a questionnaire was administered to the students. This questionnaire contained several items pertaining to perceived media richness, group-norm perception, and perception of online collaboration and communication. In total, the questionnaire contained 48 items. Students expressed their opinions using a 5-point scale ranging from 1 (=*completely disagree*) to 5 (=*completely agree*). Due to absence or sickness, 20 students did not complete the posttest questionnaire. Thus, the total number of respondents was 97 students.

3.3.5 Measures

Perceived Media Richness of the Chat tool

To measure perceived media richness of the Chat tool, the questionnaire contained a 15-item scale that addressed various aspects pertaining to the media richness of the Chat tool. The items addressed whether students were aware of agreements and disagreements during online discussions, and whether they could explain things easily to group members (Dennis et al., 1999). "During discussion in the Chat it is clear whether there is agreement among group members" is a sample item from the scale. Students' ratings were averaged to create a score for "*perceived media richness*" (a = .92). Higher scores indicate higher perceived media richness.

Group-norm Perception

To measure students' perceptions of group-norms, the questionnaire completed by the students contained three scales. The first scale consisted of three items, and asked students whether they perceived their group as having *critical group-norms*. The items were based on the work of Postmes et al. (2001). A sample item of this scale was: "Our group is critical". The second scale investigated whether students perceived their group as having *consensual group-norms*. This scale also consisted of three items based on the work of Postmes et al. (2001). An example from this scale is: "In this group people generally adapt to each other". Finally, the third scale examined whether group members perceived their group as having *exploratory group-norms*. This scale consisted of seven items, modeled after the ground rules for exploratory talk formulated

(Wegerif et al., 1999). "During discussions, criticism and counterarguments were accepted" is a sample item of this scale.

For all three scales, students' responses to the items were averaged to obtain scores for *critical* (α = .84), *consensual* (α = .59), and *exploratory* (α = .74) *group-norm perception*. Higher scores on these variables correspond to a more critical, consensual, or exploratory *group-norm* perception.

Perception of Online Collaboration and Communication

To measure how students perceived their online collaboration and communication the questionnaire contained three scales. The first scale consisted of seven items (e.g., "We helped each other during collaboration") and addressed *positive group behavior* , such as equal participation of group members, and helping group members (Webb, 1995).

Five items formed the second scale. These items (e.g., "There were conflicts in our group") addressed occurrences of *negative group behavior* such as conflicts and free riding behavior (O'Donnell & O'Kelly, 1994).

The final scale addressed students' *perceived effectiveness of their group's task strategies*. This scale was based on the work of Saavedra, Early, and Van Dyne (1993) and consisted of eight items that assessed the choices made and the strategies chosen by the group members. An example from this scale is: "We planned our group work effectively".

Again, for all three scales students' responses were averaged to compute a score for *positive* (α = .82) and *negative group behavior* (α = .68), as well as for *perceived effectiveness of group task strategies* (α = .81).

Collaborative Activities

In order to examine the influence of the SS on students' collaborative activities, a coding scheme was used (Janssen, Erkens, Kanselaar et al., 2007). The aim of this coding scheme was to provide insight into the task- and group-related processes during students' online collaboration.

Description of the coding scheme. The coding scheme consists of four different dimensions. Each dimension contains two or more coding categories. Furthermore, the scheme includes several additional categories (e.g., technical aspects) that did not belong to any of the four dimensions. In total, the scheme consists of 19 categories. The first dimension referred to *performance of task-related activities* aimed at carrying out the task (Jehn & Shah, 1997). This dimension contained two categories pertaining to the discussion of relevant task-related information: exchanging and sharing task-related information (*TaskExch)* and asking task-related questions (*TaskQues*). The abbreviations of the codes are given between parentheses.

The second dimension referred to *regulation and coordination of task-related activities*, containing four categories. Metacognitive activities that regulate task performance (e.g., making plans, monitoring task progress), are considered important to successful performance in electronic learning environments (Narciss, Proske, & Koerndle, 2007; Van der Meijden & Veenman, 2005). First, planning (*MTaskPlan*) involved discussion of strategies necessary to complete the task, and delegation of task responsibilities. Second, monitoring

(*MTaskMoni*) involved exchange of information that could be used to monitor task performance and progress, and assessing the amount of time available. Finally, evaluation involved appraisal and discussion of task performance and progress, which could be either positive (*MTaskEvl+*) or negative (*MTaskEvl-*).

Since group members also have to attend to the social and emotional element of collaboration to successfully complete a group task (Rourke et al., 1999) *performance of social activities* constituted the third dimension. This dimension contained five categories. First, greetings (*SociGree*) contribute positively to group atmosphere (Rourke et al., 1999). Second, social support remarks (*SociSupp*) referred to comments that contributed positively to group atmosphere, such as exchanging positive comments, and disclosure of personal information. Third, social resistance remarks (*SociResi*) referred to behaviors that contributed negatively to group atmosphere, such as insults and displaying negative emotions. Fourth, shared understanding (*SociUnd+*) referred to confirmations and indications of agreement, which serve to reach and maintain joint understanding. Similarly, loss of shared understanding (*SociUnd-*) referred to denials, and expressions of incomprehension.

The fourth dimension referred to *regulation and coordination of social activities*. These might be called meta-social activities. For example, group members need to discuss collaboration strategies or reflect on the manner in which they collaborated. This dimension contained four categories. First, planning (*MSociPlan*) involved discussion of collaboration strategies, such as helping each other, or proposals to work together on certain tasks. Second, monitoring (*MSociMoni*) referred to the exchange of information that could be used to monitor group processes. Finally, evaluation involved appraisal and discussion of group processes and collaboration, which could be positive (*MSociEvl+*) or negative (*MSociEvl-*).

Statements that addressed neutral, negative, or positive technical aspects of the CSCL-environment were also included in the coding scheme (*TechNeut*, *TechNega*, and *TechPosi*). Finally, statements that did not fit into any of the previously mentioned categories were coded as *Other*, referring to nonsense and off-task remarks.

Segmentation and coding procedure. During online collaboration some students only send one sentence per message, while others type several sentences that combine multiple clauses. Furthermore, even within in a single sentence, multiple concepts or statements may be expressed (Strijbos, Martens, Prins, & Jochems, 2006). Thus, it may be necessary to segment a chat message into smaller parts that are meaningful in their selves. Therefore, the chat messages were segmented into *dialogue acts* (Erkens et al., 2005). One dialogue act corresponds to a sentence or a part of a compound sentence that can be regarded meaningful in itself and has a single communicative function.

Segmentation and coding were done using the *Multiple Episode Protocol Analysis* (MEPA) computer program (Erkens, 2005). Messages were segmented into dialogue acts using a *segmentation filter*. A filter is a program, which can be specified and used in MEPA for automatic rule based coding or data manipulation. The segmentation filter automatically segments messages into dialogue acts, using over 150 decision rules. Punctuation marks (e.g., full stop, exclamation mark, question mark, comma) and connecting phrases (e.g., “and if”, or “but if”) are used to segment messages into dialogue acts. Using filters speeds up segmentation, and ensures segmentation rules are applied consistently. After the segmentation process, the dialogue acts were subsequently coded using the coding scheme.

Interobserver reliability. In an earlier study (Janssen, Erkens, Kanselaar et al., 2007), a satisfactory overall Cohen’s Kappa of .86 was found. The category Kappa’s (Cicchetti et al., 1978) ranged from .67 to 1.00. For the purpose of the current study, one rater coded 796 collaborative activities from four random protocols from the previous study. The results of this coding were compared to the results of the previous study. An overall Cohen’s Kappa of .94 was found (category Kappa range: .78 - 1.00).

Group Performance Scores

In order to measure the effect of the SS on group performance, an assessment form was developed for each part of the inquiry task. The assessment form for the first part addressed (1) the *conceptual content and the quality of the argumentation* of the answers, and (2) the *quality of the presentation* of the answers (Scarloss, 2002). *Conceptual content and quality of argumentation* were assessed using one item on a 4-point scale. For example, an answer that received one point, contained little or no relevant historical concepts and little or no argumentation, whereas a an answer that received four points contained all relevant historical concepts and adequate argumentation. *Quality of the presentation* was assessed using five items (e.g., correctness of the language used, copy-pasting of sources directly into the text, structure of the written answer) that were rated on a 3-point scale. For example, concerning correctness of the language used, groups received zero points if their answer contained seven or more language errors, one point if it contained four to six errors, and two points if it contained three or less errors.

During part two of the task, group members needed to study and categorize sources into categories, construct a diagram of their categorization in the *Diagrammer*, and explain how and why they categorized the sources. The assessment form for this part consisted of three items, which assessed the quality and completeness of the constructed diagram and the quality of the explanation. These items were rated on a 3-point scale. For instance, concerning quality of the explanation, groups received zero points if they did not formulate an explanation, one point if they wrote an explanation that did not explain completely why certain categories had been chosen, and two points if the explanation addressed why categories had been chosen and why sources had been placed in certain categories.

For the last part of the inquiry task, group members needed to collectively write an essay. Comparable to part one, *conceptual content and quality of argumentation* were assessed using three items rated on a 3-point scale. *Quality of the presentation* of the essay was assessed using five items (e.g., structure of the essay, correctness of language used, correct use of historical sources) on a 3-point scale. This was done in a similar fashion as for part one of the inquiry task.

To check the objectivity of the assessment procedure, two researchers scored seven inquiry tasks. The results of reliability analysis are presented in Table 3.1. The percentages agreement and Cohen's kappa's indicate the assessment procedure was reliable.

Table 3.1: Interrater reliability of the group performance assessment.

Part of the inquiry task	% Agreement (range)	Cohen's κ (range)
Part one		
- Conceptual content & argumentation	89.29% (85.71% - 100.00%)	.85 (.73 - 1.00)
- Presentation	84.52% (82.14% - 85.71%)	.74 (.71 - .76)
Part two	91.67% (87.50 – 100.00%)	.87 (.80 - 1.00)
Part three		
- Conceptual content & argumentation	95.24% (85.71 - 100.00%)	.90 (.59 - 1.00)
- Presentation	85.71% (85.71 - 85.71%)	.72 (.61 - .73)

3.3.6 Data Analysis

To investigate the effects of the SS on students' collaborative activities during CSCL, an independent samples *t* test with condition (SS or no SS) as an independent variable could be used. However, students' collaborative activities are *nonindependent* (Kashy & Kenny, 2000), which makes a *t* test inappropriate. This is caused by *mutual influence* (Kenny et al., 2002). That is, what one group member says, is influenced by, and influences the contributions of other group members. Therefore, students who are in the same group behave in more or less similar ways. Thus, it is expected that students who are, for example, in a group with group members who are focused on task-related activities, will also be stimulated to focus on task-related activities. To address the problem of nonindependence, *multilevel analysis* was used to examine the effects of the SS.

The multilevel analyses involved estimating two models: an empty model and a model including one or more predictor variables. By comparing the deviance of the latter model to the empty model, a decrease in deviance can be calculated. When this decrease in deviance is significant (tested with a χ^2-test), the latter model is considered a better model. Additionally, the estimated parameters of the predictor variables can be tested for significance by dividing the regression coefficient, β, by its standard error, yielding a *t*-value. A significant *t*-value indicates a significant effect of the predictor.

The line of reasoning concerning nonindependence of students' collaborative activities can be extended to the other individual measures used in this study (research questions one, two, and three). Thus, the effects of the SS on students' perceived media richness, *group-norm* perception, and perception of online collaboration and communication, will be examined using multilevel analysis as well.

3.4 Results

3.4.1 Perceived Media Richness of the Chat Tool

On average, treatment group students perceived higher media richness (M = 3.26, SD = 0.80) compared to control group students (M = 3.01, SD = 0.76). Furthermore, the effect of the SS was significant, $t(95) = 1.59$, $p = .03$. However, the associated χ^2-value was only marginally significant, $\chi^2 = 2.41$, $p = .06$.

Table 3.2: Means for group-norm perception and results of multilevel analyses.

	Treatment group students (N = 48)		Control group students (N = 49)				
	M	SD	M	SD	β	SE	χ^2
Critical group-norm perception	3.24	0.86	3.25	0.65	0.000	0.085	0.000
Consensual group-norm perception	3.50	0.89	3.46	0.66	0.022	0.062	0.123
Exploratory group-norm perception	3.82	0.53	3.60	0.53	0.108*	0.053	3.933*

Note. Mean scores along scales ranging from 1 (=*completely disagree*) to 5 (=*completely agree*). * $p < .05$.

3.4.2 Group-norm Perception

It was expected that treatment group students would perceive their *group-norms* as more critical and less consensual. The results presented in Table 3.2 show that this expectation was only partially confirmed. Regarding critical group-norm perception, no statistically significant differences between treatment and control group students were found, $t(95) = 0.00$, $p = 1.00$. Similarly, no differences were found regarding consensual group-norm perception, $t(95) = 0.35$, $p = .18$. However, the multilevel analyses revealed a significant effect of the SS on exploratory group-norm perception, $t(95) = 2.03$, $p = .01$. This indicates that treatment group students perceived that their groups were engaged more in critical but constructive online discussions, compared to control group students.

3.4.3 Perception of Online Collaboration and Communication

The results for research question three are presented in Table 3.3. Treatment group students reported more occurrences of positive group behavior compared to control group students, $t(95) = 2.31$, $p = .01$. Moreover, treatment group students perceived their group's task strategies to be more effective, compared to control group students, $t(95) = 2.53$, $p = .00$. However, treatment group students reported similar levels of negative group behavior compared to control group students, $t(95) = -1.25$, $p = .05$.

Table 3.3: Means for perception of online collaboration and communication and results of multilevel analyses.

	Treatment group students ($N = 48$)		Control group students ($N = 49$)				
	M	*SD*	*M*	*SD*	β	*SE*	χ^2
Positive group behavior	3.93	0.54	3.62	0.58	0.155^{**}	0.067	4.909^{*}
Negative group behavior	2.34	0.72	2.54	0.68	-0.100	0.079	1.532
Effectiveness of group task strategy	3.73	0.56	3.42	0.62	0.165^{**}	0.065	6.066^{**}

Note. Mean scores along scales ranging from 1 (=*completely disagree*) to 5 (=*completely agree*). $^{*}p < .05$. $^{**}p < .01$.

3.4.4 Collaborative Activities

In Table 3.4 the mean frequencies for collaborative activities are presented. The numbers in parentheses indicate the percentages of the total number of collaborative activities that were devoted to a specific activity. To examine the effect of the SS on students' collaborative activities, multilevel analyses were used as well. In this case, two predictors were added to the multilevel models. Besides condition (SS or no SS), number of dialogue acts typed was included in the model. This was done to account for the fact that some groups typed more dialogue acts and were generally more active than others. By including this predictor, the effect of the SS could be investigated independent of number of dialogue acts typed by students.

Table 3.4 also lists the results of the multilevel analyses. Number of dialogue acts typed was a significant predictor for all collaborative activities, except negative evaluations of social activities (*MSociEvl-*). Thus, in most cases participation was related to collaborative activities. For example, the more a student participated during online discussions, the more questions he or she asked (*TaskQues*).

Table 3.4: Mean frequencies and standard deviations of collaboration acts and multilevel analyses of the effects of condition.

	Treatment group students (*N* = 59)			Control group students (*N* = 58)			Total (*N* = 117)			Effect of condition		
	M Freq.	(%)	*SD*	*M* Freq.	(%)	*SD*	*M* Freq.	(%)	*SD*	Coeff.	*SE*	χ^2
Performing task-related activities												
- Info exchange (*TaskExch*)	19.75	(7.43)	18.89	24.84	(6.63)	27.89	22.27	(7.02)	23.81	-3.45*	2.41	2.00
- Asking questions (*Taskques*)	8.17	(4.11)	6.99	12.33	(3.01)	10.42	10.23	(3.56)	9.06	-2.18**	.81	6.62**
Coordinating task-related activities												
- Planning (*MTaskPlan*)	61.32	(21.61)	41.84	63.71	(21.47)	45.35	62.50	(21.54)	43.44	-2.62	4.29	.37
- Monitoring (*MTaskMoni*)	36.31	(12.93)	22.25	36.48	(12.64)	22.62	36.39	(12.78)	22.34	.12	-.85	.23
- Positive evaluations (*MTaskEvl+*)	6.12	(1.87)	5.51	5.72	(1.98)	6.14	5.92	(1.92)	5.80	.12	.63	.03
- Negative evaluations (*MTaskEvl-*)	5.92	(2.32)	5.66	6.31	(1.77)	5.07	6.11	(2.04)	5.36	-.32	.35	.85
Performing social activities												
- Greetings (*SociGree*)	8.71	(3.31)	7.10	10.31	(3.06)	10.31	9.50	(3.18)	8.68	-.98	1.12	.76

	Treatment group students (*N* = 59)			Control group students (*N* = 58)			Total (*N* = 117)			Effect of condition		
	M Freq.	(%)	*SD*	*M* Freq.	(%)	*SD*	*M* Freq.	(%)	*SD*	Coeff.	*SE*	χ^2
- Social support (*SociSupp*)	31.93	(8.64)	24.05	30.03	(0.51)	30.03	30.99	(9.58)	30.49	.28	2.88	.01
- Social resistance (*SociResi*)	10.69	(2.91)	9.87	8.67	(3.66)	8.67	9.69	(3.29)	10.12	.90	1.08	.69
- Mutual understanding (*SociUnd*+)	54.15	(21.69)	31.74	65.26	(19.09)	65.26	59.66	(20.38)	41.59	-6.95*	3.68	3.44*
- Loss of mutual understanding (*SociUnd*-)	11.68	(4.12)	8.66	11.29	(4.07)	11.29	11.49	(4.09)	8.32	-.04	.73	.00
Coordinating social activities												
- Planning (*MSociPlan*)	4.98	(1.29)	4.98	3.98	(1.56)	3.59	4.49	(1.42)	4.36	.46	.46	.98
- Monitoring (*MSociMoni*)	14.25	(3.83)	10.68	12.36	(5.07)	11.89	13.32	(4.45)	11.29	.76	1.07	.50
- Positive evaluations (*MSociEvl*+)	.76	(.07)	1.81	.34	(.32)	1.21	.56	(.20)	1.55	.20*	.14	2.06
- Negative evaluations (*MSociEvl*-)	.63	(.15)	.96	.48	(.25)	1.14	.56	(.20)	1.05	.06	.09	.47

	Treatment group students (*N* = 59)			Control group students (*N* = 58)			Total (*N* = 117)			Effect of condition		
	M Freq.	(%)	*SD*	*M* Freq.	(%)	*SD*	*M* Freq.	(%)	*SD*	Coeff.	*SE*	χ^2
Technical												
- Neutral technical (*TechNeut*)	4.02	(1.41)	3.85	4.45	(1.38)	4.92	4.23	(1.39)	4.40	-.33	.46	.54
- Negative technical (*TechNega*)	2.14	(1.09)	2.65	3.24	(.64)	4.50	2.68	(.86)	3.71	-.61*	.32	3.39*
- Positive technical (*TechPosi*)	.49	(.10)	.88	.34	(.13)	.78	.42	(.12)	.83	.06	.06	.81
Other	6.20	(1.14)	10.36	3.84	(2.77)	6.98	5.03	(1.96)	8.89	1.12	.91	1.50

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

Additionally, condition was a significant predictor for five collaborative activities. First, the SS had a negative effect on the number of task-related remarks (*TaskExch*) made by students, $t(114) = -1.43$, $p = .04$. This effect should be interpreted with caution however, as the associated χ^2-value was only marginally significant ($p = .08$). Second, having access to the SS was negatively related to the number of task-related questions (*TaskQues*) asked by a student, $t(114) = -2.69$, $p = .00$. Third, the β for *SociUnd+*, indicates a negative effect of the SS on the number of messages which were aimed at reaching and maintaining mutual understanding, $t(114) = -1.89$, $p = .02$. Fourth, the SS had a positive effect on students' use of positive evaluations of group activities, $t(114) = 1.44$, $p = .04$. Again, this effect should be interpreted with caution, as the χ^2-value was only marginally significant ($p = .08$). Finally, the SS had a negative effect on the number of negative technical remarks (*TechNega*) made, $t(114) = -1.89$, $p = .02$. Treatment group students typed less negative comments about the program.

3.4.5 Group Performance Scores

The performance scores for the different parts of the inquiry task are given in Table 3.5. Note that, since the performance scores were calculated for groups, not individual students, the number of observations is different from, for example, Table 3.4. As group performance is a *group level* variable, one-tailed *t* tests for independent samples were used to examine differences between treatment and control groups. From Table 3.5 it becomes clear that treatment groups received significantly higher scores for conceptual content and quality of argumentation for part one of the task, $t(38) = 1.88$, $p = .03$, $d = .59$. Furthermore, the quality of the presentation of part one of the inquiry task was significantly higher for treatment groups, $t(38) = 2.52$, $p = .01$, $d = .80$. No significant differences were found for the other parts of the task.

To determine the effect of the SS on group members' contribution to the written group products, additional analyses were performed. For each student, the number of characters typed in the Cowriter was calculated. Subsequently, for each group a Gini coefficient was calculated (Kiesler & Sproull, 1992). This coefficient indicated the equality of contributions to the Cowriter (0 = perfect equality, 1 = perfect inequality). On average, students contributed rather equally to the group products ($M = .22$, $SD = .11$). Furthermore, no significant differences were found between treatment and control groups, $t(38) = -1.06$, $p = .30$.

3.5 Conclusions and Discussion

During online conversations in CSCL environments, group members often experience communication difficulties (Fjermestad, 2004). Furthermore, the discussions conducted by group members are often shallow and uncritical. The present study addressed these problems by investigating the effect of visualizing agreement and discussion between group members during CSCL. A visualization called Shared Space (SS) was developed and implemented in an existing CSCL-environment. Based on an in-depth, automated analysis of chat messages typed by group members, the SS visualizes whether there is agreement or debate amongst them. It was hypothesized that giving students access to the SS visualization would be beneficial to online collaboration. In order to examine this hypothesis, a posttest-only design with a treatment and a control group was used. Treatment group students had access to the SS visualization, whereas control group students did not.

Table 3.5: Means and standard deviations for treatment and control groups for group performance scores.

	Treatment groups (N = 20)		Control groups (N = 20)				
Group performance scores	*M*	*SD*	*M*	*SD*	*t*	*p*	*d*
Part one							
- Conceptual content & argumentation	.58	.14	.50	.11	1.88	.03*	0.59
- Presentation	.63	.11	.54	.09	2.52	.01*	0.80
Part two	.65	.24	.64	.24	0.11	.46	0.03
Part three							
- Conceptual content & argumentation	.68	.27	.68	.33	0.00	1.00	0.00
- Presentation	.53	.25	.45	.23	1.05	.15	0.33

Note. To increase comparability, scores for the different parts were standardized, with 0 being the minimum amount of points and 1 the maximum. * $p < .05$.

The first research question examined whether treatment group students perceived online communication to be easier and more efficient than control group students. This hypothesis was confirmed, because questionnaire data confirmed treatment group students perceived the medium as more media rich.

The second research question considered whether the SS had an impact on students' group-norm perceptions. No differences were found regarding critical and consensual group perception. However, treatment group students reported their group held a more exploratory group-norm perception. Therefore, it may be concluded the SS helped group members to value critical but constructive online discussions.

The third research question investigated the effect of the SS on students' perceptions of online collaboration. Treatment group students reported more occurrences of positive group behavior and perceived their group's task strategies to be more effective, compared to control group students. However, reported levels of negative group behavior were similar for both groups. In conclusion, the SS seemed to have a positive effect on students' perceptions of their online collaboration.

The fourth research question addressed whether treatment group students were engaged in different collaborative activities than control group students. For some activities this hypothesis was confirmed. For example, treatment group students exchanged relatively less task-related information or questions than control group students. Furthermore, they typed fewer messages aimed at reaching and maintaining mutual understanding. This may indicate the SS helped students to correctly interpret whether there was agreement or discussion during their online discussions, thereby decreasing the need to maintain mutual understanding. Finally, the SS influenced treatment group students to send more positive evaluations of their collaboration. This indicates that SS stimulated students to engage in group processing, which is important for effective collaboration (Yager et al., 1986).

The last research question concerned whether treatment groups performed better on the inquiry group task. Treatment groups performed significantly better on the first part of the task, but not on the second and third part. This partial effect of the SS may be explained by the limited effect of the SS on students' collaborative activities. Students were mostly busy regulating their task performance by making plans and monitoring task progress. Consequently, they devoted less effort to content-related activities, such as exchanging task-related information. In fact, treatment group students typed less task-related informative remarks and questions. This may explain why the SS only had a partial impact on group performance.

Finally, some possible limitations of this study should be borne in mind. First, this study employed a complex, open-ended group task. More research is needed to determine whether the results of this study can be replicated using different types of group tasks. For other types of tasks, CMC may be better suited (Mennecke et al., 2000), thus decreasing the need for visualization of agreement and discussion. For example, CMC has been found to a suitable medium for idea generation (Fjermestad, 2004). Thus, when working on idea generation tasks, communication difficulties may have a less detrimental effect, decreasing the need for visualization of agreement and discussion. On the other hand, CMC seems to be less suitable for groups that need to negotiate conflicts of interest (Mennecke et al., 2000). As Munneke et al. (2007) argued, in such situations the task of working in a CSCL environment, studying information sources, co-writing texts, and conducting online discussions, may a very difficult one, thereby increasing the need for a tool such as the SS. Future research should explore the effects of the SS for different types of task (e.g., idea generation versus negotiation).

Secondly, this study did not take into account the influence of individual or group-level factors on the effects of the SS. For example, in groups composed of students who were familiar with each other, the effects of the SS may have been smaller since group member familiarity facilitates online communication (Adams, Roch, & Ayman, 2005). Thus, future research should investigate the possible differential effects of visualizing agreement and discussion for different types of groups (e.g., familiar versus non-familiar).

In sum, the present study found several positive effects of visualizing agreement and discussion during online collaboration. At first glance, the effects of the SS seem to concentrate on students' perceptions (e.g., treatment group students perceived their collaboration to be more positive). However, the present study also found some effects of the SS on students' behavior: treatment group students collaborated somewhat differently and performed better on some parts of the inquiry group task. Thus, the effects of the SS seem to extend beyond merely influencing students' perceptions.

3.6 Acknowledgement

The authors wish to thank Marcel Broeken and Jos Jaspers for their technical assistance.

4. Co-constructing a Representation of a Historical Debate: Effects of Representational Guidance during Computer-Supported Collaborative Learning[1]

This research investigates the role of representational guidance by comparing the effects of co-constructing two different argumentative diagrams. We used a design with two different groups defined by the type of argumentative diagram students co-constructed. The Graphical Debate-tool offered more representational guidance than the Textual Debate-tool. The results show that groups that worked with the Graphical Debate-tool constructed representations of higher quality and wrote essays that were better in terms of grounds quality. Furthermore, working with the Graphical Debate-tool was found to have a positive effect on students' learning as measured by a knowledge post-test. It can be concluded that representational guidance has an impact on group and individual performance and should therefore be taken into account during instructional design.

Keywords: Computer-supported collaborative learning; Representational tools; Argumentative tasks; Visualizations; Secondary education.

External representations (ERs) are graphical or diagrammatic representations that represent information about a topic (Van Bruggen, Kirschner, & Jochems, 2002; Zhang, 1997). A growing number of researchers advocates not only the use, but also the construction of ERs by students. The active construction of ERs, is thought to – among other things – stimulate deeper and more mindful cognitive processing, resulting in larger learning gains (Ainsworth, 2006; Cox, 1999; De Vries, 2006; Schnotz & Lowe, 2003; Stern, Aprea, & Ebner, 2003). These ERs can be constructed individually, or by groups of students. Collaborative construction – or co-construction – of ERs may link the benefits of collaborative learning, such as fostering positive social interaction, peer support, and critical discussion, to aforementioned the benefits of constructing ERs (Munneke, Andriessen, Kanselaar et al., 2007; Schwarz, Neuman, Gil, & Ilya, 2003; Van Drie et al., 2005). Unfortunately, this has not always been shown to be the case. Several researchers have reported mixed or even negative results when students collaboratively construct ERs (e.g., Suthers & Hundhausen, 2003; Toth, Suthers, & Lesgold, 2002; Van Bruggen et al., 2002).

---

[1] Janssen, J., Erkens, G., Kirschner, P. A., & Kanselaar, G. (2007). Effects of representational guidance during computer-supported collaborative learning. *Manuscript submitted for publication.*

Representational guidance refers to the fact that different ERs are capable of expressing different information (e.g., cause and effect relationships versus temporal relationships), make different information salient, or stimulate different cognitive processes than other ERs (Suthers & Hundhausen, 2003) and may account for the mixed findings on the benefits of co-constructing ERs. Several studies investigating the effects of *argumentative diagrams* showed representational guidance can influence students' behavior and learning process (Schwarz et al., 2003; Suthers, 2001; Van Amelsvoort, 2006). Argumentative diagrams constitute a specific type of ERs. They are usually embedded in computer-supported collaborative learning (CSCL) environments and are used to stimulate students to explicate their claims and arguments. Students can often add arguments, counter-arguments, evidence, and so on to the diagram. Furthermore, students can draw relationships between elements of the diagram such as causality and dependence relationships. There is evidence that representational guidance influences the effects that argumentative diagrams have on students' reasoning, use and quality of argumentation, and learning (e.g., Schwarz et al., 2003; Suthers & Hundhausen, 2003; Van Drie et al., 2005).

Some researchers maintain however that more research is needed to determine the optimal representational guidance for students' co-construction of argumentative diagrams (Erkens et al., 2005; Munneke, Andriessen, Kanselaar et al., 2007; Suthers, 2001; Toth et al., 2002). Furthermore, since most research has tended to focus on (a) the argumentative processes students engage in while co-constructing these diagrams, and (b) the quality of the co-constructed diagrams, little is known about the effects that representational guidance may have on individual learning outcomes (Suthers, 2003). The aim of this research is therefore to investigate the role of representational guidance by comparing the effects of two argumentative diagrams that differ with respect to the guidance they offer on (a) the quality of the representations constructed, (b) the quality of the group products (an essay), (c) students' performance on a knowledge post-test, (d) students' online collaborative activities, (e) students' perceptions of the usefulness of using the representation, and (f) students' perceptions of the collaborative process.

Before describing the research questions and method of research, we describe the potentials and pitfalls of the co-construction of ERs in more detail. Furthermore, we describe the two different argumentative diagrams (one offering more representational guidance than the other) used in this study, and explain why we expect representational guidance to have a positive effect on performance, learning, collaboration, and perceptions.

4.1 Co-construction of ERs: Potentials and Pitfalls

Student co-construction of argumentative diagrams combines the construction of ERs with CSCL. Although both approaches seem promising and may complement each other, each has its own pitfalls. In this section we describe the potentials and pitfalls of constructing ERs, how construction of ERs and CSCL can complement each other, and how co-construction of ERs in CSCL environments also has its own pitfalls.

4.1.1 Potentials and Pitfalls of Constructing ERs

The *benefits* of constructing ERs have been ascribed to several different causes. First, several authors have pointed to the fact that processing and constructing ERs stimulates students to engage in mental activities that facilitate learning and transfer. The construction of ERs helps students refine their understanding of complex phenomena, because it stimulates them to engage in the processes of self-explanation, externalization, elicitation, and elaboration of knowledge (e.g., Cox, 1999; Fischer et al., 2002). Furthermore, constructing ERs can help students reach higher levels of abstraction(Ainsworth, 2006; Schwartz, 1995) and thus deeper understanding of the subject matter. ERs can also help students engage in types of reasoning that are beneficial for their understanding of the subject matter (Cox, 1999). Van Drie et al. (2005), for example, found an effect of the type of representation constructed on students' use of historical reasoning: matrices elicited more talk about historical change than did lists or diagrams. Finally, ERs can help students think about concepts and theories on a deeper level by focusing them on the important concepts and relationships in the domain of study (Quintana et al., 2004).

Second, ERs can also reduce the cognitive effort needed to solve a complex problem. Argumentative diagrams for example, can help students determine whether their argumentation is balanced and whether it is supported by evidence (Munneke, Andriessen, Kanselaar et al., 2007) since there is a pictorial trace of arguments, supports, and counterarguments. In this way, ERs can support computational offloading (Ainsworth, 2006): students need to invest less effort to determine whether their argumentation is balanced (i.e., they can see it in the diagram), and they can thus devote their cognitive resources to other issues. Similarly, ERs can direct attention to unsolved parts of the problem (e.g., students notice they have mostly addressed arguments in favor of a position and have neglected contrary arguments). As such, an ER conveys students' progress through the problem (Cox, 1999; Suthers, 2001). Students can use this information to their benefit, for example by deciding whether they need to work more on a specific aspect of the representation. Finally, the information in an ER may serve as an extended working memory, helping students to recall their knowledge and ideas more easily (Löhner, Van Joolingen, & Savelsbergh, 2003; Van Bruggen et al., 2002).

Although the construction of ERs can have benefits for students' learning process, it can have *drawbacks* as well. One drawback is that there is always the possibility that students construct incorrect representations which foster misconceptions (Ainsworth, 2006). Furthermore, problems may arise when students misinterpret the format of the ER or when there is a mismatch between the ER and the students' own conceptual representation. An example of this might be when students interpret the meaning of an arrow in an argumentative diagram incorrectly (e.g., as a parent-child relationship as opposed to a causal relationship). Prior knowledge may also play a role here. When students lack prior knowledge, they may have more difficulties comprehending and successfully completing ERs (Bodemer & Faust, 2006; Cox, 1999). Finally, the cognitive demand imposed by the construction of ERs might become too high, thus preventing learning (Van Bruggen et al., 2002). For example, when students are asked to construct a representation of a topic, they have to find and process relevant information, place this information correctly in the representation, and take into account the representational format. This task may be so complex for learners that it prevents them from learning about the subject or may create extraneous cognitive load that is deleterious to learning (Kirschner, Sweller, & Clark, 2006). Indeed, research by Moreno and Valdez (Moreno & Valdez, 2005) showed students to experience higher cognitive load while constructing ERs compared to processing given ERs.

4.1.2 How CSCL and Construction of ERs Can Complement Each Other

CSCL offers students an environment in which their group tasks and communication can be supported by specific tools. Several researchers suggest that embedding tools in CSCL environments that allow the co-construction of ERs can *enhance the effects of both constructing ERs and CSCL* (Munneke, Andriessen, Kanselaar et al., 2007; Suthers & Hundhausen, 2003). First, co-construction of an ER in a CSCL environment stimulates further cognitive processing of the learning material. Compared to individually constructing an ER, students are more stimulated to externalize their thinking, to ask questions, give clarifications, and engage in a process of interactive argumentation during co-construction of ERs (Chinn & Anderson, 1998; Teasley & Roschelle, 1993; Van der Meijden & Veenman, 2005). These processes have been shown to be beneficial to students' learning processes. Second, CSCL environments offer tools that can support the cognitive and social activities that students carry out during the process of co-construction. Some environments for example, offer tools that stimulate careful planning or reflective activities (e.g., Erkens et al., 2005). Other environments try to support the social process by scripting the collaborative process (e.g., Dillenbourg, 2002). These tools may facilitate the process of creating ERs. Furthermore, in a CSCL environment ERs can fulfill an important social and communicative function (Cox, 1999; Löhner et al., 2003). When students co-construct an ER, it represents their current ideas and conceptions about the problem they are solving. In their discussions students can refer to the ER, thereby helping them to create a common frame of reference (this is called deictic referencing, c.f. Mühlpfordt & Stahl, 2007; Suthers & Hundhausen, 2003). This facilitates the creation of a joint problem space, and the process of negotiation of meaning and coordination (Erkens et al.; Teasley & Roschelle), thereby making the collaborative process easier. Thus, constructing ERs and CSCL are educational approaches that are able to complement each other.

4.1.3 Pitfalls of Co-constructing ERs in CSCL Environments

Unfortunately, research investigating the co-construction of ERs in CSCL environments shows that *this approach has its own disadvantages*. First, it is known that collaborating in a CSCL environment is not an easy task. In such environments, group members often experience communication and collaboration problems. Students have difficulties interpreting group members' messages, for example due to the reduced number of cues (e.g., intonation of voice, gestures) available in such environments (Adrianson & Hjelmquist, 1991). Furthermore, ERs in CSCL environments often take the form of argumentative diagrams and research has often found that students have difficulties engaging in interactive argumentation (Kuhn & Udell, 2003; Munneke, Andriessen, Kanselaar et al., 2007) leading to mixed outcomes of co-constructing argumentative diagrams (Van Bruggen et al., 2002). Finally, the fact that students have to construct ERs together instead of individually may make an already challenging task even more difficult. Students have to coordinate their contributions to the representation. This means they have to make sure their partners understand their additions to the representation (Cox, 1999). This is likely to increase the cognitive demands placed on learners (Van Bruggen et al., 2002). If the cognitive demands become too high, chances are that students will not benefit from co-constructing ERs (Zhang, 1997). The question is whether the benefits of co-constructing ERs outweigh their demands. If this is not the case, students' learning processes may suffer.

Thus, co-constructing ERs is a promising educational approach, but it has its own pitfalls. To help avoid these pitfalls, an important aspect to consider is the representational guidance an ER offers to students. Previous research has demonstrated that representations that differ with respect to the guidance they offer, stimulate different forms of reasoning and argumentation, and lead to different learning outcomes (Suthers & Hundhausen, 2003; Van Drie et al., 2005). In the next section we will therefore introduce two versions of a tool used for the construction of ERs in the domain of history. These tools differ with respect to the representational guidance they offer to students. We will describe why these tools can have different effects on the process of co-constructing external representations.

4.2 Co-constructing a Historical Debate Using the Debate-tool

This study involves students co-constructing an argumentative diagram in a CSCL environment about a topic in the domain of history. Group members (group sizes range from two to four students) are given a number of historical sources and are asked to use the ER to collaboratively construct a representation of the topic. This activity precedes a writing task where students have to co-author an essay on a specific historical topic that includes their findings. To aid them in this process the Debate-tool was developed. In this section we describe two different versions of the Debate-tool, which differed in the representational guidance they offered to the students.

An argument or debate involves participants taking different positions on a topic or position. This also applies to historical debates. In this study, students studied the first centuries of Christianity, and more particularly the way Christians were treated or maltreated in the Roman Empire, including historical accounts of the persecution of Christians by the Romans. However, there is some debate concerning the magnitude of these persecutions. On the one hand, some historians and the Catholic Church maintain that the persecutions were severe and a great number of Christians became martyrs because they were killed for religious reasons (*Martyrs position*). On the other hand, there is evidence this number was greatly and deliberately exaggerated by the church and Christians as propaganda meant to inspire Christians and to increase the number of converts (*Propaganda position*).

In this study, the participating students were given a number of historical and contemporary information sources and were asked to analyze and reconstruct this debate by co-constructing a representation of it with a Debate-tool and by collectively writing an argumentative essay based on their findings. In the Debate-tool and the essay, students were required to mention and discuss the arguments put forth by proponents of both positions. This required students to process and synthesize the information given in the sources since some arguments had to be constructed from information from multiple sources, because some sources contradicted each other, and because some sources contained archaic language and complex sentences. Thus, constructing an adequate representation required students to do more than simply cut-and-paste information from the sources to the tool.

Figure 4.1 and Figure 4.2 show screenshots of both versions of the Debate-tool. We will refer to the version displayed in Figure 4.1 as the *Graphical Debate-tool* and to the version displayed in Figure 4.2 as the *Textual Debate-tool*. Both versions are used to reconstruct the historical debate. In Figure 4.1, the boxes labeled Martyrs and Propaganda on the edge of the inner circle represent both positions of the debate. While working with the Graphical Debate-tool, students can add arguments (the boxes on the edge of the second circle) to either of the positions. These arguments can be found in, or inferred from the given sources. The sources also contain information that supports or refutes the arguments students add to the tool. This is represented by the boxes on the edge of the outer circle. Elements that represent supporting information have a white background, while elements that represent refuting information have a grey background. In the actual tool, these elements were color coded (i.e., supporting information using a green color and refuting information using a red color). The lines between the boxes indicate relationships between the elements of the representation.

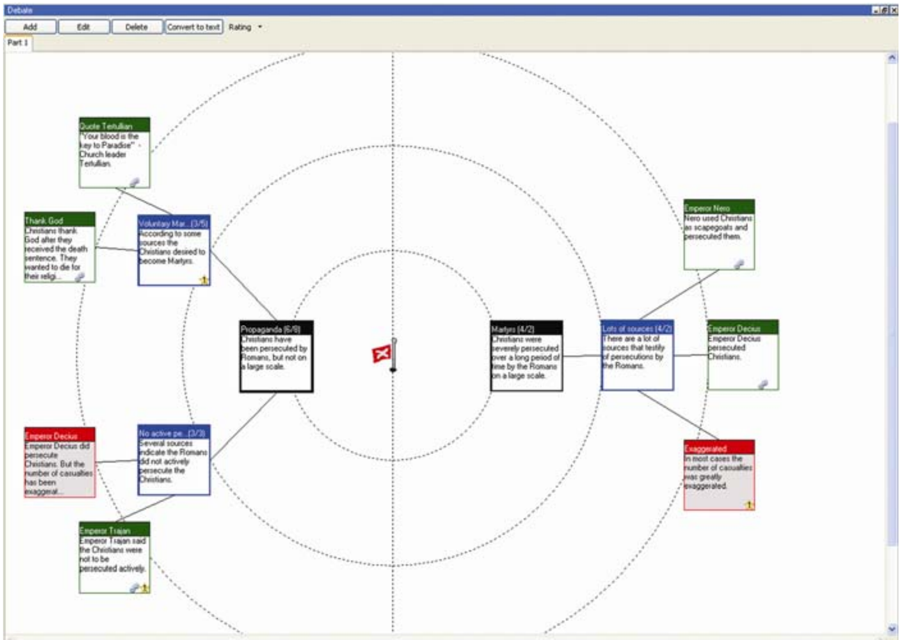


Figure 4.1: Screenshot of the Graphical Debate-tool (translated from Dutch).

In the Textual Debate-tool displayed in Figure 4.2, both positions are given as well. In this version of the tool, students also add arguments to the corresponding positions. No distinction is made however, between arguments, supports, and refutations. Instead, information is added to the Textual Debate-tool in a listwise manner (cf., Erkens et al., 2005; Van Drie et al., 2005).

In both versions of the tool, students can add elements by clicking on the Add-button. This opens a new window in which students can specify to which position the element has to be added. In addition, they can give the new element a title (which is displayed in the tool), type in an explanation why they added the element to the chosen position, and specify from which source the information came. In the Graphical Debate-tool students can also specify whether the new element is an argument, a support, or a refutation. The process of co-constructing representations (i.e., reading and processing historical sources, extracting relevant information, placing this information in the appropriate place in the representation) is almost the same for both versions of the Debate-tool.

The biggest difference between the Graphical and the Textual Debate-tool concerns the representational guidance they offer. Compared to the Textual Debate-tool, the Graphical Debate-tool uses more visualization techniques to make information salient and to help students complete the representation more effectively and efficiently. These techniques include making the strength of support of a position visible through automatic repositioning of a position and accentuation of the complexity of the argumentation. These techniques will be discussed in the next paragraphs.

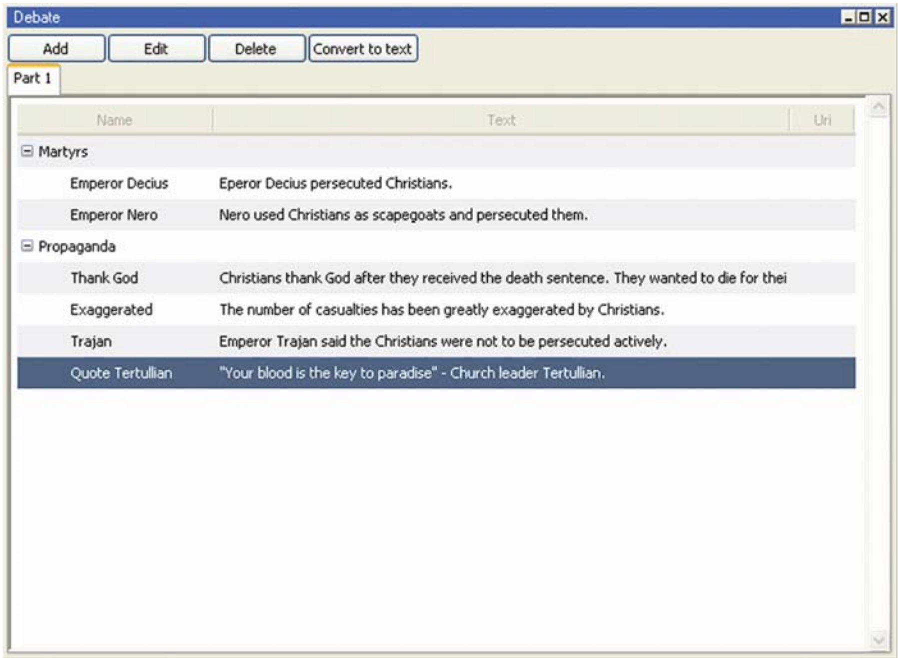


Figure 4.2: Screenshot of the Textual Debate-tool (translated from Dutch).

A first, rather obvious, difference between the two versions concerns their modality. While the Textual Debate-tool is mostly text based, the Graphical Debate-tool uses both visual and textual elements to represent information.

Secondly, the Graphical Debate-tool discerns between arguments, supports, and refutations. This feature may guide students more to find supporting and refuting information, and to formulate arguments since it is immediately clear to them if this information is present or not. Furthermore, previous research (e.g., Kuhn & Udell, 2003; Munneke, Andriessen, Kanselaar et al., 2007) has shown that students often have difficulties finding and formulating counterarguments. The option to distinguish between arguments, supports, and refutations may stimulate students to search more for counterevidence. Furthermore, by guiding students in this direction, students may also be more likely to formulate arguments, supports, and refutations in their essays. On the other hand, it could be argued that, because of its increased complexity, the Graphical Debate-tool is less suited to support the *linearization process* that occurs when writers have to transfer the content of the ER to an essay.

Third, the Graphical Debate-tool visualizes how well positions are supported by arguments and supporting information. Each time an argument is added to a position, it moves two steps closer to the central flag. Moreover, whenever a supporting piece of evidence is added to an argument, the position moves one step closer to the flag. Finally, when a refutation is added to an argument, the position moves one step away from the flag. Consequently, when a position is located closer to the flag, it is better supported by arguments: the argumentation is more strongly in favor of the position. In Figure 4.1, the Propaganda position is located nearer to the flag than is the Martyrs position, thus in its current form the Propaganda position is better supported with arguments and supporting information. It is more difficult to infer this from the Textual Debate-tool because no distinction is made between arguments, supports, and refutations although the lengths of the lists corresponding to the positions do give clues of this. The embedded representational guidance the Graphical Debate-tool offers may help students to draw a conclusion (Which position is supported best by arguments and information?), and thus may contribute to computational offloading (Ainsworth, 2006).

Fourth, the Graphical Debate-tool visualizes students' progress through the problem (Cox, 1999) better than does the Textual Debate-tool. Although both versions give an idea about the balance between both positions (e.g., Has one position been given more attention than the other?) the Graphical Debate-tool makes it clearer whether students tended to focus on certain elements (i.e., arguments, supports, or refutations). Students can infer this from the number and type of boxes that have been added to the representation. Furthermore, the boldness of the lines around the position and argument boxes also serves as an indication for their elaborateness and complexity. For example, in Figure 4.1, the lines around the Propaganda box are bolder compared to the lines around the Martyrs box, indicating the Propaganda position is more elaborate in terms of number of arguments, supports, and refutations. This may thus draw group members' attention to unaddressed parts of the problem.

Finally, in the Graphical Debate-tool students have the option to rate the quality of arguments, supports, and refutations. Sometimes an argument may be more important or valuable than another argument. Students can express this by giving ratings to arguments, positions, and refutations. In Figure 4.1, the elements entitled 'Voluntary martyrs', 'Emperor Trajan', and 'Exaggerated' received a rating once (indicated by the star in the corresponding boxes), while the other elements did not receive a rating. This rating indicates the students thought these elements were more important than other elements. A rating influences the distance of the position from the flag. When a rating is given to an argument or support, its corresponding position moves one extra step closer to the flag. On the other hand, when a rating is given to a refutation, the corresponding position moves one extra step away from the flag. It is important to note that students can rate an argument more than once, to further differentiate the importance of elements. Furthermore, the total of number of ratings that can be given by the group is limited (i.e., dependent on the size of the group). The rating functionality of the Graphical Debate-tool serves two purposes. It stimulates students to think about and discuss the importance of arguments and may help them to see which arguments are more important than others.

4.3 Research Questions

The Graphical Debate-tool offers more representational guidance to students compared to the Textual Debate-tool by using visualization techniques. Aim of this research is to investigate whether this contributes to the quality of the representations students' construct, the quality of the essays written by students, their learning as measured by their performance on a knowledge test, their collaborative activities, and their perceptions of the tool as well as their perceptions of their online collaboration and communication. In sum, the following research questions are addressed: Do students who work with the Graphical Debate-tool (Graphical Debate condition), compared to students who work with the Textual Debate-tool (Textual Debate condition), …

1. … construct representations of higher quality?
2. … write essays of higher quality in terms of (a) how well all the topics in the given sources are covered, (b) quality of the grounds used in the essays, and (c) conceptual quality of the arguments used in the essays?
3. … perform better on a knowledge post-test?
4. … collaborate in a different way online?
5. … perceive a higher usefulness of the Debate-tool?
6. … perceive their online collaboration and communication more positively?

4.4 Method and Instrumentation

4.4.1 Participants

The participants were students from five different history classes from two secondary schools. The total sample consisted of 124 eleventh-grade students (55 male, 69 female), with an average age of 16.24 years of age ($SD = 0.57$). Their teachers randomly assigned them to different groups. Due to uneven class sizes and student drop-out, this resulted in one 2-person group, 30 3-person groups, and eight 4-person groups. Because the 2- and 4-person groups collaborated in the same manner as the 3-person groups, they were not removed from the analyses.

4.4.2 Design

To investigate the effects of the Debate-tool, we used a single-factor, between subjects design with two different groups defined by the type of representation used: Graphical Debate-tool or Textual Debate-tool. We randomly assigned three classes to the Graphical Debate condition group, and two classes to the Textual Debate condition. Classes rather than groups were chosen as the unit of assignment, because having two versions of the tool in the same class was considered undesirable (e.g., students would notice some groups had another version of the tool and this might influence their behavior). In total, 79 students in 24 groups worked in the Graphical Debate condition, and 45 students in 15 groups formed the Textual Debate condition.

4.4.3 Tasks and Materials

CSCL Environment: VCRI

Students worked in a CSCL environment named *Virtual Collaborative Research Institute* (VCRI, Broeken, Jaspers, & Erkens, 2006; see Figure 3). The VCRI is a groupware program developed to support collaborative learning on inquiry tasks and is composed of several separate, but interrelated tools. Students use the *Chat* tool to synchronously communicate with other group members. To read the description of their group task or to search and read relevant information, students can use the *Sources* tool. This tool lists a number of sources which can be opened and read from the screen. Group members use the *Cowriter* as a shared word processor. Using the *Cowriter*, group members can simultaneously work on different parts of their texts. To collaboratively construct (argumentative) diagrams, students can use the *Diagrammer*. VCRI contains several other tools designed to support, among other things, the planning of the task (*Planner*), and the monitoring of group members' participation during the collaborative process (Participation-tool, see Chapter 2 or Janssen, Erkens, Kanselaar et al., 2007).

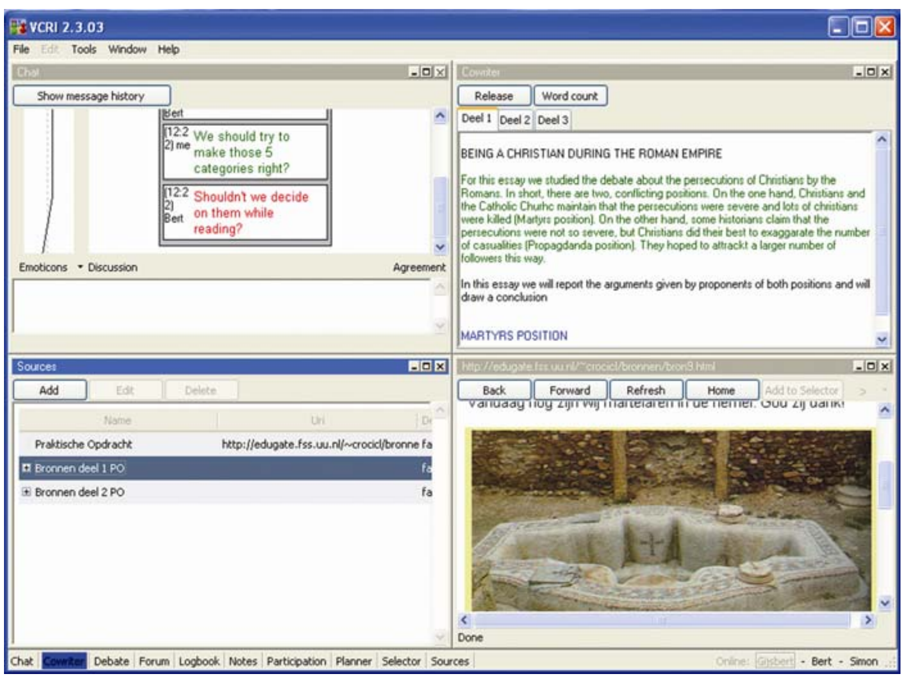


Figure 4.3: Screenshot of VCRI.

Task

Students collaborated on an inquiry group task in the domain of history. This task was adapted from an earlier study by Janssen, Erkens, and Kanselaar (2007). Topic of the task was "The first four centuries of Christianity". This inquiry task consisted of three different parts and lasted for eight 50-minute lessons.

During the first part of the task, students were given 14 historical and contemporary information sources and were asked to explore and discuss the different sources with respect to the Martyrs versus Propaganda debate. Students were required to co-construct a representation of this debate in either the Graphical Debate- or Textual Debate-tool. After they had completed their representation, they had to co-author an argumentative essay based on their findings. Students had three lessons of 50 minutes each to complete this part of the task and submit their products to their teachers. The quality of the representations constructed and the essays written during this part of the task were analyzed to answer research questions 1 and 2 (see below). This part of the task can be characterized as an open and complex task. The 14 sources were quite difficult and students often had to synthesize information from multiple sources to construct arguments that pertained to one of the two positions. Furthermore, information in some sources contradicted information in other sources. This required students to discuss, reason, and argue about the information in the sources, the arguments they presented in the ER, and the information they included in their essays.

To complete the second part of the group task, the groups had to study 25 new sources about the topic. These sources needed to be categorized into up to five different categories by the group members. Students needed to decide on these categories themselves. Furthermore, group members were instructed to construct a diagram of their categorization using the Diagrammer. For the final part of the inquiry task, group members had to collaboratively write a second essay of at least 1200 words explaining why and how Christianity developed from a small 'cult' into the main religion of the Roman Empire. It is important to note that parts 2 and 3 of the inquiry task were the same for both conditions. Furthermore, the instruction and task for part 1 was also the same for both conditions. The only difference concerned the explanation that students had to use either the Graphical or Textual Debate-tool to construct a representation of the historical debate in part one of the inquiry task.

4.4.4 Research Question 1: Quality of Representations

To determine whether groups in the Graphical Debate condition constructed representations of higher quality than the groups in the Textual Debate condition, we rated all of the items (arguments, supports, and refutations) placed in the tool by the students on a 5-point scale (ranging from 0 - 4). To facilitate the rating of the items, we developed a flowchart that contained a series of yes/no questions. These questions addressed quality aspects, such as (a) appropriateness of the item for the position that it was supposed to represent, (b) correctness and elaborateness of the explanation given to justify its inclusion, (c) use of additional information to mitigate or counter the main argument, and (d) use of references to historical persons, eras, or events to explain the historical context. Based on the answers given, we gave a score to each item, such that items that were clearly not appropriate for the position they were supposed to represent received a rating of 0. We gave a rating of 4 to items that contained an elaborate explanation, information to nuance or counter the main argument, and a historical context. Sometimes however, the items contained information that could be considered a repetition or paraphrase of information that had already been given in an earlier item. If this was the case, we marked the item to be a repetition.

Because some groups constructed representations with more items than others, we calculated a *mean quality score* for each group instead of a summed quality score for each group. If a group's representation contained repetitions of a topic, we only included the item that had received the highest rating in the mean quality score. Otherwise groups that had often repeated themselves several times would have an advantage over groups that used fewer repetitions.

Two raters (the first and second author) independently rated the quality of 85 Debate-items from four different groups. The agreement percentage between the two raters was 76%, while Cohen's κ was .69. The five categories of the rating scheme can be considered to represent an order (with 0 being the lowest, and 4 being the highest category), and thus a weighted κ was also calculated. The weighted κ takes the size of the discrepancies between raters into account, meaning that if rater A gives a rating of 0 to a certain item, while rater B gives a rating of 4, this discrepancy is given more weight than if rater A gives a rating of 2, while rating B gives a rating of 3. The weighted κ was found to be .73. Based on these findings, we concluded that the rating procedure was sufficiently reliable.

4.4.5 Research Question 2: Quality of Essays

After constructing a representation of the two positions in their Debate-tools for the first part of the inquiry group task, the students collaboratively wrote an essay detailing their findings. To determine whether groups in the Graphical Debate condition wrote better essays than groups in the Textual Debate condition, we analyzed the quality of these essays with respect to (a) how well the essay covered all the topics that could be found in the historical sources, (b) quality of grounds used, and (c) conceptual quality of the argumentation.

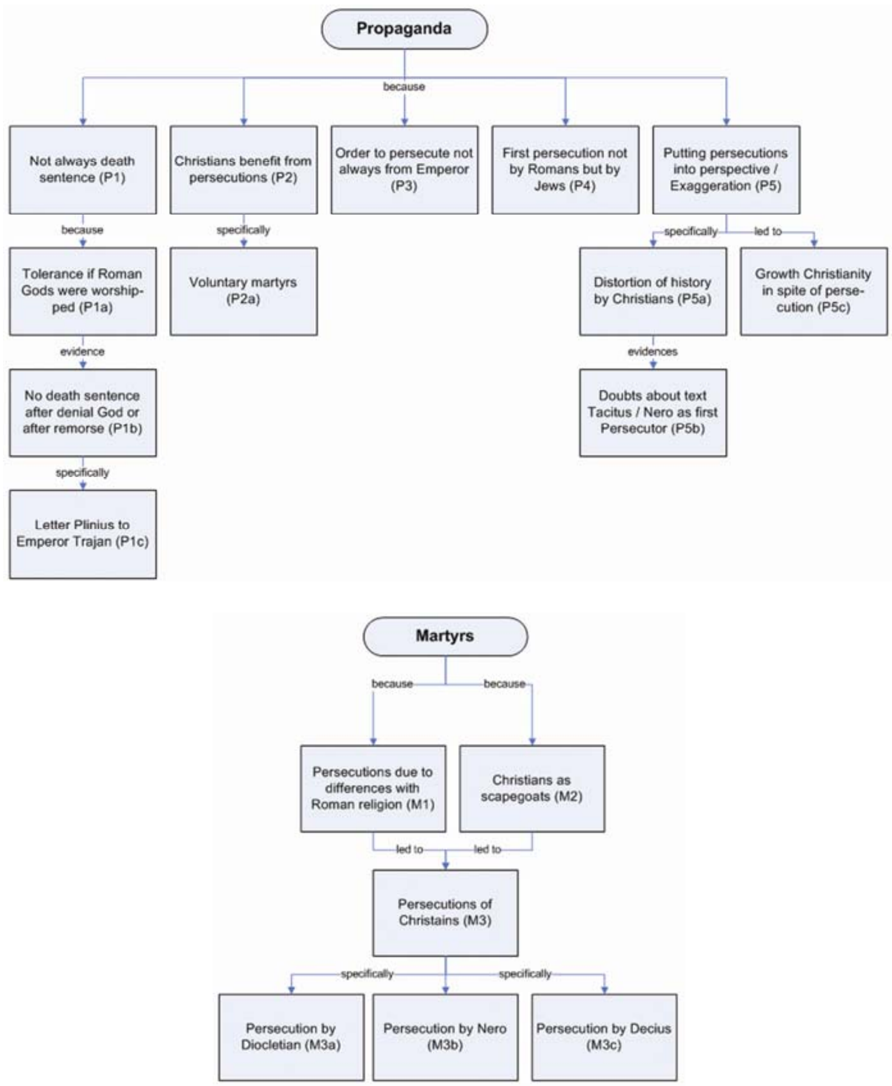


Figure 4.4: Graphical representation of the topics in the available historical sources.

Topics Covered in the Essays

To determine how well the essays covered all topics presented in the sources, we had to segment them into smaller units since each text contained several topics. Based on a detailed analysis of the historical sources available to the students, we distinguished 12 topics that related to the Propaganda position, and 6 topics that related to the Martyrs position. A graphical representation of the 18 different topics and the relationships between them is given in Figure 4.4. These topics were the starting point for the segmentation process. During the development of the segmentation process, we noted that students used two additional “topics” in their texts, which could not be found in the historical sources. The first topic (O1) concerned the assessment of arguments or positions and the drawing of conclusions (e.g., "We think the Propaganda position is supported better by arguments"). The second topic (O2) concerned text fragments, such as introductory sentences which did not belong to any of the other topics (e.g., "This essay will describe our findings on the Martyrs versus Propaganda debate"). For all 20 topics, a detailed description was given in order to facilitate segmentation.

Segmentation involved reading each sentence and determining whether it signaled the beginning of a new topic. If the sentence marked the beginning of a new topic, it also marked the end of the previous segment. Thus, a segment could vary in length from a single sentence to an entire paragraph. After segmentation, we assigned a code to each segment corresponding to its topic. Examples of five segments and their topics can be found in Figure 4.4. After the segmenting and coding of all of the topics, we determined the total number of different topics covered in text for each group.

Grounds and Conceptual Quality of Segments

The distinction between *grounds quality* (how well students used evidence or examples to support their claims) and *conceptual quality* (how conceptually correct the arguments were) is made by several authors (Clark, Sampson, Weinberger, & Erkens, 2007; e.g., Clark & Sampson, 2005; Munneke, Andriessen, Kanselaar et al., 2007; Sampson & Clark, 2006; Sandoval & Millwood, 2005). These authors argue that only examining grounds quality gives important information about how well students develop and support their written arguments with evidence and explanations, but provides no details about the validity and conceptual quality of these arguments. In theory, it is possible that students receive a high score in terms of grounds quality when they write an argument that contains several explanations and includes multiple pieces of evidence. However, at the same time the conceptual quality of the same argument may be lower, because the evidence students include or the explanations they give are (partially) flawed or incorrect (e.g., examples 1 and 4 in Table 4.1). Thus, we decided to code each segment in terms of grounds quality as well as conceptual quality.

Grounds quality. The evidence provided by students to back up the claims and opinions in their written texts formed the starting point for the analyses of grounds quality. Each text segment was judged in terms of how well and how elaborately it was supported by evidence or explanations (Sandoval & Millwood, 2005). Grounds quality of each segment was rated on a 4-point scale, ranging from 0 to 3. Segments that included no grounds in the form of explanations or evidence received a score of 0 (e.g., example 3 in Table 4.1). When only a short explanation was given to support the main claim of the segment, a score of 1 was given (e.g., example 5). Segments that included an elaborate explanation or example to support the claim of the segment, received a score of 2 (e.g., example 1). Finally, segments that included an elaborate explanation or example with one or more references to historical sources were given a score of 3 (e.g., examples 2 and 4). This analysis was not applied to segments assigned to one of the two "Other" topics (O1 and O2).

Table 4.1: Examples of coded text segments with respect to topic, grounds quality, and conceptual quality.

	Segment	Topic (see Figure 4.4)	Grounds quality	Conceptual quality
1	"It is a fact that many Christians were murdered. They were used as scapegoats, and they got the blame for the economic recession. But when the Christians were executed, the economic problems remained. Therefore it was not a successful measure. Thus, it is doubtful whether they were used as propaganda for the emperor."	M2 (Christian as scapegoats)	2	1
2	"There was never a massive, full-scale persecution. The persecution was often restricted to Rome, while it was only applied in the rest of the Empire when Christians were brought before the local ruler. In the rest of the empire there has never been a systematic persecution. Also, the persecutions never lasted very long. There were always periods of peace and quiet. For example, source 13 states 'the persecution during the rule of Decius lasted for one year, at most'."	P5 (Putting the persecutions into perspective)	3	3
3	"Christians were often persecuted."	M3 (Persecution of Christians)	0	1
4	"From source 9 it becomes clear that many Christians turned themselves over to the Romans to be executed. They hoped to follow the example of Jesus Christ by sacrificing themselves. In this respect it can be concluded that many were voluntarily persecuted and this sheds a different light on the image we have of the persecutions of Christians."	P2a (Voluntary martyrs)	3	2
5	"During the Roman era, Christians were persecuted because their religion conflicted with the Roman religion. Romans were polytheistic, whereas Christians were monotheistic."	M1 (Persecution due to differences with Roman religion)	1	1

Conceptual quality. The conceptual adequacy of the arguments given by the students, constituted the basis for the analyses of conceptual quality (Clark & Sampson, 2005). Each text segment was judged in terms of its conceptual correctness; thus segments that contained, for example, flawed conclusions, misinterpretations, or incorrect statements, received lower scores for conceptual quality than did segments that contained no errors. Conceptual quality was also rated on a 4-point scale (0 - 3). Segments that consisted of erroneous statements received a score of 0. Segments that contained some incorrect statements but also some correct ones received a score of 1 (e.g., examples 1, 3, and 5 in Table 4.1), whereas segments that contained statements that were mainly correct but that were sometimes a bit oversimplified received a score of 2 (e.g., example 4). Finally, segments that contained only conceptually correct and valid statements received a score of 3 (e.g., example 2).

As some groups covered more topics in their essays than others, we calculated the mean grounds and mean conceptual quality of the topics mentioned in their texts for each group. These mean scores were used in the comparisons between the Graphical Debate and Textual Debate conditions.

Table 4.2: Interrater reliability coefficients for all variables involved in the analysis of quality of essays.

	N	Agreement %	Cohen's *κ*	Weighted *κ*	Category *κ* range
Segmentation	4 texts	Lower bound: 81.6% Upper bound: 86.8%	–	–	–
Topic analysis	6 texts, 104 segments	87.5%	.86	–	.66 – 1.00
Grounds quality	7 texts, 113 segments	89.4%	.85	.90	.80 – .89
Conceptual quality	7 texts, 107 segments	92.2%	.88	.92	.82 – 1.00

Reliability Analyses

The first author and a research assistant were involved the reliability analyses of the segmentation procedure, topic analysis, and analyses of grounds and conceptual quality (see Table 4.2). The reliability analysis of the *segmentation procedure* involved a comparison of the boundaries of the segments as determined by the two researchers. Whenever these boundaries differed between the two researchers, this was counted as a disagreement. Because the researchers could differ in the number of segments they discerned, the percentage of agreement was determined from the perspective of both researchers. This gave an upper and a lower bound agreement percentage (see Strijbos et al., 2006 for a detailed description). As can be seen in Figure 4.4, some of the *topics* bore a close relationship to each other, thus we decided it was necessary to establish the objectivity of this coding procedure as well. Agreement percentage, Cohen's *κ*, and category kappa's (Cicchetti et al., 1978) were calculated to judge the objectivity of the topic analysis. Besides agreement percentage, Cohen's *κ*, and category kappa's, a weighted Cohen's *κ* was also calculated for grounds quality and conceptual quality. The moderate to high reliability coefficients shown in Table 4.2 indicate the analysis of the quality of the written essays was sufficiently reliable.

4.4.6 Research Question 3: Pre- and Post-test Knowledge Measures

To investigate whether students in the Graphical Debate condition learned more compared to those who worked in the Textual Debate condition, knowledge pre- and post-tests were developed. Both tests consisted of the same 15 multiple-choice items addressing topics covered in the inquiry group task. The test contained factual (Why did Christians refuse to worship the Roman Emperor?) as well as comprehension or insight questions (Explaining why a historical source was written by a Roman or a Christian). We used Cronbach's alpha to determine test reliability. Pre-test reliability was .51. This indicates low item homogeneity, probably due to the fact that students had low to average prior knowledge, causing them to guess the correct answer. Post-test reliability was .67, which can be considered acceptable.

Table 4.3: Collaboration acts (abbreviation) and category Kappa's (κ_c) of the coding scheme.

	Task-related activities		*Social activities*	
	Codes	K_c	Codes	K_c
Perfor-mance	Info exchange (*TaskExch*)	.89	Greetings (*SociGree*)	.84
	Asking questions (*TaskQues*)	.84	Social support (*SociSupp*)	.93
			Social resistance (*SociResi*)	.73
			Mutual understanding (*SociUnd+*)	.95
			Loss of mutual underst. (*SociUnd-*)	.92
Coordi-nation / regulation	Planning (*MTaskPlan*)	.89	Planning (*MSociPlan*)	.92
	Monitoring (*MTaskMoni*)	.93	Monitoring (*MSociMoni*)	.95
	Positive evaluations (*MTaskEvl+*)	.89	Positive evaluations (*MSociEvl+*)	1.00
	Negative evaluations (*MTaskEvl-*)	.87	Negative evaluations (*MSociEvl-*)	.67
Other	Neutral technical (*TechNeut*)	.90	Other / nonsense (*Other*)	.67
	Negative technical (*TechNega*)	.90		
	Positive technical (*TechPosi*)	1.00		

4.4.7 Research Question 4: Online Collaboration

To complete the task, students had to communicate and collaborate with other members of their group. There is evidence that representational guidance can have an impact on the discussions that unfold between group members as they construct a representation. For example, Van Drie et al. (2005) found that students used more historical reasoning in their discussions when they used a Matrix representation compared to students who constructed a Diagram or List. Furthermore, the representation used also had an effect on regulative and coordinative processes: matrix users discussed more about how they were going to approach the task. There is thus, reason to believe students in the Graphical Debate condition will collaborate differently than students in the Textual Debate condition. For example, the Debate-tool offers more information about students' progress through the task (i.e., Have both positions been given equal attention?) and thus might decrease the need for active monitoring of task-progress. To examine the effect of representational guidance on students' collaborative activities, we analyzed their online collaboration with a coding scheme (see Janssen, Erkens, & Kanselaar, 2007; Janssen, Erkens, Kanselaar et al., 2007).

When students collaborate on an inquiry task, they need to exchange their ideas and opinions, ask questions, and work towards producing a group product (Dennis & Valacich, 1999; McGrath, 1991). On the other hand, collaboration also involves a social-relational aspect. Students have to perform social and communicative activities that establish group well-being (Kreijns et al., 2003; Massey et al., 2003). Therefore, the coding scheme not only contains codes relating to content, but also several codes that refer to the social and communicative aspects of collaboration, such as greeting each other or engaging in activities that contribute a positive group climate (e.g., joking, social talk).

On the other hand, collaboration requires considerable coordination and regulation of these activities (Erkens et al., 2005). Metacognitive activities that regulate task performance (e.g., making plans, monitoring task progress, and evaluating plans or ideas) are considered important for online collaboration (De Jong, Kollöffel, Van der Meijden, Kleine Staarman, & Janssen, 2005). Moreover, social activities have to be coordinated and regulated as well (Manlove, Lazonder, & De Jong, 2006). For instance, students have to discuss and plan their collaboration, monitor their collaboration, and evaluate their collaborative process. Thus, the coding scheme also contained codes that referred to the regulation and coordination of task-related and social activities.

In total, the scheme contains four dimensions: *task-related activities, regulation of task-related activities*, *social activities*, and *regulation of social activities*. Each dimension contains two or more codes or so-called collaborative activities. Furthermore, the scheme included several additional categories (e.g., technical aspects) that did not belong to any of the four dimensions. In total, the scheme consisted of 19 categories (see Table 4.3).

Segmentation and coding procedure. Before the start of the coding process, it was necessary to segment students' messages into smaller parts that were meaningful themselves. This was done because multiple concepts or ideas were often conveyed within one message. Therefore, the chat messages were segmented into smaller units.

Segmentation and coding were done using the *Multiple Episode Protocol Analysis* (MEPA) computer program (Erkens, 2005). Messages were segmented using a *segmentation filter*. A filter is a program, which can be specified and used in MEPA for automatic rule based coding or data manipulation. Punctuation marks (e.g., full stop, exclamation mark, question mark, comma) and phrases connected through a conjunction (e.g., "and", "but", "or") are used to segment messages into dialogue acts. Using filters speeds up segmentation, and ensures segmentation rules are applied consistently. After the segmentation process, the segments were coded using the coding scheme.

Interrater reliability. Two researchers were involved in the interrater reliability analysis. Ten protocols were chosen at random, and from each of these protocols 100 segments were randomly selected. Thus, in total 1,000 segments were independently coded by both researchers. The overall Cohen's κ was .90 (see Table 4.3).

4.4.8 Research Question 5: Perceived Usefulness of Debate-tool

To check whether students in the Graphical Debate condition perceived higher usefulness of the tool compared to students in the Textual Debate condition, the post-test questionnaire contained a scale, which tapped into their perceptions of its usefulness. The scale consisted of eight items rated on a five-point scale. Sample items of this scale are "The Debate-tool made it easy for us to write the essay" and "The Debate-tool gave us a good overview of the two positions." The eight items formed a homogeneous scale, as indicated by a Cronbach's alpha of .80.

4.4.9 Research Question 6: Perception of Online Communication and Collaboration

Students' perceptions of their online communication and collaboration were measured with three different scales in the post-test questionnaire. The first scale addressed *positive group behavior* and consisted of seven items rated on a 5-point scale. Behaviors such as helping each other and equal participation among group members are indications of positive group behavior (Webb & Palincsar, 1996). A sample item from this scale is: "We helped each other during collaboration." The second scale tapped into *negative group behavior* and consisted of five items rated on a 5-point scale. Conflicts and free riding behavior (O'Donnell & O'Kelly, 1994) are indications of negative group behavior. "There were conflicts in our group" is an example from this scale. The final scale addressed students' *perceived effectiveness of their group's task strategies*. This scale was based on the work of Saavedra, Early, and Van Dyne (1993) and consisted of eight items that assessed the choices made and the strategies chosen by the group members. Again, these items were rated on a 5-point scale. An example from this scale is: "We planned our group work effectively." All three scales had adequate reliability coefficients (.83, .70, and .74 respectively).

4.4.10 Procedure

Before the start of the inquiry task, students completed the knowledge pre-test and received information about the task and the group compositions from their teachers. During the lessons, each student worked on a separate computer in a computer lab. The role of the teacher was similar in both conditions: they were online to answer task-related questions during all lessons. Students were also allowed to work on the inquiry group task during free periods. Thus, they could work on the task in the school's media center when they had spare time in their timetable. After the third lesson, students submitted the final version of their representations in the Debate-tool and their essays for the first part of the task. These were subsequently used in the analyses conducted for research questions 1 and 2. After the eighth and final lesson the knowledge post-test and questionnaires on perceived usefulness of the Debate-tool and perception of online communication and collaboration were administered to the students.

4.4.11 Data Analyses

When analyzing the effects of representational guidance, the data-analytical problem of *nonindependence* had to be taken into account (Kenny, Kashy, & Cook, 2006). Because students worked in groups in this study, they influenced each other. This violates the assumption of nonindependence of observations of individuals, making the results of traditional analytical techniques unreliable. Multilevel analysis can cope with nonindependence and is therefore a more appropriate technique (Snijders & Bosker, 1999). Thus, in the case where individual level variables were used (research questions 3, 4, 5, and 6), multilevel analyses were conducted. However, when the group was the unit of analysis, nonindependence is not a problem and thus in these cases (research questions 1 and 2) independent samples *t*-tests were conducted.

4.5 Results

4.5.1 Tool Use

Table 4.4: Descriptive statistics of frequencies of adding, editing, or deleting elements in Graphical Debate and Textual Debate conditions.

	Graphical Debate		Textual Debate	
	M	*SD*	*M*	*SD*
Adding	9.82	5.82	9.80	4.32
Editing	3.53	5.43	3.07	4.85
Deleting	1.96	2.30	2.11	2.41

Before proceeding with the analyses of the differences between Graphical Debate and Textual Debate conditions, we analyzed students' activities in both conditions. This was done to rule out the possibility that differences between conditions concerning quality of representations, quality of the written essays, and so on could be attributed to students being more active in one version of the Debate-tool than in the other. Three types of activities were distinguished: adding elements, editing elements, and deleting elements. Table 4.4 shows the descriptive statistics for these activities, distinguishing between the two conditions. No significant effects of condition on adding (β = 0.00, p = .99), editing (β = 0.17, p = .80), and deleting (β = -0.10, p = .72) elements in the Debate-tool were found using multilevel analyses. Thus, Graphical Debate students were not more active in their version of the tool than Textual Debate students.

4.5.2 Quality of Representations

Table 4.5 shows the average quality of the representations for groups in the Graphical Debate and Textual Debate condition. Groups in the Graphical Debate condition made representations of significantly higher quality than groups in the Textual Debate condition. This applied to the Martyrs as well as the Propaganda position. The corresponding effect sizes (Cohen's *d*) were large.

Table 4.5: Group differences for average quality of constructed representations for groups in the Graphical Debate and Textual Debate conditions.

	Graphical Debate		Textual Debate			
	M	*SD*	*M*	*SD*	*t* (37)	*d*
Average quality	2.71	0.52	2.12	0.32	3.90**	1.28
• Quality Martyrs	2.75	0.57	1.98	0.54	4.17**	1.38
• Quality Propaganda	2.70	0.54	2.28	0.44	2.49*	0.83

Note. Quality was scored on a five-point scale, ranging from 0 (=low quality) to 4 (=high quality).
$^{*}p < .05$. $^{**}p < .01$.

4.5.3 Quality of Essays

The quality of the essays written by the groups was examined by determining how well the topics given in the sources were covered in the text, as well as the grounds and conceptual quality of the texts. Table 4.6 shows the means and standard deviations for treatment and control groups. Independent sample *t*-tests were used to determine whether the conditions differed significantly from each other. As can be seen from this Table, no significant differences were found with respect to *number of topics covered* in the essays and *average conceptual quality* of the essays; although for the latter variable the corresponding effect size was moderate. However, Table 4.6 shows *average grounds quality* was significantly higher in the Graphical Debate condition. The corresponding effect size can be considered large.

Additional analyses revealed a significantly positive correlation between grounds quality and conceptual quality (r = .47, p = .00), meaning that groups that received high scores for grounds quality also received high scores for conceptual quality. Interestingly, a significantly negative correlation was found between the number of topics covered in the essays and their conceptual quality (r = -.34, p = .04). Thus, when groups attempted to cover a large number of topics in their essays, this seemed to have a negative effect on the conceptual quality of their essays.

Table 4.6: Group differences for quality of written essays for groups in the Graphical Debate and Textual Debate condition.

	Graphical Debate		Textual Debate				
	M	*SD*	*M*	*SD*	*t*	*df*	*d*
Topics covered	6.79	1.74	7.20	2.91	-0.55	20[a]	-0.18
Grounds quality	2.38	0.30	2.10	0.37	2.57*	37	0.85
Conceptual quality	2.06	0.22	1.82	0.42	2.00	19[a]	0.77

Note. Quality was scored on a five-point scale, ranging from 0 (=low quality) to 4 (=high quality).
[a] Equal variances not assumed.* $p < .05$.

4.5.4 Post-test Performance

Unfortunately, on both measurement occasions several students were unable to complete the test due to various reasons. For the knowledge pre-test, test scores from 59 (75%) of the Graphical Debate and 43 (96%) of the Textual Debate students were available. For the post-test, these numbers were 59 (75%) and 42 (93%) respectively. An analysis of the students with missing values shows that students that were unable to complete the post-test performed nearly as well on the pre-test ($M = 10.88$, $SD = 1.87$) as students that were able to complete the post-test ($M = 11.31$, $SD = 2.23$). Furthermore, students that were unable to complete the pre-test performed equally well on the post-test ($M = 12.47$, $SD = 2.01$) as students that were able to complete the pre-test ($M = 12.76$, $SD = 2.05$).

Table 4.7: Descriptive statistics for pre- and post-test scores for Graphical Debate and Textual Debate conditions.

	Graphical Debate ($N = 59$)		Textual Debate ($N = 45$)	
	M	*SD*	*M*	*SD*
Pre-test	11.39	1.92	11.02	2.51
Post-test	13.08	1.76	12.22[a]	2.27

Note. Pre- and post-test performance was measured on a scale from 0 to 15.
[a] $N = 42$.

Descriptive statistics for Graphical Debate and Textual Debate students' pre- and post-test performances are given in Table 4.7. As can be seen, student performance increased from pre- to post-test in both conditions.

Since post-test performance was measured at the student level, multilevel analysis was used to answer the question whether students in the Graphical Debate condition learned more compare to students in the Textual Debate condition. The first step in this analysis was to examine the results of a model without any independent variables, the so-called null model (Model 0, cf. Snijders & Bosker, 1999). As can be seen in Table 4.8, a considerable part ($1.33 / (1.33 + 2.83) = .32$) of the total variance in post-test scores can be attributed to group-level factors. This indicates that not only individual student characteristics such as prior knowledge and motivation, but also group characteristics such as nature and quality of the online collaboration and group composition have an effect on post-test performance.

To determine the effect of condition, while controlling for prior knowledge, condition and pre-test score were added to the model (Model 1). As the results show, both pre-test performance and condition had a significant effect on students' post-test performance. As expected, a higher pre-test score contributed to a better post-test performance. Furthermore, condition contributed significantly to post-test performance, indicating a positive effect of working with the Graphical Debate-tool. As can be seen in Table 4.8, adding these two predictors resulted in a significant decrease in the deviance of the model indicating a better model fit. To check whether the Graphical Debate-tool had a differential effect for learners with low or high prior knowledge, we added a condition by pre-test interaction term to the multilevel model. No indication for such a differential effect of the Graphical Debate-tool was found, $\beta = 0.03$, $p = .39$; $\chi^2 = 0.07$, $df = 1$, $p = .40$.

Table 4.8: Multilevel analyses of the effect of condition and pre-test performance on post-test performance.

	Model 0		Model 1	
	β	*SE*	β	*SE*
Intercept	12.78	0.27	9.60	1.07
Pre-test score			0.28**	0.10
Condition			0.42*	0.22
Variance				
Group level	1.33	0.65	0.44	0.45
Individual level	2.83	0.56	2.98	0.58
Deviance	355.70		344.63	
Decrease in deviance			11.07**	

Note. Condition was effect coded with Graphical Debate as +1 and Textual Debate as -1.
* $p < .05$. ** $p < .01$.

4.5.5 Online Collaboration

Table 4.9 shows the effect of condition on students' online collaborative activities. Out of the 19 activities, the experimental condition was found to have an effect on four collaborative activities. Condition had positive effects on *shared understanding* (SociUnd+), *negative social evaluations* (MSociEvl-), and *neutral technical remarks* (TechNeut), meaning that students in the Graphical Debate condition engaged in these activities more than did student in the Textual Debate condition. A negative effect of condition was found on *loss of shared understanding* (SociUnd-), meaning that Graphical Debate students experienced less misunderstandings and disagreements than Textual Debate students.

Table 4.9: Multilevel analyses of the effect of condition on students' collaborative activities.

	β	*SE*	χ^2
Task-related activities			
TaskExch	-7.68	6.32	1.45
TaskQues	-0.76	0.75	1.01
Regulation of task-related activities			
MTaskPlan	1.11	1.73	0.40
MTaskMoni	1.84	1.62	1.25
MTaskEvl+	0.13	0.39	0.11
MTaskEvl-	-0.07	0.47	0.02
Social activities			
SociGree	0.40	0.50	0.66
SociSupp	0.13	2.83	0.00
SociResi	1.08	0.95	1.30
SociUnd+	3.12*	1.67	3.38*
SociUnd-	-1.22*	0.58	4.15*
Regulation of social activities			
MSociPlan	-0.06	1.30	0.00
MSociMoni	1.42	0.98	2.08
MSociEvl+	0.14	0.09	2.40
MSociEvl-	0.19*	0.10	3.92*
Other codes			
TechNeut	0.81*	0.48	2.78*
TechNega	-0.66	0.50	1.74
TechPosi	-0.13	0.08	2.29
Other	0.05	0.26	0.04

Note. Condition was effect coded with Graphical Debate as +1 and Textual Debate as -1.
* $p < .05$.

4.5.6 Perceived Usefulness of Debate-tool

Overall, Graphical Debate ($M = 3.32$, $SD = 0.62$) as well as Textual Debate students ($M = 3.20$, $SD = 0.67$) perceived their respective versions of the Debate-tool as useful. No effect of condition on perceived usefulness was found, $\beta = 0.06$, $p = .17$; $\chi^2 = 0.89$, $df = 1$, $p = .17$.

4.5.7 Perception of Online Communication and Collaboration

On average, students in the Graphical Debate condition ($M = 3.69$, $SD = 0.57$) reported lower levels of *positive group behavior* compared to students in the Textual Debate condition ($M = 3.91$, $SD = 0.51$). Indeed, multilevel analyses indicated an effect of condition on perceived positive group behavior, $\beta = -0.11$, $p = .04$; $\chi^2 = 2.90$, $df = 1$, $p = .04$. This means working with the Graphical

Debate-tool had a negative effect on perceived positive group behavior, whereas working with the Textual Debate-tool had a positive effect.

Concerning *negative group behavior*, students in the Graphical Debate condition (*M* = 2.39, *SD* = 0.69) reported slightly higher levels than did students in the Textual Debate condition (*M* = 2.26, *SD* = 0.58). No significant effect of condition on perceived negative group behavior was found, $\beta = 0.07$, $p = .18$; $\chi^2 = 0.84$, $df = 1$, $p = .18$.

Finally, students in the Graphical Debate condition (*M* = 3.52, *SD* = 0.47) reported somewhat lower *effectiveness of their group's task strategies* compared to students in the Textual Debate condition (*M* = 3.68, *SD* = 0.54). However, the effect of condition on perceived effectiveness of group task strategies was not significant, $\beta = -0.09$, $p = .07$; $\chi^2 = 2.17$, $df = 1$, $p = .07$.

4.6 Conclusion and Discussion

Aim of this study was to examine the effects of representational guidance on students' performance of a group task, learning of historical concepts and facts, collaborative activities, and perceptions of each other, their collaboration, and of the tools used.

The first research question addressed the *quality of the representations* made by the participating groups. Our analyses show that groups working with the Graphical Debate-tool constructed diagrams of higher quality than groups working with the Textual Debate-tool. An explanation for this finding may lie in the cognitive effort students needed to devote to the construction of the diagram. Because the Graphical Debate-tool offered more information about the balance of the argumentation and unsolved parts of the problem than the Textual Debate-tool, this may have made the students' task relatively easier (Cox, 1999). The Graphical Debate-tool offered students more representational guidance, thus supporting computational offloading (Ainsworth, 2006). As a result, students may have been able to devote more effort to the conceptual content of their diagrams. This finding is in line with previous research. Suthers and Hundhausen (2003) for example, found differences between Text, Graph, and Matrix representations with respect to number of hypotheses and evidential relations students expressed. Also, Van Drie et al. (2005) found effects of the type of representation on the number of arguments used, the balance of arguments for and against a position, and the number of historical sources used in the representation. Finally, Fischer et al. (2002) examined the differences between a content-specific and a content-unspecific representation. They found the solutions in the content-specific condition to be of higher quality than solutions in the content-unspecific condition. Our results therefore add to a growing body of research indicating the effects of representational guidance on the quality of the representations constructed

Our second research question addressed the *quality of the essays* written by the groups. We expected Graphical Debate groups to outperform Textual Debate groups with respect to the quality of their essays. The results show that the Graphical Debate and Textual Debate groups covered the same number of topics in their essays. However, the arguments put forth by Graphical Debate groups were better developed in terms of giving evidence and examples (i.e., they were better in terms of grounds quality). Furthermore, a near significant difference (with a moderate effect size) in favor of Graphical Debate groups was found for conceptual quality of the essays. This finding is somewhat different from other research findings, where limited effects of representational guidance on collaborative writing were found (e.g., Suthers & Hundhausen, 2003; Toth et al., 2002; Van Drie et al., 2005). An explanation may lie in the *representational guidance* offered by the Graphical Debate-tool compared to the guidance offered by the tools in the work of other researchers. Our tool directs students' attention to the distinction between arguments, supports, and refutations, and this may stimulate students to incorporate these elements in their ERs and essays. It has been argued that tools that support *linearization*, that is the ordering of content and arguments into an essay, may be better supported by ERs specifically designed to support the planning of the linear structure of essays (e.g., an Outline tool as described in Erkens et al., 2005). Although the Graphical Debate was not specifically designed to support the process of linearization, it may be the case that stimulating students to systematically address all arguments, supports, and refutations of a position also facilitates the process of converting a representation into an essay.

Interestingly, we found a negative correlation between the number of topics the groups covered in their essays and the conceptual quality of their essays. When groups tried to cover a large number of topics in their essays, this negatively affected the conceptual quality of their essays. An explanation might be that when groups try to cover a lot of topics in their essays, this limits the time and effort they can devote to ensuring their claims are conceptually correct. Furthermore, this also makes writing an essay more complex, further decreasing the cognitive resources students can devote to writing conceptually correct claims. The more complex and elaborate groups' essays become, the harder it will be for them to clearly reason and structure the essay. There might be a trade-off between the completeness and the quality of the essay. Also, when groups try to address many topics (i.e., they adopt a more is better strategy), this constrains the time available to check the correctness of their claims and to create a clearly structured essay. Indeed, groups needed to work very hard to complete the task in the allotted time (3 lessons of 50 minutes each for this subtask) and thus there was quite some time pressure. Considering the complexity of the task and the existing time pressure, trying to include too many topics in the essay may not have been an adequate strategy.

After examining the effects of representational guidance on group performance, we turned our attention to its effects on students' *individual learning outcomes*. We expected students working with the Graphical Debate-tool to outperform students working with the Textual Debate-tool on a knowledge post-test. This hypothesis was confirmed. Additionally, it is important to note that students in both conditions improved their performance from pre- to post-test. Unfortunately, the number of students that were unable to complete either the pre-test, the post-test, or both was higher in the Graphical Debate condition (about 25%) than in the Textual Debate condition (about 6%). This may have negatively affected the comparability of both conditions, although missing value analyses indicated no systematic effects of missing data on pre- or post-test performance.

Previous studies have reported mixed effects of representational guidance on students' post-test performance. Van Drie et al. (2005) found differences on post-test performance between different representations, while Suthers and Hundhausen (2003) did not. An explanation for this finding may lie in the time students' spent working on the representations. In this study, students worked for three 50-minute lessons in the Graphical Debate- or Textual Debate-tool. Moreover, they worked another five lessons on the other two parts of the inquiry task. Similarly, in the Van Drie et al. study, students worked in the collaborative environment during six 50-minute lessons. In contrast, in the Suthers and Hundhausen study for example, students worked on their representations for only one session of approximately 45 minutes. Suthers and Hundhausen hypothesized that this might not have been sufficient for learning outcomes to fully develop. This may explain why in our and the Van Drie et al. study, differences between representations were found, while these were not found in the studies by Suthers and Hundhausen. In our case, students possibly worked long enough in the CSCL environments for learning outcomes to develop. Moreover, differences in learning outcomes between the Graphical Debate- and Textual Debate-tool also had sufficient time to develop.

Another explanation may lie in the type of problem. In this study, the goal was to provide a complete as possible representation of the historical debate. This is possibly in conflict with what is normally understood under argumentation, namely to persuade someone (e.g., a group member or a fictional audience) of one's own point of view. These different goals may lead to a different approach of the task by students and subsequently to different learning outcomes. For example, when Nussbaum (2005) instructed students to generate as many reasons as possible, they engaged more, and more deeply in argumentation than students who were instructed to persuade someone. Thus goal instruction may be another important factor to consider when studying the effects of representational guidance.

In previous research, representational guidance has been found to affect students' *collaborative activities*. However, guidance was only found to have an effect on 4 of the 19 collaborative activities investigated. The Graphical Debate-tool was found to have an impact on *shared understanding* during collaboration, as students' discussions showed more indications of shared understanding, and less indications of loss of shared understanding. Furthermore, these groups also typed more *negative evaluations of the social process* and made more *neutral technical remarks*. This may be an indication that collaborating while using the Graphical Debate-tool is more complex. Furthermore, we found no effect of condition on metacognitive activities (i.e., planning, monitoring, and evaluating the task process). Therefore our study offers no support for the hypothesis that representational guidance decreases group members' need to coordinate and regulate their task performance in the online discussions. Moreover, the

Graphical Debate-tool did not trigger students to question each other more or to generate more ideas about the topic. This last finding may be explained by the fact that students could use the representations in both the Graphical Debate-tool and the Textual Debate-tool to exchange information (Van Amelsvoort, Andriessen, & Kanselaar, in press; Van Drie et al., 2005). Because both tools were shared, adding an element to the representation equates to exchanging information with group members. Thus, there might be less need to engage in extensive information exchange during the chat discussions.

Our last two research questions addressed the impact of representational guidance on *students' perceptions*. Contrary to our expectations, guidance did not have an impact on students' perceived usefulness of the Debate-tool. Students in both conditions perceived their tool as useful for the inquiry task. When looking at group and student performance, these results are not so surprising. In both conditions, groups constructed representations and wrote essays of at least average quality and student performance increased from pre- to post-test. In this sense, both representations were useful to the students. Furthermore, neither tool was designed to be aversive, thus equal levels of perceived usefulness in both conditions were perhaps to be expected. Also, students were not able to compare the two versions of the Debate-tool directly, because they only worked with one version of the tool. Therefore, we were not able to ask which tool they thought was more useful, but were forced to ask them to rate the usefulness of the version of the tool they used. Finally, one can question the validity of measures that ask students to rate the usefulness of something. Often students' perceptions of the usefulness of something do not correspond to its effectivity or efficiency. Students may be more concerned with 'getting the job done', than with constructing the best possible representation and engaging in argumentative discussion. Thus what is useful for students may be very different from what we as educational designers consider useful (e.g., a tool that enables them to complete the task as fast as possible versus a tool that triggers deep and mindful study of the topic at hand).

We expected the Graphical Debate-tool to have positive effects on students' perceptions of their online communication and collaboration. This was not confirmed. On the contrary, students in the Textual Debate condition reported significantly higher levels of positive behavior. However, the negative effect of condition on perceived positive group behavior should not be interpreted to mean that Graphical Debate students did not enjoy their collaborative experience. On a scale from 1 to 5, students in the Graphical Debate condition rated positive group behavior with an average of 3.69. This indicates they too experienced the collaborative experience positively. On the other hand, when we examined students' online collaborative activities, we found they more often voiced negative evaluations of the collaborative process and that they made more remarks about technical aspects of the environment. Possibly, students experienced the Graphical Debate-tool as more a difficult and complex tool for collaboration. This may explain why they perceived their collaboration less positively than students in the Textual Debate condition. Furthermore, as stated above, students often approach argumentation in a quick and easy manner and are often aimed at maintaining consensus (Felton & Kuhn, 2001). Therefore, a tool which forces them to argue more elaborately and critically may lead to more negative feelings toward the collaborative process. This sometimes called the *paradox of collaborative learning*: we assume students learn from elaborating, arguing, and criticizing, while students themselves are often aimed at reaching consensus to complete their task as fast and efficient as possible (Erkens et al., 2006).

This study was conducted to investigate how representational guidance affects co-construction of ERs and employed two conditions. In general, the results are in favor of the Graphical Debate condition. However, our design does not permit us to conclude which elements of the Graphical Debate-tool contributed to these positive effects. Was it the graphical nature of the tool? Was it the visualizations that gave feedback about strength of support of a position and the complexity of argumentation? Or was it a combination of both? To be able to answer these questions a more elaborate research design is needed. On the other hand, based on the results of this study it could be argued that the *representational guidance as a whole* offered by the Graphical Debate-tool was more beneficial than the guidance offered by the Textual Debate-tool. More research is needed to disentangle the effects that specific elements of representational guidance have on the co-construction of ERs.

This study shows the advantage of the Graphical Debate-tool over the Textual Debate-tool. We were able to demonstrate the effectiveness of the Graphical Debate-tool in terms of helping groups of students to construct meaningful and correct representations and clearly argued and conceptually correct essays. Furthermore, while working with the Debate-tool, students acquired more knowledge about the topic of study. These are promising and important results, since increasing group and individual achievement is an important hallmark for establishing the effectiveness of an educational design (Kirschner, 2002; Kirschner, Strijbos et al., 2004). One direction for future research on the effects of representational guidance may be to conduct more elaborate experiments to determine which specific aspects (and combinations of aspects) of representational guidance are most effective. Yet another direction may be to combine the use of ERs with instruction in proper forms of argumentation as secondary education students often lack the necessary skills to engage in argumentation (Kuhn & Udell, 2003; Nussbaum, 2005). Finally, the possible interaction between the goal of the task (e.g., to generate as much arguments as possible or to persuade) and representational guidance deserves further study. Whichever direction future research takes, the results of this study firmly point to the need for teachers and researchers to carefully consider the representational guidance they offer to students when using representation tasks in their teaching or research.

4.7 Acknowledgements

The authors would like to thank Jos Jaspers and Marcel Broeken for their technical assistance, Ilse Mink for her help during the data collection, and Renske de Kleijn for her assistance during the data analyses.

5. General Conclusion and Discussion

In Chapter 1, we started by describing some problems that may occur during computer-supported collaborative learning (CSCL) such as a lack of awareness, communication problems between group members, coordination problems, and a lack of quality discussions. We also outlined how visualizations could help to address these problems by (1) making complex information easier to interpret thereby decreasing the cognitive demands placed on group members, (2) giving feedback about important aspects of the collaborative process, (3) raising group members' awareness of the collaborative process, (4) facilitating coordination of online collaboration, and (5) providing a motivational incentive. In the chapters that followed, we described three tools that were developed to address some of these problems. We also examined the effects of these three tools on the process of collaboration (e.g., increase in awareness, facilitation of coordination) and the outcomes of collaboration (e.g., group performance, perceptions of the collaborative process). Table 5.1 gives an overview of the developed tools and the processes and outcomes that were examined. In the next sections, we give a summary of each study and use the results of the three studies to answer the general research question.

5.1 Summary of the Studies and General Conclusion

Before answering the general research question formulated in Chapter 1, we first summarize the results of the studies described in Chapter 2, 3, and 4.

5.1.1 Study 1: Visualizing Participation

Chapter 2 described the impact of the Participation-tool, a tool developed to visualize group members' relative contributions to a chat discussion, on online collaboration. When working in CSCL environments, it is often difficult to determine what other group members are doing and whether everyone is contributing equally to the group task (Bødker & Christiansen, 2006; Kreijns, 2004). Furthermore, coordination problems during CSCL sometimes occur. When students work together, they have to coordinate their activities. This not an easy task, because there are many different activities that group members must carry out during teamwork. To coordinate collaboration, group members must ensure that everyone participates in the collaborative process (e.g., by proposing strategies or externalizing ideas and knowledge). Furthermore, participation needs to be distributed equally among group members: no group member should dominate the collaboration, nor should group members engage in free riding behavior (Erkens, 2004; Teasley, 1995). The Participation tool was designed to specifically address the lack of information about group members' participation in CSCL environments. Additionally, it is intended to facilitate the coordination of collaboration, more specifically the participation and equality of participation.

To examine the effects of the Participation tool, 17 groups of secondary education students (*N* = 52) were given access to the Participation tool while working in the VCRI environment, while 5 groups (*N* = 17) were not given access to the tool. These groups worked on a historical inquiry task about witches and the persecution of witches. We examined the impact of the Participation tool on awareness, coordination, and group performance.

The Participation tool is a tool that gives feedback to students about their participation and the participation of the other group members. We hypothesized that this would increase group members' awareness of the group processes and activities taking place during their collaboration. This assumption was largely unconfirmed: the Participation tool was not found to have an effect on students' awareness. It is worth noting however, that group members with access to the Participation tool reported they knew better when another group member was *taking a free ride*. The Participation tool therefore enabled students to determine free riding behavior more easily.

The results of the first study showed the *Participation tool to have an effect on participation*: students who had the option to inspect the Participation tool were found to have higher participation rates. *No effect of the Participation tool on equality of participation* was found however. On the other hand, the intensity with which the Participation tool was used by students was found to have an effect on equality of participation: more equal participation was found in groups that used the tool more often. This indicates that giving group members access to the Participation tool and use of the Participation tool stimulates participation and equality of participation.

We also examined the groups' collaborative activities to determine whether the Participation tool had an effect on coordination strategies. We demonstrated that students who could use the Participation tool *engaged more in coordination and regulation of social activities* by typing more statements that addressed the planning of their collaboration. This indicates that visualizing participation stimulates group members to also devote effort to coordinating their collaboration.

Finally, we expected that through stimulating participation and equality of participation, raising awareness, and facilitating coordination group performance would increase. To test this assumption, we examined the quality of the products made by the participating groups. *No effect of the Participation tool on the quality of group products was found*, however.

5.1.2 Study 2: Visualizing Agreement and Discussion

Chapter 3 described a visualization called Shared Space. This visualization was designed mainly to address two different problems that sometimes occur during online collaboration: (1) communication problems (e.g., misunderstandings or difficulties conveying opinions) due to the leanness of chat communication and (2) absence of critical and constructive discussions. The

Shared Space visualizes whether there is agreement or discussion among group members during their discussions. We hypothesized that by giving group members feedback about the types of discussions they are conducting (e.g., consensual discussions or critical discussions) their awareness about their conversational strategies would increase. This could help group members to adopt group norms that are aimed at critical and constructive discussion instead of consensual discussion. Furthermore, the Shared Space could stimulate students to discuss and evaluate their discussion strategies and thus facilitate coordination of the collaborative process.

We examined the effects of the Shared Space by giving 20 groups of students access to chat-tool with Shared Space visualization, while 20 other groups had access to the regular chat-tool of the VCRI environment. Like the students in Study 1, these students worked on a historical inquiry task in the VCRI environment. This time, subject of the task was the first centuries of Christianity.

The results of Study 2 suggest that the Shared Space was effective in *increasing the richness* of the VCRI environment for online communication. Students who communicated through a Chat-tool with Shared Space visualization perceived the medium as more suitable for communication than students who communicated through a version of the Chat-tool without Shared Space visualization. These students reported they experienced less communication problems.

Our analyses of students' collaborative activities during their chat conversations confirm this finding and indicate the Shared Space *facilitated coordination*. Students with access to the Shared Space were less busy exchanging information and typed fewer messages aimed at reaching and maintaining mutual understanding. This indicates that for these students it may have been *easier to understand* their group members' contributions. In other words, the process of negotiation of meaning was facilitated.

Concerning the quality of the online discussions, the Shared Space was found to have an *effect on exploratory group norm perception*. This indicates that groups that had access to the Shared Space perceived their discussions as more critical and constructive than did groups without access to the Shared Space.

With respect to the outcome of collaboration, analyses of questionnaires tapping into group members' perceptions of their collaboration show the Shared Space to have *an effect on these perceptions*. Students with access to the Shared Space reported more positive group behavior (e.g., helping behavior, positive group atmosphere) and more effective group task strategies than students without access to the Shared Space. In other words, they perceived their collaboration as more positive and effective.

Lastly, we examined the effect of the Shared Space on group performance. We hypothesized that by stimulating students to adopt exploratory group norms and by decreasing communication problems, the Shared Space would also increase group performance. This hypothesis was only partly confirmed as groups with access to the Shared Space only *outperformed groups without access on one part of the group task* (out of a total of three parts).

5.1.3 Study 3: Increasing Representational Guidance using Visualizations

Unlike the Studies 1 and 2 which focused on visualizing social aspects of the collaborative process, Study 3 focused on visualizing task-related aspects of this process. Chapter 4 described the Graphical Debate-tool and the Textual Debate-tool and their effects on the quality of representations co-constructed by groups, essays written by groups, learning of historical concepts and facts, and students' perceptions of the tool and the collaborative process.

Both tools were used by groups of students to co-construct a representation of a historical debate about the magnitude of the persecution of Christians in the Roman Empire. The Graphical Debate-tool visualizes how well positions are supported by arguments and supporting information through automatic repositioning of the position: each time an argument or a piece of supporting information is added to the representation the corresponding position is moved closer to the center of the representation, indicating it is better supported. Adding refuting information moves a position away from the center. This way, students can easily see which position is best supported. The Textual Debate-tool does not provide such support. Furthermore, the Graphical Debate-tool allows group members to quickly see which position has been given more attention and whether there has been a tendency to focus on a certain type of information (e.g., supporting or refuting information). This way, the group's progress through the problem is visualized (Cox, 1999). In sum, in Study 3 visualizations of task-related aspects were used to increase the *representational guidance* offered by the Graphical Debate-tool (Suthers, 2003; Van Drie et al., 2005).

We hypothesized that by visualizing task-related aspects of the collaboration in thus increasing representational guidance, the collaborative process would become more effective and efficient. First, the Graphical Debate-tool makes complex information easier to interpret. Without feedback about the balance between two positions, it is much harder to decide which position is better supported by arguments and supporting information. The Graphical Debate-tool may facilitate the drawing of such a conclusion. Furthermore, coordination of the collaboration may also be easier because there is for example less need to coordinate, regulate, and monitor task performance.

Compared to the groups that worked with the Textual Debate-tool ($N = 45$), the groups that worked with the Graphical Debate-tool ($N = 79$) performed rather well. *Their representations and essays were of significantly higher quality.* In addition, they performed *significantly better on a knowledge post-test*. This shows visualization of task-related aspects of the collaboration has an impact on group and individual performance. On the other hand, we found *no evidence that the Graphical Debate-tool facilitated coordination* in the groups' chat protocols. Students that worked with the Graphical Debate-tool were engaged in the planning, monitoring, evaluating their task progress as much as students who worked with the Textual Debate-tool. In sum, we *did not find evidence that the Graphical*

Debate-tool facilitated the coordination of collaboration. Strikingly, in the Graphical Debate condition, students *perceived their collaboration as less positive*.

Table 5.1: Overview of the studies reported on in this thesis, the tools developed and the processes and outcomes examined.

Study	*Tool*	*Problem(s)*	*Process(es)*	*Outcome(s)*
1	Participation-tool	• Lack of awareness • Coordination problems, specifically participation and equality of participation	• Awareness (+/-) • Coordination, specifically (equality of) participation (+), and coordinative activities (+)	• Group performance (+/-)
2	Shared Space	• Lack of awareness • Communication problems • Coordination problems • Lack of quality discussions	• Media richness, decrease of communication problems (+) • Group norm perception (+) • Coordination (+)	• Perceptions of collaborative process (+) • Group performance (+)
3	Graphical Debate	• Lack of quality discussions • Coordination problems	• Decreasing effort to understand information (+/-) • Coordination (+/-)	• Perceptions of collaborative process (-) • Group performance (+) • Individual learning (+)

Note. Signs in parentheses indicate effects of visualizations on processes and outcomes of collaboration: (+) indicates a positive effect, (+/-) no or mixed effect, and (-) a negative effect.

5.1.4 Answering the General Research Question

We now turn our attention to answering the general research question of this thesis, which was formulated as follows:

> *How and why does giving students access to visualizations of the collaborative process affect the collaborative process, students' perceptions of the process, and group or individual performance?*

The first part of the general research question addresses *how* visualizations affect the outcomes of collaboration (see Table 5.1). The Shared Space was found to have positive effects on students' perceptions of the collaborative process, and to increase group performance on one part of the group task. The Graphical Debate-tool was found to increase group performance (e.g., groups constructed better representations and wrote better essays) and individual learning of historical facts and concepts. On the other hand, the Participation tool was not found to have an effect on group performance. In sum, based on the studies in

Chapters 2, 3, and 4, there is evidence that the visualizations can, at least in part, positively affect the outcomes of online collaboration.

The second part of the general research question addresses *why* visualizations may affect collaboration (see Table 5.1). Study 1 offers some evidence that by giving feedback about participation levels and equality of participation, (social) *awareness* is raised (students who worked with the Participation tool reported they knew better when there was free riding behavior in their group). In Study 3 the Graphical Debate-tool was developed to raise students' awareness of several task-related aspects of the collaborative process (e.g., Have both positions been given equal attention? Is one position supported better by arguments? Has refuting information been neglected?). Although groups that worked with this tool outperformed groups that worked with the Textual Debate-tool, students did not report a higher usefulness of the tool (e.g., they did not perceive that the tool helped them to assess whether one position was better supported by arguments than the other). Although awareness was identified as an important aspect of the collaborative process, the importance of awareness received mixed support in both studies.

Communication problems were addressed by the Shared Space. In CSCL environments, students sometimes have difficulties interpreting each others' messages due to the leanness of the medium (Daft & Lengel, 1986; Dennis et al., 1999). The Shared Space successfully increased the *media richness* of the VCRI environment, by visualizing agreement and discussion during online discussions. The results of Study 2 show this visualization to have the intended effect: students perceived the medium as richer and reported less communication problems. This was also confirmed by the analyses of students' online collaborative activities. Chat protocols of groups with access to the Shared Space showed these groups to devote less effort to reaching and maintaining mutual understanding. In sum, our research shows that visualizations can be effectively used to decrease communication problems during online collaboration, by increasing the richness of the environment for communication.

Quality discussions are considered important for successful collaborative learning (Mercer, Wegerif, & Dawes, 1999; Van der Linden et al., 2000). However, students do not always engage in quality discussions during collaboration. This problem was addressed in Studies 2 and 3. In Study 2, the Shared Space gave group members feedback about the types of discussions they were conducting. The students who had access to the Shared Space during their online conversation, perceived their discussions to be more exploratory than students who did not have access to the Shared Space, meaning they perceived their discussions as critical and constructive. These types of discussions have been found to be of importance for learning to occur during collaboration (Wegerif et al., 1999).

While the Shared Space tried to stimulate quality discussions by visualizing agreement and discussion, the Graphical Debate-tool tried to foster these kinds of discussions by focusing students on the difference between arguments, supporting information, and refuting information. We expected that this would lead to a more in-depth discussion of the historical facts, concepts, and opinions found in the historical sources. Our analyses of the online discussions between group members did not confirm this. Students who worked with the Graphical Debate-tool did not exchange more task-related information, nor did they ask more task-related questions than their counterparts in the Textual Debate condition. An explanation for this finding may be that when students are given a representation like the Debate-tool, there is less need to discuss things in the chat, because the representation itself serves as an externalization of knowledge (Van Drie et al., 2005). Through the representation itself, students may engage in "discussion" by adding arguments and supporting or refuting information. Our analyses did show that groups that worked with the Graphical Debate-tool constructed representations of higher quality than groups working with the Textual Debate-tool. This may be interpreted to indicate that the "discussions" through the Graphical Debate-tool were of higher quality than those through the Textual Debate-tool. The conclusion that may be drawn from this is that visualization of social aspects of collaboration (i.e., visualizing the types of discussions group members are conducting) can have a direct effect on the (perceived) quality of group members' discussions. On the other hand, visualizing cognitive aspects of collaboration may not have direct effects on the discussions through a chat tool, but may still improve "discussion" through a representation.

The final types of problems we sought to address with the developed visualizations were *coordination problems*. Collaboration is a difficult task, because it involves the enactment of many different activities. Furthermore, these activities are carried out by different group members, making the task even more difficult. Without effective coordination, group members are likely to engage in conflicting or repetitive activities, thus disrupting the collaboration itself (Malone & Crowston, 1992). In Chapter 1 we outlined several different strategies that group members can apply to ensure successful coordination of their collaboration, namely (a) activation of knowledge and skills, (b) grounding, and (c) negotiation and coming to agreement.

In Study 1 the Participation tool tried to facilitate coordination and specifically participation and equality of participation (i.e., each group member has the opportunity to apply his/her knowledge or skills) by visualizing group members' contribution to the online discussion. We expected that by giving group members access to such a visualization they would be stimulated and motivated to increase their participation and to strive for equal participation. Furthermore, we also expected that these visualizations would trigger discussions about the collaboration itself. In other words, we expected group members to discuss about their collaboration: How were to going to proceed with their collaboration? Was everyone contributing sufficiently to the collaboration? These expectations were mostly confirmed. We were able to show that students who used the Participation tool participated more by sending more elaborate messages than students who did not use the Participation tool. Furthermore, our analyses of the online dialogues also show that this increase could not be attributed to group members trying to deliberately manipulate the visualization (e.g., by typing nonsense or irrelevant remarks to increase the size of their spheres). In fact, our analyses showed that students who had access to the Participation tool engaged in less social talk, but discussed more about the planning of their collaboration. This indicates that the Participation tool not only stimulated students to activate their knowledge and skills by increasing their participation in the online discussions, but also encouraged them to discuss more about the coordination of their collaboration. In sum, the Participation tool was found to have an effect on coordination of group members' collaboration.

The Shared Space, described in Study 2, focused on other aspects of coordination, namely grounding and negotiation and coming to agreement. By visualizing agreement and discussion among group members we hoped to facilitate the grounding process among group members. In Chapter 1 we proposed that checking could be considered an important activity for group members to establish common ground (Erkens, 2004). When group members wish to create a mutual understanding and common frame of reference (i.e., grounding), they need to check whether their teammates understand the information or proposals being expressed. The Shared Space facilitates this process by analyzing each chat contribution and updating the visualization when the contribution can be considered an expression of agreement or discussion. Thus, we hypothesized that by giving group members access to a Chat-tool with Shared Space visualization, the process of coming to mutual understanding would become easier. Our analyses of group members' online discussions indicate that this might have been the case. Students that had access to the Shared Space were less busy with reaching and maintaining mutual understanding. By receiving feedback about their group members' contributions (e.g., did they agree or disagree?) their grounding process was facilitated and therefore they needed to communicate their understanding less of group members' contributions. On the other hand, we identified activation of knowledge and skills as a strategy to coordinate collaboration. However, the Shared Space was found to have a negative effect on the amount of information exchange within the group and the amount of task-related question that were being asked. Thus, the grounding process was facilitated by the Shared Space, but the activation process was not. In sum, the visualization used in Study 2 had positive effects for the grounding process (i.e., the process was facilitated, group members devoted less effort to it), but had negative effects for the activation process.

The Graphical Debate-tool, described in the final study, was developed - among other things – to facilitate the coordination process of co-construction of external representations. We hypothesized that by visualizing task-related aspects of the collaborative process, group members would be stimulated to engage more deeply in the process of argumentation, negotiation, and coming to agreement. Furthermore, we hypothesized that these students would not have to devote as much effort to coordination and regulation of their task-performance, since the Graphical Debate-tool gives students feedback about their progress through the task (Cox, 1999). These hypotheses were not confirmed. Groups that worked with the Graphical Debate-tool did not engage less in the coordination and regulation of their task-related activities than groups that worked with the Textual Debate-tool. Furthermore, we did not find that these groups exchanged more task-related information, which indicates the argumentation and negotiation process was not improved. In sum, the Graphical Debate-tool was found to have positive effects on group and individual performance, but these effects could not be attributed to an improved coordination process within these groups.

In conclusion, the studies described in this thesis highlight the potential of using visualizations in CSCL-environments. The Graphical Debate-tool and - to a lesser extent – the Shared Space were found to have positive effects on the outcomes of collaboration in terms of better group and individual performance. Additionally, the Shared Space was found to have positive effects on group members' perceptions of the collaborative process. Furthermore, these studies also shed light on the processes through which visualizations can enhance the outcomes of collaboration. We demonstrated that visualizations can be used to raise group members' awareness, increase the richness of CSCL-environments, stimulate quality discussions among group members, and facilitate the process of coordination.

5.2 General Discussion

In the previous paragraph we answered the general research question of this thesis by stating that visualizations can be used to enhance the collaborative process. We stated that increased awareness and richness of the online environment and on improved discussion quality and coordination process may account for these effects of visualizations. However, some questions and issues concerning the use of visualizations in CSCL-environments and CSCL in general remain unaddressed. We conclude this Chapter by raising some of these issues.

5.2.1 Measuring How Visualizations Affect Collaboration

The three studies reported on in this thesis all try to determine how visualizations affect collaboration. For this purpose we developed several instruments. These instruments have their strengths and weaknesses. For example, in Study 1 we measured the effects of using the Participation tool on students' awareness of group processes and activities using a questionnaire. One could argue that questionnaires may have their limitations when measuring awareness. Rather, one could also use think-aloud or stimulated recall techniques to identify whether students' verbalizations or recalled experiences showed indications of increased awareness due to their use of the Participation tool (Ericsson & Simon, 1993; King, 1993; Meijer, Zanting, & Verloop, 2002). Furthermore, students' online chat discussions also showed indications of students using the Participation tool to increase their awareness (e.g., students discussing the size of their spheres). An integrated approach to analyzing awareness using multiple methods might shed additional light on the effects of visualizations on group members' awareness during online collaboration (cf., Hmelo-Silver, 2003).

Similar arguments could be made about the way we tried to study the effects of visualizations on the quality of students' discussions. In Study 2 for example, we examined the effects of the Shared Space on students' *perceived* group norms using questionnaires. Students answered questions about the nature of the discussions (critical, consensual, or exploratory) that were held in their groups. These group norm perceptions have been found to influence the collaborative process (Postmes et al., 2001). Although this gives an indication about how students perceived the quality of their discussion, an analysis of their actual discussions is needed to determine whether students' perceptions correspond to their behavior. Although we did analyze students' discussions, our coding scheme was not specifically designed to assess the argumentative quality of online discussions. With our coding scheme we can identify for example, whether students engage in information exchange and questioning. However, exchanging information and asking questions are necessary but not sufficient prerequisites for quality argumentation. An analysis of group members' argumentation in terms of for example breadth (number of topics addressed) and depth (sequence of formulating arguments and counterarguments) of the discussion (cf., Munneke, Andriessen, Kanselaar et al., 2007) would shed additional light on the impact of visualizations on the argumentative quality of group discussions. An analysis of the breadth and depth of argumentation could also have been useful to more deeply examine the effects of the Graphical Debate-tool in Study 3 on the quality of students' discussions.

In Chapter 1 we identified facilitation of coordination as an important mechanism through which visualizations can have an impact on online collaboration. To determine the effects of the three developed visualizations we used a coding scheme which identified the different activities (task-related activities, coordination of task-related activities, social, and coordination of social activities) students engage in during collaboration. This coding scheme proved useful to determine to what extent students coordinate task-related and social aspects of their collaboration (e.g., Do they engage in the planning, monitoring, and evaluation of task-related and social aspects of collaboration?). In Chapter 1 we also identified three different coordination processes: activation of knowledge and skills, grounding, and negotiation and coming to agreement. Although our coding scheme provides general information about these processes, more detailed analyses could also have been valuable. For example, our coding scheme addresses the process of grounding and creating a common frame of reference by assessing whether the online discussions contain indications of whether group members are busy reaching and maintaining mutual understanding or whether there are indications mutual understanding has not been reached. More detailed analyses of the grounding process could for example include analyses of whether group members respond appropriately to proposals (this is an indication of joint attention, cf. Barron, 2003), whether group members check whether the information that is being communicated fits with the common frame of reference (cf., Beers, Boshuizen, Kirschner, & Gijselaers, 2007; Erkens et al., 2005), or whether the ideas and suggestions offered by one group member are used by other group members in the remainder of the discussion and whether these ideas and suggestions are used in the group product (cf., Beers, 2005). In short, more detailed analyses of the effects visualizations on coordination processes could provide valuable information about how visualizations facilitate coordination of collaboration.

5.2.2 Visualizations: Task-related versus Social

In Chapter 1 we distinguished between visualizations that visualize task-related aspects of the collaboration and visualizations that visualize social aspects of the collaboration. The Participation tool and Shared Space are examples of visualizations of social aspects of collaboration, while the Graphical Debate-tool visualizes task-related aspects of this process. The question could arise which kind of visualizations are more valuable. We think both kinds of visualizations have their merits. As we have seen in Studies 1 and 2, visualizing social aspects of the collaboration mainly influences the *process of collaboration* and enhances *students' perceptions of the collaborative process*. On the other hand, in Study 3 visualizing task-related aspects of collaboration was found to have mostly positive effects on *group and individual performance*. Thus, one could argue that, based on the results of this thesis research, visualizing task-related aspects is mainly beneficial for the effectiveness of collaboration (group and individual performance may be enhanced) while visualizing social aspects is mainly beneficial for how enjoyable the collaborative experience is. Although educational research tends to focus on making the collaborative process more effective instead of making it more enjoyable (Kirschner, 2004), we think that the affective dimension of collaboration is also important. Thus, it is a promising finding in its own right that visualizing social aspects of the collaboration seem to make the collaborative process more enjoyable. Clearly, both goals (increasing the effectiveness and the enjoyability of collaboration) are worth striving for.

5.2.3 Design and Use of Visualizations

Another question that may arise is how to proceed with implementing visualizations in our CSCL-environments. Should we try to incorporate as much different visualizations in these environments? Should we force group members to use a visualization or should we let them decide for themselves whether they want to use it? Should we try to design visualizations that combine several task-related and/or social aspects in one visualization so as not to overload the user with too many different tools? What other aspects of collaboration should we try to visualize? Of course the answers to these questions depend on many things.

First, it depends on the users of these visualizations and the tasks they are working on. If students are not co-constructing external representations it makes no sense to offer them the Graphical Debate-tool. It would probably only distract them. Furthermore, student and group characteristics may also influence the decision about which types of visualizations are implemented. For example, the Shared Space was designed – amongst other things - to increase the richness of the VCRI-environment to prevent communication problems. However, if the collaborative process involves groups that share a social history (e.g., group members have collaborated before on projects or groups that consist of friends), group members may not have difficulties interpreting group members' messages and may not find it difficult to determine whether there is a form of shared understanding within the group (Adams et al., 2005). It could be hypothesized that such groups would not benefit from using the Shared Space.

Similarly, group characteristics may also influence the usefulness and effectiveness of visualizations. The Participation tool may for example be more useful in large groups. In large groups, there is a greater risk that free riders remain unnoticed (Bonito, 2000); it is easier to "hide in the crowd" in large groups. Thus, it might not be very useful to give a dyad access to the Participation tool because dyad members may be more aware of whether there is equal participation (i.e., the visualization contains redundant information). On the other hand, even in small groups it may be beneficial to stimulate discussion about the collaborative process itself.

Another question is whether students should be forced or coerced to use the visualizations that are implemented in CSCL-environments. In Study 1, use of the Participation tool was not mandatory, thus students could ignore or close the tool if they wanted to. Therefore, we found considerable differences between students regarding the extent to which they used the Participation tool. On the other hand, in Study 2, the Shared Space was integrated in the Chat-tool. Because the Chat-tool was the primary mode of communication for group members, use of the Shared Space was somewhat coerced. Although it is of course possible that some group members chose to ignore the visualization of agreement and discussion offered by the Shared Space. Finally, use of the Graphical Debate-tool was necessary to complete the group task. Group members needed to construct a representation of the historical debate in the tool to complete the task, thus using the visualization was mandatory. The question remains which strategy works better: pressuring students to use a visualization or not? More research is needed to answer this question.

In our last study group members worked with the Graphical Debate-tool, but they also had access to the Participation tool and Shared Space. One may ask whether by giving students access to three different visualizations, students do not become overloaded with all the information they need to interpret. This is another important question that designers face: how to optimize CSCL-environments by using visualizations without placing to great a burden on group members' information processing capabilities (Kirschner et al., 2006; Van Bruggen et al., 2002).

Finally, the question remains what other aspects of collaboration should be visualized in online learning environments. In our opinion this question should be answered by carefully considering the users of our learning environments (Kirschner, 2002; Kirschner, Martens, & Strijbos, 2004). This involves studying what group members do in these environments and what problems they encounter. It also involves determining what kind of support offered by visualizations is actually needed by group members. Furthermore, research should be conducted to determine how group members use and perceive the visualization, and whether the visualization increases the effectiveness, efficiency, or enjoyability of the collaborative experience. By adopting such an approach to the design and development of visualizations we likely prevent ourselves from designing tools that are developed merely because they *can* be developed.

5.2.4 Domain and Tasks

The studies conducted for this thesis were carried out for the subject of history. In all three studies, students worked on an inquiry group task. These tasks required group members to study historical sources to gather enough information to write essays, answer questions, or construct representations. Obviously, these types of tasks are very complex and demanding. When using these kinds of tasks, there is the possibility that students become overburdened by its complexity, and thus student learning is hampered (Kirschner et al., 2006). The question therefore arises whether (a) the tasks used in our studies were too complex and (b) task complexity might have influenced our results.

The tasks used in the three studies were definitely complex. Students had to read many sources, and interpret and synthesize them. Research on how students search information and process information on the World Wide Web for example, shows students often have difficulties finding good information, determining the usefulness of the information found, and processing the information to answer research questions (cf., Kuiper, Volman, & Terwel, 2004). Additionally, group members' learning process was self-directed: group members had to discover the correct way to approach such a task for themselves (cf., De Jong & Van Joolingen, 1998; Swaak, Van Joolingen, & De Jong, 1998). Furthermore, students had to manage the complex task of coordinating their collaboration. Moreover, students were expected to engage in interactive argumentation about historical concepts and opinions. Finally, students were required to synthesize all of the information in collectively written essays. It is therefore not surprising the participating students worked intensely in the VCRI for several lessons. But were tasks too complex? We can of course not be sure, but the results reported in Chapter 2, 3, and 4 concerning group performance seem to indicate most groups performed at least adequately. Additionally, the tasks we developed for the reported studies were created in collaboration with the participating teachers to ensure they were within students' capabilities. Furthermore, inquiry tasks are not uncommon in the second stage of secondary education. Students were therefore reasonably familiar with the demands these tasks placed on them. The tasks may have therefore been complex, but for most groups they were not too complex.

Could task complexity then have influenced the effectiveness of the visualizations used in our studies? This question can be approached from two different viewpoints. First, one could argue that if the task is complex, students can benefit the most from visualizations. After all, they are designed to make the students' task a little bit easier. On the other hand, it could also be argued that when a task is already complex, adding visualizations that also require students' attention to an online learning environment makes the task even more complex. This may be especially true if the visualization is not properly designed or its intention is not clear to its users. Obviously, future research should consider the question under which degrees of task complexity visualizations are most effective.

Our studies were all carried out in the subject of history. This subject is very different from for example subjects in the science domain (e.g., physics or chemistry). For example, in the science domain it may be easier to determine whether group members' solutions, explanations, or answers are correct or incorrect. In learning environments for the domain of science it may therefore be easier to visualize task-related aspects of the collaborative process. In scientific simulations students or groups of students therefore often work on assignments and the simulation often provides students with feedback about the correctness of group members' answers and their progress through the task (cf., Gijlers & De Jong, 2005). For the subject of history, which is in the argumentative domain, using visualizations to give feedback about task-related aspects of the collaboration remains a challenge. Study 3 however, illustrates that such visualizations may also be valuable for the subject of history.

A final remark concerns the influence of the domain on the types of discussion and argumentation group members engage in. For all three tasks, discussion and argumentation were considered important. However, students were mostly required to discover and discuss arguments that could be found in or inferred from the available historical sources. Their goal was therefore to get a clear as possible picture of the arguments pertaining to the subject of the task. In the argumentative domain, students often have to persuade someone (e.g., a group member or a fictional audience) of their own point of view. These different goals may lead to a different approach of the task by students, a different form of argumentation, and subsequently to different learning outcomes (Munneke-de Vries, 2008; Nussbaum, 2005). In other subjects, where other types of argumentative tasks are used, the argumentation process is probably different and therefore the impact of visualizations on this process may also be different.

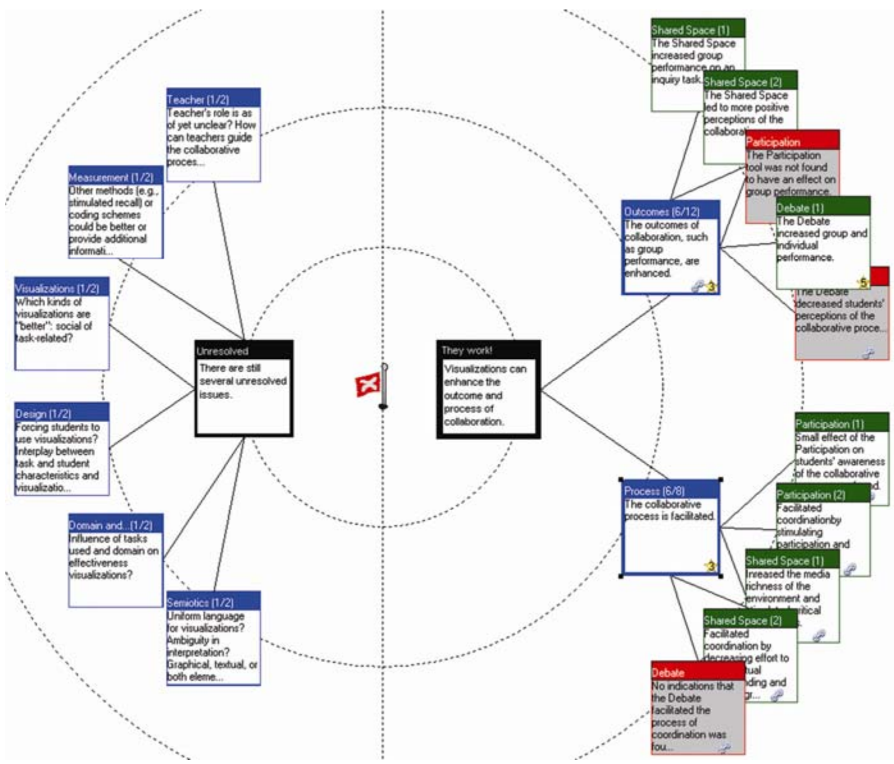
5.2.5 The Language of Visualizations

A final remark we wish to place her concerns the "language" (semiotics) of visualizations. As was noted in Chapter 1, the research into the effects of visualizations in CSCL-environments has just begun. It is therefore not surprising that there is no uniform way to visualize task-related or social aspects of collaboration. The examples shown in Chapter 1 highlight this observation. Several visualizations aimed to visualize participation, but they all differed in how they visualize participation (our Participation tool described in Chapter 2 is no exception to this). Considering the complexity of the aspects of collaboration that can be visualized in CSCL-environments, it is not so surprising that there is no uniform way of visualizing them. However, this means that users probably also have no uniform way of perceiving and interpreting these visualizations. Most users probably have never come across such visualizations before. How then do our users know how to interpret these visualizations? And do they interpret them correctly? Is it a problem if users interpret them incorrectly? And what happens if the users do not agree with the visualization? All of these questions need to be addressed if we are to gain a deeper understanding of how visualizations may or may not help to improve online collaboration. Furthermore, many visualizations (and certainly the visualizations developed in this thesis) use mostly graphical elements to visualize task-related and social aspects of the collaborative process. It is also worthwhile investigating whether some aspects can be better visualized using textual elements, graphical elements, or a combination of both.

5.2.6 Role of the Teachers

The research in this thesis focused on the effects of visualizations on students' behavior, perceptions, and performance. Up to now, we did not discuss the role of the teacher in this process. The teacher's role during collaboration, however, can be very important: The teacher can discover and address misconceptions, can model appropriate collaborative behavior, ask critical questions, provide scaffolding, and monitor task progress (Berge, 1995; De Laat, Lally, Lipponen, & Simons, 2006; Hmelo-Silver, 2004; Paulsen, 1995; Salmon, 2000). Furthermore, students often copy or model many of the activities carried out by their teachers (Gillies, 2006; Webb, Nemer, & Ing, 2006). CSCL-research tends to overlook or neglect the role of the teacher in guiding groups of students during online collaboration (exceptions are for example Anderson, Rourke, Garrison, & Archer, 2001; De Laat, Lally, Lipponen, & Simons, 2007; Lockhorst, 2004; Veldhuis-Diermanse, 2002). This is unfortunate since the teacher can have a large impact on the effectiveness of collaborative learning. On the other hand, it is also known that guiding and facilitating the process of collaboration is not an easy task for teachers. Monitoring, guiding, and controlling small group learning, be it face-to-face or in a distributed setting, is a very work-intensive process (Bolhuis, 2000; Krol-Pot, 2005; Ros, 1994). For example, it can be difficult for teachers to determine whether students understand the subject matter and whether they are on-task or not.

Could the teachers participating in our studies then have had an effect on our results? Although we cannot entirely rule this out, we think this has not been the case. Although teachers had access to their own version of the VCRI-environment in which they could monitor the progress their groups were making and kinds of discussions that were being held, the activities of the teachers were mainly aimed at communicating deadlines or organizational matters. In only a few cases did they ask critical questions or make remarks about the quality of the collaboration in some groups. We therefore think that the teachers had very little effect on group members' collaboration. This was probably due to two factors. First, most participating teachers had limited experience with facilitating and moderating online discussion and collaboration. This may have made it difficult for them decide what kind of guidance was needed for each group. Second, as we mentioned earlier, guiding and facilitating collaboration is a complex task. In our studies teachers often had classrooms of over 25 students. This meant that they often were responsible for guiding up to 10 different groups. In the teacher environment of the VCRI, teachers can receive information about all groups: They have access to their chat discussions, can see the current state of the group products, and so on. It is not difficult to imagine that is a daunting task for even the most experienced teacher to process and interpret all of this information and then to decide what actions need to be taken to ensure successful collaboration for all groups. In our future research we will therefore investigate and address the problems teachers face when they work in CSCL-environments.



6. Epilogue: Using Multilevel Analysis in CSCL Research

CSCL researchers are often interested in the processes that unfold between learners in online learning environments and the outcomes that stem from these interactions. However, studying collaborative learning processes is not an easy task. Researchers have to make quite a few methodological decisions such as how to study the collaborative process itself (e.g., develop a coding scheme or a questionnaire), on the appropriate unit of analysis (e.g., the individual or the group), and which statistical technique to use (e.g., descriptive statistics, analysis of variance, correlation analysis). Recently, several researchers have turned to multilevel analysis (MLA) to answer particular research questions (e.g., De Wever, Van Keer, Schellens, & Valcke, 2007; Dewiyanti, Brand-Gruwel, Jochems, & Broers, 2007; Schellens et al., 2005; Strijbos, Martens, Jochems et al., 2004). However, CSCL studies that apply MLA analysis still remain relatively scarce. Instead, many CSCL researchers continue to use 'traditional' statistical techniques (e.g., analysis of variance, regression analysis), although these techniques may not be appropriate for what is being studied.

In the previous chapters of this thesis, MLA was used to answer several research questions. In these chapters we were rather brief in our explanation of why MLA was used; therefore we feel the need to explain more elaborately why MLA is often necessary to correctly answer the questions CSCL researchers address. Furthermore, we wish to highlight the consequences of failing to use MLA when this is called for, using data from the studies reported on in previous chapters. It should be noted that the discussion in this chapter concerns only research questions that involve a *dependent variable measured at the individual level* (e.g., satisfaction with the collaborative process, trust of group members, post-test knowledge measures, number of argumentative utterances produced). The data-analytical problems that are outlined in this epilogue do not apply to those research questions that involve a dependent variable measured at the group level (e.g., a grade assigned to a group product).

6.1 Multilevel Analysis: A 'New' Methodological Approach in CSCL Research

Over the last 5 years, multilevel analysis (MLA) has been adopted by several CSCL researchers to answer their research questions, because MLA is especially suited to "appropriately grasp and disentangle the effects and dependencies on the individual level, the group level, and sometimes the classroom level" (Strijbos & Fischer, 2007, p. 391). Although MLA is a relatively 'new' technique, especially to CSCL researchers, its development began in the

1980's (Snijders & Bosker, 1999) and has since then been used in several research disciplines.

ML analysis was initially embraced by educational researchers interested in school effectiveness research because it is well suited to the type of datasets they analyze. Consider, for example, an educational researcher interested in the effect of class size (i.e., the independent variable) on student achievement (i.e., the dependent variable). To investigate this effect, he or she would collect data about class sizes in different schools, as well as data on student achievement (e.g., standardized test scores on language, mathematics, and so on). However, such a research question poses several problems. First, this research question yields a hierarchically *nested dataset* with students nested in classrooms, and with classrooms nested within schools (and if this was an international study, even with schools nested within countries). Furthermore, the researcher would encounter the problem of *nonindependence* of his/her dependent variable of interest. Many statistical techniques (e.g., *t*-test, regression analysis), assume that the achievement scores (or other dependent variables) of the students in the dataset are independent from each other (Kashy & Kenny, 2000; Snijders & Bosker, 1999). In the example just provided, this will probably not be the case. Due to a common experience for example (e.g., the teaching they receive by their teacher), the scores within one classroom may not be independent at all since the overall classroom environment will affect all children in the class and even the behavior of individuals in the class will affect the others. Finally, the imaginary educational researcher would have to take into account that his or her variables of interest, class size and student achievement, are measured at different levels. Class size is measured at the class level, while achievement is measured at the student level. The number of available observations for both variables differs (i.e., the number is smaller for class size than for achievement. To properly address these issues, MLA was developed, and since then it became an important technique for school effectiveness research (Bosker & Snijders, 1990; De Leeuw & Kreft, 1986).

Social psychologists have also acknowledged the analytical problems described above. They are frequently interested in how individuals' thoughts and behaviors are influenced by other people. Many social psychological concepts involve two or more persons (e.g., attraction, interactive behavior, marital satisfaction) and thus the behavior of individuals *within a group* is often the focus of study (Kashy & Kenny, 2000). A social psychologist might for example be interested in how the division of household chores (i.e., the independent variable) affects marital satisfaction (i.e., the dependent variable). To answer this question, the researcher would have to observe married couples and record who did which chore (and calculate for example a ratio), and to administer to both spouses a questionnaire to measure their marital satisfaction. From this example it becomes clear the social psychologist encounters the same problems as the educational researcher. Both encounter the problem of hierarchically nested datasets (in this case individuals nested within couples), both involve variables at different levels of measurement (in this case the household chores ratio is measured at the level of the couple, whereas marital satisfaction is measured at the level of the individual), and in both cases the observations of the dependent variable are probably not independent (in this case there might even be a negative relationship: the husband may be more satisfied if he does little housework, while this may negatively affect the marital satisfaction of his wife).

The problem of statistical nonindependence of dependent variables (i.e., group members exerting a psychological influence on each other) has especially received considerable attention in social psychology since the 1980's (e.g., Bonito, 2002; Kenny, 1995, 1996; Kenny & Judd, 1986; Kenny & Judd, 1996). It is therefore not surprising that social psychologists frequently use MLA to deal with these issues (cf., Bonito & Lambert, 2005; Campbell & Kashy, 2002; Kenny et al., 2002).

6.2 The Problems CSCL Researchers Encounter

Similar to other research disciplines, CSCL researchers encounter the abovementioned problems of hierarchically nested datasets, nonindependence of dependent variables, and differing units of analysis. We explain these problems below.

6.2.1 Hierarchically Nested Datasets

In CSCL-environments, students work in groups, therefore studying online collaboration often involves investigating group processes and how these processes are affected by contextual factors (e.g., the environment itself, the composition of the group, prior knowledge and experiences of the group members). It is not difficult to understand this leads to *hierarchically nested datasets*, since groups consist of two or more individuals and thus in these cases individuals are nested within groups. In many cases, CSCL researchers will encounter at least two levels: the group and the individual. The group is then the macro- or level-2 unit and the individual the micro- or level-1 unit (Snijders & Bosker, 1999). CSCL researchers may also use datasets that have even more levels of analysis. A researcher might for example be interested in the effects of the teacher's experience with CSCL on the way his or her students collaborate online. This researcher will have a dataset with three levels: students are nested within groups, while groups are nested within teachers' classrooms.

Another CSCL researcher might be interested in the development of students' online interactive behavior over time. This researcher would therefore collect data about students' interactive behavior on different measurement occasions. This would also lead to a dataset with three levels: measurement occasions are nested within students, and students are nested within groups (Kenny et al., 2006; Snijders & Bosker, 1999). Whenever researchers encounter datasets with hierarchically nested data, MLA is needed to appropriately model this data structure since it can appropriately disentangle the effects of the different levels on the dependent variable(s) of interest (Snijders & Bosker).

6.2.2 Nonindependence of Dependent Variables

Because their participants work in groups, CSCL researchers also encounter the problem of *nonindependence* of their dependent variables. This means students within a group may be more similar to each other than are persons from different groups (Kenny et al., 2002). In the case of CSCL, the main source of this nonindependence is the mutual influence group members have on each other (Bonito, 2002; Kenny, 1996). In the studies described in chapters 2, 3, and 4 for example, students could discuss with each other through a Chat-window and a Forum. Through these discussions, students influenced each other. In some cases for example, a student displayed negative behavior, and this prompted the other group members to respond negatively as well. Furthermore, some students in our studies were very active in the Chat conversations (e.g. they proposed a lot of strategies and asked a lot of questions). This could have triggered the other group members to also become more active in the chat.

This reciprocal influence of group members is not necessarily positive, it can also be negative. In the previously mentioned example concerning active group members stimulating other group members to be more active, the reverse could also happen: When one group member is very active in the learning environment, this may trigger other group members to "sit back" and do little since that group member is doing so much (O'Donnell & O'Kelly, 1994; Webb & Palincsar, 1996). Kenny et al. (2002) therefore noted that mutual influence can not only cause students to behave more similarly, but may also cause students to behave differently from their group members. This is called the *boomerang effect*. Another example is that when group members behave negatively, a student may decide to counter this by displaying more positive behavior. Role assignment (cf., Schellens et al., 2005; Strijbos, Martens, Jochems et al., 2004; Strijbos, Martens, Jochems, & Broers, 2007) may also lead to differential behavior. If one group member, for example, is given the task to ask critical questions, while the other group member has to monitor task progress, this may lead to differing behavior (e.g., the first student will ask many questions, but will display less metacognitive behavior, while the second student may display high levels of metacognitive behavior but may ask fewer questions). Kenny et al. therefore make a distinction between *positive nonindependence* where group members influence each other in such a way that they behave more similarly and *negative nonindependence* where group members influence each other to behave differently. Thus, since group members influence each other in a group context, this will likely lead to either positive or negative nonindependence of the dependent variables that are being investigated which in turn has to be dealt with during data analysis.

The degree of nonindependence can be estimated using the *intraclass correlation coefficient*[8] (ICC, cf. Kashy & Kenny, 2000; Kenny et al., 2002). Values of the ICC can range from -1 to +1. An ICC +1 for satisfaction with the collaborative process (scored on a 4-point scale ranging from 1 to 4) for example,

[8] For an excellent description on how to compute the ICC for a specific dataset, the reader is referred to Kenny et al. (2006).

indicates that when a group member has a score of 4 on this measure, the other group members will also have a score of 4. Conversely, an ICC of -1 for the same measure, indicates that when one student has a score of 4 on this measure, his or her partners will have a score of 1.

An alternative interpretation of the ICC is in terms of the *amount of variance that is accounted for by the group* (Kenny et al., 2006). When the ICC for satisfaction with the collaborative process is found to be .40 for example, this means the 40% of the variance in this measure is accounted for by the group, and thus that 60% is accounted for by other (e.g., individual) factors.

The dependent measures that CSCL researchers are interested in will often be nonindependent. Strijbos et al. (2004) for example, studied the effect of roles on perceived group efficiency. They found an ICC of .47, meaning 47% of this measure is accounted for by the group. Group members displayed rather similar levels of perceived group efficiency, probably due to their common experiences in the CSCL environment. In our own research we found similar values of the ICC. In Chapter 2 for example, we studied the effect of the Participation tool (PT) on students' perceptions of positive group behavior. For this measure we found an ICC of .41, indicating a considerable influence of the group on this measure. The effect of the Graphical Debate-tool on students' post-test performance was studied in Chapter 4. For this dependent variable we found that 32% of the total variance could be attributed to the group level. On the other hand, not all researchers find similar amounts of variance accounted for by the group. De Wever et al. (2007) for example, report only 3% of the students' level of knowledge construction was linked to the group level. However, these examples still illustrate the presence of nonindependence in datasets of CSCL researchers.

Nonindependence also needs to be addressed when conducting statistical analyses, because it distorts estimates of error variances, thus making standard errors, p-values, and confidence intervals invalid when this distortion is not taken into account (Kenny, 1995; Kenny et al., 2006). Traditional statistical techniques such as t-tests, analyses of variance, and regression analyses cannot cope with this distortion because they assume the variables are independent. Therefore CSCL researchers using these types of analyses run an increased risk of committing Type I or Type II errors (Kashy & Kenny, 2000). Whether the chance to falsely reject (Type I error) or falsely accept (Type II error) the null hypothesis is increased, depends on the sign of the ICC (either positive or negative), and the type of dependent variable for which the ICC was calculated (see Kashy & Kenny for a detailed discussion).

The ICC can also be tested for significance. If the ICC is significant, its effect is large enough to bias statistical tests (Kenny et al., 2006). However, because sample sizes are often small in CSCL research, the ICC may not be significant, while it is actually still large enough to bias standard errors, p-values and so on. Kenny et al. (2002) therefore propose assuming group data are nonindependent even though the ICC is not significant.

6.2.3 Differing Units of Analysis

A final problem that CSCL researchers encounter concerns the *differing units of analysis* their datasets often contain. This has to do with the abovementioned hierarchical structure of their datasets. Some variables that CSCL researchers are interested in are measured at the individual level (e.g., gender, interactive behavior, familiarity with other group members), whereas other variables are measured at the group level (e.g., gender group composition, group performance, group consensus). Savicki and Kelly (2000) for example, studied the effect of gender and gender group composition (male-only, female-only, or mixed) on satisfaction with online collaboration. Their dependent variable was measured at the individual level (satisfaction), while their two independent variables were measured at the both the individual (gender) and group (gender group composition) level. Thus, their dataset contained variables with differing units of analysis.

Chapter 2 of this thesis provides another example. Amongst other things, we examined the effect of experimental condition (whether groups used the PT or not) on how often students' engaged in different collaborative activities. Furthermore, we wanted to control for students' level of participation, since some students were more active in the online discussions than others. Thus our dependent variable was measured at the individual level (collaborative activities), while the independent variable was measured at the group level (PT or no PT), and the control variable (level of participation) was measured at the individual level. Again, to be able to cope with these different units of analysis, MLA is needed. Traditional statistical techniques cannot properly take these differing units of analysis properly into account (Snijders & Bosker, 1999).

6.3 Common Analysis Strategies

In this section we describe three strategies that researchers can use to deal with the data analytical problems described in the previous sections, namely ignoring nonindependence of dependent variables, aggregating or disaggregating data, and MLA. We illustrate the different analysis strategies by referring to two different examples.

The first example comes from the data collected for the study described in Chapter 3 of this thesis. In this chapter, we examined – amongst others - the effect of experimental condition on the number of times students evaluated the social interaction positively during their online conversations. In one condition students used the Shared Space (SS) to communicate online, while in the other condition (No SS) the students communicated through a regular chat-tool.

The data collected for Chapter 4 of this thesis are used as a second example. In this Chapter we investigated the effects of representational guidance on students' performance on a knowledge post-test. Our design used two conditions: In one condition students used the Graphical Debate-tool to construct external representations of a historical debate, while in the other condition students used the Textual Debate-tool to construct such representations. Both versions of the tool differed with respect to the representational guidance they offered to the students. The Graphical Debate-tool made extensive use of visualization techniques to visualize certain aspects of the collaborative problem solving process (e.g., Was there a balance between the number of arguments pertaining to both positions?). In Chapter 4 we also used MLA to investigate the effects of condition (Graphical versus Textual Debate-tool) on students' post-test performance.

6.3.1 Ignoring Nonindependence

A first strategy, and also still the most common practice during the analysis of group data (Kenny et al., 2002), is to ignore the hierarchical structure of the dataset, the nonindependence, and the differing units of analysis and perform statistical techniques such as *t*-tests or (M)ANOVA's. As we discussed previously, this biases significance tests of inferential statistics (e.g., *t* or *F*-values), sometimes making tests too liberal and, and at other times, too conservative.

If we would choose this strategy for the dataset from Chapter 3, we could use regression analysis to answer the question whether the Shared Space had an effect on the number of times students evaluated the collaboration positively. In this regression analysis, we include number of positive evaluations of the collaboration a student typed in the Chat-tool as a dependent variable and condition (effect coded with Shared Space as +1 and No Shared Space as -1) as an independent variable. Furthermore, we also include participation (e.g., the total number of messages students sent) in the regression equation to control for the fact that some students were more active during the online collaboration than others. As can be seen in Table 6.1, we find no effect of condition (Shared Space or No Shared Space) using this regression model on positive evaluations of the collaboration, B = 0.20, SE = 0.14, $p = .08$ (one-tailed significance). Thus, if we adopt a strategy that ignores nonindependence, we would conclude that the Shared Space does not have an influence on the number of times students type positive evaluations of their collaboration. As we will see in section 6.3.3, analyzing the same dataset MLA yields a different conclusion.

Table 6.1: Regression analysis of the effect of the Shared Space on number of positive evaluations of the collaborative process typed in the Chat-tool.

	B	*SE B*	*β*
Condition (-1 = No SS, +1 = SS)	0.202	0.142	.131
Participation	0.001	0.001	.140

If we adopt this strategy to our analyses of the effects of the Graphical Debate-tool on students' post-test performance, we could use analysis of covariance (ANCOVA). The ANCOVA model would include post-test performance as the dependent measure of interest, condition (Graphical Debate-tool versus Textual Debate-tool) as the independent variable, and pre-test performance as a covariate. The results of this analysis can be found in Table 6.2. As can be seen in this Table, condition had a significant impact on post-test performance, $F(1, 82) = 3.98$, $p = .05$. In conclusion, if we adopt a strategy that ignores nonindependence, we would conclude that condition has a significant impact on post-test performance.

Table 6.2: Analysis of covariance for condition (Graphical versus Textual Debate-tool).

	df	*MS*	*F*	η^2
Pretest performance (Covariate)	1	39.14	11.11**	.12
Condition (Graphical or Textual Debate)	1	14.07	3.98*	.05
Error	82	3.55		

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

Ignoring nonindependence is a frequently encountered approach in CSCL research. Francescato et al. (2006) for example, studied differences in students' evaluations of collaboration in online and face-to-face learning groups. Among other things, they investigated whether online learning groups perceived differing levels of social presence and satisfaction with the collaborative process than face-to-face groups. However, they found no differences between online and face-to-face groups using analyses of variance, although there was a tendency for online groups to be more satisfied with the collaborative process ($p = .17$). Because this study involves students working in groups, the evaluations of Francescato et al.'s students are most likely nonindependent. However, their analyses fail to take nonindependence into account, and thus the *p*-values reported by the authors might be biased. As we have seen with our own data, this might lead to a larger *p*-value, and thus a false rejection of the null hypothesis (i.e., no differences between face-to-face and online learning groups). Using a more appropriate statistical technique, MLA, Francescato et al. might have been able to demonstrate significant differences between face-to-face and online learning groups.

Another example comes from the work of Guiller and Durndell (2007) who studied the effect of gender on students' linguistic behavior in online discussion groups. Guiller and Durndell studied for example, whether male more absolute adverbials (i.e., strong assertions such as 'obviously') and imperatives (i.e., giving commands) than female students. In order to answer this question they coded students' messages and classified each message in terms of the linguistic behavior shown by the students. Guiller and Durndell then used X^2-analyses to determine whether male and female students differed with respect to these behaviors. Although the authors found male students to use more absolute adverbials and imperatives, the corresponding X^2-values were not significant. However, by using X^2-analyses they too ignored the nonindependence of their dependent variables. Again, group members communicated and discussed with each other, so therefore they likely influenced each other. Using MLA, Guiller and Durndell, might have been able to detect statistically significant differences between male and female students on use of certain linguistic behaviors.

6.3.2 Aggregating or Disaggregating Data

Another strategy to deal with the problems described in the previous section is to *aggregate* individual data to the level of the group (Snijders & Bosker, 1999). If we apply an aggregation strategy to our data from Chapter 3, we could calculate the sum of positive evaluations of the collaboration for each group. On average, Shared Space groups exchanged 2.25 (SD = 3.16) positive evaluations of the collaboration, while No Shared Space groups exchanged only 1.00 (SD = 1.95) of these messages. Again, we could then use a regression analysis to examine the effects of condition (Shared Space or No Shared Space) on the number of positive evaluations of the collaborative process exchanged by the group. This regression analysis includes number of positive evaluations as the dependent variable, condition as the independent variable, and again level of participation (i.e., the total number of messages sent by the whole group) as a control measure. The results of the regression analysis are displayed in Table 6.3. As can be seen, condition was not found to have a significant impact on the number of positive evaluations of the collaborative process sent, $B = 0.60$, $SE = 0.42$, $p = .08$. This yields a conclusion comparable to the previously described strategy of ignoring nonindependence: the Shared Space does not have an influence on the amount of positive evaluations of the collaborative process exchanged during online collaboration. This conclusion is different however from the conclusion reported in Chapter 3 where we used MLA to investigate the effects of the Shared Space.

Table 6.3: Regression analysis of the effect of the Shared Space on number of positive evaluations of the collaborative process typed in the Chat-tool by the group.

	B	*SE B*	*β*
Condition (-1 = No SS, +1 = SS)	0.202	0.142	.131
Participation	0.001	0.001	.140

This strategy can also be applied to our data from Chapter 4. Applying an aggregation strategy involves computing, for each group, the average post- and pre-test score of the individual group members. Using such a strategy, we find Graphical Debate groups to attain, on average, a post-test score of 13.02, while Textual Debate groups attain a an average score of 12.24. To test the effect of condition on post-test performance, we could again conduct an analysis of covariance, using post-test performance as the dependent variable, condition as the independent variable, and pre-test performance as a covariate. The results of this analysis are displayed in Table 6.4. As can be seen, the effects of condition are not significant if we adopt an aggregation strategy in this case, $F(1, 37) = 0.28$, $p = .61$.This means we would conclude, in contrast to the previous strategy of ignoring nonindependence, the Graphical Debate-tool does not have positive effects on students' post-test performance. Again, this is a different conclusion than the one we reported in Chapter 4 where we used MLA to study the effects of the Graphical Debate-tool.

Table 6.4: Analysis of covariance for condition (Graphical versus Textual Debate-tool) on group level variables

	df	*MS*	*F*	η^2
Pretest performance (Covariate)	1	11.39	6.27	.28
Condition (Graphical or Textual Debate)	1	0.50	0.28	.02
Error	37	1.82		

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

This strategy is also used in a study described by Van der Meijden and Veenman (2005). Van der Meijden and Veenman compared dyads using face-to-face (FTF, *N* = 20) and computer-mediated communication (CMC, *N* = 22) with respect to exchange of high-level elaboration (e.g., elaborate explanations or requests for help). Students' collaboration was coded using a coding scheme. However, the percentages of high-level elaboration "were calculated by summing the individual code frequencies" (p. 843) and dividing these by the total number of utterances. An independent samples *t*-test was then used to establish whether FTF and CMC conditions differed significantly with respect to high-level elaboration. High-level elaboration was thus treated as a group level variable. Such an analysis however, ignores the fact that high-level elaboration is in essence an individual level variable (although it may be affected by group level variables). Furthermore, by aggregating to the group level, this analysis uses fewer observations for high-level elaboration than are available. For this variable only 20 + 22 = 42 observations are used, while in effect there are 42 * 2 = 84 observations, therefore Van der Meijden and Veenman run the risk of committing a Type II error. Fortunately, in their study the differences between FTF and CMC were large enough to detect a significant difference between FTF and CMC groups with respect to the percentage of high-level elaborations exchanged.

The reverse strategy can also be applied: treating group level data as if they were measured at the individual level. This is called *disaggregation*. Consider for example, the study by Savicki, Kelley, and Lingenfelter (1996) about the effects of gender group composition on students' satisfaction with the collaborative process. Group composition was measured at the group level (all male, all female, or mixed groups), while satisfaction was measured at the individual level (students completed a questionnaire individually). In total, their sample consisted of 6 groups and 36 students. Savicki et al. conducted an analysis of variance to examine whether group composition affected satisfaction. However, this analysis does not take into account that group composition was measured at the group level. Thus, Savicki et al.'s analysis mistakenly uses 36 observations for the group composition variable, while in fact there are only 6 observations for this variable. This led to an exaggeration of the actual sample size for this variable and increased the chance of committing a Type I error (Snijders & Bosker, 1999).

The dangers of such a strategy are highlighted when we reanalyze our data from Study 3. In this study, we examined the effects of the Graphical Debate-tool on the quality of the essays written by groups. Essay quality was measured in terms of number of topics covered in the essay and grounds and conceptual quality of the essay. Because the essays were written by groups, this variable was a group-level measure: each group received one score for number of topics covered, grounds quality, and conceptual quality. If we would have adopted a disaggregation strategy, we would have given each student within the group the same score for these three quality indicators. This leads to an increase of the sample size from 39 groups to 124 students. In Chapter 4, using *t*-tests, we found no significant differences with respect to the number of topics covered, $t = -0.55$, $p = .59$, and the conceptual quality of the essay, $t = 2.00$, $p = .06$. However, when we disaggregate our data, and then use *t*-tests to examine the differences between the Graphical and Textual Debate tool, we find different *t*- and *p*-values, namely for number of topics covered, $t = -1.38$, $p = .17$, and for conceptual quality, $t = 3.24$, $p = .00$. This example shows that using a disaggregation strategy might lead to biased *t*- and *p*-values and even different conclusions (i.e., in the case of conceptual quality the conclusion would be different).

6.3.3 Multilevel Analysis

MLA was designed specifically to cope with hierarchically nested data (Snijders & Bosker, 1999). Furthermore, it is a useful technique when researchers use datasets that have different units of analysis, such as group and individual level variables (Kenny et al., 2006). Finally, MLA can deal with the nonindependence of observations that results from the mutual influence group members have on each other (Snijders & Bosker).

At present, MLA is slowly finding its way to the CSCL research community: more and more CSCL researchers are using MLA to analyze their data. Strijbos et al. (2004) for example, studied the effect of role assignment on perceived group efficiency. Their study used two conditions: a condition in which specific roles (e.g., project planner, editor) were assigned to students and a condition without role assignment. Thus, condition was a group level independent variable. Perceived group efficiency was measured using several questionnaires, and was therefore an individual level dependent variable. It is not difficult to see that in the Strijbos et al. study hierarchically nested data were collected since students were nested in groups. Furthermore, their study employed variables measured at different units of analysis. Finally, in their dataset nonindependence was present, since they reported an ICC of .47 for perceived group efficiency. Strijbos et al. therefore constructed a ML model with perceived group efficiency as dependent variable and condition (role or non-role assignment) as an independent variable. Using MLA, they were able to model the nonindependence in their datasets and to analyze their dependent and independent variable at their appropriate levels of analysis.

In Chapter 3 of this thesis we also used MLA to study the effects of the Shared Space on students' use of positive evaluations of the collaboration. We constructed a ML model that included number of times a student typed a positive evaluation of the collaboration as a dependent variable and condition (Shared Space or No Shared Space) as an independent variable. Furthermore, we included participation (e.g., total number of messages sent) again to control for the fact that some students typed more messages than other students. This way, the effects of the Shared Space could be studied independent of students' participation. As can be seen in Table 6.5, we found a significant effect of the Shared Space on the number of positive evaluations of the collaboration students typed, $\beta = 0.20$, SE = .14, $p = .04$ (one-tailed significance). Although the differences in p-values are small (see Table 6.7), this analysis strategy leads to a different conclusion than the previous two strategies (ignoring nonindependence and aggregating or disaggregating data), namely that the Shared Space affects the number of positive evaluations of the collaboration. Thus, in this case MLA prevented us from making a Type II error (i.e., falsely accepting the null hypothesis).

Table 6.5: Multilevel analysis of the effect of condition (Shared Space or No Shared Space) on number of positive evaluations of the collaborative process exchanged.

	β	*SE*
Participation	0.01*	0.00
Condition (SS or No SS)	0.20*	0.14
Deviance	429.14	
Decrease in deviance	2.06*	

* $p < .05$.

Table 6.6: Multilevel analysis of the effect of condition (Graphical or Textual Debate-tool) on post-test performance.

	β	*SE*
Pre-test performance	0.28**	0.10
Condition (Graphical or Textual)	0.42*	0.22
Deviance	344.63	
Decrease in deviance	11.07**	

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

To analyze the effects of the Graphical Debate-tool on students' post-test performance we used MLA (see Chapter 4). Our ML model included students' pre-test performance and condition (Graphical versus Textual Debate-tool). The results of this analysis can be found in Table 6.6. We found a significant effect of condition on post-test performance, indicating that the Graphical Debate-tool helped students to perform better on the knowledge post-test, $\beta = 0.42$, SE = .22, $p = .03$.

Table 6.7 and Table 6.8 summarize the results of the different analysis strategies. As can be seen, the p-values are somewhat different if one strategy is chosen rather than another strategy. Although the differences in p-values may not seem spectacular (with the exception if an aggregation strategy is chosen for the evaluation of the effect of the Graphical Debate-tool), they nevertheless lead to different conclusions in some cases.

Table 6.7: Summary of differing results for effects of Shared Space on students' positive evaluations of the collaborative process.

	Ignoring nonindependence	Aggregating or disaggregating data	Multilevel analysis
Statistical analysis	Regression analysis	Regression analysis	MLA
Significance of effect of condition	Not significant, p = .08	Not significant, p = .08	Significant, p = .04
Conclusion	No effect of Shared Space on positive evaluations of collaborative process	No effect of Shared Space on positive evaluations of collaborative process	Positive effect of Shared Space on positive evaluations of collaborative process

Table 6.8: Summary of differing results for effects of Graphical Debate-tool on students' post-test performance.

	Ignoring nonindependence	Aggregating or disaggregating data	Multilevel analysis
Statistical analysis	Analysis of covariance	Analysis of covariance	MLA
Significance of effect of condition	Significant, p = .05	Not significant, p = .61	Significant, p = .03
Conclusion	Positive effect of Graphical Debate-tool on post-test performance	No effect of Graphical Debate-tool on post-test performance.	Positive effect of Graphical Debate-tool on post-test performance

6.4 Conclusion and Discussion

In this chapter we discussed the data analytical problems CSCL researchers frequently encounter, namely hierarchically nested datasets, nonindependence of dependent variables, and differing units of analysis. We argued that, in order to take these problems into account, MLA should be used. We also demonstrated that alternative analysis strategies such as ignoring nonindependence or aggregating or disaggregating data can lead to different results and possibly to mistakes regarding the significance or non-significance of these results. We therefore strongly advocate the use of MLA in CSCL research. Fortunately, more and more CSCL researchers are beginning to use this technique to answer their research questions.

It should be noted that we do not claim that in the cases where CSCL researchers used other analyses than MLA their conclusions are wrong. This need not be the case. However, these researchers do have an increased chance of committing Type I or Type II errors. We hope this chapter will contribute to an increased awareness of the risks of using traditional statistical techniques such as *t*-tests and ANOVAs, and future CSCL research will use ML analysis when this is appropriate.

Of course not all data-analytic problems that CSCL researchers encounter are solved by using ML analysis. Furthermore, MLA has its own limitations. First, MLA is mostly used when the dependent variable is measured at the interval level of measurement. Sometimes however, researchers may be interested in dichotomous (e.g., success or failure of group work) or categorical dependent variables (e.g., levels of knowledge construction). Although MLA techniques have been developed to incorporate these kinds of dependent variables (multilevel logistic regression, see Snijders & Bosker, 1999), they are rarely adapted to CSCL data.

Second, CSCL researchers are often interested in data over time. An example might be how familiarity with group members affects trust-development in CSCL environments over time. To investigate this question a researcher would collect data about trust levels on different occasions. This adds even more problems to analyzing CSCL data. The effects of familiarity on trust may not be the same at every measurement occasion (e.g., its effects may be greater at the beginning of the collaboration). Furthermore, the level of trust at measurement occasion 1 may also have an effect on the level of trust at occasion 2 (if trust was high at occasion 1, this may affect trust at occasion 2). This creates a new type of nonindependence: autocorrelation (Kenny et al., 2006). Again, MLA techniques have been developed to analyze time-series data (cf., Chiu & Khoo, 2003, 2005; Kenny et al., 2006), but they are not often used in CSCL research. CSCL researchers should therefore begin to investigate the possibilities of using MLA for time-series data.

Finally, MLA will not be a suitable technique to answer all research questions. Quite a lot CSCL research focuses on capturing the interactive processes that unfold between group members. In some cases researchers are interested in providing "thick" or "rich" descriptions of the collaborative process (Baker, 2003; Hmelo-Silver & Bromme, 2007). In such cases, MLA is obviously useless. Furthermore, it has been argued that studying intersubjective meaning making or group cognition should be the focus of CSCL research (Stahl, 2006; Suthers, 2006). This involves studying "how people make sense of situations and of each other" (Suthers, p. 321). Researchers with such a perspective on CSCL research could object to disentangling group and individual aspects of collaborative learning. They would argue that in order to understand the collaborative process, the group should be the unit of analysis, not the individual. Again, if one has such an approach to studying CSCL, using MLA will not be a sensible strategy.

CSCL research can still make progress by incorporating MLA in its repertoire of analysis techniques. It is an encouraging development that CSCL researchers are turning toward MLA more often. It is our hope and expectation that this development will continue and that CSCL researchers are going to find new ways to deal with the complex data analytical problems they are faced with.

References

Abrami, P. C., & Chambers, B. (1996). Research on cooperative learning and achievement: Comments on Slavin. *Contemporary Educational Psychology, 21*, 70-79.

Adams, S. J., Roch, S. G., & Ayman, R. (2005). Communication medium and member familiarity: The effects on decision time, accuracy, and satisfaction. *Small Group Research, 36*, 321-353.

Adrianson, L. (2001). Gender and computer-mediated communication: Group processes in problem solving. *Computers in Human Behavior, 17*, 71-94.

Adrianson, L., & Hjelmquist, E. (1991). Group processes in face-to-face and computer mediated communication. *Behaviour and Information Technology, 10*, 281-296.

Adrianson, L., & Hjelmquist, E. (1999). Group processes in solving two problems: Face-to-face and computer-mediated communication. *Behaviour and Information Technology, 18*, 179-198.

Ainsworth, S. (2006). DeFT: A conceptual framework for considering learning with multiple representations. *Learning and Instruction, 16*, 183-198.

Alker Jr., H. R. (1965). *Mathematics and politics*. New York: The MacMillan Company.

Anderson, T., Rourke, L., Garrison, D. R., & Archer, W. (2001). Assessing teaching presence in a computer conferencing context, *Journal of Asynchronous Learning Networks* (Vol. 5).

Andriessen, J., Erkens, G., Van de Laak, C., Peters, N., & Coirier, P. (2003). Argumentation as negotiation in electronic collaborative writing. In J. Andriessen, M. Baker, & D. Suthers (Eds.), *Arguing to learn: Confronting cognitions in computer-supported collaborative learning environments* (pp. 79-115). Dordrecht, The Netherlands: Kluwer Academic Publishers.

Artzt, A. F., & Armour-Thomas, E. (1997). Mathematical problem solving in small groups: Exploring the interplay of students' metacognitive behaviors, perceptions, and ability levels. *Journal of Mathematical Behavior, 16*, 63-74.

Baker, K., Greenberg, S., & Gutwin, C. (2001, May). *Heuristic evaluation of groupware based on the mechanics of collaboration.* Paper presented at the Engineering for Human-Computer Interaction: 8th IFIP International Conference, EHCI 2001, Toronto, Canada.

Baker, M. (2003). Computer-mediated argumentative interactions for the co-elaboration of scientific notions. In J. Andriessen, M. Baker, & D. Suthers (Eds.), *Arguing to learn: Confronting cognitions in computer-supported collaborative learning environments* (pp. 47-78). Dordrecht, The Netherlands: Kluwer Academic Publishers.

Baltes, B. B., Dickson, M. W., Sherman, M. P., Bauer, C. C., & LaGanke, J. (2002). Computer-mediated communication and group decision making: A meta-analysis. *Organizational Behavior and Human Decision Processes, 87*, 156-179.

References

Barkhi, R., Jacob, V. S., & Pirkul, H. (1999). An experimental analysis of face-to-face versus computer mediated communication channels. *Group Decision and Negotiation, 8*, 325-347.

Barron, B. (2000). Achieving coordination in collaborative problem-solving groups. *Journal of the Learning Sciences, 9*, 403-436.

Barron, B. (2003). When smart groups fail. *Journal of the Learning Sciences, 12*, 307-359.

Beers, P. J. (2005). *Negotiating common ground: Tools for multidisciplinary teams.* Unpublished PhD thesis, Open University of The Netherlands.

Beers, P. J., Boshuizen, H. P. A., Kirschner, P. A., & Gijselaers, W. H. (2007). The analysis of negotiation of common ground in CSCL. *Learning and Instruction, 17*, 427-435.

Benbunan-Fich, R., Hiltz, S. R., & Turoff, M. (2003). A comparative content analysis of face-to-face vs. asynchronous group decision making. *Decision Support Systems, 34*, 457-469.

Berge, Z. L. (1995). Facilitating computer conferencing: Recommendations from the field. *Educational Technology, 15*, 22-30.

Bodemer, D., & Faust, U. (2006). External and mental referencing of multiple representations. *Computers in Human Behavior, 22*(1), 27-42.

Bødker, S., & Christiansen, E. (2006). Computer support for social awareness in flexible work. *Computer Supported Cooperative Work (CSCW), 15*, 1-28.

Bolhuis, S. M. (2000). *Naar zelfstandig leren: Wat doen en denken docenten [Toward independent learning: What do teachers think and do?].* Unpublished PhD thesis, University of Nijmegen, The Netherlands.

Bonito, J. A. (2000). The effect of contributing substantively on perceptions of participation. *Small Group Research, 31*, 528-553.

Bonito, J. A. (2002). The analysis of participation in small groups: Methodological and conceptual issues related to interdependence. *Small Group Research, 33*, 412-438.

Bonito, J. A., & Lambert, B. L. (2005). Information similarity as a moderator of the effect of gender on participation in small groups: A multilevel analysis. *Small Group Research, 36*, 139-165.

Bosker, R. J., & Snijders, T. A. (1990). Statistische aspecten van multi-niveau onderzoek [Statistical aspects of multilevel research]. *Tijdschrift voor Onderwijsresearch, 15*, 317-329.

Broeken, M., Jaspers, J., & Erkens, G. (2006). Virtual Collaborative Research Institute (VCRI). Version 2.3. Utrecht, The Netherlands: Utrecht University.

Butler, D. L., & Winne, P. H. (1995). Feedback and self-regulated learning: A theoretical synthesis. *Review of Educational Research, 65*, 245-281.

Buunk, B. P., Collins, R. L., Taylor, S. E., Van Yperen, N. W., & Dakof, G. A. (1990). The affective consequences of social comparison: Either direction has its ups and downs. *Journal of Personality and Social Psychology, 59*, 1238-1249.

Campbell, L., & Kashy, D. A. (2002). Estimating actor, partner, and interaction effects for dyadic data using PROC MIXED and HLM: A user-friendly guide. *Personal Relationships, 9*, 327-342.

Carlson, J. R., & Zmud, R. W. (1999). Channel expansion theory and the experiential nature of media richness perceptions. *Academy of Management Journal, 42*, 153-170.

Carlson, K. D., & Schmidt, F. L. (1999). Impact of experimental design on effect size: Findings from the research literature. *Journal of Applied Psychology, 84*(6), 851-862.

Cavanaugh, C. S. (2001). The effectiveness of interactive distance education technologies in K-12 learning: A meta-analysis. *International Journal of Educational Telecommunications, 7*, 73-88.

Chinn, C. A., & Anderson, R. C. (1998). The structure of discussions that promote reasoning. *Teachers College Record, 100*, 315-368.

Chiu, M. M., & Khoo, L. (2003). Rudeness and status effects during group problem solving: Do they bias evaluations and reduce the likelihood of correct solutions? *Journal Of Educational Psychology, 95*, 506-523.

Chiu, M. M., & Khoo, L. (2005). A new method for analyzing sequential processes: Dynamic multilevel analysis. *Small Group Research, 36*(5), 600-631.

Cicchetti, D. V., Lee, C., Fontana, A. F., & Dowds, B. N. (1978). A computer program for assessing specific category rater agreement for qualitative data. *Educational and Psychological Measurement, 38*, 805-813.

Clark, D. B., Sampson, V., Weinberger, A., & Erkens, G. (2007). Analytic frameworks for assessing dialogic argumentation in online learning environments. *Educational Psychology Review, 19*, 343-374.

Clark, D. B., & Sampson, V. D. (2005). Analyzing the quality of argumentation supported by personally-seeded discussions. In T. Chan & D. D. Suthers (Eds.), *Computer-supported collaborative Learning 2005: The next 10 years!* (pp. 76-85). Taipei, Taiwan: Lawrence Erlbaum Associates.

Clark, H. H., & Brennan, S. (1991). Grounding in communication. In L. B. Resnick, J. M. Levine, & S. Teasley (Eds.), *Perspectives on socially shared cognition* (pp. 127-149). Washington, DC: American Psychological Association.

Cohen, E. G. (1994). Restructuring the classroom: Conditions for productive small groups. *Review of Educational Research, 64*, 1-35.

Cole, P., & Nast-Cole, J. (1992). A primer on group dynamics for groupware developers. In D. Marca & G. Bock (Eds.), *Groupware: Software for computer-supported cooperative work* (pp. 44-56). Los Alamitos, CA: IEEE Computer Society Press.

Cox, R. (1999). Representation construction, externalised cognition and individual differences. *Learning and Instruction, 9*, 343-363.

Daft, R. L., & Lengel, R. H. (1986). Organizational information requirements, media richness and structural design. *Management Science, 32*, 554-571.

Damon, W., & Phelps, E. (1989). Critical distinctions among three approaches to peer education. *International Journal of Educational Research, 13*, 9-19.

De Groot, R., Drachmann, R., Hever, R., Schwarz, B. B., Hoppe, U., Harrer, A., et al. (2007). Computer supported moderation of E-discussions: The ARGUNAUT approach. In C. A. Chinn, G. Erkens, & S. Puntambekar (Eds.), *Mice, minds, and society: The Computer Supported Collaborative Learning (CSCL)*

Conference 2007 (Vol. 8, pp. 165-167). New Brunswick, NJ: International Society of the Learning Sciences.

De Jong, F., Kollöffel, B., Van der Meijden, H., Kleine Staarman, J., & Janssen, J. (2005). Regulative processes in individual, 3D and computer supported cooperative learning contexts. *Computers in Human Behavior, 21*, 645-670.

De Jong, T., & Van Joolingen, W. R. (1998). Scientific discovery learning with computer simulations of conceptual domains. *Review of Educational Research, 68*, 179-202.

De Laat, M., Lally, V., Lipponen, L., & Simons, R.-J. (2006). Analysing student engagement with learning and tutoring activities in networked learning communities: A multi-method approach. *International Journal of Web Based Communities, 2*, 394-412.

De Laat, M., Lally, V., Lipponen, L., & Simons, R.-J. (2007). Online teaching in networked learning communities: A multi-method approach to studying the role of the teacher. *Instructional Science, 35*, 257-286.

De Leeuw, J., & Kreft, I. (1986). Random coefficient models for multilevel analysis. *Journal of Educational Statistics, 11*, 57-85.

De Lisi, R., & Golbeck, S. L. (1999). Implications of Piagetian theory for peer learning. In A. M. O'Donnell & A. King (Eds.), *Cognitive perspectives on peer learning* (pp. 3-37). Mahwah, NJ: Lawrence Erlbaum Associates.

De Vries, E. (2006). Students' construction of external representations in design-based learning situations. *Learning and Instruction, 16*, 213-227.

De Vries, J., Havekes, H., Aardema, A., & Van Rooijen, B. (Eds.). (2004). *Actief historisch denken: Opdrachten voor activerend geschiedenisonderwijs [Active historical thinking: Assignments for activating history education]*. Stichting Geschiedenis, Staatsinrichting en Educatie.

De Wever, B., Van Keer, H., Schellens, T., & Valcke, M. (2007). Applying multilevel modelling to content analysis data: Methodological issues in the study of role assignment in asynchronous discussion groups. *Learning and Instruction, 17*, 436-447.

Dennis, A. R., Kinney, S. T., & Hung, Y. T. C. (1999). Gender differences in the effects of media richness. *Small Group Research, 30*, 405-437.

Dennis, A. R., & Valacich, J. S. (1999). *Rethinking media richness: Towards a theory of media synchronicity.* Paper presented at the 32nd Hawaii International Conference on Information Systems (HICSS), Kohala Coast, HI.

Dewiyanti, S., Brand-Gruwel, S., Jochems, W., & Broers, N. J. (2007). Students' experiences with collaborative learning in asynchronous Computer-Supported Collaborative Learning environments. *Computers in Human Behavior, 23*, 496-514.

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*, 1017-1076.

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

Dimant, R. J., & Bearison, D. J. (1991). Development of formal reasoning during successive peer interactions. *Developmental Psychology, 27*, 277-284.

Donath, J. (2002). Visualizing online conversations. *Communications of the ACM, 45*(4), 45-49.

Dönmez, P., Rosé, C., Stegmann, K., Weinberger, A., & Fischer, F. (2005). Supporting CSCL with automatic corpus analysis technology. In T. Koschmann, D. D. Suthers, & T.-W. Chan (Eds.), *Proceedings of the 2005 conference on Computer support for collaborative learning: The next 10 years!* (pp. 125-134). Mahwah, NJ: Lawrence Erlbaum Associates.

Dourish, P., & Bellotti, V. (1992, November). *Awareness and coordination in shared work spaces.* Paper presented at the ACM Conference on Computer Supported Cooperative Work (CSCW'92), Toronto, Canada.

Dubrovsky, V. J., Kiesler, S., & Sethna, B. N. (1991). The equalization phenomenon: Status effects in computer-mediated and face-to-face decision-making groups. *Human-Computer Interaction, 6*, 119-146.

Ellis, C. A., Gibbs, S. J., & Rein, G. (1992). Groupware: Some issues and experiences. In D. Marca & G. Bock (Eds.), *Groupware: Software for computer-supported cooperative work* (pp. 23-43). Los Alamitos, CA: IEEE Computer Society Press.

Ellis, S. (1997). Strategy choice in sociocultural context. *Developmental Review, 17*, 490-524.

Ericsson, K. A., & Simon, H. A. (1993). *Protocol analysis*. Cambridge, MA: MIT Press.

Erkens, G. (1997). *Coöperatief probleemoplossen met computers in het onderwijs: Het modellering van coöperatieve dialogen voor de ontwikkeling van intelligente onderwijssystemen [Cooperative problem solving with computers in education: Modelling cooperative dialogues for the development of intelligent educational systems].* Unpublished PhD. thesis, Utrecht University, The Netherlands.

Erkens, G. (2003). Multiple Episode Protocol Analysis (MEPA). Version 4.9. Utrecht University, The Netherlands.

Erkens, G. (2004). Dynamics of coordination in collaboration. In J. van der Linden & P. Renshaw (Eds.), *Dialogic learning: Shifting perspectives to learning, instruction, and teaching* (pp. 191-216). Dordrecht, The Netherlands: Kluwer Academic Publishers.

Erkens, G. (2005). Multiple Episode Protocol Analysis (MEPA). Version 4.10. Utrecht University, The Netherlands.

Erkens, G., Jaspers, J., Prangsma, M., & Kanselaar, G. (2005). Coordination processes in computer supported collaborative writing. *Computers in Human Behavior, 21*, 463-486.

Erkens, G., Prangsma, M., & Jaspers, J. (2006). Planning and coordinating activities in collaborative learning. In A. M. O'Donnell, C. E. Hmelo-Silver, & G. Erkens (Eds.), *Collaborative learning, reasoning, and technology* (pp. 233-263). Mahwah, NJ: Lawrence Erlbaum Associates, Publishers.

Felton, M., & Kuhn, D. (2001). The development of argumentive discourse skill. *Discourse Processes, 32*, 135-153.

References

Fischer, F., Bruhn, J., Gräsel, C., & Mandl, H. (2002). Fostering collaborative knowledge construction with visualization tools. *Learning and Instruction, 12*, 213-232.

Fjermestad, J. (2004). An analysis of communication mode in group support systems research. *Decision Support Systems, 37*, 239-263.

Fletcher-Flinn, C. M., & Gravatt, B. (1995). The efficacy of computer assisted instruction (CAI): A meta-analysis. *Journal of Educational Computing Research, 13*, 219-241.

Forman, E. A., & Cazden, C. B. (1985). Exploring Vygotskian perspectives in education: The cognitive value of peer interaction. In J. V. Wertsch (Ed.), *Culture, communication, and cognition: Vygotskian perspectives* (pp. 323-347). Cambridge: University Press.

Francescato, D., Porcelli, R., Mebane, M., Cuddetta, M., Klobas, J., & Renzi, P. (2006). Evaluation of the efficacy of collaborative learning in face-to-face and computer-supported university contexts. *Computers in Human Behavior, 22*, 163.

Fuks, H., Pimentel, M., & Lucena, C. J. P. (2006). R-U-Typing-2-Me? Evolving a chat tool to increase understanding in learning activities. *International Journal of Computer-Supported Collaborative Learning, 1*, 117-142.

Gijlers, H., & De Jong, T. (2005). The relation between prior knowledge and students' collaborative discovery learning processes. *Journal of Research in Science Teaching, 42*, 264-282.

Gillies, R. M. (2006). Teachers' and students' verbal behaviours during cooperative and small-group learning. *British Journal of Educational Psychology, 76*, 271-287.

Guiller, J., & Durndell, A. (2007). Students' linguistic behaviour in online discussion groups: Does gender matter? *Computers in Human Behavior, 23*, 2240-2255.

Gutwin, C., & Greenberg, S. (1999). The effects of workspace awareness support on the usability of real-time distributed groupware. *ACM Transactions on Computer-Human Interaction, 6*, 243-281.

Gutwin, C., & Greenberg, S. (2002). A descriptive framework of workspace awareness for real-time groupware. *Computer Supported Cooperative Work (CSCW), 11*, 411-446.

Gutwin, C., & Greenberg, S. (2004). The importance of awareness for team cognition in distributed collaboration. In E. Salas & S. M. Fiore (Eds.), *Team cognition: Understanding the factors that drive processes and performance* (pp. 177-201). Washington: APA Press.

Gutwin, C., Stark, G., & Greenberg, S. (1995). Support for workspace awareness in educational groupware. In J. L. Schnase & E. L. Cunnius (Eds.), *The first international conference on Computer support for collaborative learning* (pp. 147-156). Mahwah, NJ: Lawrence Erlbaum Associates Inc.

Henry, R. A. (1995). Improving group judgment accuracy: Information sharing and determining the best member. *Organizational Behavior and Human Decision Processes, 62*, 190-197.

Hertz-Lazarowitz, R., & Bar-Natan, I. (2002). Writing development of Arab and Jewish students using cooperative learning (CL) and computer-mediated communication (CMC). *Computers & Education, 39*, 19-36.

Hetzel, R. D. (1996). A primer on factor analysis with comments on patterns of practice and reporting. In B. Thompson (Ed.), *Advances in social science methodology* (Vol. 4, pp. 175-206). Stamford, CT: JAI Press.

Hever, R., De Groot, R., De Laat, M., Harrer, A., Hoppe, U., McLaren, B., et al. (2007). Combining structural, process-oriented and textual elements to generate awareness indicators for graphical E-discussions. In C. A. Chinn, G. Erkens, & S. Puntambekar (Eds.), *Mice, minds, and society: The Computer Supported Collaborative Learning (CSCL) Conference 2007* (Vol. 8, pp. 286-288). New Brunswick, NJ: International Society of the Learning Sciences.

Hightower, R., & Sayeed, L. (1995). The impact of computer-mediated communication systems on biased group discussion. *Computers in Human Behavior, 11*, 33-44.

Hmelo-Silver, C. E. (2003). Analyzing collaborative knowledge construction: Multiple methods for integrated understanding. *Computers & Education, 41*, 397-420.

Hmelo-Silver, C. E. (2004). Problem-based learning: What and how do students learn? *Educational Psychology Review, 16*, 235-266.

Hmelo-Silver, C. E., & Bromme, R. (2007). Coding discussions and discussing coding: Research on collaborative learning in computer-supported environments. *Learning and Instruction, 17*, 460-464.

Hobman, E. V., Bordia, P., Irmer, B., & Chang, A. (2002). The expression of conflict in computer-mediated and face-to-face groups. *Small Group Research, 33*, 439-465.

Hollingshead, A. B., & McGrath, J. E. (1995). Computer-assisted groups: A critical review of the empirical research. In R. A. Guzzo, E. Salas, & Associates (Eds.), *Team effectiveness and decision making in organizations* (pp. 46-78). San Francisco: Jossey-Bass.

Howell-Richardson, C., & Mellar, H. (1996). A methodology for the analysis of patterns of participation within computer mediated communication courses. *Instructional Science, 24*, 47-69.

Janssen, J., Erkens, G., & Kanselaar, G. (2007). Visualization of agreement and discussion processes during computer-supported collaborative learning. *Computers in Human Behavior, 23*, 1105-1125.

Janssen, J., Erkens, G., Kanselaar, G., & Jaspers, J. (2007). Visualization of participation: Does it contribute to successful computer-supported collaborative learning? *Computers & Education, 49*, 1037-1065.

Jaspers, J., Broeken, M., & Erkens, G. (2004). Virtual Collaborative Research Institute (VCRI) (Version 2.0). Utrecht: Onderwijskunde Utrecht, ICO/ISOR.

Jaspers, J., Broeken, M., & Erkens, G. (2005). Virtual Collaborative Research Institute (VCRI). Version 2.2. Utrecht, The Netherlands: Utrecht University.

Jehn, K. A., & Shah, P. P. (1997). Interpersonal relationships and task performance: An examination of mediation processes in friendship and acquaintance groups. *Journal of Personality and Social Psychology, 72*, 775-790.

Jermann, P. (2004). *Computer support for interaction regulation in collaborative problem-solving.* Unpublished PhD thesis, University of Geneva, Switzerland.

Johnson, D. W., & Johnson, R. T. (1999). *Learning together and alone: Cooperative, competitive, and individualistic learning* (5th ed.). Boston: Allyn and Bacon.

Johnson, D. W., Johnson, R. T., Roy, P., & Zaidman, B. (1985). Oral interaction in cooperative learning groups: Speaking, listening, and the nature of statements made by high-, medium-, and low-achieving students. *Journal of Psychology: Interdisciplinary and Applied, 119*(4), 303-321.

Johnson, D. W., Johnson, R. T., & Stanne, M. B. (1990). Impact of group processing on achievement in cooperative groups. *Journal of Social Psychology, 130*, 507-516.

Kashy, D. A., & Kenny, D. A. (2000). The analysis of data from dyads and groups. In H. T. Reis & C. M. Judd (Eds.), *Handbook of research methods in social and personality psychology* (pp. 451-477). Cambridge: Cambridge University Press.

Keller, T., & Tergan, S. O. (2005). Visualizing knowledge and information: An introduction. In S. O. Tergan & T. Keller (Eds.), *Knowledge and information visualization: Searching for synergies* (Vol. 3426, pp. 1-23). Berlin: Springer-Verlag Berlin.

Kenny, D. A. (1995). The effect of nonindependence on significance testing in dyadic research. *Personal Relationships, 2*, 67-75.

Kenny, D. A. (1996). Models of non-independence in dyadic research. *Journal of Social and Personal Relationships, 13*, 279-294.

Kenny, D. A., & Judd, C. M. (1986). Consequences of violating the independence assumption in analysis of variance. *Psychological Bulletin, 99*, 422.

Kenny, D. A., & Judd, C. M. (1996). A general procedure for the estimation of interdependence. *Psychological Bulletin, 119*, 138-148.

Kenny, D. A., Kashy, D. A., & Cook, W. L. (2006). *Dyadic data analysis*. New York/London: The Guilford Press.

Kenny, D. A., Mannetti, L., Pierro, A., Livi, S., & Kashy, D. A. (2002). The statistical analysis of data from small groups. *Journal of Personality and Social Psychology, 83*, 126-137.

Kiesler, S., Siegel, J., & McGuire, T. W. (1984). Social psychological aspects of computer-mediated communication. *American Psychologist, 39*, 1123-1134.

Kiesler, S., & Sproull, L. (1992). Group decision making and communication technology. *Organizational Behavior and Human Decision Processes, 52*, 96-123.

King, A. (1994). Guiding knowledge construction in the classroom: Effects of teaching children how to question and how to explain. *American Educational Research Journal, 31*, 338-368.

King, L. H. (1993). High and low achievers' perceptions and cooperative learning in two small groups. *Elementary School Journal, 93*, 399-416.

Kirschner, P. A. (2002). Can we support CSCL? Educational, social and technological affordances for learning. In P. A. Kirschner (Ed.), *Three worlds of CSCL: Can we support CSCL?* (pp. 7-47). Heerlen, The Netherlands: Open Universiteit.

Kirschner, P. A. (2004). Design, development, and implementation of electronic learning environments for collaborative learning. *Educational Technology Research and Development, 52*(3), 39-46.

Kirschner, P. A., Beers, P. J., Boshuizen, H. P. A., & Gijselaers, W. H. (in press). Coercing shared knowledge in collaborative learning environments. *Computers in Human Behavior*.

Kirschner, P. A., Martens, R. L., & Strijbos, J. W. (2004). CSCL in higher education? A framework for designing multiple collaborative environments. In J. W. Strijbos, P. A. Kirschner, & R. L. Martens (Eds.), *What we know about CSCL, and implementing it in higher education* (pp. 3-30). Boston: Kluwer Academic Publishers.

Kirschner, P. A., Strijbos, J. W., Kreijns, K., & Beers, P. J. (2004). Designing electronic collaborative learning environments. *Educational Technology Research and Development, 52*(3), 47-66.

Kirschner, P. A., Sweller, J., & Clark, R. E. (2006). Why minimal guidance during instruction does not work: An analysis of the failure of constructivist, discovery, problem-based, experiential, and inquiry-based teaching. *Educational Psychologist, 41*, 75-86.

Kreijns, K. (2004). *Sociable CSCL environments: Social affordances, sociability, and social presence.* Unpublished PhD thesis, Open Universiteit, Heerlen, The Netherlands.

Kreijns, K., & Kirschner, P. A. (2001, October). *The social affordances of computer-supported collaborative learning environments.* Paper presented at the 31st Annual Frontiers in Education Conference, Reno, NV.

Kreijns, K., Kirschner, P. A., & Jochems, W. (2003). Identifying the pitfalls for social interaction in computer-supported collaborative learning environments: A review of the research. *Computers in Human Behavior, 19*, 335-353.

Kreijns, K., Kirschner, P. A., Jochems, W., & Van Buuren, H. (2004). Measuring perceived quality of social space in distributed learning groups. *Computers in Human Behavior, 20*, 607-632.

Krol-Pot, K. (2005). *Toward interdependence: Implementation of cooperative learning in primary schools.* Unpublished PhD thesis, Radboud University, Nijmegen, The Netherlands.

Krol, K., Janssen, J., Veenman, S., & Van der Linden, J. (2004). Effects of a cooperative learning program on the elaborations of students working in dyads. *Educational Research and Evaluation, 10*, 205-237.

Kuhn, D. (1991). *The skills of argument.* Cambridge, MA: Harvard University Press.

Kuhn, D., & Udell, W. (2003). The development of argument skills. *Child Development, 74*, 1245-1260.

Kuiper, E., Volman, M., & Terwel, J. (2004). Internet als informatiebron in het onderwijs: Een verkenning van de literatuur [Using the World Wide Web (WWW) as an information resource in K-12 education: A review of the literature]. *Pedagogische Studiën, 81*, 423-443.

Kumpulainen, K., & Mutanen, M. (1999). The situated dynamics of peer group interaction: An introduction to an analytic framework. *Learning and Instruction, 9*(5), 449-473.

Lazonder, A. W., Wilhelm, P., & Ootes, S. A. W. (2003). Using sentence openers to foster student interaction in computer-mediated learning environments. *Computers & Education, 41*, 291.

Lee, A., Girgensohn, A., & Zhang, J. (2004). Bowsers to support awareness and social interaction. *IEEE Computer Graphics and Applications, 24*(5), 66-75.

Lipponen, L., Rahikainen, M., Lallimo, J., & Hakkarainen, K. (2003). Patterns of participation and discourse in elementary students' computer-supported collaborative learning. *Learning and Instruction, 13*, 487-509.

Lockhorst, D. (2004). *Design principles for a CSCL environment in teacher training.* Unpublished PhD thesis, IVLOS, Utrecht University, The Netherlands.

Löhner, S., Van Joolingen, W. R., & Savelsbergh, E. R. (2003). The effect of external representation on constructing computer models of complex phenomena. *Instructional Science, 31*(6), 395-418.

Lou, Y., Abrami, P. C., & d'Apollonia, S. (2001). Small group and individual learning with technology: A meta-analysis. *Review of Educational Research, 71*, 449-521.

Malone, T. W., & Crowston, K. (1992). What is coordination theory and how can it help design cooperative work systems? In D. Marca & G. Bock (Eds.), *Groupware: Software for computer-supported cooperative work* (pp. 100-113). Los Alamitos, CA: IEEE Computer Society Press.

Manlove, S., Lazonder, A. W., & De Jong, T. (2006). Regulative support for collaborative scientific inquiry learning. *Journal of Computer Assisted Learning, 22*, 87-98.

Massey, A. P., Montoya-Weiss, M. M., & Hung, Y. (2003). Because time matters: Temporal coordination in global virtual teams. *Journal of Management Information Systems, 19*, 129-155.

McGrath, J. E. (1991). Time, interaction, and performance (TIP). *Small Group Research, 22*, 147-174.

Meijer, P. C., Zanting, A., & Verloop, N. (2002). How can student teachers elicit experienced teachers' practical knowledge? Tools, suggestions, and significance. *Journal of Teacher Education, 53*, 406-419.

Mendoza-Chapa, S., Romero-Salcedo, M., & Oktaba, H. (2000, October). *Group awareness support in collaborative writing systems.* Paper presented at the 6th International Workshop on Groupware (CRIWG'00), Madeira, Portugal.

Mennecke, B. E., Valacich, J. S., & Wheeler, B. C. (2000). The effects of media and task on user performance: A test of the task-media fit hypothesis. *Group Decision and Negotiation, 9*, 507-529.

Mercer, N., Wegerif, R., & Dawes, L. (1999). Children's talk and the development of reasoning in the classroom. *British Educational Research Journal, 25*, 95-111.

Michinov, N., & Primois, C. (2005). Improving productivity and creativity in online groups through social comparison process: New evidence for asynchronous electronic brainstorming. *Computers in Human Behavior, 21*, 11-28.

Moreno, R., & Valdez, A. (2005). Cognitive load and learning effects of having students organize pictures and words in multimedia environments: The role of student interactivity and feedback. *Educational Technology Research and Development, 53*(3), 35-45.

Mühlpfordt, M., & Stahl, G. (2007). The integration of synchronous communication across dual interaction spaces. In C. A. Chinn, G. Erkens, & S. Puntambekar (Eds.), *Mice, minds, and society: The Computer Supported Collaborative Learning (CSCL) Conference 2007* (Vol. 8, pp. 518-527). New Brunswick, NJ: International Society of the Learning Sciences.

Munneke-de Vries, E. L. (2008). *Arguing to learn: Supporting interactive argumentation through computer-supported collaborative learning.* Unpublished PhD thesis, Utrecht University, The Netherlands.

Munneke, L., Andriessen, J., Kanselaar, G., & Kirschner, P. (2007). Supporting interactive argumentation: Influence of representational tools on discussing a wicked problem. *Computers in Human Behavior, 23*, 1072-1088.

Munneke, L., Andriessen, J., Kirschner, P., & Kanselaar, G. (2007). Effects of synchronous and asynchronous CMC on interactive argumentation. In C. A. Chinn, G. Erkens, & S. Puntambekar (Eds.), *Mice, minds, and society: The Computer Supported Collaborative Learning (CSCL) Conference 2007* (Vol. 8, pp. 528-537). New Brunswick, NJ: International Society of the Learning Sciences.

Munneke, L., Van Amelsvoort, M., & Andriessen, J. (2003). The role of diagrams in collaborative argumentation-based learning. *International Journal of Educational Research, 39*, 113-131.

Narciss, S., Proske, A., & Koerndle, H. (2007). Promoting self-regulated learning in web-based learning environments. *Computers in Human Behavior, 23*, 1126-1144.

Niederhoffer, K. G., & Pennebaker, J. W. (2002). Linguistic style matching in social interaction. *Journal of Language and Social Psychology, 21*, 337-360.

Nussbaum, E. M. (2005). The effect of goal instructions and need for cognition on interactive argumentation. *Contemporary Educational Psychology, 30*, 286-313.

O'Donnell, A. M. (2001). Group processes in the classroom. In N. J. Smelser & B. B. Baltes (Eds.), *International encyclopedia of the social & behavioral sciences* (pp. 6413-6417). Amsterdam: Elsevier.

O'Donnell, A. M., & O'Kelly, J. (1994). Learning from peers: Beyond the rhetoric of positive results. *Educational Psychology Review, 6*, 321-349.

Paulsen, M. F. (1995). Moderating educational computer conferences. In Z. L. Berge & M. P. Collins (Eds.), *Computer mediated communication and the online classroom. Volume three: Distance Learning* (Vol. 3, pp. 81-89). Cresskill, NJ: Hampton Press.

Postmes, T., Spears, R., & Cihangir, S. (2001). Quality of decision making and group norms. *Journal of Personality and Social Psychology, 80*, 918-930.

Prasolova-Forland, E. (2002, August). *Supporting awareness in education: Overview and mechanisms.* Paper presented at the International Conference on Engineering Education, Manchester, U.K.

Quintana, C., Reiser, B. J., Davis, E. A., Krajcik, J., Fretz, E., Duncan, R. G., et al. (2004). A scaffolding design framework for software to support science inquiry. *Journal of the Learning Sciences, 13*, 337-386.

Romero-Salcedo, M., Osuna-Gómez, C., A., Sheremetov, L., Villa, L., Morales, C., Rocha, L., et al. (2004). Study and analysis of workspace awareness in CDebate: A groupware application for collaborative debates. *Interacting with Computers, 16*, 657-681.

Ros, A. A. (1994). *Samenwerking tussen leerlingen en effectief onderwijs [Collaboration between students and effective education]*. Unpublished PhD thesis, Groningen, The Netherlands.

Rourke, L., Anderson, T., Garrison, D. R., & Archer, W. (1999). Assessing social presence in asynchronous text-based computer conferencing. *Journal of Distance Education, 14*(2), 50-71.

Saab, N., Van Joolingen, W. R., & Van Hout-Wolters, B. (2007). Supporting communication in a collaborative discovery learning environment: The effect of instruction. *Instructional Science, 35*, 73-98.

Saavedra, R., Earley, P., & Van Dyne, L. (1993). Complex interdependence in task-performing groups. *Journal of Applied Psychology, 78*, 61-72.

Salmon, G. (2000). *E-moderating: The key to teaching and learning online*. London: Kogan Page Limited.

Salomon, G., & Globerson, T. (1989). When teams do not function the way they ought to. *International Journal of Educational Research, 13*, 89-99.

Salomon, G., & Perkins, D. N. (1998). Individual and social aspects of learning. In P. D. Pearson & A. Iran-Nejad (Eds.), *Review of Research in Education* (Vol. 23, pp. 1-24). Washington, DC: American Educational Research Association.

Sampson, V. D., & Clark, D. B. (2006). Assessment of argument in science education: A critical review of the literature. In S. A. Barab, K. E. Hay, & D. T. Hickey (Eds.), *Proceedings of the 7th International Conference of the Learning Sciences (ICLS)* (Vol. 2, pp. 655-661). Mahwah, NJ: Lawrence Erlbaum Associates.

Sandoval, W. A., & Millwood, K. A. (2005). The quality of students' use of evidence in written scientific explanations. *Cognition and Instruction, 23*, 23-55.

Savicki, V., & Kelley, M. (2000). Computer mediated communication: Gender and group composition. *CyberPsychology & Behavior, 3*, 817-826.

Savicki, V., Kelley, M., & Ammon, B. (2002). Effects of training on computer-mediated communication in single or mixed gender small task groups. *Computers in Human Behavior, 18*(3), 257-270.

Savicki, V., Kelley, M., & Lingenfelter, D. (1996). Gender and group composition in small task groups using computer-mediated communication. *Computers in Human Behavior, 12*, 209-224.

Scaife, M., & Rogers, Y. (1996). External cognition: How do graphical representations work? *International Journal Of Human-Computer Studies, 45*, 185-213.

Scarloss, B. A. (2002, April). *Assessing complex academic performance at the group level.* Paper presented at the American Education Research Association (AERA) annual meeting, New Orleans, LO.

Schellens, T., & Valcke, M. (2006). Fostering knowledge construction in university students through asynchronous discussion groups. *Computers & Education, 46*, 349-370.

Schellens, T., Van Keer, H., & Valcke, M. (2005). The impact of role assignment on knowledge construction in asynchronous discussion groups: A multilevel analysis. *Small Group Research, 36*, 704-745.

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

Schmidt, K. (2002). The problem with 'Awareness': Introductory remarks on 'Awareness in CSCW'. *Computer Supported Cooperative Work (CSCW), 11*, 285-298.

Schnotz, W., & Lowe, R. (2003). External and internal representations in multimedia learning - Introduction. *Learning and Instruction, 13*, 117-123.

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

Schwarz, B. B., Neuman, Y., Gil, J., & Ilya, M. (2003). Construction of collective and individual knowledge in argumentative activity. *Journal of the Learning Sciences, 12*, 219-256.

Shepperd, J. A. (1993). Productivity loss in performance groups: A motivation analysis. *Psychological Bulletin, 113*, 67-81.

Slavin, R. E. (1996). Research on cooperative learning and achievement: What we know, what we need to know. *Contemporary Educational Psychology, 21*, 43-69.

Snijders, T. A. B., & Bosker, R. J. (1999). *Multilevel analysis: An introduction to basic and advanced multilevel modeling*. London: Sage Publications.

Spiegel, D. (2001). *Coterie: A visualization of the conversational dynamics within IRC.* Unpublished Master's thesis, MIT.

Stahl, G. (2004). Building collaborative knowing: Elements of a social theory of CSCL. In J. W. Strijbos, P. A. Kirschner, & R. L. Martens (Eds.), *What we know about CSCL, and implementing it in higher education* (pp. 53-85). Boston: Kluwer Academic Publishers.

Stahl, G. (2006). *Group cognition: Computer support for building collaborative knowledge*. Cambridge, MA: MIT Press.

Stahl, G., Zemel, A., Sarmiento, J., Cakir, M., Weimar, S., Wessner, M., et al. (2006). Shared referencing of mathematical objects in online chat. In S. A. Barab, K. E. Hay, & D. T. Hickey (Eds.), *Proceedings of the 7th International Conference of the Learning Sciences (ICLS)* (Vol. 2, pp. 716-722). Mahwah, NJ: Lawrence Erlbaum Associates.

Stern, E., Aprea, C., & Ebner, H. G. (2003). Improving cross-content transfer in text processing by means of active graphical representation. *Learning and Instruction, 13*, 191-203.

Straus, S. G. (1997). Technology, group process, and group outcomes: Testing the connections in computer-mediated and face-to-face groups. *Human-Computer Interaction, 12*, 227-266.

Straus, S. G., & McGrath, J. E. (1994). Does the medium matter? The interaction of task type and technology on group performance and member reactions. *Journal of Applied Psychology, 79*, 87-97.

Strijbos, J.-W., & Fischer, F. (2007). Methodological challenges for collaborative learning research. *Learning and Instruction, 17*, 389-393.

Strijbos, J. W., Martens, R. L., & Jochems, W. M. G. (2004). Designing for interaction: Six steps to designing computer-supported group-based learning. *Computers & Education, 42*, 403-424.

Strijbos, J. W., Martens, R. L., Jochems, W. M. G., & Broers, N. J. (2004). The effect of functional roles on group efficiency: Using multilevel modeling and content analysis to investigate computer-supported collaboration in small groups. *Small Group Research, 35*, 195-229.

Strijbos, J. W., Martens, R. L., Jochems, W. M. G., & Broers, N. J. (2007). The effect of functional roles on perceived group efficiency during computer-supported collaborative learning: A matter of triangulation. *Computers in Human Behavior, 23*, 353-380.

Strijbos, J. W., Martens, R. L., Prins, F. J., & Jochems, W. M. G. (2006). Content analysis: What are they talking about? *Computers & Education, 46*, 29-48.

Suthers, D. D. (2001). Towards a systematic study of representational guidance for collaborative learning discourse. *Journal of Universal Computer Science, 7*, 254-277.

Suthers, D. D. (2003). Representational guidance for collaborative inquiry. In J. Andriessen, M. Baker, & D. Suthers (Eds.), *Arguing to learn: Confronting cognitions in computer-supported collaborative learning environments* (pp. 27-46). Dordrecht: Kluwer Academic Publishers.

Suthers, D. D. (2006). Technology affordances for intersubjective meaning making. *International Journal of Computer Supported Collaborative Learning, 1*, 315-337.

Suthers, D. D., & Hundhausen, C. D. (2003). An experimental study of the effects of representational guidance on collaborative learning processes. *Journal of the Learning Sciences, 12*, 183-218.

Swaak, J., Van Joolingen, W. R., & De Jong, T. (1998). Supporting simulation-based learning: The effects of model progression and assignments on definitional and intuitive knowledge. *Learning and Instruction, 8*, 235-253.

Sweller, J., & Chandler, P. (1994). Why some material is difficult to learn. *Cognition and Instruction, 12*, 185-233.

Teasley, S. D. (1995). The role of talk in children's peer collaborations. *Developmental Psychology, 31*, 207-220.

Teasley, S. D., & Roschelle, J. (1993). Constructing a joint problem space: The computer as a tool for sharing knowledge. In S. P. Lajoie (Ed.), *Computers as cognitive tools: Technology in education* (pp. 229-258). Hillsdale, NJ, England: Lawrence Erlbaum Associates, Inc.

Thompson, L., & Fine, G. A. (1999). Socially shared cognition, affect, and behavior: A review and integration. *Personality and Social Psychology Review, 3*, 278-302.

Thompson, L. F., & Coovert, M. D. (2003). Teamwork online: The effects of computer conferencing on perceived confusion, satisfaction and postdiscussion accuracy. *Group Dynamics, 7*, 135-151.

Toth, E. E., Suthers, D. D., & Lesgold, A. M. (2002). Mapping to know: The effects of representational guidance and reflective assessment on scientific inquiry. *Science Education, 86*, 264-286.

Tudge, J. R. H. (1989). When collaboration leads to regression: Some negative consequences of socio-cognitive conflict. *European Journal of Social Psychology, 19*(2), 123-138.

Van Amelsvoort, M. (2006). *A space for debate: How diagrams support collaborative argumentation-based learning.* Unpublished PhD-thesis, Utrecht University, Utrecht, The Netherlands.

Van Amelsvoort, M., Andriessen, J., & Kanselaar, G. (in press). How students structure and relate argumentative knowledge when learning together with diagrams. *Computers in Human Behavior*.

Van Bruggen, J. M., Boshuizen, H. P., & Kirschner, P. A. (2003). A cognitive framework for cooperative problem solving with argument visualization. In P. A. Kirschner, S. J. Buckingham-Shum, & C. S. Carr (Eds.), *Visualizing argumentation: Software tools for collaborative and educational sense-making* (pp. 25-47). London: Springer.

Van Bruggen, J. M., Kirschner, P. A., & Jochems, W. (2002). External representation of argumentation in CSCL and the management of cognitive load. *Learning and Instruction, 12*, 121-138.

Van der Linden, J., Erkens, G., & Nieuwenhuysen, T. (1995). Gemeinsames Problemlösen in Gruppen. *Unterrichtswissenschaft, 23*(4), 301-315.

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. 1-19). Dordrecht: Kluwer Academic Publishers.

Van der Meijden, H., & Veenman, S. (2005). Face-to-face versus computer-mediated communication in a primary school setting. *Computers in Human Behavior, 21*, 831-859.

Van der Pol, J. (2007). *Facilitating online learning conversations: Exploring tool affordances in higher Education.* Unpublished PhD thesis, Utrecht University, The Netherlands.

Van Drie, J. (2005). *Learning about the past with new technologies: Fostering historical reasoning in computer-supported collaborative learning.* Unpublished PhD thesis, Utrecht University, Utrecht, The Netherlands.

Van Drie, J., Van Boxtel, C., Jaspers, J., & Kanselaar, G. (2005). Effects of representational guidance on domain specific reasoning in CSCL. *Computers in Human Behavior, 21*, 575-602.

Van Meter, P., & Stevens, R. J. (2000). The role of theory in the study of peer collaboration. *Journal of Experimental Education, 69*, 113-127.

Veenman, S., Denessen, E., Van den Akker, A., & Van der Rijt, J. (2005). Effects of a cooperative learning program on the elaborations of students during help seeking and help giving. *American Educational Research Journal, 42*(1), 115-151.

Veerman, A. (2003). Constructive discussions through electronic dialogue. In J. Andriessen, M. Baker, & D. Suthers (Eds.), *Arguing to learn: Confronting cognitions in computer-supported collaborative learning environments* (pp. 117-143). Dordrecht, The Netherlands: Kluwer Academic Publishers.

Veldhuis-Diermanse, A. E. (2002). *CSCLearning? Participation, learning activities and knowledge construction in computer-supported collaborative learning in higher education.* Unpublished PhD thesis, Wageningen University, The Netherlands.

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

Ware, C. (2005). Visual queries: The foundation of visual thinking. In S. O. Tergan & T. Keller (Eds.), *Knowledge and information visualization: Searching for synergies* (Vol. 3426, pp. 27-35). Berlin: Springer-Verlag Berlin.

Warschauer, M. (1996). Comparing face-to-face and electronic discussion in the second language classroom. *CALICO Journal, 13*, 7-26.

Webb, N. M. (1995). Group collaboration in assessment: Multiple objectives, processes, and outcomes. *Educational Evaluation and Policy Analysis, 17*, 239-261.

Webb, N. M., & Mastergeorge, A. M. (2003). The development of students' helping behavior and learning in peer-directed small groups. *Cognition and Instruction, 21*, 361-428.

Webb, N. M., Nemer, K. M., & Ing, M. (2006). Small-group reflections: Parallels between teacher discourse and student behavior in peer-directed groups. *Journal of the Learning Sciences, 15*, 63-119.

Webb, N. M., & Palincsar, A. S. (1996). Group processes in the classroom. In D. C. Berliner (Ed.), *Handbook of educational psychology* (pp. 841-873). New York: Simon & Schuster Macmillan.

Webb, N. M., Troper, J. D., & Fall, R. (1995). Constructive activity and learning in collaborative small groups. *Journal of Educational Psychology, 87*(3), 406-423.

Wegerif, R., Mercer, N., & Dawes, L. (1999). From social interaction to individual reasoning: An empirical investigation of a possible socio-cultural model of cognitive development. *Learning and Instruction, 9*, 493-516.

Weinberger, A., Ertl, B., Fischer, F., & Mandl, H. (2005). Epistemic and social scripts in computer-supported collaborative learning. *Instructional Science, 33*, 1-30.

Weinberger, A., & Fischer, F. (2006). A framework to analyze argumentative knowledge construction in computer-supported collaborative learning. *Computers & Education, 46*, 71-95.

Wheeler, R., Suls, J., & Martin, R. (2001). Psychology of social comparison. In N. J. Smelser & P. B. Baltes (Eds.), *International encyclopedia of the social & behavioral sciences* (Vol. 21, pp. 14254-14257). Amsterdam: Elsevier.

Yager, S., Johnson, R. T., Johnson, D. W., & Snider, B. (1986). The impact of group processing on achievement in cooperative learning groups. *Journal of Social Psychology, 126*, 389-397.

Zhang, J. J. (1997). The nature of external representations in problem solving. *Cognitive Science, 21*, 179-217.

Zumbach, J., Hillers, A., & Reimann, P. (2004). Distributed problem-based learning: The use of feedback mechanisms in online learning. In T. S. Roberts (Ed.), *Online collaborative learning: Theory and practice* (pp. 86-102). Hershey, PA: Idea Group, Inc.

Zumbach, J., & Reimann, P. (2003). Influence of feedback on distributed problem based learning: Enhancing online communities in group functioning and well-being. In B. Wasson, S. Ludvigsen, & U. Hoppe (Eds.), *Designing for change in networked learning environments* (pp. 219-228). Dordrecht, The Netherlands: Kluwer.

Samenvatting

Dit proefschrift richt zich op computerondersteund samenwerkend leren (computer-supported collaborative learning, afgekort CSCL). In een CSCL-omgeving werken kleine groepen leerlingen aan complexe, uitdagende taken. Hierbij bestuderen en verwerken zij zelfstandig, maar onder begeleiding van de docent, informatiebronnen, plannen en bewaken zij het leerproces, overleggen zij met elkaar over de lesstof en zijn zij gezamenlijk verantwoordelijk voor het groepsproduct (vaak een werkstuk of essay). Dit hele proces verloopt via de computer. Leerlingen werken met, en communiceren via de computer.

Hoewel er in de afgelopen jaren veel onderzoek gedaan is naar de doeltreffendheid van CSCL en de processen die tijdens CSCL plaatsvinden, blijkt dat er zich tijdens CSCL, net als tijdens 'gewone' face-to-face samenwerking (i.e., zelfde tijd en plaats), diverse problemen voor kunnen doen. Sommige van deze problemen zijn niet uniek voor CSCL, maar komen ook in andere samenwerkingssituaties voor, terwijl andere problemen juist verergerd lijken te worden door het samenwerken in een elektronische leeromgeving. In *Hoofdstuk 1* van dit proefschrift worden vier van deze problemen beschreven.

Het eerste probleem dat besproken wordt, heeft betrekking op de problemen die groepsleden ervaren bij het verkrijgen en interpreteren van informatie over de activiteiten van de rest van de groepsleden. Het *bewustzijn of besef (awareness) van wat de ander aan het doen is*, is vaak laag in CSCL-omgevingen (Gutwin & Greenberg, 2004; Gutwin et al., 1995). In 'gewone' samenwerkingssituaties is het vaak niet moeilijk om te weten waar groepsleden mee bezig zijn. Door het gedrag en de opmerkingen van groepsleden in de gaten te houden, is het vrij eenvoudig om te weten wat zij aan het doen zijn. In een CSCL-omgeving is het moeilijker om deze informatie te verzamelen. Groepsleden zijn niet direct zichtbaar en er is vaak weinig informatie voorhanden over wat zij doen. Dit beperkte besef van wat groepsleden aan het doen zijn, kan voor problemen zorgen. Wanneer leerlingen samenwerken, dienen zij bijvoorbeeld de eigen activiteiten maar ook de activiteiten van groepsleden in de gaten te houden. Verantwoordelijkheden zijn vaak gedeeld in een groepscontext, waardoor het van belang is om te bewaken dat iedereen zijn of haar bijdrage levert aan het groepsproces. Zonder informatie over wat groepsleden aan het doen zijn, is dit een moeilijke opgave.

Communicatieproblemen worden ook regelmatig gesignaleerd. Diverse onderzoekers dat het elektronisch communiceren voor problemen zorgt (zie Fjermestad, 2004). Elektronische communicatie wordt als minder rijk gezien dan bijvoorbeeld face-to-face communicatie. Dit komt bijvoorbeeld doordat tijdens elektronische communicatie, non-verbale communicatiemiddelen zoals gezichtsuitdrukkingen of gebaren niet beschikbaar zijn. Hierdoor hebben groepsleden regelmatig problemen om hun ideeën en opinies duidelijk te maken of komen deze ideeën verkeerd over (je kunt bijvoorbeeld niet eenvoudig vast stellen of een opmerking als grap bedoeld was). Daarnaast hebben zij soms moeite om de berichten van andere groepsleden te begrijpen en om te weten of hun argumenten en redeneringen begrepen en geaccepteerd worden door groepsleden (Adrianson & Hjelmquist, 1999; Kiesler et al., 1984; Thompson & Coovert, 2003). Omdat leerlingen in CSCL-omgevingen aan complexe, open problemen en groepstaken werken, kunnen dergelijke communicatieproblemen voor een verminderde taakprestatie zorgen (Kiesler & Sproull, 1992; Van der Meijden & Veenman, 2005).

Tijdens het samenwerken dienen leerlingen diverse complexe activiteiten uit te voeren. Deze activiteiten hebben enerzijds betrekking op het uitvoeren van de taak en het plannen, bewaken en evalueren van de voortgang van de taak en anderzijds op het sociale aspect van het samenwerken. Anderzijds dienen zij ook aandacht te schenken aan het creëren van een positief klimaat om in samen te werken en aan het plannen, bewaken en evalueren van het groepsproces (Cole & Nast-Cole, 1992; Kreijns et al., 2003; McGrath, 1991). Doordat groepsleden zoveel verschillende activiteiten uit dienen te voeren tijdens het samenwerken, ontstaat de noodzaak om deze te coördineren. Wanneer er sprake is van *coördinatieproblemen*, dan voeren groepsleden bijvoorbeeld activiteiten uit die conflicteren of overlappen met de activiteiten van andere groepsleden. Dit kan leiden tot irritatie en een verminderde groepsprestatie (Malone & Crowston, 1992).

Om samenwerking succesvol te coördineren kunnen groepsleden diverse strategieën gebruiken. Allereerst is het van belang dat zij hun kennis en vaardigheden activeren. Tijdens het samenwerken is het nodig dat groepsleden hun ideeën, kennis en oplossingen inbrengen in de groepsdiscussie door deze te externaliseren of uit te spreken (Erkens, 2004; Kirschner et al., in press). Daarnaast is het van belang dat er een zekere mate van evenwichtigheid is in de mate waarin groepsleden kennis en ideeën inbrengen in de discussie (Cohen, 1994; Damon & Phelps, 1989). Ten tweede dienen groepsleden te zorgen dat er tijdens het samenwerken sprake is van wederzijds begrip en een gedeeld referentiekader (Clark & Brennan, 1991; Van der Pol, 2007). Ten slotte dienen groepsleden te onderhandelen over de te kiezen oplossingen en strategieën en dienen zij hierover overeenstemming te bereiken (Di Eugenio et al., 2000; Erkens et al., 2006). Doordat groepsleden diverse activiteiten dienen te coördineren en deze coördinatie het gebruik van complexe strategieën vereist, bestaat de kans dat tijdens het samenwerken coördinatieproblemen ontstaan. Doordat groepsleden in CSCL-omgevingen meer moeite hebben om te weten wat de overige groepsleden uitvoeren en zij mogelijk communicatieproblemen ervaren, is de kans aanwezig dat in deze omgevingen de coördinatieproblemen vaker voorkomen of verergerd worden.

Het laatste probleem dat onderscheiden wordt, heeft te maken met het *gebrek aan diepgaande discussies tijdens samenwerkend leren*. Wanneer groepsleden aan een groepsopdracht werken, is het van belang dat zij op een kritische en constructieve met elkaar in discussie gaan. Dit betekent dat zij kritisch zijn op de eigen ideeën en argumenten maar ook op die van groepsleden, dat er een klimaat in de groep is waarin het geven van kritiek wordt gewaardeerd en er ook verklaringen en argumenten gegeven worden (Chinn & Anderson, 1998; Munneke, Andriessen, Kanselaar et al., 2007; Wegerif et al., 1999). Uit onderzoek blijkt echter dat dit soort discussies lang niet altijd gevoerd wordt tijdens face-to-face en computer-ondersteunde samenwerking (Kuhn & Udell, 2003; Munneke, Andriessen, Kanselaar et al., 2007; Van der Meijden & Veenman, 2005).

In *Hoofdstuk 1* wordt het gebruik van visualisaties in CSCL-omgevingen beschreven als een mogelijke oplossing voor de bovenstaande problemen. Visualisaties kunnen gebruikt worden om taakgerelateerde en sociale aspecten van het groepsproces visueel weer te geven en deze zo voor groepsleden inzichtelijk te maken. Zo kunnen visualisaties bijvoorbeeld gebruikt worden om participatie en gelijkwaardigheid in participatie visueel weer te geven (zie Jermann, 2004).

Het gebruik van visualisaties in CSCL-omgevingen heeft mogelijk enkele belangrijke voordelen. Op de eerste plaats kan door het gebruik van visualisaties het *verwerken en interpreteren van grote hoeveelheden complexe informatie vergemakkelijkt worden* (Keller & Tergan, 2005). Deze visualisaties voorzien groepsleden op de tweede plaats van *feedback of terugkoppeling* over het verloop van het groepsproces (i.e., Wordt er goed samengewerkt? Draagt iedereen in gelijke mate bij aan de samenwerking?). Deze feedback kan door groepsleden gebruikt worden om het groepsproces te bespreken en te evalueren en het communiceren over het groepsproces te vergemakkelijken. Door deze terugkoppeling wordt mogelijk ook het *besef van groepsleden over het verloop van het samenwerkingsproces vergroot* (Gutwin & Greenberg, 2004). Omdat visualisaties feedback geven en het bewustzijn van groepsleden over het samenwerkingsproces mogelijk vergroten, kan het *coördineren van de samenwerking ook vergemakkelijkt worden*. Door bijvoorbeeld participatie en gelijkwaardigheid van participatie te visualiseren, kunnen groepsleden achterhalen of er groepsleden zijn die het groepsproces domineren of zich aan het proces onttrekken. Dit kan ertoe leiden dat er besloten wordt om een andere samenwerkingsstrategie te kiezen. Ten slotte kunnen visualisaties groepsleden ook een *motivationele stimulans* bieden (Jermann, 2004). Visualisaties kunnen bijvoorbeeld gebruikt worden om de bijdragen van groepsleden aan het groepsproces identificeerbaar te maken. Dit kan ertoe leiden dat groepsleden zich met elkaar gaan vergelijken en kan hen aansporen om te streven naar een betere prestatie (Michinov & Primois, 2005).

In dit proefschrift wordt onderzocht of het gebruik van visualisaties inderdaad ertoe leidt dat het samenwerkingsproces in CSCL-omgevingen beter en effectiever verloopt. De centrale vraagstelling van dit proefschrift luidt dan ook: *Op welke manier heeft het voor groepsleden beschikbaar maken van visualisaties van het proces van samenwerking, invloed op het samenwerkingsproces, op de percepties van groepsleden van de samenwerking en op groeps- en individuele prestaties?* Om deze vraagstelling te beantwoorden zijn drie verschillende studies uitgevoerd die de effecten van drie verschillende visualisaties onderzochten.

Studie 1: Effecten van visualisatie van participatie

In *Hoofdstuk 2* van dit hoofdstuk wordt de *Participation-tool* beschreven. Deze applicatie is ontwikkeld om de relatieve bijdrage van groepsleden aan de groepsdiscussie in een CSCL-omgeving te visualiseren. De Participatie visualiseert de relatieve bijdrage van de groepsleden aan de groepsdiscussie in CSCL-omgeving. Tijdens het samenwerken is het voor het coördineren van de samenwerking van belang dat alle groepsleden een bijdrage leveren aan de groepsdiscussie door mening te geven, argumenten te geven en oplossingen aan te dragen. De Participation-tool tracht het besef van groepsleden over de bijdragen van groepsleden aan het groepsproces te verhogen en discussie over het samenwerkingsproces te stimuleren. Onderzocht is of de Participation-tool (a) het besef van groepsleden van het samenwerkingsproces verhoogde, (b) leidde tot een verhoogde participatie van groepsleden en een meer gelijkwaardige participatie, (c) de samenwerkingsactiviteiten die groepsleden uitvoerden beïnvloedde, en (d) leidde tot betere groepsprestaties.

Om deze onderzoeksvragen te beantwoorden werkten 52, 5-VWO leerlingen in 17 groepen in de VCRI-omgeving (Virtual Collaborative Research Institute, Broeken et al., 2006) waarbij zij de beschikking hadden over de Participation-tool. Daarnaast werkten 17 leerlingen in 5 groepen in dezelfde omgeving, maar dan zonder toegang tot de Participation-tool. In de VCRI-omgeving werkten de leerlingen in groepen van 2, 3 of 4 leerlingen aan een Praktische Opdracht voor het vak geschiedenis. Voor deze opdracht dienden zij diverse historische bronnen te bestuderen en deze informatie te integreren in een eindwerkstuk. De groepsleden verrichtten deze activiteiten in de VCRI-omgeving. Zo konden zij bijvoorbeeld de bronnen via het programma raadplegen, met elkaar overleggen via een chat-applicatie en een discussie forum en gezamenlijk schrijven aan het werkstuk via een gedeelde tekstverwerker.

Uit de resultaten van deze studie bleek dat de Participation-tool slechts een beperkt effect had op het besef van groepsleden van het samenwerkingsproces. Leerlingen die de beschikking hadden over de Participation-tool gaven alleen aan dat zij beter wisten wanneer een groepslid meeliftgedrag vertoonde; dat wil zeggen dat deze leerlingen beter wisten wanneer een groepslid geen of weinig werk verrichtte. De Participation-tool had echter wel een duidelijk effect op het participatieniveau van groepsleden. Wanneer leerlingen de beschikking hadden over deze tool, namen zij actiever deel aan de groepsdiscussies (i.e., zij verstuurden meer berichten) dan leerlingen die geen beschikking hadden over deze tool. In beide condities was er echter sprake van een redelijke mate van gelijkwaardigheid in participatie tussen groepsleden. Door de chat-discussies tussen groepsleden te bestuderen, konden we aantonen dat de Participation-tool een effect had op het samenwerkingsproces. Groepsleden die de beschikking hadden over de tool waren, in vergelijking tot groepsleden die niet de beschikking hadden over deze tool, meer bezig met het coördineren van hun samenwerking. In het bijzonder bespraken zij vaker hoe het samenwerkingsproces aangepakt en gepland diende te worden. Ondanks dat de Participation-tool invloed had op het participatieniveau van groepsleden en op de manier waarop samengewerkt werd, konden we geen effect van de tool op de kwaliteit van de groepsproducten aantonen.

Studie 2: Effecten van visualisatie van overeenstemming en discussie

De *Shared Space* wordt beschreven in *Hoofdstuk 3*. De Shared Space, geïmplementeerd in de *Chat-tool* van de eerder beschreven VCRI-omgeving, analyseert de berichten die groepsleden versturen en probeert vast te stellen of deze berichten indicaties van overeenstemming en discussie tussen groepsleden bevatten. Na deze analyse visualiseert de Shared Space de *overeenstemming* of *het verschil van mening* door de chat geschiedenis (history) naar links (discussie) of naar rechts (overeenstemming) te verplaatsen. Een verschil van mening wordt opgevat als een teken dat er discussie gevoerd wordt. Op deze manier geeft de Shared Space feedback over de soort discussies die groepsleden (kritische discussies of discussies gericht op consensus) aan het voeren zijn en wordt mogelijk het besef van groepsleden hierover vergroot. Dit leidt er mogelijk toe dat groepsleden op een andere, meer kritische en diepgaande manier gaan discussiëren. Daarnaast leidt dit er mogelijk toe dat groepsleden de manier waarop zij discussiëren en samenwerken meer gaan bespreken. Anderzijds verhoogt de Shared Space ook de rijkheid (media richness, zie Daft & Lengel, 1986) van de VCRI-omgeving, omdat leerlingen inzicht wordt geboden in de mate waarin groepsleden het eens of oneens zijn met hun ideeën en voorstellen.

De effecten van de Shared Space zijn onderzocht door 59 leerlingen in 20 groepen te laten samenwerken in de VCRI-omgeving waarin zij communiceerden via een Chat-tool met Shared Space visualisatie te vergelijken met 58 leerlingen in 20 groepen communiceerden via een Chat-tool zonder Shared Space visualisatie. Ook hier betrof het 5-VWO leerlingen die in kleine groepen aan een Praktische Opdracht voor het vak geschiedenis werkten. De groepstaak was vergelijkbaar met de taak die voor Studie 1 gebruikt is.

De resultaten van deze studie laten zien dat de Shared Space er inderdaad in slaagt om de VCRI-omgeving tot een rijkere omgeving voor communicatie te maken. Leerlingen die beschikten over de Shared Space beoordeelden de VCRI-omgeving als een meer geschikte omgeving voor communicatie en gaven aan dat zij minder communicatieproblemen ervoeren. Analyses van het samenwerkingsproces in de VCRI-omgeving lieten echter niet zien dat groepsleden met de Shared Space vaker de manier waarop zij discussieerden en samenwerkten bespraken. De Shared Space bleek wel op een andere manier het coördinatieproces te vergemakkelijken. Groepen die met de Shared Space werkten, waren minder bezig met uitwisselen van informatie en het bereiken en in stand houden van wederzijds begrip. Dit wijst erop dat het voor leerlingen wellicht makkelijker was om de berichten van hun groepsleden te begrijpen. Met andere woorden, het bereiken van wederzijds begrip en een gedeeld referentiekader (EN: grounding) wordt mogelijk vergemakkelijkt door de Shared Space. In vergelijking met leerlingen die niet met de Shared Space werkten, gaven leerlingen die wel werkten met de Shared Space bovendien aan dat zij de discussies in hun groep als kritisch en constructief ervoeren. In deze groepen was er meer ruimte voor kritische, constructieve discussie in de beleving van de leerlingen. In de beleving van deze leerlingen was de samenwerking in deze groepen ook positiever en effectiever. Ten slotte bleek de Shared Space een beperkte invloed te hebben op de kwaliteit van de groepsproducten. De groepstaak bestond uit drie onderdelen. Analyses wezen uit dat groepen die werkten met de Shared Space beter presteerden op het eerste onderdeel van deze taak.

Studie 3: Effecten van visualisatie van taakgerelateerde aspecten van de samenwerking

In *Hoofdstuk 4* werd de *Graphical Debate-tool* besproken. In tegenstelling tot de Participation-tool en de Shared Space is deze tool vooral ontwikkeld om taakgerelateerde aspecten van het samenwerkingsproces te visualiseren. In deze studie gebruikten leerlingen de Graphical Debate-tool om een representatie van een historisch debat samen op te bouwen. Dit debat had betrekking op de schaal waarop christenen in het Romeinse Rijk vervolgd zijn. In dit debat zijn twee posities te onderscheiden, namelijk de Propaganda positie die stelt dat de schaal van de vervolgingen door de kerk en christenen is overdreven om zo nieuwe gelovigen te trekken, terwijl de Martelaren positie stelt dat christenen daadwerkelijk op grote schaal vervolgd zijn. De Graphical Debate-tool visualiseert hoe goed beide posities ondersteund zijn met argumenten en ondersteunende informatie. Telkens wanneer er een argument of een ondersteunend informatie-element wordt toegevoegd, wordt de betreffende positie dichter naar het centrum van de representatie verschoven. Wanneer er weerleggende of tegensprekende informatie aan een positie wordt toegevoegd, wordt deze positie juist meer naar de buitenkant of periferie van de representatie verschoven. Op deze manier kunnen leerlingen gemakkelijk bepalen welke positie het best door argumenten en informatie ondersteund wordt. Daarnaast kunnen leerlingen in de Graphical Debate-tool gemakkelijk bepalen welke positie de meeste aandacht heeft gekregen en of er vooral geconcentreerd is op een bepaald type informatie (ondersteunende of weerleggende informatie). Op deze manier wordt taakvoortgang gevisualiseerd en inzichtelijk gemaakt (Cox, 1999).

Verondersteld werd dat de Graphical Debate-tool, door taakgerelateerde aspecten van de samenwerking te visualiseren, het samenwerkingsproces effectiever en efficiënter zou laten verlopen. Op de eerste plaats maakt de Graphical Debate-tool complexe informatie inzichtelijker en gemakkelijker te begrijpen. Zonder terugkoppeling over de balans tussen twee posities is het bijvoorbeeld moeilijker vast te stellen welke positie beter door argumenten en ondersteunende informatie wordt onderbouwd. De Graphical Debate-tool vergemakkelijkt op deze wijze mogelijk het trekken van conclusies over het debat. Daarnaast vergemakkelijkt de Graphical Debate-tool wellicht ook de coördinatie van de samenwerking omdat er bijvoorbeeld minder noodzaak is om de taakvoortgang te bewaken en reguleren. Ten slotte zoude tool wellicht ook de discussie en argumentatie over het historische debat tijdens de samenwerking kunnen stimuleren.

Deze veronderstellingen zijn onderzocht door 79 5-VWO leerlingen te laten werken met de Graphical Debate-tool, terwijl 45 leerlingen werkten met een Textual Debate-tool. Deze Textual Debate-tool werkte op dezelfde manier als de Graphical Debate-tool, maar maakte geen gebruik van visualisatietechnieken om leerlingen feedback te geven over taakgerelateerde aspecten van de samenwerking. Bovendien maakte deze tool geen expliciet onderscheid tussen argumenten, ondersteuningen en weerleggingen. In beide condities werkten de groepen aan het samen maken van een representatie over het historische debat op basis van dezelfde historische bronnen. Na het maken van de representatie schreven de groepen in beide condities een essay over hun bevindingen.

De resultaten van deze studie laten zien dat groepen die met de Graphical Debate-tool werkten goed presteerden; over het algemeen presteerden deze groepen beter dan groepen die met de Textual Debate-tool gewerkt hadden. De representaties die zij maakten in hun versie van de Debate-tool waren bijvoorbeeld van significant betere kwaliteit dan de representaties van groepen die werkten met de Textual Debate-tool. Bovendien waren de essays die zij gezamenlijk schreven ook beter, vooral wat betreft de kwaliteit van de argumentatie van de essays. Ten slotte presteerden leerlingen die werkten met de Graphical Debate-tool ook beter op een kennistoets over het onderwerp dan leerlingen die werkten met de Textual Debate-tool. Dit toont aan dat het visualiseren van taakgerelateerde aspecten van de samenwerking de effectiviteit van de samenwerking kan verhogen. Overigens moet hierbij opgemerkt worden dat in beide condities de leerlingen vooruitgang tonen op de kennistoets: in beide condities presteerden de leerlingen na afloop van de samenwerking beter op de kennistoets dan voorafgaan aan de samenwerking.

Daarentegen was er geen bewijs voor de veronderstelling dat de Graphical Debate-tool de coördinatie van de samenwerking vergemakkelijkte. Leerlingen die met deze tool werkten, besteedden net zo veel aandacht aan het plannen, monitoren en evalueren van de taakuitvoering dan leerlingen die met de Textual Debate-tool werkten. Bovendien bleken leerlingen die met de Graphical Debate-tool niet meer te discussiëren en argumenteren over onderwerpen die aan de taak gerelateerd waren. Ten slotte bleek de tool een negatieve invloed te hebben op de percepties van de samenwerking. Leerlingen die met de Graphical Debate-tool werkten ervoeren hun samenwerking als minder positief.

Algemene conclusie en discussie

Na het bespreken van de resultaten van de drie uitgevoerde studies wordt in *Hoofdstuk 5* de algemene onderzoeksvraag beantwoord. Op basis van deze studies kan geconcludeerd worden dat visualisaties inderdaad *de uitkomsten* van computer-ondersteunde samenwerking kunnen beïnvloeden. In Hoofdstuk 4 bleek de Graphical Debate-tool te leiden tot betere groepsprestaties (groepen construeerden betere representaties en schreven betere essays) en individuele leerprestaties. Daarnaast had het gebruik van de Shared Space tijdens de online samenwerking een positieve invloed op de kwaliteit van één onderdeel van de groepstaak. De Participation-tool had echter geen invloed op de groepsprestaties. Wat betreft de percepties van leerlingen van het samenwerkingsproces kan geconcludeerd worden dat de Shared Space inderdaad een positieve invloed had op deze percepties. Dat gold anderzijds niet voor de Graphical Debate-tool. Deze had een negatief effect op de percepties van leerlingen van het samenwerkingsproces. Overigens kan hieruit niet geconcludeerd worden dat de leerlingen die werkten met de Graphical Debate-tool een negatief oordeel over de samenwerking hadden: op een schaal van 1 tot 5 beoordeelden zij de samenwerking gemiddeld met bijna een 4. Door de complexiteit van de Graphical Debate-tool ten opzichte van de Textual Debate-tool ervoeren de leerlingen deze tool mogelijk als een moeilijkere tool voor online samenwerking. Bovendien hebben veel leerlingen vaak de neiging om snel tot consensus te komen zo de samenwerking zo snel en efficiënt mogelijk af te ronden (Erkens et al., 2006; Felton & Kuhn, 2001). De Graphical Debate-tool dwingt leerlingen om juist wél te argumenteren en te bekritiseren, waardoor ze mogelijk negatievere opvattingen over de samenwerking krijgen.

De studies van dit proefschrift geven ook inzicht in *hoe* visualisaties invloed hebben op computer-ondersteunde samenwerking. Allereerst lijkt Studie 1 (Hoofdstuk 2) er op te wijzen dat visualisaties het besef en bewustzijn van groepsleden van het samenwerkingsproces deels kunnen verhogen. Leerlingen die de Participation-tool konden gebruiken, wisten beter wanneer er binnen hun groep meeliftgedrag vertoond werd.

De Shared Space probeerde communicatieproblemen in de elektronische leeromgeving te verminderen door de rijkheid van de omgeving te vergroten. De resultaten uit Hoofdstuk 3 suggereren dat visualisaties gebruikt kunnen worden om de rijkheid van CSCL-omgevingen te vergroten. Leerlingen die de Shared Space gebruikten, ervoeren de omgeving als een betere omgeving voor communicatie en ervoeren minder communicatieproblemen. De analyse van het samenwerkingsproces in de CSCL-omgeving ondersteunt deze bevinding. Leerlingen die met de Shared Space werkten, besteedden minder tijd aan het bereiken en handhaven van een gedeeld referentiekader. Visualisaties kunnen dus gebruikt worden om communicatieproblemen tussen leerlingen te verminderen door de richness van de omgeving te vergroten.

In de Hoofdstukken 3 en 4 werd geprobeerd door visualisaties in de Shared Space en de Graphical Debate-tool de kwaliteit van de groepsdiscussies te verhogen. De resultaten in deze hoofdstukken tonen aan dat dit slechts voor een deel gelukt is. In Studie 2 rapporteerden leerlingen die de Shared Space gebruikten inderdaad frequenter dat hun groepsdiscussies kritisch en constructief waren dan leerlingen die de Shared Space niet gebruikten. De Shared Space visualiseerde een sociaal aspect van de samenwerking, namelijk de mate waarin groepsleden het met elkaar eens waren of met elkaar discussieerden. De Graphical Debate-tool daarentegen visualiseerde een taakgerelateerd aspect van de samenwerking, namelijk de balans tussen posities die onderbouwd worden met argumenten, ondersteuningen en weerleggingen. Daarnaast wijst deze tool leerlingen op het verschil tussen argumenten, ondersteuningen en weerleggingen. Het bleek echter niet dat groepsleden die met deze tool werkten meer over taakinhoudelijke onderwerpen discussieerden dan leerlingen met de Textual Debate-tool. Een verklaring hiervoor kan zijn dat wanneer leerlingen de beschikking hebben over een externe representatie als de Debate-tool, er minder noodzaak is om informatie via de chat te bespreken omdat de representatie zelf al als een externalisatie van deze informatie fungeert (Van Drie et al., 2005). Dit lijkt ondersteund te worden door het gegeven dat groepen die via de Graphical Debate-tool werkten wel betere representaties co-construeerden dan groepen die via de Textual Debate-tool werkten. De "discussies" via de Graphical Debate-tool waren dus van een betere kwaliteit dan de "discussies" via de Textual Debate-tool.

Visualisaties kunnen ook gebruikt worden om de coördinatie van samenwerking te vergemakkelijken. Op basis van de resultaten in Hoofdstuk 2, 3 en 4 lijkt de conclusie gerechtvaardigd dat visualisaties inderdaad coördinatieproblemen kunnen verminderen. In Hoofdstuk 2 bleken leerlingen die toegang hadden tot de Participation-tool in grotere mate participeerden in de groepsdiscussies dan leerlingen die geen toegang hadden tot deze tool. Bovendien bleek de Participation-tool leerlingen te stimuleren om meer te discussiëren over de planning van de samenwerking. Beide bevindingen laten zien dat de Participation-tool invloed had op het coördineren van de samenwerking.

De Participation-tool legde de nadruk op participatie van groepsleden tijdens groepsdiscussie, de Shared Space op andere aspecten van coördinatie, namelijk grounding, onderhandeling en het bereiken van overeenstemming. Door groepsleden de beschikking te geven over de Shared Space werd getracht om het proces van grounding, onderhandeling en het bereiken van overeenstemming te faciliteren. De resultaten uit Hoofdstuk 3 laten zien dat dit deels gerealiseerd is. Groepsleden die met de Shared Space werkten, besteedden minder tijd aan het bereiken en handhaven van een gedeeld referentiekader. Anderzijds bleken deze groepsleden minder bezig te zijn met het uitwisselen van inhoudelijke informatie en het elkaar bevragen over deze informatie.

Ten slotte werd door de Graphical Debate-tool geprobeerd de coördinatie van het samen maken van een externe representatie te faciliteren. We veronderstelden dat het visualiseren van taakgerelateerde aspecten van de samenwerking (i.e., Welke positie is beter onderbouwd met argumenten? Is er vooral aandacht besteed aan ondersteunende informatie en niet aan weerleggende informatie?), het coördineren en reguleren van een dergelijke taak minder nodig zou zijn. Analyses van het samenwerkingsproces in de CSCL-omgeving lieten echter niet zien dat dit ook daadwerkelijk het geval was. Bovendien bleken groepen die met de Graphical Debate-tool werkten, niet meer dan in de andere conditie, te discussiëren en argumenteren over de taakinhoud.

Samengevat benadrukken de resultaten uit Hoofdstukken 2, 3 en 4 de mogelijkheid van visualisaties om het samenwerkingsproces in CSCL-omgevingen te vergemakkelijken. De Shared Space en vooral de Graphical Debate-tool hadden positieve effecten op de kwaliteit van de geproduceerde groepsproducten. Daarnaast blijken visualisaties ook belangrijke effecten te hebben op het proces van samenwerking, met name op het coördinatieproces.

Bij het interpreteren van de resultaten van dit proefschrift dient met enkele beperkingen rekening gehouden te worden. De eerste heeft betrekking op de manier waarop geprobeerd is de invloed van visualisaties op de samenwerking in kaart te brengen. Zo is bijvoorbeeld geprobeerd om het besef van leerlingen van de activiteiten van hun groepsleden te meten via een vragenlijst. Andere meetinstrumenten zijn ook geschikt om dit besef vast te stellen (bijvoorbeeld hardop-denk protocollen of stimulated recall interviews), en hadden zodoende aanvullende of andere informatie opgeleverd. Hetzelfde geldt voor de manier waarop geprobeerd is om in Studie 2 en 3 de invloed van visualisaties op de kwaliteit van de discussie te meten. Het gehanteerde codeerschema had vooral betrekking op het uitwisselen van informatie en het bevragen van groepsleden over deze informatie en geeft daardoor weinig inzicht in de soorten argumenten die groepsleden gebruiken of de inhoudelijke en argumentatieve kwaliteit van de informatie die uitgewisseld wordt. Andere codeerschema's bieden een mogelijkheid om deze gegevens wel te verzamelen (vgl., Clark et al., 2007; Munneke, Andriessen, Kanselaar et al., 2007).

Een tweede beperking van dit onderzoek heeft betrekking op het vakgebied waarin de studies zijn uitgevoerd, namelijk het vak geschiedenis. Dit vakgebied leent zich wellicht minder goed voor het visualiseren van taakgerelateerde aspecten van de samenwerking. In andere vakgebieden, vooral de gebieden in het empirische domein van de bètawetenschappen, is het eenvoudiger en eenduidiger om deze aspecten te visualiseren aangezien er in dit domein duidelijkere criteria zijn voor bijvoorbeeld goede of foute antwoorden. Anderzijds laten de resultaten in Hoofdstuk 4 dat wanneer taakgerelateerde aspecten van de samenwerking gevisualiseerd worden, dit belangrijk effecten op de prestaties van groepen en leerlingen kan hebben. Toekomstig onderzoek zou zich kunnen richten op het verder onderzoeken welke sociale en taakgerelateerde aspecten van samenwerking effectief gevisualiseerd kunnen worden. Daarnaast is het voor het vak geschiedenis vaak de bedoeling dat leerlingen een zo duidelijk en compleet mogelijk beeld schetsen van een historische situatie, fenomeen of discussie. Het is vaak de bedoeling dat zij, op basis van historische bronnen, de argumenten interpreteren, reconstrueren en weergeven en daarnaast dat zij de waarde van deze argumenten beoordelen. Dit is een ander doel dan wanneer leerlingen bijvoorbeeld iemand moeten overtuigen van een bepaald standpunt zoals bij een discussie of een debat. Deze verschillende doelen kunnen ertoe leiden dat leerlingen de taak anders benaderen, op een andere manier argumenteren en daardoor ook andere leeruitkomsten behalen (Munneke-de Vries, 2008; Nussbaum, 2005)

Een laatste mogelijke beperking heeft betrekking op de rol van de docenten in deze studies. Aangezien de docenten in deze onderzoeken een beperkte rol hadden (i.e., zij boden beperkte begeleiding aan de groepen), is de rol van de docent niet onderzocht. Docenten hadden een beperkte begeleidende rol omdat het begeleiden van samenwerking in een elektronische leeromgeving een complex proces is. Onderzoek laat echter ook zien dat docenten een belangrijke rol kunnen vervullen bij het begeleiden van samenwerking. Zij kunnen kritische vragen stellen, misconcepties herkennen en aan de orde stellen, het leerproces ondersteunen, en het verloop van de samenwerking bewaken (Berge, 1995; De Laat et al., 2006; Hmelo-Silver, 2004; Paulsen, 1995; Salmon, 2000). Daarnaast blijken leerlingen vaak het gedrag van docenten te kopiëren of te modelleren (Gillies, 2006; Webb et al., 2006). CSCL-onderzoek zou meer aandacht moeten hebben voor de rol van de docent tijdens het begeleiden van samenwerking via de computer, aangezien de docent aanzienlijke invloed kan hebben op leerlingen. In ons toekomstige onderzoek zullen we daarom de docent meer centraal stellen: Welke problemen ervaren zij wanneer ze groepen begeleiden in een elektronische leeromgeving en hoe kunnen zij ondersteund worden bij dit begeleidingsproces?

Epiloog: Het gebruik van Multilevel analyse bij CSCL-onderzoek

Tot slot is een epiloog aan dit proefschrift toegevoegd waarin het gebruik van multilevel analyse bij diverse analyses van dit proefschrift is toegelicht. CSCL-onderzoekers worden vaak geconfronteerd met complexe datasets: deze datasets hebben vaak een hiërarchische structuur (leerlingen genest binnen groepen) en bevatten vaak variabelen die op verschillende niveaus gemeten zijn (zowel op het groepsniveau als op het niveau van de leerling). Bovendien zijn observaties van de afhankelijke variabelen waarin CSCL-onderzoekers geïnteresseerd zijn vaak niet onafhankelijk van elkaar. Dit komt doordat leerlingen in groepen werken en elkaar op die manier beïnvloeden. Om goed met deze complexe datasets om te kunnen gaan is het gebruik van multilevel analyse aan te bevelen. Door gebruik te maken van de data die in de Hoofdstukken 2, 3 en 4 is verzameld, wordt in deze epiloog gedemonstreerd dat het gebruik van multilevel analyse onderzoekers kan behoeden voor het maken van Type I en Type II fouten en dus voor het trekken van verkeerde conclusies.

List of Publications

Submitted journal articles

Erkens, G., & Janssen, J. (submitted). Automatic coding of online collaboration protocols.

Janssen, J., Erkens, G., Kirschner, P. A., & Kanselaar, G. (submitted). Influence of group member familiarity on online collaborative learning.

Janssen, J., Erkens, G., Kirschner, P. A., & Kanselaar, G. (submitted). Effects of representational guidance during computer-supported collaborative learning.

Janssen, J., Erkens, G., & Schep, A. (submitted). Gender differences in online collaboration styles.

Journal articles, refereed

De Jong, F., Kollöffel, B., Van der Meijden, H., Kleine Staarman, J., & Janssen, J. (2005). Regulative processes in individual, 3D and computer supported cooperative learning contexts. *Computers in Human Behavior, 21*, 645-670.

Janssen, J., Erkens, G., & Kanselaar, G. (2007). Visualization of agreement and discussion processes during computer-supported collaborative learning. *Computers in Human Behavior, 23*, 1105-1125.

Janssen, J., Erkens, G., Kanselaar, G., & Jaspers, J. (2007). Visualization of participation: Does it contribute to successful computer-supported collaborative learning? *Computers & Education, 49*, 1037-1065.

Krol, K., Janssen, J., Veenman, S., & Van der Linden, J. (2004). Effects of a cooperative learning program on the elaborations of students working in dyads. *Educational Research and Evaluation, 10*, 205-237.

Journal articles, non-refereed

Janssen, J. (2005). "Ze werken zelfstandig, lang en geconcentreerd": Ervaringen met samenwerken via computers. *Kleio: Tijdschrift van de vereniging van docenten in geschiedenis en staatsinrichting in Nederland (VGN), 46*(2), 22-25.

Janssen, J., Munneke, L., & Kanselaar, G. (2005). Boekbespreking van: *Design principles for a CSCL environment in teacher training*. *Pedagogische Studiën, 82*, 414-416.

Janssen, J., & Van der Meijden, H. (2002). Een virtuele wereld in de praktijk: Samenwerken in Active Worlds. *Kleio: Tijdschrift van de vereniging van docenten in geschiedenis en staatsinrichting in Nederland (VGN), 43*(6), 7-8.

Book chapters

Veenman, S., Krol, K., Van der Meijden, H., Denessen, E., Janssen, J., Van den Akker, A., et al. (2004). Uitleggedrag van leerlingen in een coöperatieve onderwijsleersituatie. In L. Verhoeven & R. Voeten (Eds.), *Onderwijskunde in theorie en praktijk: Bijdragen aangeboden aan Cor*

Aarnoutse bij zijn afscheid als hoogleraar aan de Katholieke Universiteit Nijmegen (pp. 14-26). Tilburg: Zwijsen.

Conference presentations

Damsa, C. I., Erkens, G., & Janssen, J. (2007). Discourse synthesis: A research study on the use of semantic information for collaborative writing in secondary education. In B. Csapo & C. Csikos (Eds.), *Proceedings of the 12th conference of the European Association for Research on Learning and Instruction (EARLI)* (pp. 713). Budapest, Hungary: Graduate School of Education, University of Szeged, Faculty of Arts.

Erkens, G., & Janssen, J. (2006). Automatic coding of communication in collaboration protocols. In S. A. Barab, K. E. Hay & D. T. Hickey (Eds.), *Proceedings of the 7th International Conference of the Learning Sciences (ICLS)* (Vol. 2, pp. 1063-1064). Mahwah, NJ: Lawrence Erlbaum Associates.

Erkens, G., Janssen, J., Jaspers, J., & Kanselaar, G. (2006). Visualizing participation to facilitate argumentation. In S. A. Barab, K. E. Hay & D. T. Hickey (Eds.), *Proceedings of the 7th International Conference of the Learning Sciences (ICLS)* (Vol. 2, pp. 1095-1096). Mahwah, NJ: Lawrence Erlbaum Associates.

Janssen, J., Erkens, G., Jaspers, J., & Broeken, M. (2005a). Effecten van visualisatie van participatie tijdens computerondertsteund samenwerkend leren (CSCL). In M. Valcke, K. De Cock, D. Gombeir & R. Vanderlinde (Eds.), *Meten en onderwijskundig onderzoek. Proceedings van de 32e Onderwijs Research Dagen 2005* (pp. 560-562). Gent, België: Vakgroep Onderwijskunde - Universiteit Gent.

Janssen, J., Erkens, G., Jaspers, J., & Broeken, M. (2005b). Effects of visualizing participation in computer-supported collaborative learning. In C. P. Constantinou, D. Demetriou, A. Evagorou, M. Evagorou, A. Kofteros, M. Michael, C. Nicolaou, D. Papademetriou & N. Papadouris (Eds.), *11th European Conference for Research on Learning and Instruction: Multiple perspectives on effective learning environments* (pp. 484-485). Nicosia, Cyprus: University of Cyprus.

Janssen, J., Erkens, G., Jaspers, J., & Broeken, M. (2005c). *Effects of visualizing participation in computer-supported collaborative learning.* Paper presented at the first annual ICO research meeting, Utrecht University, The Netherlands.

Janssen, J., Erkens, G., Jaspers, J., & Broeken, M. (2006a). Visualization of agreement and discussion processes during online collaborative learning. In G. Clarebout & J. Elen (Eds.), *Avoiding simplicity, confronting complexity: Advances in studying and designing powerful (computer-based) learning environments* (pp. 175-184). Rotterdam, The Netherlands: Sense Publishers.

Janssen, J., Erkens, G., Jaspers, J., & Broeken, M. (2006b). Visualisatie van constructieve argumentatie in computer-ondersteund samenwerkend leren. In J. Beishuizen, P. Van den Bos, Y. Leeman, M. Lunenberg, B. van Oers, W. van Os, C. Peters, J. van der Schee, K. van der Veer, A. Verkleij,

M. Volman & W. Wardekker (Eds.), *Samen kennis ontwikkelen: Proceedings Onderwijs Research Dagen 2006* (pp. 280-282). Amsterdam: Onderwijscentrum VU.

Janssen, J., Erkens, G., Kirschner, P., & Kanselaar, G. (2007a). Influence of group member familiarity on online collaborative learning. In C. A. Chinn, G. Erkens & S. Puntambekar (Eds.), *Mice, minds, and society: The Computer Supported Collaborative Learning (CSCL) Conference 2007* (Vol. 8, pp. 301-310). New Brunswick, NJ: International Society of the Learning Sciences.

Janssen, J., Erkens, G., Kirschner, P., & Kanselaar, G. (2007b). Online visualization of agreement and discussion during computer-supported collaborative learning. In C. A. Chinn, G. Erkens & S. Puntambekar (Eds.), *Mice, minds, and society: The Computer Supported Collaborative Learning (CSCL) Conference 2007* (Vol. 8, pp. 310-313). New Brunswick, NJ: International Society of the Learning Sciences.

Janssen, J., Erkens, G., & Schep, A. (2006). De invloed van sekse en groepssamenstelling tijdens computer-ondersteund samenwerkend leren. In J. Beishuizen, P. van den Bos, Y. Leeman, M. Lunenberg, B. van Oers, W. van Os, C. Peters, J. van der Schee, K. van der Veer, A. Verkleij, M. Volman & W. Wardekker (Eds.), *Samen kennis ontwikkelen: Proceedings Onderwijs Research Dagen 2006* (pp. 283-285). Amsterdam: Onderwijscentrum VU.

Janssen, J., Krol, K., & Veenman, S. (2003). Student elaborations in cooperative learning dyads: Effects of a cooperative learning program on the interactions of students working in dyads. In L. Mason, S. Andreuzza, B. Arfè & L. Del Favero (Eds.), *10th European conference for research on learning and instruction: Biennial meeting, Padova, Italy, August 26-30, 2003* (pp. 747-748). Padova, Italty: Cooperativa Libraria Editrice Università di Padova.

Janssen, J., Van der Meijden, H., & Ligorio, M. B. (2003). Collaborating in a 3D virtual world for history, culture, and arts. In L. Mason, S. Andreuzza, B. Arfè & L. Del Favero (Eds.), *10th European conference for research on learning and instruction: Biennial meeting, Padova, Italy, August 26-30, 2003* (pp. 310). Padova, Italty: Cooperativa Libraria Editrice Università di Padova.

Janssen, J., Veenman, S., & Van Wensveen, N. (2002b). Leergerichte interacties: Een onderzoek naar de effecten van het scholingsprogramma Coöperatief Leren op de interacties van basisschoolleerlingen. In F. Daems, R. Rymenans & G. Rogiest (Eds.), *Onderwijsonderzoek in Nederland en Vlaanderen: Proceedings van de 29ste Onderwijs Research Dagen 2002 te Antwerpen* (pp. 108-109). Antwerpen, Belgium: Universiteit Antwerpen.

Curriculum Vitae

Jeroen Janssen werd geboren op 7 juni 1977 in Venlo. In 1996 behaalde hij het VWO-diploma aan het Blariacum College te Blerick. In datzelfde jaar begon hij met de studie Pedagogiek aan de Radboud Universiteit Nijmegen (voorheen Katholieke Universiteit Nijmegen). Na het behalen van de propedeuse Pedagogiek, begon hij in 1997 aan de doctoraalstudie Onderwijskunde. Zijn afstudeerscriptie ging over het interactiegedrag van basisschoolleerlingen tijdens coöperatief leren. Na het behalen van zijn bul in 2001 werkte hij als junior onderzoeker bij de sectie Onderwijs & Educatie en het Academisch Centrum van de Faculteit Sociale Wetenschappen. In 2004 begon hij aan zijn promotieonderzoek over computer-ondersteund samenwerkend leren aan de Afdeling Onderwijskunde van de Universiteit Utrecht. Gedurende deze periode was hij ook als docent betrokken bij enkele vakken van de Opleiding Onderwijskunde. Momenteel is hij werkzaam als postdoc onderzoeker en docent bij de Afdeling Onderwijskunde.