Facilitating Online Learning Conversations[1]

[1] Exploring tool affordances in higher education

Jakko van der Pol

Facilitating online learning conversations

Exploring tool affordances in higher education

Jakko van der Pol

IVLOS Series

The IVLOS-series is published by IVLOS Institute of Education of Utrecht University (Instituut voor Lerarenopleiding, Onderwijsontwikkeling en Studievaardigheden). The purpose of this series is the dissemination of results of research to enhance the quality of education.

The members of the editorial board are:
Prof. dr. G. J. Westhoff
Prof. dr. A. Pilot
Prof. dr. J. D. H. M. Vermunt
Prof. dr. R. J. Simons

Recent publications in this series are:
S. F. Akkerman. Strangers in dialogue: Academic collaboration accross organizational boundaries.
I. C. H. Mathijsen. Denken en handelen van docenten.
J. C. de Jong. Uitgesproken complex. Interactie tussen scriptieschrijvers en begeleiders.
M. F. de Laat. Networked Learning.
F. Evelein. Psychologische basisbehoeften van docenten-in-opleiding. Een onderzoek naar het verband tussen de basisbehoeftevervulling van docenten-in-opleiding, hun interpersoonlijk functioneren en de inzet van kernkwaliteiten.
D. Lockhorst. Design principles for a CSCL environment in teacher training.
D. M. L. Verstegen. Iteration in instructional design: An empirical study on the specification of training simulators.
C. Driessen. Analyzing textbook tasks and the professional development of foreign language teachers.
I. van den Berg. Peer assessment in universitair onderwijs. Een onderzoek naar bruikbare ontwerpen.
B. Koster. Lerarenopleiders onder de loep. De ontwikkeling van een beroepsprofiel voor lerarenopleiders en het effect van het kennisnemen daarvan op hun zelf-diagnose.
P. den Brok. Teaching and student outcomes: a study on teachers' thoughts and actions from an interpersonal and a learning activities perspective.

CIP-GEGEVENS KONINKLIJKE BIBLIOTHEEK, DEN HAAG
Van der Pol, J.
Facilitating online learning conversations. Exploring tool affordances in higher education.
Het faciliteren van elektronische leerconversaties. Een onderzoek naar de eigenschappen en effecten van discussiefora in het hoger onderwijs.
Proefschrift Universiteit Utrecht – Met literatuuropgave – Met samenvatting in het Nederlands.
ISBN 978-90-393-45832
Keywords: computer supported collaborative learning, anchored discussion, annotation, collaborative literature processing, peer feedback
Trefwoorden: computerondersteund samenwerkend leren, gecontextualiseerde discussie, annotatie, gezamelijke literatuurverwerking, peer feedback
Print: Budde-Elinkwijk Grafische Producties, Nieuwegein
Cover design: Jakko van der Pol
Copyright Jakko van der Pol, Utrecht, 2007
Language correction: Bill Burki (Chapter 3); Djuddah A.J. Leijen (Chapters 1, 4, 5, and 6)
© 2007 All rights reserved
© 2007 Alle rechten voorbehouden

Facilitating Online Learning Conversations

Exploring tool affordances in higher education.

Het faciliteren van elektronische leerconversaties

Een onderzoek naar de eigenschappen en effecten van discussiefora in het hoger onderwijs.
(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. W. H. Gispen, ingevolge het besluit van het college voor promoties in het openbaar te verdedigen op

dinsdag 3 juli 2007 des middags te 4.15 uur

door

Jakko van der Pol
geboren op 6 april 1976, te Utrecht

Promotor: Prof. dr. P. R. J. Simons
Co-promotor: Dr. W. F. Admiraal

Acknowledgements

It all started on a rainy evening during a philosophical meeting of social scientists in the attic of an old house along a canal in Amsterdam, where I decided to try to write a dissertation. The seed for this decision, however, had already been laid a long time before. For this, I want to thank my teachers at the Educational Science department in Utrecht, who laid the foundations that I have since long profited from. Jerry in particular, for mentoring me during my graduation and sending me to a CSCL conference for the first time. But also the "old gang" of PhD students there: Lisette, Marije, Joris, Arja, Chiel, Janette, Eline, and Maaike for their stimulating example and refreshing walks through the Uithof meadows.

Next, it was the IVLOS that gave me a home and the possibility to realize this PhD ambition. IVLOS not only provided me with a project of my own and the opportunity to deepen my expertise on online learning conversations, but also with a pleasant and fertile environment to grow as a researcher. I foremost want to thank Robert-Jan and Wilfried, for collaboratively giving shape to this project and for guiding it to a successful finish. Wilfried for being a devils advocate, a sparring partner in methodological issues, and for always having an open door and the time to help. Robert-Jan for his confidence, the skill of perspective taking, and constructive guidance.

During this project I have been blessed with the company offered by many enthusiastic colleagues. I want to thank Sanne and Jurjen, for making it a shared endeavor and for coming up with nice plans on alternative work locations. Machteld, Äli, Rosanne, Jacobiene, Annemarieke, Maaike, Mai, Ilya, Marjolijn, Inne, Patricia and all my other IVLOS colleagues, for creating an enjoyable and interesting experience over the years. Mieke, for providing the best R&O talks ever, Perry for offering never ending statistical advice, Willemien for providing APA support, and Ineke for pleasantly and productively conducting a joint research project.

There were also many others outside of the IVLOS to brighten and deepen my experience as a PhD student. The shared understanders PJ, Ingrid, Piet and Sanne and all the guys and girls of the CSCL-analysis club. Kati for inspiring mails about CSCL and of course all my fellow VPO PhD students for their company and our special gatherings over the years.

I furthermore want to thank Wendy Selzer for supplying the Annotation Master and tailoring it to my special wishes. The IVLOS, that has charted new territory in engaging in the further development of the Annotation system and Will and INFI for making this adventure into a concrete success. Off course, an important part of this successful implementation is also due to all the teachers and students that were willing to start using the tool, thereby making this research project possible.

For providing me with a great international experience I want to thank Pierre, for welcoming me at the EPFL in the most beautiful city in Switzerland and ensuring an optimal use of my time. My thank goes to NWO as well for making this trip financially feasible. Also the guys and girls from CRAFT, for taking me in for a few fun and educational months and introducing me to some interesting literature along the way. Especially Mauro, for our enjoyable collaboration and for harassing me as long as was needed for me to start my first blog.

Of course in the Netherlands there were also many friends to keep me grounded and socialized. Licht '95, for cultivating a love for bad movies, scientific discussions and continuously growing stories about the good old days, the "Per & Annemieke gang", for offering the required amount of social gatherings and gaming nights, and Animalus festus for the good parties and non-conversations.

Finally, I want to thank my family. Elke, Willem and Mariana for their support and understanding, Eef and Bill for their extra effort on the layout and on keeping the "asses" out, and off course Max for keeping things in perspective by always offering me glimpses of the real world out there, that I will soon be joining again.

Jakko van der Pol
Utrecht, May 2007

Contents

*"*Collaborative learning is a complex process where a lot can go wrong on the aspects of sharing knowledge, coordination, alignment and responsibilities. It also presents a great demand on teachers investments in time and knowledge. Although it is known that online facilitation can be a great aid to this complex process and can provide an added value, success cannot be taken for granted*."*

(Los, 2006)*

* Los, G.-J. (2006). *1 + 1 is more than 2. Get more out of your education by letting students learn with digital facilitation. [1+1 is meer dan 2. Haal meer uit uw onderwijs door studenten samenwerkend te laten leren met digitale ondersteuning.]* Retrieved March, 22, 2007 from http://www.onderwijscentrum.vu.nl/ICT_en_O/index.cfm/home_subsection.cfm/subsectionid/2D634DAB-1279-D040-8BB8C52C2E708E06.

Chapter 1

Introduction*

Imagine a teacher in higher education: A teacher with a progressive state of mind and an enthusiasm for social constructivist theories who is continuously trying to improve the learning of his or her students, but is nevertheless still struggling to create successful and sustainable social constructivist teaching practices in his or her everyday life. Especially when trying to fulfill the promise of Computer Supported Collaborative Learning (CSCL), this teacher may have trouble realizing an added value and depth to students' learning processes and to maintain students' initial levels of motivation and engagement after a fading first wave of technology-related enthusiasm. The needs and problems of these - real or imaginary - teachers form the inspiration of the academic exercise we will report about in this dissertation, setting its standards and determining its direction in several ways.

First, this research project regards the use of CSCL as a means to reach a certain goal, studying CSCL with good, deep, or effective learning processes of students in mind. As indicated by Säljö (2003), introducing technologies in institutionalized forms of learning has not (always) lived up to the initial expectations in reaching these goals. In response to this observation, Säljö proposes to abandon the question whether new technologies can improve learning altogether, because learning does not become better or more efficient, just different. However, we think it is still possible and worthwhile to investigate if it is possible to enhance learning through CSCL and how this can be done. Not only do we think that educational research is not neutral when it comes to learning, we feel it also has a certain responsibility. As the positive expectations about the advantages of CSCL for educational practice have partly led to its widespread introduction, we think educational science should also investigate if and how these expectations can be realized. We continue to

---

* The introduction of this dissertation is partly based on a paper written by Cherubini, Van der Pol, and Dillenbourg (2005, July), a collaboration which was made possible by funding from the Dutch organization for scientific research NWO.

believe in the potential of CSCL for student learning, albeit depending on many related variables and conditions. We therefore propose to direct research towards studying how to shape and develop the different (specialized) uses of CSCL in such a way that their learning potential can be better realized.

Second, seeing CSCL as a means and not as a goal in itself, we will be aiming for something more than the presence of social interaction, knowledge sharing, or community building. Andriessen, Baker, and Suthers (2003) distinguish between learning from a certain dialogue and learning to conduct that dialogue. Although the second learning goal will usually need to be reached to a certain degree as a necessary condition for reaching the first, it alone may not be sufficient. In order to justify the investments in students' and teachers' time and the efforts that are required, we feel that it is important that students' efforts are invested as much as possible in activities that lead to and deepen learning, in relation to their content-related learning goals. Just as one can distinguish different kinds of loads on students' cognitive resources that have different, if not contradictory, effects on their learning processes on an individual level (Paas & Van Merriënboer, 1994), it is also possible to distinguish different effort-demanding processes in collaborative learning. Parallel to the aim of decreasing the extrinsic and optimizing the germane cognitive load, we can transfer this thought to collaborative learning situations (see Dillenbourg & Bétrancourt, 2006) and strive for the same optimal allocation of students' efforts. The aim then is to let students invest as much of the available effort as possible into the production of rich interactions, resulting in an "optimal collaborative load".

Third, our concerns for educational practice have led to the choice for a form and area of CSCL where we expected to achieve the greatest gains. The area we will primarily focus on, is students' processing of academic texts, using *online asynchronous discussion* in blended learning situations. As stated by Lapadat (2002) asynchronous online discussion is particularly suitable for the collaborative construction of meaning and presents a great potential for conceptual change. We will argue that online discussion is especially useful for the appropriation of abstract knowledge, which can still be seen as an essential activity in higher education and as forming the basic ingredient for the development of many required competences. The medium presents students with an open learning environment that allows interaction on a conceptual level. This "discursive" interaction offers the possibility for education to

connect to the way students (initially) make sense of the material (Laurillard, 1993), while at the same time interactively stimulating them to expand this understanding to a higher level (Petraglia, 1998). By offering the potential for students to engage in an active, thoughtful and personally meaningful processing of content, it can facilitate the appropriation of new ideas and the deepening of existing understanding in a self-discovering way. Thus, the kind of online discussion we wish to realize is not so much a critical and opinion-centered debate, but a more constructive conversation that is aimed at the processing of content and a deepening of understanding. We will refer to this kind of discussion as "learning conversations", described by Bellamy (1997) as conversations that lead to learning because they allow participants to make connections between previously unrelated ideas, see old ideas in a new way, and lead to conceptual change. However, this great potential also comes with a great challenge and the described reflective learning conversations can also be one of the most difficult to realize. Being in a learning situation means that students are not only expected to share the knowledge they already have, but also collaboratively construct new knowledge and understanding. Collaboratively developing new meaning and creating new understanding that surpasses that of the individual participants, may be paradoxical in itself, and will be difficult and complex to say the least. As such, the development of semantic understanding presents a fragile process that can only progress in a gradual way (Alargamot & Andriessen, 2002). This challenge is further increased by the fact that asynchronous online learning conversations can also be hindered by certain medium-related constraints, such as its limited degree of interactivity and its natural tendency to diverge.

Fourth, our approach to tackle these challenges will be to investigate and manipulate the characteristics of different tools that can be used. Like Kirschner, Sweller, and Clark (2006), we believe students do need to be guided in creating successful constructivist and discovery-based learning practices, especially when they only possess low levels of prior knowledge and experience. However, we will not investigate how to provide this guidance by means of direct instruction, but how to do this by using the "affordances", or *"aspects of an artifact's design that suggests how it is to be used"* (Johnson, 1995, p. 219), of different tools. Although we will make use of the vast amount of knowledge that already exists on how to successfully create and implement online learning conversations (see for instance Pilkington, 2004; Salmon, 2002; Van der Pol &

Admiraal, 2003), the research in this dissertation will focus on how the affordances of different tools can be of further use. As stated by Pilkington, *"It is also important to meet students 'where they are at' and not where we want or expect them to be if we are to effectively facilitate online discussion"* (p. 163). In our case this means that, as several ways to create a sufficient quantity of responses already exist, we will direct our attention to further enhancing their content and quality. This focus on affordances of tools means we will investigate how the functional characteristics of a tool can facilitate *and stimulate* the production of rich and constructive interactions. Providing training and instruction presents an additional burden on a teacher's available time and, while known to be effective, this effect may be temporarily and fade over time. In addition, an advantage of using the affordances of a tool to facilitate students' interaction is that it does not directly steer or constrain students' actions too explicitly or strongly. As expressed by Innes (2004), using criteria to direct students' collaboration too explicitly is not desirable, as *"they sometimes shape the students' behavior without expanding their understanding"* (p. 247). It has been stated (Dillenbourg, 2002) and demonstrated (Beers, 2005) that constraining students' collaboration too much by "overscripting" it, can have disruptive effects.

Resulting from these four incentives, we can formulate the following problem definition that will be the central focus of this dissertation: How can we use the affordances of tools to facilitate students' online learning conversations? While being grounded in general social constructivist theory, this project will use a mix of communication theories and other existing CSCL approaches in the development of tools and in investigating students' learning conversations.

In the next sections, we will examine the notion of *online learning conversations* in more detail to understand better why they can be difficult to create in academic education. First, we describe how people in conversation aim to reach mutual understanding based upon the available common ground and grounding. Next, we will identify the "effort after shared understanding" (Schwartz & Lin, 2000), or "grounding" (Clark & Brennan, 1991) as an essential aspect of learning conversations. Especially in a university setting grounding-for-learning, or "semantic grounding", (see Baker, Hansen, Joiner, and Traum, 1999) is suggested to differ from grounding as found in regular, everyday conversations. Then, we will discuss the additional affordances and constraints of conducting learning conversations *online*.

Conducting conversations and the role of common ground

Before exploring the specific characteristics of conversation for learning and conversation online, it is useful to look at some of the basic mechanisms of conversation in general.

Conversation can be seen as a specific form of communication, in that it is generally interactive, for a large part language based, and conducted by a limited number of participants. Regarding the latter, this paragraph will mainly cover dyadic conversations, as this is often the focus of traditional face-to-face conversation theories which we will start with. In most kinds of conversation, the goal usually lies in reaching a certain degree of mutual understanding, meaning that both partners understand to a certain degree what the other means by his or her utterances, in order to communicate effectively. Two main elements that play a role in achieving this, are *grounding* and *common ground* (see Clark & Brennan, 1991; Clark, 1992). Roughly, grounding can be described as an active process which takes place during the conversation itself, whereas common ground can be seen more as presenting an available resource on which the conversation is based. As explained by Clark, the notion of common ground lets speakers use the information they have about the knowledge, ideas, intentions, or (shared) history of one's communication partner to omit redundant information from their utterances and communicate more efficiently. In instances where relying on the available common ground is not sufficient, people can resort to the use of grounding activities. In a broad sense, grounding can be identified as anything a person does to increase the degree of (mutual) understanding. In a more limited sense, grounding can be identified as certain explicit actions one can undertake to reach this goal, such as checking, acknowledging, or repairing (see for instance Poesio & Traum, 1997).

In sum, we interpret common ground as something that conversation partners start with and base their grounding efforts on, in order to achieve a certain degree of mutual understanding. Using the terms offered by Akkerman et al. (in press), we view common ground as an "overlap" in the range of meaning that is available for the different partners in a certain conversational situation, and mutual understanding as a degree of "similarity" in the meaning that these partners attribute to specific utterances during the conversation itself. To clarify this interpretation, we have schematically depicted it in Figure 1 on the next page.

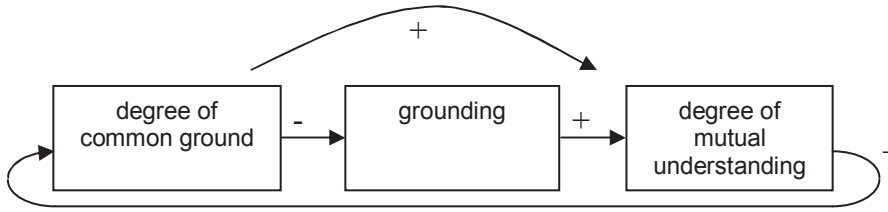

Figure 1. Schematic view of the conceptual placement of common ground, grounding and mutual understanding and the relations between these notions.

In Figure 1, arrows indicate the respective relations between common ground, grounding and mutual understanding[1]. As indicated by the positive arrow from mutual understanding back to common ground, it is possible for the mutual understanding to become part of the common ground of certain participants for the future conversation. Another self-reinforcing relation, (indicated by the negative arrow from common ground to grounding), is formed by the functional nature of grounding, which means that the less common ground there is to start with, the more grounding activities will occur (as, for instance, found by Van der Pol, 2002). Finally, the success and ease of grounding itself also depends on the available common ground, as indicated by the positive arrow at the top of Figure 1. Kraus and Fussell (1991) express this reciprocal relation between grounding and common ground as follows: *"It is hard to find some if you don't have some already and you don't have any unless you find it"* (p. 4). Thus, although it is possible to (partially) correct and prevent misunderstanding by performing grounding activities, the available amount of common ground also plays an important role in determining the efficiency of communication. Having explained our basic model of common ground, grounding and mutual understanding, two additional comments should be made regarding the fact that common ground is not a fail-proof resource that simply accumulates and leads directly to mutual understanding, and regarding the context-sensitive nature of common ground.

[1] In our description of conversations so far, we have not mentioned the concepts of shared understanding and group cognition. We see both terms as indicating even more than mutual understanding, in the sense that they may only belong to a situation or group as a whole. In addition, as indicated by Akkerman, Beers, Van den Bossche, Van der Pol, and Mulder (2003), they may also imply a certain degree of agreement.

First, the fuzzy and ambiguous nature of common ground makes using common ground less straightforward than it may seem. While in Clark's (1992) reasoning piece information X is either "known" or "unknown" to person A or B, there can also be many stages in between, and many different ways of knowing piece of information X. Presenting an alternative perspective on common ground, Sperber and Wilson (1995) view it rather as a "mutual cognitive environment", which they define as the overlap in the participants' *manifest* meaning. Thus, we can identify a person's manifest meaning as the range of possible meaning that is evoked or triggered by the presented evidence, in a particular situation. As stated by Croft and Cruse (2004), this body of conceptual content (also known as the purport), does not correspond to any specific interpretation but presents wide range of meaning which, as demonstrated by Billig (1988), can even still contain contradictory meanings as well. Viewing common ground as an overlap in people's manifest meaning means that a high degree of common ground does not ensure people will attribute the same meaning to a certain utterance, but that they will be more likely to. This perspective on common ground is also more in line with the "mutual knowledge paradox" as described by Clark & Marshall (1981). As someone's knowledge, intentions or experiences can never be accessed directly and can only be inferred from one's communicative interactions (Draper & Anderson, 1991), mutual understanding can never be reached completely but only to a certain *degree*.

Second, because a person's range of triggered manifest meaning highly depends on his/her perception of the collaborative situation, so does the amount of common ground of multiple conversation partners. As emphasized by Bereiter (2002), the individual process of meaning-making is intrinsically connected to a particular context and cannot be seen in isolation. Because we have defined common ground as overlap in individual's range of manifest meaning, this degree of overlap is strongly influenced by the strength or the focusing effect of the conversational context. As expressed by Sperber and Wilson, *"a mismatch between the context envisaged by the speaker and the one actually used by the hearer may result in a misunderstanding"* (1995, p. 16). Consequently, because participants' perceptions of their conversational context are not stable but dynamic, so is the available common ground (see Koschmann & Le Baron, 2003). As stated by Baker et al. (1999), participants' perceptions of the communicative situation (such as the collaborative goal and the medium that is

used) also influence their grounding process. In the next two paragraphs, we will further investigate the nature of conversation with a *learning* goal and of conversations which are conducted *online*.

The special nature of learning conversations

In conversation for learning, one of the main goals is to engage students in the collaborative processing of meaning in order to deepen their levels of understanding. Therefore, as expressed by Schwartz and Lin (2000), the process of grounding itself, also described as the "effort after shared understanding", plays a central role. However, three additional remarks to this observation should be made: First, although it is the activity of grounding itself that is closely related to learning, not *all* or *any* kind of grounding will automatically produce the desired results. Second, the kind of grounding that is most strongly associated with learning cannot be expected to be as self-regulatory and taken for granted as the grounding-for-conversation we know from conversation in everyday life. Third, in learning situations where participants have only a limited understanding of the content they are talking about, successful semantic grounding becomes even more difficult.

Our first remark, that not any kind of grounding will suffice, applies especially to the setting and the goal that we focus on in this research project. In university education, learning conversations will often concern abstract knowledge which cannot be experienced directly but only through the descriptions of others, making learning essentially a mediated phenomenon (Laurillard, 1993). Especially when the goal lies in collaboratively developing a better understanding of abstract subject matter, it is mainly conceptual or semantic grounding (see Baker et al., 1999), that is most strongly associated with learning. Besides the semantic level of grounding, there also exists a pragmatic level of grounding, which is more directly concerned with recognizing other students' communicative intentions and with ensuring the continuation of the collaboration. As stated by Baker et al., "*Semantic level grounding however, relates to attaining mutual understanding of what is meant by certain terms and expressions; it thus relates more closely to learning in a specific knowledge domain by means of interpersonal interaction"* (1999, p. 42).

Regarding the second remark, we think that students' semantic grounding processes may often not progress as far as we would want them to, as this may not be necessary to continue the conversation at a seemingly acceptable level.

Being a primal mechanism in everyday communication, the process grounding is driven by a principle of "least collaborative effort" (Clark & Wilkes-Gibbs, 1986), meaning that participants will limit the efforts they invest in grounding activities to what is sufficient to continue the conversation. Similarly, Sperber and Wilson's (1995) relevance-theoretic comprehension procedure indicates that a hearer will also follow a path of least effort in deriving the meaning and implications of a certain utterance. Sperber and Wilson also mention that *"a communicative intention can be fulfilled without the corresponding informative intention being fulfilled"* (p. 30). In conversations in general, a high degree of semantic grounding may not always be necessary to continue the conversation (which can, for example, be the case in a competitive debate). Kuhn and Goh (2005, p. 347), state that one of the mistakes teachers make when conducting classroom discussions is

> "… to allow the activity to relapse into nothing but consecutive self-expression, first on the part of one student, then another. It does not matter much what each student says, and no student need listen to another. In this worst-case scenario, the only attention the next student pays to the speaker is to wait to observe a signal that this speaker is about to finish, so that he or she can begin. As long as everyone gets their share of turns to speak and no one speaks too long, there is a wealth of opportunity for self-expression. Yet, no further purpose is fulfilled."

Similarly, Schwarz and Glassner (2003) conclude that people are generally highly skilful in challenging, counterchallenging, justifying or agreeing during everyday conversation, but that the arguments are often mediocre according to analytical criteria. In sum, what is "sufficient to continue the conversation", might not be the same as what is "sufficient for learning" (Baker et al., 1999), and instead of striving for a *minimal* collaborative effort, participants in learning conversations should strive for an *optimal* collaborative effort (see Dillenbourg, Traum, & Schneider, 1996). Finally, even when people do aim for a high degree of semantic grounding and mutual knowledge, they may succumb to the "illusion of shared knowledge", which, as shown by Ross, Greene and House (1977), means that people are likely to overestimate their amount of mutual knowledge.

Our third remark on the additional difficulty of semantic grounding for learners refers to the influence of students' understanding of the subject matter on the effectiveness of their communication. As identified by Innes (2004), language has a double function in dialogic collaborative learning, as both a means to conduct disciplinary discourse and as the goal of increasing students' conceptual understanding. According to him, these two functions of dialogue *"are enmeshed because students' inability to communicate with each other impacts the inquiry process and consequently affects their ability to acquire useful knowledge"* (p. 192). Or, in the words of Baker et al.: "Learning from grounding - collaborative learning - can thus be viewed as appropriation of semiotic tools, mediated by those very tools" (1999, p. 12). Thus, similarly to the reciprocal relation between grounding and common ground that we described in the previous paragraph, the ability of students to construct knowledge together also largely depends on the amount of prior knowledge they have available. Especially in learning situations where students are still wrestling with the subject matter, it may become more difficult for them to properly understand others and explain themselves, which in turn affects the collaborative deepening of understanding.

Conducting learning conversations online

As we have described in the previous section, the effectiveness of students' learning conversations cannot always be taken for granted, as a high degree of semantic grounding may not be required to continue the conversation and even if this is the case, the process may be hindered by students' low levels of prior knowledge. In this section, we will show that conducting learning conversations online (thereby focusing on the use of asynchronous discussion boards), can present some additional problems.

First, compared to communication face to face, the distance and freedom that is associated with online forums seem to present users with a reduced expectancy of conversational coherence and a preference for expressing and exchanging existing opinions and experiences. As such, online discussion often can be seen to resemble a sequence of monologues rather than a true dialogue (Fay, Garrod, & Carletta, 2000). Although, as expressed by Herring (1999), this reduced interactional coherence is usually not a serious impediment to users' recreational use of computer-mediated communication (CMC), it can hinder them in conducting effective online learning conversations. As indicated by Andriessen (2006), more constructive forms of argumentation have more to

offer to education than the aggressive forms that can be found in talk shows or the political sphere. Elaborating on this distinction, we think that a further subdivision can be made. When students interact in a constructive and collaborative way, two different kinds of argumentation can emerge. First, students may argue about the value of certain statements in a text and second, they may argue about the meaning they attribute to it. As a result of our focus on students' collaborative processing of text, we have a special interest in the second, meaning-oriented kind. This preference is also shared by Klemm (1998), who expresses it in the following way:

> "Don't settle for just opinions. Everybody has opinions. They are like knee jerk reflexes, occurring with little thought once they have been formed. Thus, it is not surprising that many classroom discussion groups online are dominated by opinion messages, rather than rigorous analysis and creative thought." (p. 3)

Second, the reciprocal relationship between grounding and common ground and the double function of language become even more emphasized when conducting learning conversations online. The context-reduced nature of communicating at a distance (this aspect will be elaborated on in Chapter 2) may leave students with a greater range of manifest meaning and presenting a less focused area of common ground. Thus, online conversations may leave more room for miscommunication than communication face-to-face. In addition, the reduced level of interactivity and absence of non-verbal communication (see Chapter 3 for a more elaborate description) make miscommunication or misalignment both harder to detect and to repair. While we have seen that people will usually invest as little effort as necessary to continue the conversation and repair any miscommunications *after* they arrive, asynchronous conversation places more emphasis on preventive strategies, like perspective taking (Järvelä & Häkkinen, 1999) and audience design (Clark, 1992). As for these processes the existence of a largely shared frame of reference is an important prerequisite, asynchronous conversations depend less on grounding and more on common ground. A complicating factor in collaborative learning conversations is that these processes of audience design and perspective taking become more complex in "many-to-many"

communication as it then becomes more complex to decide who to design for, or whose perspective to take.

The following example (see Figure 2) shows two messages that were collected in a study of using asynchronous discussion groups for the collaborative processing of academic texts (Van der Pol, 2002). The data collected in this study often showed messages that in form resembled successful online discussions with successive messages containing new knowledge and argumentations. Although on a pragmatic level students´ actions were sufficiently grounded, their mutual understanding on a semantic level, however, was sufficient to ensure a successful collaborative construction of knowledge. In the given example we see that is can be very difficult to determine how the different messages relate to one another and what the logical implications of a reaction is for the messages that precede it, which makes it difficult for students to build upon each others' contributions towards a deeper understanding of the subject matter.

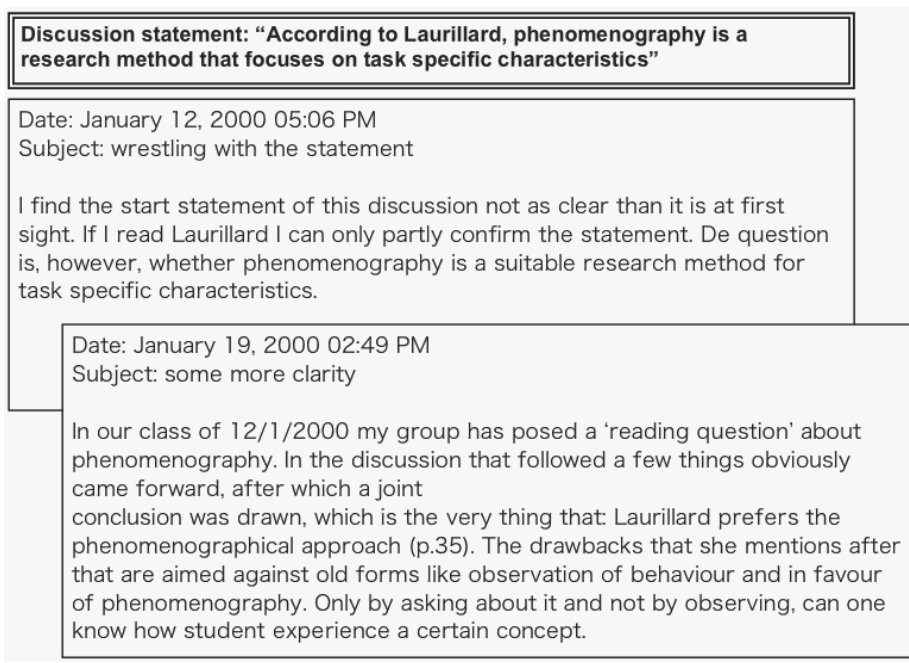

Figure 2. Sample online learning conversation between university students.

When we look more closely at the second message, we see that it aims to provide "some more clarity" in response to the first message. However, this

reply presents a gradual shift in topic as it discusses the question whether Laurillard is in favor of phenomenography or not, instead of the question whether, according to Laurillard, "phenomenography focuses on task specific characteristics". Although both messages are related, the subtle change in topic of the second one means it presents no direct implications to the first, which hinders the drawing of a new conclusion. This short fragment makes clear that students not only have to guess for the meaning of Laurillard's book (see Flower, 1990 on text comprehension), but also for the meaning of the discussion statement and for the meaning of each other's messages. It also illustrates that collaboratively developing semantic understanding is a difficult and, as mentioned by Alargamot and Andriessen (2002), a gradual process.

In the last three sections we have described conversations in general, conversations for learning, and conversations online. We have identified several obstacles for conducting effective online learning conversations, such as the fact that the less common ground and prior knowledge students have to start with, the more difficult the process of conducting online learning conversations becomes. These obstacles underlined the need for facilitation, as expressed in the central problem definition of this dissertation: How can we use the affordances of tools to facilitate students' online learning conversations?, and identified the degree in which these conversations are meaning-oriented and interactionally coherent as important elements. The remainder of this dissertation will present several studies that attempt to facilitate and investigate these elements in the challenging process of creating successful online learning conversations.

Overview of the dissertation

This dissertation consists of several theoretical and empirical chapters, which relate to the problem that have been identified in the preceding sections. These chapters will build upon one another in the following way.

In Chapter 2 we explore how the distance of online communication affects students' collaboration by offering decreased levels of co-intentionality and co-reference. From this, we launch the idea of enhancing the context of online communication to support students' collaboration by *anchoring* it in their study material. Not only could this provide a natural focus of students' collaborative intentions, it also may prevent miscommunication as *"the most effective way of clarifying the meaning of messages is to relate them to a shared context"* (Riva,

2001, p. 217 in George & Labas, in press). Having identified several potential benefits of anchoring online learning conversations, this chapter describes the preliminary results of an experimental study that compares anchored discussion with the more commonly used threaded discussion for the collaborative processing of academic texts.

In Chapter 3, we examine the asynchronous nature of online discussion and its effects on online learning conversations. From this perspective, we see that a limited degree of interactivity and the increased emphasis on audience design and interpretation again point to several potential advantages of anchored discussion. In this chapter, the data from the experimental study reported on in the previous chapter will be analyzed in more detail. Empirical evidence is collected on how anchoring online learning conversations influences the degree to which students are engaged in processing the meaning of the text (as a form of co-intentionality), the efficiency of communication (by, among other things, the existence of a shared frame of reference), and the presence of several "constructive activities" (see Veerman, 2000).

Because anchoring online learning conversations proves indeed to hold several benefits, the next study, that is described in Chapter 4, continues to the use anchored discussion to facilitate students online learning conversations. In this second study, we attempt to further enhance the quality of students' online learning conversations by shifting our focus towards the quality of the responses that students receive. We will look at the relations between individual messages and create a coding scheme for the *local relevance* of replies. In this experimental design, we study the use of a system for anchored discussion with a functionality for peer-evaluation that aims to increase the relevance of students' replies.

Chapter 5 further investigates the characteristics of the feedback that students provide each other. This time, we study the use of online peer feedback. By tracking the revisions in students' products and relating these to the received feedback, we will examine how the feedback is received and when it is used by the receiver to revise his or her writing product. The data in this study have been collected in two different educational contexts, using three different tools: the discussion board in Blackboard, the anchoring tool as described in Chapters 3 and 4, and a virtual learning environment with a specific functionality for peer feedback.

Finally, in Chapter 6, we will reflect upon the theory and the concepts that were used, the methodology of this research project, the implemented tools, the empirical results, and the implications for educational research and practice.

References

Akkerman, S., Van den Bossche, P., Admiraal, W., Gijselaers, W., Segers, M., Simons, R. J., & Kirschner, P. (in press). Reconsidering group cognition: From conceptual confusion to a boundary area between cognitive and socio-cultural perspectives? *Educational Research Review*.

Akkerman, S., Beers, P. J., Van den Bossche, P., Van der Pol, J., & Mulder, I. (2003, May). *Het bereiken van een gemeenschappelijke kennisbasis in teams: een zoektocht naar een gemeenschappelijk kader. [Reaching shared knowledge in teams: A search for common ground]* Paper presented at the Onderwijs Research Dagen (ORD 2003), Heerlen, The Netherlands.

Alamargot, D., & Andriessen, J. (2002). The "power" of text production activity in collaborative modelling: Nine recommendations to make a computer supported situation work. In M. Baker, P. Brna, K. Stenning, & A. Tiberghien (Eds.), *The role of communication in learning to model* (pp. 275-302). Mahwah, NJ: Lawrence Erlbaum Associates.

Andriessen, J. (2006). Arguing to learn. In: K. Sawyer (Ed.), *Handbook of the learning sciences* (pp. 443-459). Cambridge, MA: Cambridge University Press.

Andriessen, J., Baker, M., & Suthers, D. (2003). Argumentation, computer support, and the educational context of confronting cognitions. In J. Andriessen, M. Baker, & D. Suthers (Eds.), *Arguing to learn: Confronting cognitions in Computer-Supported Collaborative Learning environments* (pp. 1-25). Dordrecht: Kluwer Academic Publishers.

Baker, M., Hansen, T., Joiner, R., & Traum, D. (1999). The role of grounding in collaborative learning tasks. In P. Dillenbourgh (Ed.), *Collaborative learning: Cognitive and computational approaches* (pp. 31-63). Amsterdam: Pergamon/ Elsevier Sience.

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

Bellamy, R. (1997). *Support for learning conversations.* Abstract retrieved November 11, 2002, from http://kmi.open.ac.uk/seminars/abstract.cfm?id=10.

Bereiter, C. (2002). *Education and mind in the knowledge age*. Mahwah, NJ: Lawrence Erlbaum Associates.

Billig, M. (1988). *Ideological dillema's; a social psychology of everyday thinking*. London: Sage Publications.

Cherubini, M., Van der Pol, J., & Dillenbourg, P. (2005, July). Grounding is not shared understanding: Distinguishing grounding at an utterance and knowledge level. Paper presented at *CONTEXT'05, the Fifth International and Interdisciplinary Conference on Modelling and Using Context*, Paris.

Clark, H. H. (1992). *Arenas of language use*. Chicago: University of Chicago Press.

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

Clark, H. H., & Marshall, C. R. (1981). Definite reference and mutual knowledge. In A. K. Joshi, B. L. Webber, & I. A. Sag (Eds.), *Elements of discourse understanding* (pp. 10-63). New York: Cambridge University Press.

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

Croft, W., & Cruse, D.A. (2004). *Cognitive Linguistics*. Cambridge: Cambridge University Press.

Dillenbourg, P. (2002). Over-scripting CSCL: The risks 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 University of the Netherlands.

Dillenbourg, P., & Bétrancourt, M. (2006). Collaboration load. In J. Elen, & R. E. Clark (Eds.), *Handling complexity in learning environments: Research and theory*. Amsterdam: Elsevier.

Dillenbourg, P., Traum, D., & Schneider, D. (1996) Grounding in multi-modal task-oriented collaboration. In P. Brna, A. Paiva, & J. Self (Eds.), *Proceedings of the European Conference on Artificial Intelligence in Education* (pp. 401-407). Lisbon, Portugal: Fundacao Calouste Gulbenkian.

Draper, S. W., & Anderson, A. (1991). The significance of dialogue in learning and observing learning. *Computers & Education, 17*(1), 93-107.

Fay N., Garrod, S. C., & Carletta J. (2000). Group discussion as interactive dialogue or as serial monologue: The influence of group size. *Psychological Science, 11*(6), 481–486.

Flower, L. (1990). Negotiating academic discourse. In L. Flower (Ed.), *Reading to write: Exploring a cognitive and social process*. Oxford, UK: Oxford University Press.

George, S., & Labas, H. (in press). E-learning standards as a basis for contextual forums design. *Computers in Human Behavior*.

Herring, S. (1999, June). Interactional coherence in CMC. *Journal of Computer-Mediated Communication 4*(4). Retrieved March 4, 2004, from http://jcmc.indiana.edu/vol4/issue4/herring.html

Innes, R. B. (2004). *Reconstructing undergraduate education: Using learning science to design effective courses*. Mahway, NJ: Lawrence Erlbaum Associates.

Järvelä, S., & Häkkinen, P. (1999, March). *Web based cases in teaching and learning - reciprocal understanding and perspective taking in conversation*. Paper presented at the conference for Computer Assisted Learning, London.

Johnson, J. A. (1995, June). *A comparison of user interfaces for panning on a touch-controlled display*. Paper presented at the Conference on Human Factors in Computing Systems, Denver, CO.

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*(2), 75–86.

Klemm, W. R. (1998). Eight ways to get students more engaged in online conferences. *T.H.E. Journal*, August. Retrieved, March 27, 2007, from http://www.thejournal.com/articles/14054.

Koschmann, T., & Le Baron, C. (2003). Reconsidering common ground: Examining Clark's contribution theory in the OR. In K. Kuutti, E. Karsten, G. Fitzpatrick, P. Dourish, & K. Schmidt (Eds.), *ECSCW 2003: Proc. Eighth European Conference on Computer-Supported Cooperative Work*. Amsterdam: Kluwer.

Krauss, R. M., & Fussell, S. R. (1990). Mutual knowledge and communicative effectiveness. In J. Galegher, R. E. Kraut, C. Egido (Eds.), *Intellectual teamwork: Social and technological foundations of cooperative work*. Hillsdale, NJ: Erlbaum.

Kuhn, D., & Goh, W. (2005). Arguing on the computer. In: T. Koschmann, D. Suthers & T. W. Chan (Eds.), *Proceedings of the 2005 conference on Computer support for collaborative learning: the next 10 years!* (pp. 346 - 352). Mahwah, NJ: Lawrence Erlbaum Associates.

Lapadat, J. C. (2002). Written Interaction: A Key Component in Online Learning. *Journal of Computer Mediated Communication (JCMC), 7*(4). Retrieved April 6, 2007, from http://jcmc.indiana.edu/vol7/issue4/lapadat.html.

Laurillard, D. (1993). *Rethinking university teaching*. London: Routeledge.

Paas, F., & Van Merriënboer, J. J. G. (1994). Instructional control of cognitive load in the training of complex cognitive tasks. *Educational Psychology Review, 6*, 51-71.

Petraglia, J. (1998). *Reality by design: The rhetoric and technology of authenticity in education*. Mahwah, NJ: Lawrence Erlbaum Associates.

Pilkington, R. (2004). Developing discussion for learning. *Journal of Computer Assisted Learning, 20*(3), 161-164.

Poesio, M., & Traum, D. (1997, November). *Representing conversation acts in a unified semantic/pragmatic framework*. Paper presented at Working Notes: AAAI Fall Symposium on Communicative Action in Humans and Machines, Cambridge, MA.

Riva, G. (2001). Communicating in CMC: making order out of miscommunication. In L. Anolli et al. (Eds.), *Say not to say: new perspectives on miscommunication* (pp. 203–233). Amsterdam: IOS Press.

Ross, L., Greene, D., & House, P. (1977). The false consensus phenomenon: An attributional bias in self-perception and social perception processes. *Journal of Experimental Social Psychology*, *13*, 279-301.

Säljö, R. (2003). Representational tools and the transformation of learning. In B. Wasson, S. Ludvigsen & U. Hoppe (Eds.), *Designing for change in networked learning environments: Proceedings of the international conference on computer support for collaborative learning*. Dordrecht: Kluwer Academic Publishers.

Salmon, G. (2002). *E-tivities: The key to active online learning*. London: Kogan Page.

Schwartz, D. L., & Lin, X. D. (2000). Computers, productive agency, and the effort after shared meaning. *Journal of Computing in Higher Education, 12*(2), 3-33.

Schwarz, B. B., & Glassner, A. (2003). The blind and the paralytic: Fostering argumentation in everyday and scientific issues. In J. Andriessen, M. Baker, & D. Suthers (Eds.), *Arguing to learn: Confronting cognitions in computer-supported collaborative learning environments* (pp. 227-260). Dordrecht: Kluwer Academic Publishers.

Sperber, D. & Wilson, D. (1995). *Relevance: communication & cognition*. Oxford, UJ : Blackwell Publishing.

Van der Pol, J. (2002). *Identifying and modelling variables in complex CSCL-situations. Case study: The use of asynchronous electronic discussions*. Paper presented at the Conference on computer support for collaborative learning (CSCL2002), Boulder, CO.

Van der Pol, J., & Admiraal, W. F. (2003). Het succesvol inzetten van asynchrone elektronsiche discussies [Succesfully implementing asynchronous online discussions]. *Onderzoek van Onderwijs, 32*(2), 26–31.

Veerman, A. L. (2000). *Computer-supported collaborative learning through argumentation*. Utrecht, The Netherlands: Utrecht University.

Chapter 2

Context enhancement for co-intentionality and co-reference in asynchronous CMC*

Abstract

The regulative and semantic "distance" of electronic conferencing may impede the topical alignment and the unambiguous interpretation of messages, hindering collaborative learning processes. Compared to a face-to-face environment, in electronic conferencing this distance may be caused by a reduced strength of online context. Explicitly defining the context of messages in an electronic environment may increase the writers' co-intentionality and co-reference. An annotation tool is presented, strengthening the context by providing a document under discussion and enabling users to anchor their messages in specific passages of the document. Preliminary results indicate that the tool does indeed reinforce the context, focusing the online discussion around a certain topic (increasing co-intentionality) and providing a frame of reference for single messages (increasing co-reference).

* Van der Pol, J., Admiraal, W. F., & Simons, P. R. J. (2006). Context enhancement for co-intentionality and co-reference in asynchronous CMC. *Artificial Intelligence & Society, 22*(3), 301-313.

Computer-mediated communication (CMC) is often compared with face-to-face (F2F) conversation and regular mail. Some forms of CMC bear more resemblance to the first (e.g. chat) and some more to the latter (e.g. asynchronous discussion or email). However, all types of CMC share a fundamental characteristic that results in a new and unique way of communication. One basic characteristic, which will be studied in this article, is the "distance" that is inherently implicated: What are the communicative effects of this lack of shared physical context? With respect to synchronous communication (chat), Reid (1996) suggests that its lack of social cues provides a high degree of personal freedom. Asynchronous CMC seems to provide this communicative freedom as well. While in F2F conversations—even in groups—it would be awkward for one of the communication partners to suddenly remain silent, an email or discussion board message that goes without response will not surprise anyone. Both in synchronous and asynchronous CMC, receiving a response to a particular message seems not to be as self-evident as in F2F communication, resulting in a lower degree of "answerability" (Hunt 1996).

The distance of CMC not only influences social aspects, such as personal freedom, but also—in more subtle manners—*regulative* and *semantic* aspects of collaboration. Even if participants in an online conference respond to each other (motivated by task instruction, tools, social bonding or other factors), content analysis reveal problems with the alignment and unambiguous interpretation of messages (Van der Pol 2002). Only a few online discussion tools have been specifically designed to deal with the regulative and semantic effects, and opportunities, of distance in CMC. In this study, which is concerned with the use of CMC for educational purposes, we will investigate how and why the distance of online communication influences regulative and semantic processes, aiming to arrive at a specialized tool for asynchronous collaborative learning. We will present a preliminary analysis of the affordances of this tool for collaborative learning in academic education, as compared with a regular discussion board.

Regulative and semantic characteristics of asynchronous CMC

Closely related to the largely socially based decreased level of "answerability", as described in the introduction, a second effect of CMC is the fact that participants seem less inclined than in F2F conversation to continue subjects raised by others. Users seem to experience more freedom to engage in new topics, which causes most electronic discussions to be divergent of character (Andriessen & Sandberg 1999; Mäkitalo, Häkkinen, Salo, & Järvelä, 2001). Educational researchers have found that, due to this divergence, it is more difficult to keep an asynchronous discussion focused on topic than is the case F2F (Romiszowski 1995). Herring (1999) reports a high degree of "disrupted adjacency" and "topic decay" in CMC. Hence, many instructional techniques have been developed for "keeping the thread", such as restating the original question when responses are going in the wrong direction (Beaudin 1999). We define such problems of topic decay as a lack of *co-intentionality*. Co-intentionality concerns the degree of explicit or implicit common goal directedness ("What are we going to discuss here?" or "Do we want to talk about the same subject?") and relates largely to regulative processes.

A third and again closely related effect of distance in asynchronous CMC is that participants appear to experience difficulties in interpreting messages, caused by a lack of "perspective taking" (Järvelä & Häkkinen 1999). When studying the implementation of a bulletin board in an advanced university course, Pena-Shaff & Nicholls (2004, p. 258) found that "*students did not always reach interpretations based on the analysis of all the ideas stated in the discussions, but rather stated their own interpretations and beliefs*", which the authors identify as producing primarily a process of self-reflection rather than a dialogical process of knowledge construction. Wan and Johnson (1994) conclude that a lack of integration of other participants' ideas left the group knowledge base with a substantial amount of redundancy and inconsistency. Winiecki (1999) describes how students and instructors often found discussions difficult to follow and also observed frequent misunderstandings, at times seriously affecting the learning process. This link between the problem of interpretation and the integration of ideas seems obvious. If a student is not quite sure what someone else means (a meaning that might even be obscure for the writer himself[1]), it will be very hard to process and elaborate on those ideas. Interpretation difficulties by remaining in one's own frame of reference and not successfully grasping the meaning that someone else attributes to a certain concept or

statement can be defined as a lack of *co-reference*. Co-referencing, or "referring to the same entity" (Hirschman, Robinson, Burger, & Vilain, 1998), is a more micro-level process than co-intentionality and situated rather in a semantic dimension.

Common ground and context

Both the levels of co-intentionality and of co-reference can be seen as aspects of the "common ground" of communication partners and thus theories on common ground can help us to understand how co-intentionality and co-reference might be related to the distance of CMC. Common ground refers to the goals, information, meaning and ideas people believe they share with others, which can be used to communicate and collaborate efficiently (Clark & Wilkes-Gibbs 1986). The notion of common ground enables people to refer to objects, people or ideas without having to describe them explicitly or comprehensively. However, critics of Clark's theory have demonstrated that common ground is not a fail-proof resource that simply accumulates (Koschmann 2003). While in Clark's ideas using common ground seems to be a logical and almost mathematical process of deduction, the practice of the human mind does not behave this logically or computer-like. With the aim of minimizing the effort invested in communication the human mind rather takes shortcuts and works by association, thereby being highly context sensitive. Bereiter (2002) even states that meaning cannot be separated from context and exists only in the relation between a person and a situation. Put more generally, common ground can never be reached completely because any situation triggers a broad range of meanings that is unique for each individual. However, the individual ranges of meanings can overlap to a certain extent, resulting in a certain degree of a "shared cognitive environment" (Sperber & Wilson 1995). As it still holds a wide array of meanings, a shared cognitive environment (or common ground) does not imply that people *will* make the same assumptions, merely that they are *capable* of doing so. Hence we do not define "common ground" as equaling "shared understanding". We see it rather as a prerequisite.

Besides being largely determined by situational characteristics and contextual variables such as common background, culture or personal history, the degree to which participants' cognitive environments overlap in a certain situation can also be increased by activities such as checking or clarifying meaning, or trying to take the perspective of someone else (Fussel, 2002). These activities are also called "grounding". Baker, Hansen, Joiner, and Traum (1999) distinguish *semantic* and *pragmatic* grounding. Semantic grounding is concerned with establishing shared meanings (resulting in co-reference), whereas pragmatic grounding is concerned with understanding each other's communicative intentions (resulting in co-intentionality). Both levels of grounding can be related to respective levels of common ground. Pragmatic grounding, for instance, aims to increase the pragmatic aspects of common ground, which especially concern participants' shared perception of the discourse situation, as indicated by Poesio and Muskens (1997). Completed with a, much used, socio-emotional account of grounding (Mäkitalo, Häkkinen, Salo, & Järvelä, 2003), we would like to picture three interrelated sub-areas of grounding and common ground (Figure 1). We use this distinction to be able to show that social, pragmatic and semantic processes are parallel type of activities, which sometimes have mutually beneficial influences, but also, in the case of limited resources, might interfere.

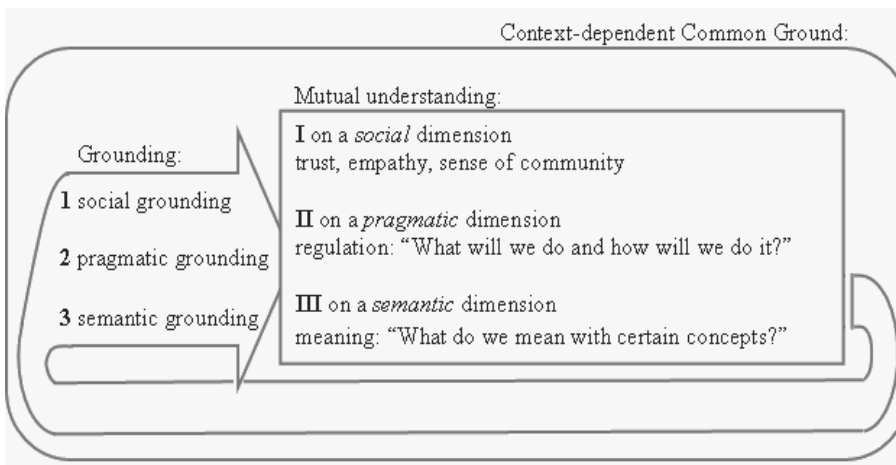

Figure 1. Distinguishing interrelated sub-areas of grounding and common ground.

Figure 1 also shows the cyclical character of grounding and common ground, since a higher degree of common ground to begin with greatly facilitates grounding, which in its turn is directed at enhancing common ground (see Cherubini, Van der Pol, & Dillenbourg, 2005). Note that "context" is mentioned as well because, apart from grounding, it also influences the degree of common ground in a certain situation. Because contexts differ in the strength of their meaning triggering effect, they influence the degree of overlap between participants´ activated cognitive environments, thus focusing co-reference and co-intentionality to a higher or less high extent.

Online context: "bridging the distance"

Especially in situations where grounding becomes problematic, as provided by asynchronous CMC's lack of non-verbal communication and low rate of "turntaking", the influence of context on common ground becomes more important. Therefore, it does seem fruitful to examine the possibilities to focus co-intentionality and co-reference (as specific elements of common ground) through the influence of context. It seems clear that, because of the absence of a shared physical context, '*"CMC may be considered to be 'context-reduced' in nature"* (Smith, 2003, p. 30). While this decreases the medium's focusing effect, it might also provide some opportunities. One could say that F2F conversation has its participants engaged in social processes to such a high degree that there is less room left for the semantic content. Therefore the absence of certain social contextual features in CMC also creates room for new elements. Instead of seeing social processes as a prerequisite for collaborative learning and developing online environments that increase social reciprocal responsibility (e.g. by introducing different communicative "roles"; Strijbos, Martens, Jochems, & Broers, 2004), improving "sociability" (Kreijns, Kirschner, & Jochems, 2002), or applying "uncertainty reducing cooperation scripts" (Mäkitalo et al. 2004), our approach will be different. Creating a stronger context for co-intentionality and co-reference could directly increase the focus on task-directed and task-effective communication in CMC, without the need to increase social, coordinative or regulative communication. As Andreasen concludes: '*"The richer the context is represented, the bigger the chance for students to develop shared understandings"* (2005, p. 11).

Increasing context for co-intentionality

As described, CMC can lack sufficient levels of co-intentionality because its communicative distance produces a variety in participants' intentions. Especially when the goal is to have a focused and detailed theoretical discussion, this variety can become problematic. Even when an original remark or question is aimed at processing some literature, the affordance of online discussion groups seems to make reactions often drift in the direction of expressing opinions and personal experiences (Preece, 2000), which decreases the coherence of the messages in a particular thread. This way, personal and opinion-oriented discussion can interfere with theoretical knowledge building (Järvelä & Häkkinen, 1999). Earlier research in asynchronous CMC (Van der Pol, 2002), on the coherence of messages within threads, has identified the domain of discussion as an important variable: more coherent discussion was found within discussions with fixed and limited topics. Thus, when the intention ("what is it that we are going to talk about") is more clear to all participants, they are less likely to drift off. This means that if a tool can naturally direct and focus the domain of discussion around a certain topic (e.g., in our case one of the actual course readings), it should be able to increase the topical alignment of the messages. The first threaded discussion system to group discussions explicitly around specific items for discussion was the BSCW system or "Basic Support for Cooperative Work" (Bently, Horstmann, & Trevor, 1997). As can be seen in Figure 2, BSCW offers discussion on a number of different items, such as a URL ("BSCW project page"), or an Excel sheet ("Project plan"). This way of focusing the discussion around particular documents sets up a context for co-intentionality, directing and focusing the overall aim of the conversation.

Figure 2. Screenshot of BSCW environment

Increasing context for co-reference

In our search for ways to enhance co-reference within learning in CMC, we follow Reyes and Tchounikine (2003), who define learning in CMC to require more than just information exchange and being directed at the integration of messages and the taking of new perspectives. They quote Bellamy (1997) who thinks learning conversations should *"allow participants to make connections between previously unrelated ideas"* (p. 84). One particular concern is the fact that traditional forum type systems only allow users to respond to entire messages as a whole. As messages often consist of multiple lines of thoughts, online discussion threads hide the true relation between replies and the thoughts they address. In agreement with Bellamy and Woolsey (1998), Reyes and Tchounikine state that this "interactional incoherence" impedes the emergence of learning conversations in two ways: both in the topical alignment and in the establishment of common ground. The forum type system developed by Reyes and Tchounikine is an attempt to increase the interactional coherence of learning conversations by introducing the "what you answer is what you link" criterion. That means they change the actual minimal unit that can be referred to from a message to a topic within a message (see Figure 3), giving the reply a more detailed frame of reference or context for co-reference. As stated by Riva (2001) *"the most effective way of clarifying the meaning of messages is to relate them to a shared context"* (p. 217). This context could draw students from their personal frame of reference, stimulate perspective taking and facilitate the interpretation and integration of ideas.

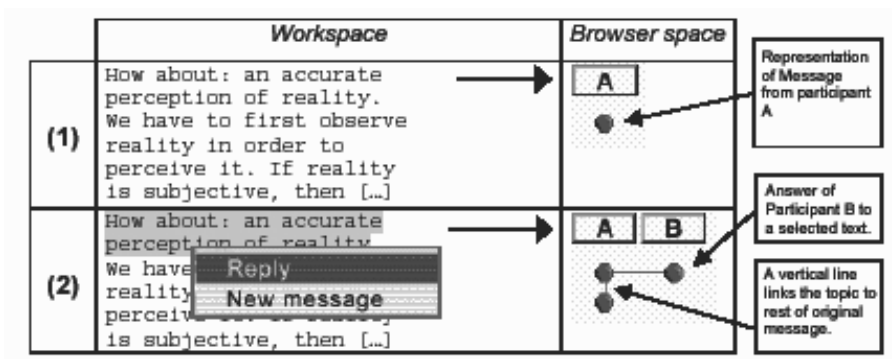

Figure 3. Reyes and Tchounikine's system for increased interactional and sequential coherence.

Integrated design: An annotation system

In our setting of university education, we are particularly interested in increasing co-intentionality by drawing collaborative learning conversations deeper into the subject matter and providing co-reference by identifying which specific frame of reference messages relate to (as miscommunication is often caused by the fact that the meaning of scientific concepts often differs from their common-sense one). Both functionalities of the two described systems, setting up a domain for discussion (context for co-intentionality) and creating a localized frame of reference (context for co-reference), are incorporated in the design of an "annotation system". The context for co-intentionality means that it concentrates the discussion around a specific document; the context for co-reference allows the user to anchor messages to specifically marked statements or ideas. Different with the two tools described earlier, an annotation system reinforces co-intentionality by integrating both document and discussion in the same screen (instead of presenting it as a downloadable file), and increases co-reference by referring to passages from the discussed document (instead of referring to topics from other messages). This makes an annotation system (see Figure 4) particularly useful for the collaborative processing of literature.

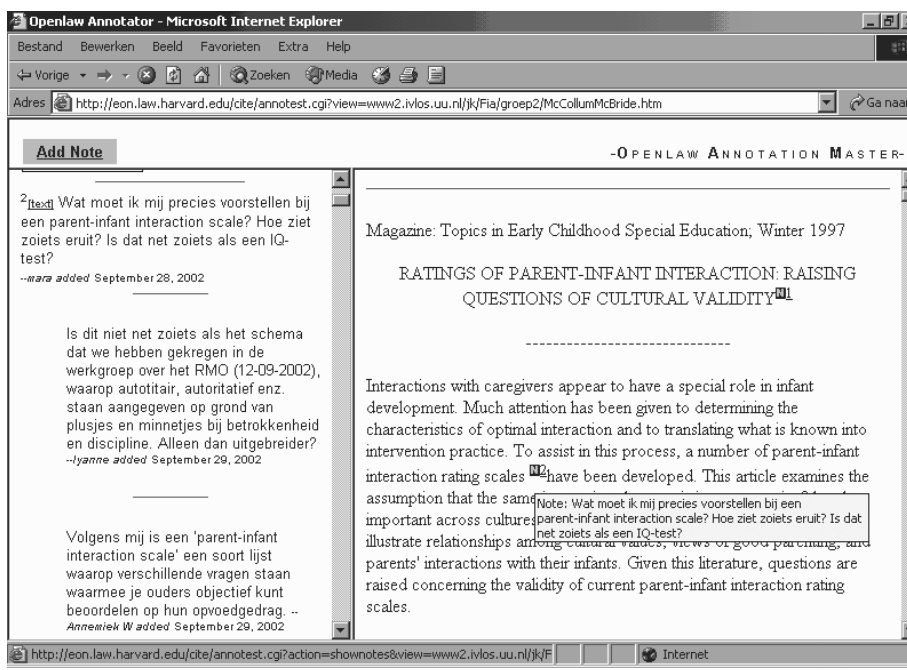

Figure 4. Openlaw Annotation Master (by Wendy Selzer).

The displayed annotation system in Figure 4 simultaneously displays both the regular "threaded" discussion and the document under discussion. Moreover, both items are cross-linked. Every new posting (in the left hand frame) can be anchored to a selection from the text (right hand frame), which gives it the term "anchored discussion" (Bernheim Brush, Bargeron, Grudin, Borning, & Gupta, 2002). The left-hand frame of an annotation system functions as a regular forum type tool, with options for creating new messages (Notes) or replying to existing ones. The left and right-hand frames are automatically kept synchronized as well: when clicking at a particular selection in the text to the right (such as the blue "N1" icon), its connected messages appear left and when clicking at a particular message on the left (the blue "text" link), the appropriate text selection is automatically displayed to the right. In the example above, the left-hand frame shows a part of the second thread and the right-hand frame shows the first page of the article, with the anchors for threads one and two. This annotation system has been used in our study into the enhancement of context to increase co-intentionality and co-reference in asynchronous learning conversations about course literature.

Comparing an annotation system with a standard discussion forum

In a Dutch first-year university course on General Pedagogics, we used two conferencing systems for the collaborative processing of the course literature. About 50% of the students used Blackboard as a standard discussion forum and the other half used an adapted version of the Openlaw Annotation Master (Selzer, 1997) as an example of an annotation system . Apart from the described differences, both systems provided the functionality of creating new messages and replying to other ones. Identical instruction in both systems has been used including that the students had to ask each other questions about text passages they did not understand. We present a typical thread of each system and describe some of the differences that became apparent after a first qualitative comparison of the interactions in both systems.

Starting with Blackboard, the regular discussion forum, messages resemble email correspondence in more than one respect, most of them carrying more information than just the student's actual question or answer about the subject matter (see item 3 in Figure 5 on the next page). First of all, most open and close with a *social* introduction and a departure phrase (1). They also contain *regulative* statements about what it is the author is going to do in this mail and

what subject he or she wants to talk about, thus creating a context for co-intentionality before introducing the core message (2). After this core message, we again find regulative (and partially social) statements in which the authors seem to want to justify their questions or give some additional information about the value of their response (2).

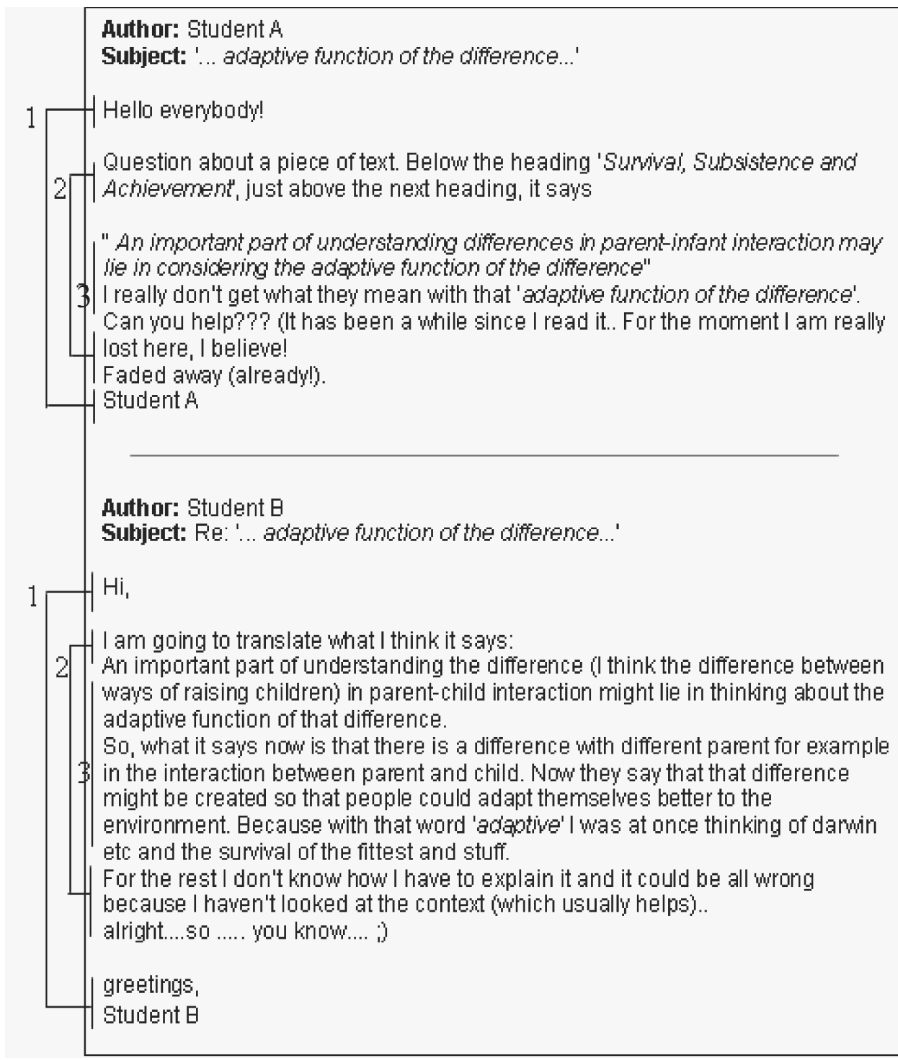

Figure 5. Thread from a traditional threaded discussion (Blackboard). The messages were translated into English, where words that already were written in English are placed between quotation marks.

This message structure we find in the regular forum discussion closely resembles the standard structure of outgoing messages on telephone answering machines. As described by Hammer and Veronesi (1999) these messages usually consist of an opening, a core section with a directive speech act (requesting the caller to leave a message) and a closing. Hammer and Veronesi explain that this format is useful to construct a social relationship between caller and "callee", with the purpose of keeping in contact and requesting a response from the caller, who is situated at such a distance that he cannot be relied upon to do so automatically.

When looking at a typical sequence from the annotation system (see Figure 6 below), the short and to-the-point nature of the messages immediately stands out. We do not find any social introductions or closings and only one regulative comment about the reason for posting a message (2). The messages consist, for the most part, solely of the actual question or remark about the content (3), and the students do not seem to feel a need to explain why they are asking a question or to mention that they would really like an answer. The fact that the discussion document is dominantly present on screen seems to strengthen student's intention to discuss the documents' meaning, thus making it clear what is going to happen and increasing co-intentionality.

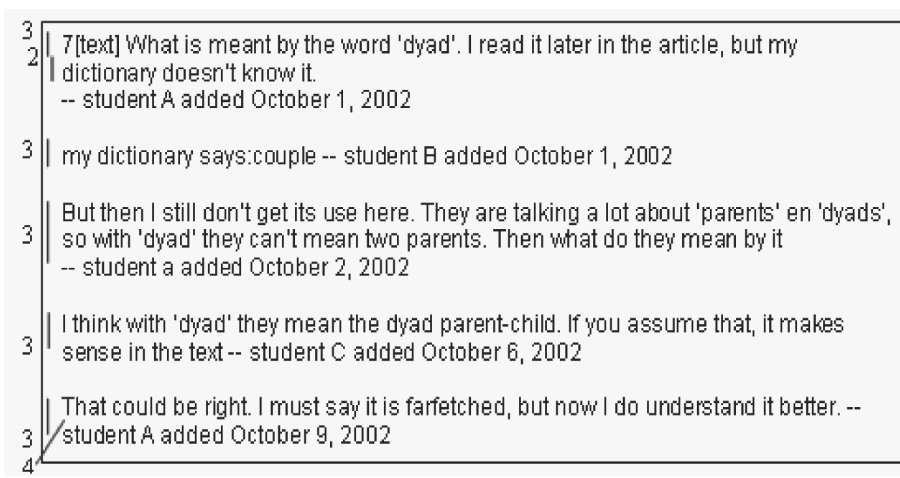

Figure 6. Thread from the Annotation Engine. The messages were translated into English, where words that already were written in English are placed between quotation marks.

An annotation system would seem to enhance co-reference as well. This is not only caused by the anchoring of the messages, but may also be influenced by the fact that its layout presents related messages more closely positioned near each other, whereas Blackboard usually displays only one message at a time. The impression of enhanced co-reference appears to be supported by the use of demonstrative pronouns to the content of other messages such as "that" (element 4 in Figure 6). Finally, the decreased length of the messages also translated itself into a larger number of messages. If we assume that shorter messages require less effort to write than longer messages, this indicates that the total amount of effort that is put into the creation of messages remains more or less stabile for both conditions. This increased rate of turn-taking could also very well facilitate grounding. As we see in Figure 6, the author of the original question (student A) is not satisfied after the first response (student B). When the thread would have ended at this point, as was the case in the Blackboard example, an important chance for further elaboration, or clarification, might have been missed.

Conclusion

Comparing the communication we found in both systems, an annotation system appears to create a different kind of communication in three ways:

1. The annotation system produces less social and regulative activities which demonstrates higher levels of answerability and co-intentionality;
2. The annotation system produces short, demonstrative referring statements which indicate the presence of a shared frame of reference for individual messages, and
3. The annotation system produces shorter messages, increasing the number of message within threads, thus increasing the amount of turn-taking and the possibilities for repairing misunderstandings. This effect was unexpected and possibly partly caused by the first two effects.

While these results indicate a stronger sense of context in the annotation system , they also seem to support our suspicion that social communication and coordination is not always necessary for a successful collaboration. As we have seen users in an annotation system put all their energy straight into the core content of their messages, we do not want to overvalue the importance of social communication for collaborative learning in CMC[2]. These findings about

the lack of a relation between social communication and successful collaboration is supported by other studies. Dewianti (2005) concluded that stimulating the regulation of group processes does not enhance the process of knowledge co-construction. This result was also found by Lockhorst (2004, p. 220), who studied asynchronous CMC used in teacher training, and writes that: *"In our research we did not find a relation between the students' online social off-task communication and their content related communication"*. Moreover, after teambuilding activities to improve the collaborative processes of the students, the social communication increased while communication about task content decreased". In a study aimed at establishing the beneficial effects of socio-emotional processes in CMC, Mäkitalo et al. (2004, p. 312) also had to conclude a negative effect of coordinative and social communication: '*The results indicate that the uncertainty reducing script in fact increased the amount of discourse and decreased information seeking"*. Especially within academic education, where the main goal of online discussion often includes a thorough processing of literature and collaboratively building a deeper understanding of it, we want to emphasize the importance of students integration of online communication with information seeking and content-related reflection.

Finally, a more general conclusion from these preliminary findings could be that, although people in a new medium will at first tend to copy known ways of communication, they are by no means set in their old ways and can—stimulated by a tool—easily adopt new ways of interaction. The observation that the contextualized design of an annotation system seems to give the tool a stronger sense of purpose, means that the functionality of a tool can influence patterns and forms of collaborative learning conversations. As predicted by the grounding model of Baker et al. (1999), tools can (without having to train "new literacy" or to direct communication in more forceful ways), influence communicative behavior by means of their "affordance", or *"the aspects of an artifact's design that suggests how it is to be used"* (Johnson, 1995, 219).

Regarding the possibilities for future research, we will use these first empirical impressions of an annotation system to develop a coding scheme to categorize the collected data and perform statistical analyses. We will then investigate whether our preliminary conclusions about the levels of increased co-intentionality and co-reference will indeed turn out to be valid and whether the statistical analyses allow us to make more definite statements about the affordances of annotation systems for the online collaborative processing of

literature. Should this turn out to be the case, we will then continue with the development of a more sophisticated annotation system in order to further increase the quality of online asynchronous learning conversations.

References

Andriessen, J., & Sandberg, J. (1999). Where is education heading and how about AI. *International Journal of Artificial Intelligence in Education, 10*(2), 130-150.

Baker, M., Hansen, T., Joiner, R., & Traum, D. (1999). The role of grounding in collaborative learning tasks. In P. Dillenbourg (Ed.) *Collaborative learning: cognitive and computational approaches* (pp. 31-63). Amsterdam: Elsevier Science.

Beaudin, B. P. (1999). Keeping online asynchronous discussions on topic. *Journal of Asynchronous Learning Networks, 3*(2), 41-53.

Bellamy, R. (1997). *Support for learning conversations*. Retrieved November 11, 2002, from http://kmi.open.ac.uk/seminars/abstract.cfm?id=10.

Bellamy R, & Woolsey, K. (1998). Learning conversations. *SIGCHI Bulletin 30*(2), 108–112.

Bently, R., Horstmann, T., & Trevor, J. (1997). The world wide web as enabling technology for CSCW: The case of BSCW. *Journal of Collaborative Computing, 6*, 111-134.

Bereiter, C. (2002). *Education and mind in the knowledge age.* Mahwah, NJ: Lawrence Erlbaum Associates.

Bernheim Brush, A. J., Bargeron, D., Grudin, J., Borning, A., & Gupta, A. (2002). *Supporting interaction outside of class: Anchored discussion vs. discussion boards.* In G. Stahl (Ed.), *Computer support for collaborative learning: Foundations for a CSCL community. Proceedings of CSCL 2002* (pp. 425-434). Hillsdale, NY: Lawrence Erlbaum Associates.

Cherubini, M., Van der Pol, J., & Dillenbourg, P. (2005, July). *Grounding is not shared understanding: Distinguishing grounding at an utterance and knowledge level.* Paper presented at CONTEXT'05, the Fifth International and Interdisciplinary Conference on Modelling and Using Context, Paris.

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

Dewianti, S. (2005). *Learning together, a positive experience.* Heerlen, The Netherlands: Open University.

Hammer, F., & Veronesi, D. (1999). Between formality and originality: Answering machine outgoing messages in comparison. In J. Verschueren (Ed.) *Pragmatics in 1998; Selected papers from the 6th International Pragmatics Conference* (pp. 198-224). Antwerp, Belgium: International Pragmatics Association.

Herring, S. (1999). Interactional coherence in CMC. *Journal of Computer-Mediated Communication 4*(4). Retrieved March 4, 2004, from http://jcmc.indiana.edu/vol4/issue4/herring.html

Hirschman, L., Robinson, P., Burger, J., & Vilain, M. (1998, March). *Automating co-reference: The role of annotated training data.* Paper presented at AAAI Spring Symposium on Applying Machine Learning to Discourse Processing, Palo Alto, CA.

Hunt, R. A. (1996, May). *Affordances and constraints of electronic discussions.* Paper presented at the 13th Inkshed Working Conference, Hecla Island, Manitoba, Canada.

Järvelä, S., & Häkkinen, P. (1999, March). *Web based cases in teaching and learning - reciprocal understanding and perspective taking in conversation.* Paper presented at the Computer Assisted Learning Conference (CAL99), London.

Johnson, J. A. (1995, June). *A comparison of user interfaces for panning on a touch-controlled display.* Paper presented at the Conference on Human Factors in Computing Systems, Denver, CO.

Koschmann, T., & Le Baron, C. (2003). Reconsidering common ground: Examining Clark's contribution theory in the OR. In K. Kuutti, E. Karsten, G. Fitzpatrick, P. Dourish, & K. Schmidt (Eds.), *ECSCW 2003: Proc. Eighth European Conference on Computer-Supported Cooperative Work*. Amsterdam: Kluwer.

Kreijns, K., Kirschner, P. A., & Jochems, W. (2002). The sociability of computer-supported collaborative learning environments. *Journal of Education Technology & Society, 5*(1), 8–22.

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

Mäkitalo, K., Häkkinen, P., Salo, P., & Järvelä, S. (2001, August). *Analysing a mechanism of the common ground in web based interaction.* Paper presented at the European Association for Research on Learning and Instruction Conference, Fribourg, Switzerland.

Mäkitalo, K., Weinberger, A., Stegmann, K., Järvelä, S., Häkkinen, P., & Fischer, F. (2004, July). *Uncertainty reducing cooperation scripts in online learning environments*. Paper presented at the Special Interest Meeting of the European Association for Research on Learning and Instruction SIG 6 and SIG 7, Tuebingen, Germany.

Pena-Shaff, J. B., & Nicholls, C. (2004). Analyzing student interactions and meaning construction in computer bulletin board discussions. *Computers & Education, 42*, 243-265.

Poesio, M., & Muskens, R, (1997). *The dynamics of discourse situations*. Retrieved September 14, 2003 from http://citeseer.ist.psu.edu/cache/papers/cs/31562/http:zSzzSzwww.cis.uni-muenchen.dezSzsilzSzworkshopzSzabstractszSzpoesio.pdf/the-dynamics-of-discourse.pdf

Preece, J. (2000). *Online communities: Designing usability, supporting sociability*. New York: John Wiley & Sons.

Reid, E. (1996). Communication and community on internet relay chat: constructing communities. In P. Ludlow (Ed.) *Conceptual issues on the electronic frontier*. Cambridge, MA: MIT Press.

Reyes, P., & Tchounikine, P. (2003). Supporting emergence of threaded learning conversations through augmenting interactional and sequential coherence. In B. Wasson, S. Ludvigsen, & U. Hoppe (Eds.) *Designing for change in networked learning environments: Proceedings of the International Conference on Computer Support for Collaborative Learning* (pp. 83-92). Dordrecht, The Netherlands: Kluwer Academic Publishers.

Riva, G. (2001). Communicating in CMC: Making order out of miscommunication. In L. Anolli, R. Ciceri, & G. Riva (Eds.) *Say not to say: New perspectives on miscommunication* (pp. 203–233). Amsterdam: IOS Press.

Romiszowski, A. (1995). Use of hypermedia and telecommunications for case-study discussions in distance education. In F. Lockwood (Ed.), *Open and distance learning today*. New York: Routledge.

Selzer, W. (1997). Annotation Engine [Computer software]. Retrieved January 15, 2002 from http://cyber.law.harvard.edu/projects/annotate.html.

Smith, B. (2003). The use of communication strategies in computer-mediated communication. *System, 31*, 29-53.

Sperber, D., & Wilson, D. (1995). *Relevance: Communication & cognition*. Oxford, UK: Blackwell Publishing.

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

Van der Pol, J. (2002, January). *Identifying and modeling variables in complex CSCL-situations. Case study: the use of asynchronous electronic discussions.* Paper presented at the Computer Support for Collaborative Learning Conference, Boulder, CO.

Wan, D., & Johnson, P. M. (1994). Experiences with Clare: A computer supported collaborative environment. *International Journal of Human Computer Studies, 41*, 851-879.

Winiecki, D. J. (1999). Keeping the thread: Adapting conversational practice to help distance students and instructors manage discussions in an asynchronous learning network. *Deosnews 9*(2). Retrieved March 23, 2004, from www.ed.psu.edu/ACSDE/deos/deosnews/deosnews9_2.asp.

---

[1] This is especially in the case of learning situations, where students start the discussion with low levels of understanding of the subject matter.

[2] Note that we are concerned with regular university education, where students do know each other. When participants do not know each other (as is the case in distance education), social processes are likely to be of higher importance.

Chapter 3

The affordance of anchored discussion for the collaborative processing of academic texts*

Abstract

A system for anchored discussion is compared with a system for traditional forum discussion (Blackboard), and their collaborative and communicative affordances for the collaborative processing of academic texts are investigated. Results show that discussion in the system for anchored discussion is more directed at processing the meaning of texts than discussion in the traditional forum, which is more oriented towards the sharing of personal opinions and experiences. This difference in orientation produces a more constructive collaboration in the system for anchored discussion, versus a more debate-like collaboration in the forum discussion. Additionally, while messages in the traditional forum resemble usual discussion or email conversation and contain social and regulative comments, discussion in the system for anchored discussion is seen to be more efficient and to-the-point. We conclude that for collaborative text comprehension by undergraduate students, anchored discussion might be more suitable than traditional forum discussion. Finally, the observed differences can be explained by the stronger defined collaborative context in the system for anchored discussion, which focuses participants' collaborative intentions and their frames of reference.

* Van der Pol, J., Admiraal, W., & Simons, P. R. J. (2006). The affordance of anchored discussion for the collaborative processing of academic texts. *International Journal of Computer Supported Collaborative Learning, 1*(3), 339-357.

In higher education, processing academic literature can be seen as a central but often quite challenging task, especially for undergraduate students. Assuming that a deep processing of the subject matter requires an active construction of knowledge by the learner (Boekaerts & Simons, 1995; Phillips, 1995) for which social interaction can be helpful (Simons, Van der Linden, & Duffy, 2000), this study concerns the facilitation of collaborative literature processing. Focusing on students' collaboration as a means to learning, this study can be labeled as being based on an "interactional constructivist epistemologya" (Suthers, 2005). In attempting to create a successful collaboration that involves students' active interaction with content, this study will make use of online asynchronous discussion. As Warschauer (1997) states, this medium offers the important possibility to link dialogue and interaction with individual study and reflection. Combining the advantages of social interaction, such as increased 'ownership' of ideas and the opportunity to connect to existing knowledge, with the possibilities of delayed communication for (re)reading, (re)writing, and reflection (Moon, 1999), online discussion should provide room for a thorough processing of students' course materials.

Clark, Weinberger, Jucks, Spitulnik, and Wallace (2003) have remarked, however, that generic tools for online discussion still have some limitations when used in educational practice and that not all if its potential is yet being realized. Many studies report a lack of collaborative knowledge construction in online discussion. Activities that are scarcely found are integrating (Wan & Johnson, 1994), transforming (Veerman, 2000), or discussing (Hewitt & Teplovs, 1999) one another's ideas. More generally, students are found not to display many "higher cognitive skills" (Sringam & Geer, 2000), or to engage much in "constructive communication" (Lipponen, 2001). Instead, Pena-Shaff and Nicholls (2004), Guzdial and Turns (2000), and Fay, Garrod and Carletta (2000) found students' communication to consist largely of independent monologues, a finding that seems consistent with that of De Laat (2002) and McLoughlin and Luca (2000), who report communication to be mainly directed at what Gunawardena, Lowe and Anderson (1997) describe as "lower-level learning activities," such as the sharing and comparing of knowledge.

Supporting collaborative knowledge construction

From the previous section we can conclude that collaborative knowledge construction is a delicate process that in many educational situations will need to be supported in order to be successful. However, deciding how or where to support it requires a thorough understanding of the process itself. Important in this respect is the study of Järvelä and Häkkinen (2000) who established a link between the depth of discussion and the subject of messages.

In their definition of different levels of discussion, they associate theory-based messages with deeper levels of constructive and cross-referenced discussion than opinion-based messages, which are associated with lower-level discussion that includes less constructive and more independent messages. Research by Van der Pol (2002) into the online collaborative processing of academic literature reports this relationship as well. In this study, it seemed especially difficult for students to provide each other with specific and relevant feedback to advance their understanding of the subject matter. Instead, students were more inclined to share existing experiences and perspectives, resulting in a more associative connection between consecutive messages.

This apparent preference of students for sharing opinions and experiences instead of building new understanding of the subject matter can be understood in the following way. To begin with, the mediated nature of abstract knowledge can cause learning in a university setting to differ from more direct learning as it might take place in other "real" settings (Laurillard, 1993). Since the somewhat 'unnatural' task of processing academic learning materials might not lie very close to the personal perspectives of the participating students, a link might need to be negotiated between students' personal and more academic perspectives (Petraglia, 1998). Furthermore, constructing new knowledge is a difficult process, as students need to 'discover' a new understanding of the subject matter that no-one yet possesses individually. As Stahl (2000) illustrates in his model of "social knowledge building," students' personal understanding –or better, their "tacit pre-understanding,"– forms an essential input in the social knowledge-building cycle. A possible effect of a limited personal understanding of the subject matter is that both articulating one's questions and interpreting those of others and providing them with specific and relevant feedback will require a high amount of mental effort. While this meaning-processing effort seems to be exactly what drives learning (Baker, Hansen, Joiner, & Traum, 1999; Schwartz & Lin, 2000), it also might sometimes exceed

students' capabilities. Instead, sharing existing opinions and experiences seems to be a less-demanding option for students to participate in online discussions.

Especially in electronic environments, it seems important to monitor the amount of effort that is required for successful collaboration as the medium is limited in supporting the development of mutual understanding, or "grounding" (as defined by Clark & Brennan, 1991). Not only does inferring the perspective of the future reader during message formulation ("audience design") seem to be more difficult in many-to-many communication than in two-party conversations (Gergle, Kraut, & Fussel, 2004), but, as argued by Fussel and Benimoff (1995), several additional features of online discussion, such as its delayed feedback, low amount of "turn-taking" and lack of non-verbal clues, make it hard to repair miscommunication and maintain a "shared communicative context." Although grounding can be seen as a functional process (Clark & Brennan, 1991; Dillenbourg, 1999), meaning that the amount of grounding activity generally will match the need for it, Gergle, Kraut and Fussel demonstrated that with regard to the communicative efficiency of computer-mediated communication (CMC), these compensations often fall short.

Although the complex processes of successful collaborative learning may require more than just a strong link between discussion and subject matter and efficient communication, they can be seen as important prerequisites for a successful collaborative processing of literature, and we will now direct our attention towards finding ways to facilitate these two processes. One way to focus the collaboration and grounding efforts of students would be to change the pedagogical approach. However, introducing additional training or elaborate instructions could increase the already high demand of time and effort in online discussion for both students and teachers. We believe, in line with Dillenbourg (1999), that grounding efforts should remain subordinated to the accomplishment of the task and the production of "rich" interactions; it would then seem important to make sure students' efforts are invested in the most optimal and productive way. To accomplish this, our study investigates how overcoming the two identified obstacles could be afforded by features of the electronic environment itself. We feel this could be a fruitful approach as the functionality of the generic discussion tools generally being used in university settings have not specifically been developed to support a collaborative processing of literature.

We will specifically investigate whether a tool's functionality can influence students' collaboration implicitly, thus preserving an 'open' learning environment that does not restrict users' actions.

A specialized design: Anchored discussion

Hunt (1998) maintained that the fact that students in older bulletin board systems appeared to respond immediately and without much reflection was the effect of a lack of context. In line with this thought, we will investigate a tool for "anchored discussion" (see Bernheim Brush, Bargeron, Grudin, Borning, & Gupta, 2002), trying to overcome some of the limitations of asynchronous electronic communication as mentioned in the introduction. Anchored discussion (see Figure 1 on the next page) finds its roots in literature processing-oriented systems for shared annotation (see Davis & Huttenlocher, 1995; Sumner & Buckingham Shum, 2001), yet takes a slightly different approach. While shared annotation starts from the notion of personal annotation made visible to peers, anchored discussion starts from the notion of collaborative discussion that is contextualized, or anchored, within a specific content. While systems for shared annotation are inclined to display individual notes within the text (as in Schoonenboom, 2002) anchored discussion, or "linked artifact-centered discourse," displays both artifact and discussion in a linked, yet independent manner (Takeda & Suthers, 2002).

An effect of this difference in origin and interface is that shared annotation might leave more room for individual processes, but is shown to have some limitations in supporting interactivity. Nokelainen, Miettinen, Kurhila, Floréen, and Tirri (2005) found a positive relation between an individual learner's activity in a system for shared annotation and their study success, but they also established a possible distracting effect of shared annotation as users viewed self-made highlights and comments as being more useful than those made by other learners. Comparably, a large-scale study by Cadiz, Gupta, and Grudin (2000) on a system for shared annotation showed that the majority of annotations did not contain any replies, whereas studies by Bernheim Brush et al. (2002) and Guzdial and Turns (2000) showed that threads in a system for anchored discussion were significantly longer than those in regular forum discussion. Although a tool's effect on the quantity of interaction highly depends on its pedagogical implementation and particular functionality (see Bernheim Brush et al., 2002), these results indicate that the

interaction-oriented design of anchored discussion could offer good possibilities for supporting students' collaborative processing of academic texts.

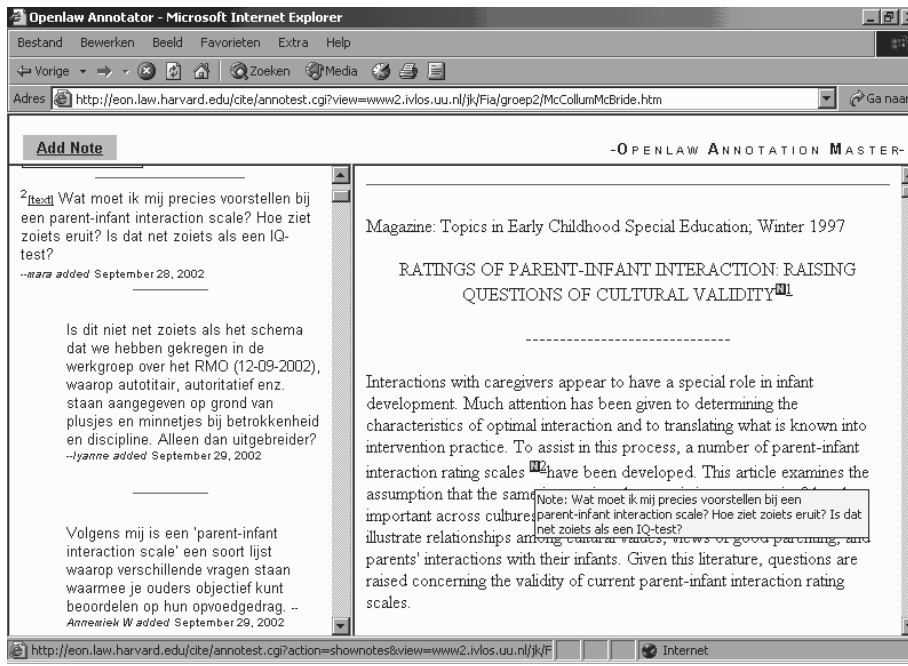

Figure 1. Adapted version of the Openlaw Annotation Master (Selzer, 2000), presenting both the threaded discussion (left) and the article under discussion (right).

While several of the studies mentioned above determined anchored discussion to be a potentially valuable medium for collaborative learning, they do not directly investigate its effects on the quality of interaction, or compare this to other tools for collaborative learning. Therefore, this study will aim to compare a system for anchored discussion with a system for traditional forum discussion and to investigate their collaborative and communicative affordances.

Anchored discussion versus forum discussion

Having presented anchored discussion as a viable alternative for regular online discussion, we will now describe how we expect their respective functional differences to influence the collaborative learning processes and to possibly support students' collaborative processing of scientific texts. Looking at the functional differences between a system for regular forum discussion (Figure 2) and the system for anchored discussion used in this study, we see that the system for anchored discussion integrates a document or text into the online discussion environment in two ways. Apart from physically presenting the text online, adjacent to the threaded discussion, it offers the possibility to anchor messages to specific passages of the text.

The visual integration of text and discussion in the system for anchored discussion might present the most obvious difference from regular forum discussion. As to the effect of this integration, Herrmann and Kienle (2003) mention that material that is provided in computer-based collaborative learning environments can and should serve as a "context for collaboration". In line with this thought, Gay, Sturgill, Martin, and Huttenlocher (1999) state that document-mediated communication can define a stronger collaborative context, setting the orientation and providing a means for effective communication. In other words, the explicit document-centeredness of anchored discussion might naturally direct users' collaborative intentions towards the processing of that text. According to Guzdial and Turns (2000), this could strengthen the link between discussion and study material and make the discussions more effective. Additionally, this automatic focus of students' intentions and perceptions of the collaborative goal might also reduce the need for coordination of the collaborative process.

The possibility of anchoring messages to specific passages of the text can provide individual messages or threads with a stronger frame of reference. As Herrmann and Kienle (2003) describe, being able to refer to a piece of available context will reduce the level of explicitness that is required. Additionally, as described in Clark and Brennan's (1991) "principle of minimal effort," this available frame of reference can be expected to reduce the "space for misunderstanding" (Dillenbourg, 1999), facilitating interpretation and requiring students to invest less effort in clarifying their messages.

Figure 2. Partial print screen of the system for regular forum discussion (Blackboard) that was used in this study.

Research questions

To investigate whether the design of the system for anchored discussion supports the collaborative processing of academic texts more than a system for traditional forum discussion this study focuses on four concrete research questions. The first two aim to assess whether the system for anchored discussion strengthens the link between discussion and text and whether it enhances the efficiency of communication. The last two questions aim to check the general suitability of the system for anchored discussion to facilitate students' online discussions.

Do the two systems:

1. Differ in their ability to strengthen the link between discussion and study material?
2. Support different levels of communicative efficiency?
3. Produce a different kind of conversation?
4. Provide different constructive activities?

Method

Research context

This study was conducted in a Dutch first-year first trimester pedagogy course titled "General Pedagogy"[1]. In this course, the students collectively received weekly lectures and participated in weekly seminars in subgroups. In total, the 193 students enrolled in the course were divided into 9 seminar groups with 6 teachers. During the course, the students had to read several English articles and discus them in the seminars. The reading for the course was planned week by week. To stimulate students' processing of the course material, the course provided an online facility for collaboration. This voluntary online collaboration took place either in a regular Blackboard discussion forum (see Figure 2), or in the system for anchored discussion described above (see Figure 1). To control the influence of the seminar teacher in the two conditions, each seminar group was randomly split into two subgroups, each of which was assigned to one of the conditions.

Two articles were covered in a two-week online discussion round (starting in week 1 and 3), prior to their face-to-face discussion in the seminars. The students were instructed to try to help each other develop a better understanding of the text by asking each other questions about difficult passages and lines of reasoning and by trying to explain to each other how these can be understood. As a reward for sufficient participation (submitting at least 2 messages per week), students were allowed to skip a question on the final exam. The discussions were moderated by the seminar teachers (most of whom were inexperienced ICT users), who received instruction by the researcher. This instruction not only served to demonstrate the electronic environments, but also to optimize the teachers' pedagogical implementation of them in the course and provide an optimal base for collaboration in both conditions. The educational advantages of asynchronous electronic discussion were discussed, and several strategies to increase the chances of creating successful discussions, as described in Pena-Shaff and Nicholls (2004) and Van der Pol and Admiraal (2003), were highlighted. The instruction emphasized the medium's potential to make students' intuitive understanding explicit, which enables the teachers to connect to them. It was also stated that to encourage students to express their intuitive and uncertain ideas and questions about the text, a constructive and helpful conversation may be more effective than a sharp and critical debate. Concerning moderation of the discussions, we

suggested the teachers not be too authoritative, but rather to respond to students' difficulties with follow-up questions and to try to support their advancement in small steps.

Operationalization of the research questions

The electronic messages of the discussions, also known as the collaboration protocols, will serve as the main research data, completed by information from questionnaires and group interviews. With 9 seminar groups, each containing two online discussion groups (one for each condition), 18 student groups were simultaneously discussing the same course materials for a 2-week period. This was done twice, resulting in a total of 36 discussions with a total of 1,133 non-accidental, student-generated, task-related messages. After finishing the discussions, students were asked to complete a questionnaire and four group interviews were conducted. In total, 111 questionnaires were fully completed, 62 for the anchored discussion condition and 49 for the regular forum discussion condition. Two discussion groups from each condition were interviewed, with approximately 10 students each.

These three data sources will be used to address the research questions in the following way. First, the subjects of messages and the number of references to the subject matter (collaboration protocols) and students' reported off-line reading activities (questionnaire), will be used as indicators for the link between discussion and study material. Second, the communicative efficiency will be measured in terms of the elaborate or demonstrative nature of references, with the need for self-clarification within individual messages and students' reported levels of mutual understanding being evaluated. Third, the nature of conversation will be investigated using the frequency of social and regulative comments and the average number and length of messages. These protocol data are completed by students' experiences with the two tools, as collected in the group interviews. Fourth, the constructive activities are based on the number and type of questions, answers and critical reactions ("message type"), as well as the number of argumentations, confirmations and clarifications. Finally, additional data from the protocols, interviews and questionnaires will be used to check whether the two systems were easily used and if they presented any major technical difficulties that might have interfered with the investigation of their hypothesized affordances.

Measures

Collaboration protocols

In order to analyze the content of the collaboration protocols, an instrument has been developed. While many instruments already exist in this domain, they were not found to optimally suit our specific research context and questions. The main reasons for this consideration were that theory-based models did not seem to fit our practical reality of students' collaborative knowledge building when possessing only low levels of expertise. As these students do not generally follow a scientific cycle of inquiry, it is important not to overlook "where students are at" (Pilkington, 2004) and at how students in practice engage in constructive conversation. Other instruments, like the coding scheme of Järvelä and Häkkinen (1999) do seem to fit better our context, but were found to present difficulties when trying to establish sufficient levels of inter-rater reliability. Therefore, a new coding scheme was developed that, instead of presenting a measure for the overall quality of a discussion, was aimed at identifying several separate and more basic characteristics of the discussions. This development was partly done in a top-down manner, based on existing coding schemes, and partly in a bottom-up approach, grounded in the collected data. Messages were used as the unit of coding. The categories that were created cover the variables "message subject," "message type," "argumentation," "confirmation," "clarifying," "referring" and "social" and "regulative" comments (see Table 1). The inter-rater reliability, in terms of Cohen's kappa, exceeded .70 for all categories, with an average of .82. Due to a low reliability, the connectivity (the relevance of a response in relation to the message it replies to) has not been included in further analysis.

Regarding the subjects of messages, we have identified three main categories that describe whether a message is concerned with the meaning of the article ("What does the author want to say in this article?"), students' opinion ("What is the student's personal opinion about certain issues in or related to this article?"), and whether the message is concerned with the task of processing the article at all (non task-related). Furthermore, since anchored discussion has been found to produce discussion that is more focused on specific concepts than regular forum discussion (Bernheim Brush et al., 2002), the messages oriented at establishing the meaning of the article are further coded on how general or local is the issue they concern: Ranging from the overall idea and structure of the article, to individual statements (with or

without argumentative reasoning), to the meaning of specific words or concepts. The category "message type" includes whether a message is identified to start a new thread ("thread-opener"), to pose a new question within an existing thread ("follow-up question"), to express a form of disagreement or doubt ("critical reaction"), or to constitute another type of reaction or answer. When a thread-opener concerns a statement, this is scored to be either expressed as being a sure thing ("definite") or in a more careful way ("open"). When it concerns a question, it is scored as containing a possible answer ("full") or not ("empty"). Having labeled the subject and type of the messages, messages are coded for the constructive activities "argumentation," "confirmation" and "clarifying," with descriptions based on the coding schemes of Veerman (2000) and Veldhuis-Diermanse (2002). These three variables contain activities that are scored as being either present at least one time in a particular message (1) or not at all (0). Thus, it is possible for a message to contain both a supporting argument ("argumentation"), an expression of agreement with someone else's earlier statement ("confirmation") and an explanation of what they mean by this ("clarification"). Note that clarification is used here to indicate when students explain what they meant by something they wrote earlier in the same message (such as paraphrasing oneself, or giving an example) and not in the sense of explaining ideas from earlier in the discussion. To provide further information about how discussion messages are linked to the article and subject matter, students' referring activities are coded as referring to other persons ("I don't think Jan is right"), as referring to the content of either messages or the text of the article ("I don't think that is true"), or as referring to a specific location in the article in either a long and comprehensive manner ("On page 2 just below the first paragraph") or a short and demonstrative manner ("Here"). The latter distinction was used because the use of demonstrative expressions can give information about what contextual information participants consider to be evident and shared with their peers. Finally, messages have been coded on the occurrence of social and regulative comments (cf., Veldhuis-Diermanse, 2002). Social comments comprise expressions such as "Hello everyone" and "Goodbye/ good luck" and regulative comments expressions like "Can anyone help me with this question?". Social comments are used to establish and support social relationships and regulative comments function to coordinate the discussion process.

Table 1

Instrument for Dialogue Analysis

Variable	Categories	Type[a]	Cohen's κ
Message subject	Meaning of the article		.79
	Overall	○	
	Structure	○	
	Statement		
	Argumentative	○	
	Non argumentative	○	
	Concept	○	
	Opinion of student	○	
	Non task-related	○	
Message type	Thread openers		.91
	Statement		
	Definite	○	
	Open	○	
	Question		
	Full	○	
	Empty	○	
	Critical reaction	○	
	Follow-up question		
	Full	○	
	Empty	○	
	Answer/ reaction	○	
Argumentation		☐	.82
Confirmation		☐	.83
Clarifying		☐	.75
(Connectivity)		☐	.36
Referring to a person		☐	.91
Referring to the content of another message		☐	.77
Referring to the content of the text		☐	.79
Comprehensively referring to a location in the article		☐	.79
Demonstratively referring to a location in the article		☐	.73
Social comments		☐	.90
Regulative comments		☐	.84

[a]"Type" shows whether only one of multiple categories is to be scored (expressed by an option button), or whether it is simply a "yes/no"-variable (expressed by a checkbox).

Questionnaires and group interviews

The questionnaire (see Table 2) was used to gather information about the usability of the tools, the perceived quality of the discussion, students' offline reading activities, their perceived levels of (mutual) understanding and their perception of the learning gains of the discussions. These processes were not directly visible in the protocols. Each item was intended to measure different information. The group interviews aim to get a feel for how the tool and study are experienced by the students, leaving room for a variety of input. They questioned students about the usability of the tools and their implementation in the course and focused on identifying possible disturbing factors for the design and results of the study.

Table 2

Questionnaire Items (for this purpose ordered in themes)

Theme	Item
Link between discussion and study material	1. How much of the article did you read before starting to participate in the discussion? [a]
	2. "Before I reply to a message, I always reread the relevant passage from the article." [a]
Communicative efficiency	3. "It was easy to understand the questions and answers of others." [b]
	4. "Others usually understood very well what I was trying to say or ask." [b]
	5. "In the discussions, there was a lot of miscommunication." [b]
Usability of the tool	6. "I find the discussion tool practical for discussing the article online." [b]
	7. "I experienced technical difficulties with the discussion tool." [b]
	8. How many of the total number of discussion messages did you read? [a]

Note. All measured as 5-point Likert type items. [a]For these items 1=0-20%; 2= 21-40%; 3=41-60%; 4=61-80%; and 5=81-100%. [b]For these items 1= "strongly disagree" to 5= "strongly agree".

Analyses

The unit of analysis is discussion level. This means that the data has been aggregated to be able to identify the characteristics of the discussion. In this, we corrected for the length of the discussion. T-tests, with an $\alpha=0.05$, tested the differences between the condition for regular forum discussion (with a score of 0) and the anchored discussion condition (with a score 1). For the variable's message type and subject, we divided this alpha by the number of categories to compensate for the larger number of statistical tests involved. In order to explore additional insights into the relation between tool (forum or anchored discussion), subject, and students' constructive activities, regression analyses have been performed.

Results

Use of the systems

Even though most students had no experience with either of the systems and their participation was voluntary, both discussion systems were frequently used, resulting in 514 messages in the system for forum discussion and 782 in the system for anchored discussion (excluding duplicate messages that are caused by accidentally clicking the send-button). The relative number of task-related messages in the two systems does not differ significantly (90% for the forum and 92% for the anchored discussion). In the 117 questionnaires that were collected, students did not report significant differences in technical difficulties with the two tools (with means of 1.34 in the forum and 1.70 in the anchored condition), or in the extent to which students found the tool practical to use (with means of 3.57 in the forum and 3.34 in the anchored condition). Furthermore, students did not report significant differences between the conditions in the percentage of messages that were read (with a mean of 4.43 in both the forum and in the anchored condition). Finally, the results from both the discussion protocols and the group interviews did reveal some technical issues with the system for anchored discussion (such as the placement of certain buttons) that should be resolved in any further development, but which did not seem to have hindered constructive use of the system.

Link between discussion and study material

First, the link between discussion and study material is indicated by the subject of conversation. The two systems show a significant difference in the extent to

which discussions are focused on the meaning of the article (see Table 3), $t(22.46)= 2.29$, $p=.032$. Second, the link between discussion and study material becomes visible in the object of referring activities. While discussions in the forum discussion more often contain referrals to persons, $t(34)= 6.10$, $p<.001$, discussions in the system for anchored discussion contain more direct referrals to actual content (ideas or statements) of other messages, $t(31.14)= -3.13$, $p<.01$ and of the text, $t(20.42)= -3.88$, $p<.01$. Third, while the questionnaires did not reveal a significant difference in the amount of literature the students had read before starting the discussion, they did show that the tool influenced students' reading activities. Students who used the system for anchored discussion ($M= 3.58$, $SD=1.02$), reported to have reread the relevant section of the article before replying to a message more often than those in the system for forum discussion ($M=2.77$, $SD=1.13$), $t(108)=-3.94$, $p<.001$.

Table 3

Percentages of Meaning-oriented Conversation and Objects of Referrals by Condition

Variable	Forum discussion	Anchored discussion
Message subject: the article's meaning	58	77
Referring to persons	34	13
Referring to content of another message	0	5
Referring to content of the text	5	12

Note. Percentages in all tables with variables from the coding scheme represent the number of messages containing the indicated activity in relation to the total number of messages within that condition. Variable numbers match the numbering as found in Table 1.

Communicative effectiveness

Regarding the clarification activities (see Table 4), we see that the students in the regular forum discussion use more clarifying statements, $t(34)= 5.48$, $p<.001$. Table 4 also shows that in the forum discussion we find more referrals that use comprehensive expressions (references that can be understood on their own without contextual information), $t(15.59)= 4.52$, $p<.001$, whereas in the system for anchored discussion we find more demonstrative referrals (that are more brief because of the use of expressions such as "here" or "that," as can be

seen in Figure 3), t(29.87)= -3.95, p<.001. Concerning the effects of these clarifications and referring efforts, the questionnaires do not present a significantly different level of mutual understanding in the two conditions.

Table 4

Percentages of Types of Referrals and Clarifications by Condition

Variable	Forum discussion	Anchored discussion
Clarifying	35	20
Comprehensive referring to a location in the article	15	2
Demonstrative referring to a location in the article	5	19

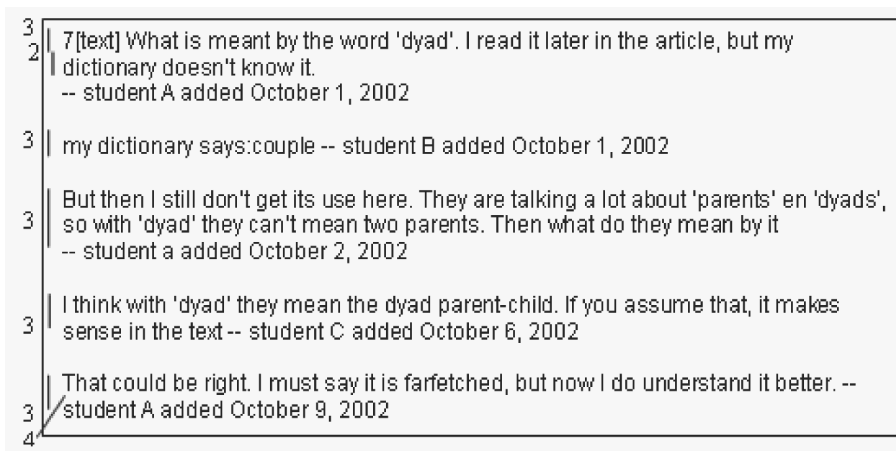

Figure 3. Thread from the system for anchored discussion, using demonstrative expressions.

Nature of discussion

In the forum discussion, the structure of messages generally resembles the structure of emails, with introductory and closing parts that include more than the actual question or remark about the subject matter (see Van der Pol, 2006, for examples and a more elaborate description). Messages in this condition also more often contain social statements, t(13.82)= 4.46, p=<.01, and regulative comments, t(14.49)= 3.91, p<.01, as can be seen in Table 5 on the next page.

Table 5

Percentages of Social and Regulative Statements by Condition

Variable	Forum discussion	Anchored discussion
Social comments	42	2
Regulative comments	25	4

The fact that the system for anchored discussion shows less social and regulative comments also expresses itself in the average message length. Taking the complete collection of messages (n=1342), we find that messages in the forum discussion have an average length of 57 words (see Table 6), whereas their counterparts in the system for anchored discussion, on average, consist of 38 words, $t(635.51)=14.17$, $p<.001$. However, a significantly larger number of messages in the system for anchored discussion ($t(28.49)= -2.13$, $p=.04$) almost completely compensates for this shorter average message length.

Table 6

Average Message Length and Number of Messages per Discussion by Condition

Variable	Forum discussion	Anchored discussion
Words per message	57.1	37.7
Messages per discussion	25.6	37.3

Finally, the group interviews revealed a slight difference between the systems with regard to the type of discussion it best supports. While the system for anchored discussion was experienced as particularly useful for the discussion of specific statements and concepts, the forum discussion was found to be better suited for more general discussion. However, the protocol analysis of the message subjects revealed no differences between the systems in the degree to which discussions are dedicated to the overall meaning of a text or to the meaning of specific concepts, nor in any of the other subcategories of the variable "message subject".

Constructive activities

Scores on the subcategories of the variable "message type" (2) did not differ significantly across conditions. However, the traditional forum discussion does show relatively more argumentations ($t(34)=3.25$, $p<.01$) and confirmations ($t(34)= 2.19$, $p=.036$) than does the system for anchored discussion (see Table 7).

Table 7

Percentages of Argumentations and Confirmations by Condition

Variable	Forum discussion	Anchored discussion
Argumentation	30	19
Confirmation	25	17

Expecting that it might be more difficult for students to support why they think that a passage from a text has a particular meaning than to support why they have a certain personal opinion, we checked the data for a relation between message subject and argumentation and found there is indeed less argumentation in meaning oriented discussion than in opinion-oriented discussion ($r=-.61$, $p<.001$, $n=36$). To examine whether the different level of argumentation in the two conditions, as established earlier, might be mediated by this relation, a linear regression analysis was conducted (as the data fit linear models) with both message subject and tool as predictors for argumentation (see Table 8 on the next page). From this, we find that the correlation between condition and argumentation decreases from .49 to .30 when controlling for message subject, which confirms a partial mediating effect of message subject on the relationship between condition and argumentation (cf. Baron & Kenny, 1986). Because this mediating effect of message subject might also exist for other dependent variables, they were checked for a possible correlation with message subject. Besides argumentation, message subject proved to correlate significantly with the percentage of confirmation ($r=-.48$, $p=<.01$, $n=36$). Additional regression analyses revealed that this correlation is no longer significant when controlling for message subject (see Table 9 on the next page), making the relationship between condition and the percentage of confirmations entirely mediated by message subject.

Table 8

Regression analysis for condition and message subject predicting argumentation (N=36)

Variable	B	SE	β	p
Step 1				
Condition	-1.09	.03	-.49	.003
Step 2				
Condition	-.07	.03	-.30	.043
Message subject	-.21	.06	-.49	.001

Table 9

Regression analysis for condition and message subject predicting confirmation (N=36)

Variable	B	SE	β	p
Step 1				
Condition	-.073	.03	-.35	.036
Step 2				
Condition	-.16	.07	-.20	.241
Message subject	-.04	.03	-.41	.017

Conclusion

Turning back to the formulated research questions, several conclusions can be drawn.

First, an increased percentage of meaning-oriented discussion, a more frequent referring to content, and a higher reported frequency of rereading relevant passages from the article indicate an affordance for anchored discussion to strengthen the link between discussion and study material. The cause for this enhanced link might be that the on-screen presence of the article, as well as the tool's specialized design, suggest to students that the discussion is to be focused on the meaning of the article. Put more generally, as conjectured in our description of anchored discussion, its design may have influenced students' perceptions of the collaborative goal and focused their collaborative intentions.

Second, the system for anchored discussion seems to offer an increased communicative efficiency. With briefer referrals (demonstrative rather than comprehensive) and messages containing fewer self-clarifications than the system for regular forum discussion, participants need fewer words to express their ideas. On a more theoretical level, these brief referrals can be said to

demonstrate the existence of a certain "frame of reference," as was expected. It is important to note, however, that this increase in communicative efficiency does not seem to lead to higher levels of mutual understanding, but rather seems to decrease the amount of effort that is required to reach this same level of mutual understanding.

Third, discussion shows a different general character in the two systems, devoting relatively more attention to establishing social relationships and regulating the collaborative processes in the regular forum discussion and being more straightforward and "to-the-point" in the system for anchored discussion (also resulting in a greater number of messages). This absence of social and regulative coordination in the system for anchored discussion can be interpreted as a reduced need for establishing a "call back pressure," (for a more elaborate description of this concept see Chapter two) which might again have been caused by a greater task-directedness as influenced by the system's functional design. Fourth, regarding the constructive activities, we found some differences between the conditions. Both argumentations and confirmations are found relatively more often in the forum discussion. However, as this can be (partially) explained by a stronger orientation towards opinion-oriented communication in the forum discussion, the relation between condition and the amount of argumentations and confirmations can be said to have been mediated by the subject of discussion.

Discussion

Since we did not find any alternative reasons for the differences between the discussion in both systems, we contribute the results to the affordances of the two systems, and in particular to their functional design. Although the presence of meaning-oriented, efficient communication does not yet guarantee collaborative learning, the presented results do seem to confirm that the practical affordances of anchored discussion make it a good starting point for supporting (the early stages of) collaborative literature processing. On the other hand, to stimulate a more personal and critical discussion that might be more beneficial for participants with sufficient levels of understanding of the subject matter, traditional forum discussion might be better suited. Furthermore, while this study focused on answering four rather concrete, theory-driven, research questions, the presented results might also have some broader implications for the field of CSCL.

First, it seems possible to distinguish two different types of discussion. While the exchange of personal opinions has been shown to be related to a more argumentative kind of discussion, a strengthened link between discussion and study material seemed to be accompanied by a more constructive discussion, moving towards seeing the subject of discussion from a third-person perspective. Since being able to stand back from one's personal viewpoints and trying to get the meaning of a message or article might be a crucial element in the early stages of collaborative text comprehension, it seems useful to distinguish "argumentation for opinion" from "argumentation for interpretation"; if trying to stimulate the latter, a more soft approach might be needed (Veerman, 2000). This means it could be important to support students to take a vulnerable position and express even their "dumb'" thoughts (Gay et al., 1999), as these can be seen as the only "building blocks" that are available to work with in trying to collaboratively advance their understanding of a text.

Second, an increased task-directedness in the system for anchored discussion was accompanied by a decreased amount of social and regulative communication. Taking a "grounding-is-functional" perspective on this finding would allow us to conclude that this condition presented students with sufficient levels of shared collaborative intentions and needed less explicit coordination of the process. Thus, although both socio-emotional and regulative processes are essential elements for successful CSCL (Mäkitalo, Salo, Häkkinen, & Järvelä, 2001), they do not necessarily need to manifest themselves in an explicit way. Especially in situations where sufficient levels of interdependency, trust and community have already been developed --for example in courses with abundant face-to-face contact between students-- facilitating social and regulative communication might not be necessary and a more direct facilitation of the process of collaborative knowledge construction might be more productive. Furthermore, this emphasis on a direct facilitation of the collaborative task might also be a more effective way to increase students' motivation. As Järvenoja and Järvelä (2005) show, social processes play a less important role in determining student's motivation and emotion, than aspects that are related to the accomplishment of the task.

Finally, we suggest that for an increased availability of anchored discussion in (university) practice[2], it could be useful to integrate functionality for anchored discussion into existing electronic learning environments.

Regarding the possibilities for future research, the obtained results encourage continuing research on and development of systems for anchored discussion. Furthermore, since the effects of anchored discussion on concrete learning results are still of a hypothesized nature, it would be useful to broaden and elaborate the techniques of analysis in order to assess the learning potential of certain patterns of interaction more closely. In particular, an operationalisation of the relevancy and usefulness of replies would make it possible to evaluate the content of online discussion more accurately. Finally, gaining some insight into a possible way of facilitating certain concrete aspects of the collaborative processing of literature in asynchronous CMC by no means "solves" all of its limitations and still leaves many opportunities for further realizing the potential of online discussion for collaborative learning.

References

Baker, M., Hansen, T., Joiner, R., & Traum, D. (1999). The role of grounding in collaborative learning tasks. In P. Dillenbourg (Ed.), *Collaborative learning: Cognitive and computational approaches* (pp. 31–63). Oxford: Pergamon.

Baron, R. M., & Kenny, D. A. (1986). The moderator-mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. *Journal of Personality and Social Psychology, 64*, 708–722.

Bernheim Brush, A. J., Bargeron, D., Grudin, J., Borning, A., & Gupta, A. (2002). Supporting interaction outside of class: Anchored discussion vs. discussion boards. In G. Stahl (Ed.), *Computer support for collaborative learning: Foundations for a CSCL community. Proceedings of CSCL 2002*. (pp. 425–434). Hillsdale, NY: Lawrence Erlbaum Associates.

Boekaerts, M., & Simons, P. R. J. (1995). Leren en instructie: Psychologie van de leerling en het leerproces [Learning and instruction: Psychology of the student and the learning process]. Assen: Van Gorcum.

Cadiz, J. J., Gupta, A., & Grudin, J. (2000). Using web annotations for asynchronous collaboration around documents. In D. G. Durand (Ed.), *Proceedings of the ACM Conference on Computer Supported Cooperative Work 2000* (pp. 309–318). New York: ACM Press.

Clark, H. H., & Brennan, S. E. (1991). Grounding in communication. In L. B. Resnick, J. M. Levine, & S. D. Teasly (Eds.), *Perspectives on socially shared*

cognition (pp. 127–149). Washington, DC: American Psychological Association.

Clark, D. B., Weinberger, A., Jucks, I., Spitulnik, M., & Wallace, R. (2003). Designing effective science inquiry in text-based computer-supported collaborative learning environments. *International Journal of Educational Policy, Research, & Practice, 4*(1), 55–82.

Davis, J. R., & Huttenlocher, D. P. (1995). Shared annotation for cooperative learning. In J. L. Schnase & E. L. Cunnius (Eds.), *Proceedings of CSCL 1995* (pp. 84–88). Bloomington. : Lawrence Erlbaum Associates.

De Laat, M. (2002). Network and content analysis in an online community discourse. In G. Stahl (Ed.), *Computer support for collaborative learning: Foundations for a CSCL community. Proceedings of CSCL 2002* (pp. 625–626). Hillsdale, NY: Lawrence Erlbaum Associates.

Dillenbourg, P. (1999). What do you mean by collaborative learning? In P.

Dillenbourg (Ed.), *Collaborative-learning: Cognitive and computational approaches* (pp.1–19). Oxford, UK: Elsevier.

Fay N., Garrod, S. C., & Carletta J. (2000). Group discussion as interactive dialogue or as serial monologue: The influence of group size. *Psychological Science, 11*(6), 481–486.

Fussell, S. R., & Benimoff, N. I. (1995). Social and cognitive processes in interpersonal communication: Implications for advanced telecommunications technologies. *Human Factors, 37,* 228–250.

Gay, G., Sturgill, A., Martin, W., & Huttenlocher, D (1999). Document-centered peer collaborations: An exploration of the educational uses of networked communication technologies. *Journal of Computer-Mediated Communication, 4*(3).

Gergle, D., Kraut, R. E., & Fussell, S. R. (2004). Language efficiency and visual technology: Minimizing collaborative effort with visual information. *Journal of Language and Social Psychology, 23,* 491–517.

Gunawardena, C. N., Lowe, C. A., & Anderson, T. (1997). Analysis of a global online debate and the development of an interaction analysis model for examining social construction of knowledge in computer conferencing. *Journal of Educational Computing Research, 17*(4), 397–431.

Guzdial, M., & Turns, J. (2000). Effective discussion through a computer-mediated anchored forum. *Journal of the Learning Sciences, 9*(4), 437–469.

Herrmann, T., & Kienle, A. (2003, January). *Integration of communication, coordination and learning material _ A guide for the functionality of collaborative learning environments.* Paper presented at the 36th Annual Hawaii International Conference on System Sciences, Hilton Waikoloa, Hawaii.

Hewitt, J., & Teplovs, C. (1999). An analysis of growth patterns in computer conferencing threads. In C. Hoadley & J. Roschelle (Eds.), *Designing new media for a new millennium: Collaborative technology for learning, education, and training. Proceedings of CSCL 1999* (pp. 232-241). Mahwah, NJ: Lawrence Erlbaum Associates.

Hunt, R. A. (1998). Electronic discussions in learning and teaching: Why they don't work, and how they might. *Connexions: The Newsletter of the International Society for the Exploration of Teaching Alternatives, 10*(2), 1–7.

Järvelä, S. & Häkkinen, P. (2000). Levels of Web-based discussion: Theory of perspective-taking as a tool for analysing interaction. In B. Fishman & S. O'Connor-Divelbiss (Eds.), *Proceedings of the Fourth International Conference on the Learning Sciences* (pp. 22–26). Mahwah, NJ: Erlbaum.

Järvenoja, H., & Järvelä, S. (2005). How students describe the sources of their emotional and motivational experiences during the learning process: A qualitative approach. *Learning and Instruction, 15*(5), 465–480.

Laurillard, D. (1993). *Rethinking university teaching.* London: Routeledge.

Lipponen, L. (2001). *Computer-supported collaborative learning: From promises to reality.* Unpublished doctoral dissertation, University of Turku, Turku, Finland.

Mäkitalo, K., Salo, P., Häkkinen, P., & Järvelä, S. (2001, August). *Analysing the mechanisms of a common ground in Web-based interaction.* Paper presented at the JURE pre-conference of the 9th European Conference for Research on Learning and Instruction, Fribourg, Switzerland.

McLoughlin, C., & Luca, J. (2000). Cognitive engagement and higher order thinking through computer conferencing: We know why but do we know how? In A. Herrmann & M. M. Kulski (Eds.), *Flexible futures in tertiary teaching. Proceedings of the Ninth Annual Teaching and Learning Forum.* Perth, Australia: Curtin University of Technology.

Moon, J. A. (1999). *Reflection in learning and professional development.* London: Stylus Publishing Inc.

Nokelainen, P., Miettinen, M., Kurhila, J., Floréen, P., & Tirri, H. (2005). A shared document-based annotation tool to support learner-centered

collaborative learning. *British Journal of Educational Technology, 36*(5), 757–770.

Pena-Shaff, J., & Nicholls, C. (2004). Analyzing student interactions and meaning construction in computer bulletin board discussions. *Computers & Education, 42*(3), 243–265.

Petraglia, J. (1998). *Reality by design: The rhetoric and technology of authenticity in education.* Mahwah, London: Lawrence Erlbaum Associates.

Phillips, D. C. (1995). The good, the bad, and the ugly: The many faces of constructivism. *Educational Researcher, 24*(7), 5–12.

Pilkington, R. (2004). Developing discussion for learning. *Journal of Computer Assisted Learning 20*(3), 161–164.

Schoonenboom, J. (2002). A template for discussing large texts on the web: The Pragglejaz site. *British Journal of Educational Technology, 33*(1), 103–107.

Schwartz, D. L., & Lin, X. D. (2000). Computers, productive agency, and the effort after shared meaning. *Journal of Computing in Higher Education, 12*(2), 3–33.

Selzer, W. (2000). Annotation Engine (Technical report). Cambridge, MA: Harvard University, Berkman Center for Internet & Society.

Simons, P. R. J., Van der Linden, J., & Duffy, T. M. (2000). New learning: Three ways to learn in a new balance. In P. R. J. Simons, J. van der Linden, & T. M. Duffy (Eds.), *New learning* (pp. 1–20). Dordrecht, The Netherlands: Kluwer Academic Publishers.

Sringam, C. & Geer, R. (2000, September). *An Investigation of an instrument for analysis of student-led electronic discussions.* Paper presented at ASCILITE 2000, Coffs Harbour, Australia.

Stahl, G. (2000). A model of collaborative knowledge-building. In B. Fishman, & S. O'Connor-Divelbiss (Eds.), *Proceedings of the Fourth International Conference of the Learning Sciences* (pp. 70–77). Mahwah, NJ: Lawrence Erlbaum Associates.

Sumner, T., & Buckingham Shum, S. (2001 + datum). JIME: An interactive journal for interactive media. *First Monday*, *6*, (2). Retrieved datum + url

Suthers, D. (2005) Technology affordances for intersubjective learning: A thematic agenda for CSCL. In T. Koschmann, D. Suthers, & T. W. Chan (Eds.), *Computer Supported Collaborative Learning 2005: The Next 10 Years* (pp. 135–144). Mahwah, NJ: Lawrence Erlbaum Associates.

Takeda, T. & Suthers, D. (2002, May). *Online workspaces for annotation and discussion of documents*. Poster session presented at WWW 2002, Honolulu, Hawaii.

Van der Pol, J. (2002, January). *Identifying and modeling variables in complex CSCL-situations. Case study: The use of asynchronous electronic discussions.* Paper presented at the Conference on Computer Support for Collaborative Learning (CSCL2002), Boulder, CO.

Van der Pol, J., Admiraal, W. F. (2003). Het succesvol inzetten van asynchrone elektronsiche discussies [Succesfully implementing asynchronous online discussions]. *Onderzoek van Onderwijs, 32*(2), 26–31.

Van der Pol, J., Admiraal, W. F., & Simons, P. R. J. (2006). Context enhancement for co-intentionality and co-reference. *AI & Society, 20*, 301-313 .

Veerman, A. L. (2000). *Computer-Supported collaborative learning through argumentation.* Utrecht, The Netherlands: Utrecht University.

Veldhuis-Diermanse, E. (2002). *CSCLearning?* Unpublished doctoral dissertation, Utrecht University, Utrecht, The Netherlands.

Wan, D., & Johnson, P. M. (1994). Experiences with Clare: A computer supported collaborative environment. *International Journal of Human-Computer Systems, 41*, 851–879.

Warschauer, M. (1997).Computer-mediated collaborative learning: Theory and practice. *Modern Language Journal, 81*(3), 470–481.

[1] The term "pedagogy," in Dutch, is used to refer to the science that is concerned mainly with educating or rearing children, and deals with normative questions such as what to learn (opposed to the more instrumental approach of educational science that is directed mainly at how to learn).

[2] For inquiries about the practical availability of systems for anchored discussion, please contact j.vanderpol@ond.vu.nl, or find a working system at http://www.annotatiesysteem.nl/.

Chapter 4

Peer evaluation in online anchored discussion for an increased local relevance of replies*

Abstract

This study investigates the use of an evaluation function to increase the local relevance of replies in online anchored discussion. Being implemented in a university course on French linguistics, a regular system for anchored discussion is compared with two versions that are enhanced with an integrated evaluation function. The function asks students to evaluate the relevance of each others' replies. To compare between experimental and control conditions, the collaboration protocols are analyzed with a newly developed coding scheme for the local relevance of replies. Statistical processing of the data is done with a multilevel approach and results indicate that an evaluation function can effectively increase the local relevance of students' replies, but only if it is actually used to a certain extent. Using the evaluation function is hypothesized to increase students' awareness of the importance of writing relevant replies.

---

* Van der Pol, J., Admiraal, W. F., & Simons, P. R. J. (2007). *Peer evaluation in online anchored discussion for an increased local relevance of replies*. Manuscript submitted for publication.

Computer supported collaborative learning (CSCL) offers students possibilities for a deep and active processing of their subject matter, especially when complementing face-to-face (F-2-F) interaction (Dietz-Uhler & Bishop-Clark, 2001; Häkkinen, Järvelä & Mäkitalo, 2002). As stated by Lapadat (2002), asynchronous online discussion in particular facilitates reflection, conceptual change and the collaborative construction of meaning, making it especially suitable for the collaborative processing of academic literature. This study - which aims to facilitate students' collaborative processing of literature - will use a specialized form of online discussion called "anchored discussion" (Bernheim Brush, Bargeron, Grudin, Borning, & Gupta, 2002), that integrates students' online discussion with the subject matter that is being discussed. Van der Pol, Admiraal, and Simons (2006) have demonstrated that anchored discussion is better suited for supporting the early stages of collaborative text processing than regular forum discussion, in that it affords a more efficient and meaning-oriented collaboration by relating the discussion more closely with the subject matter.

However, anchored discussion might still have some constraints with regard to the coherence of students' interaction and the learning potential of their peer-directed replies. In this article, we develop and investigate an enhanced tool for anchored discussion that aims to increase the quality of students' replies.

Constructive learning conversations and the relevance of replies

While the level of interactional or "cross-turn" coherence of students' online learning conversations seems to be closely related to their effectiveness for collaborative learning (Hoadley & Enyedi, 1999; Hsi, 1997), it is also often identified as a problematic element (Herring, 1999; Reyes & Tchounikine, 2003; Van der Meij, De Vries, Boersma, Pieters, & Wegerif, 2005). This is understandable, because an inherent aspect of a collaborative learning situation is that students - who usually do not yet fully master the subject matter - act as both the creators and the receivers of feedback. A limited level of understanding sets high demands for the quality of feedback students require, as it might otherwise be difficult for them to correctly interpret and use it. Yet, at the same time, it also means that providing high quality feedback is more

difficult. Webb and Mastergeorge (2003) examined the effectiveness of helping behavior in peer-directed groups more closely, and they found, amongst others, that effective help seekers ask precise questions and effective help givers provide detailed explanations of the material. In previous research on using online discussion for the collaborative processing of texts Van der Pol (2002) found that (undergraduate) students do experience difficulties with providing each other with specific feedback, especially when they are processing difficult texts in a foreign language. Students' responses were often found to be only associatively linked to each other, leading rather to a collection of different ideas than to the building and deepening of ideas. Bellamy (1997) already stated that a constructive theoretical learning conversation should consist of more than just information exchange. A concrete example of what this "more" may be can be found in Sperber and Wilson's (2004) "Relevance Theory". Viewing the pursuit of relevance as a central mechanism in human communication, Sperber and Wilson define the relevance of new information as the degree in which it has potential cognitive effects for the person who receives it[1]. According to Sperber and Wilson:

> "Something is relevant to an individual when it connects with background information he has available to yield conclusions that matter to him: say, by answering a question he had in mind, improving his knowledge on a certain topic, settling a doubt, confirming a suspicion, or correcting a mistaken impression" (2004, p. 608).

In order to look at the quality of replies in students' online learning conversations, we will use Sperber and Wilson's (2004) account of relevance to create the basic framework for a coding scheme for the relevance of replies. However, our coding scheme will take a somewhat different perspective on relevance than Sperber and Wilson in attempting to create a more intersubjective and third-person view on the relevance of replies. Because an online learning conversation usually consists of multiple parties that actively rotate the roles of senders, receivers and "overhearers" (see Clark, 1992), it needs to be semantically coherent in order for all parties to benefit from it. In following this less personal and more semantic approach, we will label our

[1] (divided by the amount of effort it takes to derive these implications).

coding scheme as indicating the "local relevance of replies" (see Herring, 1999), defined as the degree in which new ideas hold potential implications for earlier expressed ideas.

The constraints of online discussion for achieving local relevance

The apparent limited affordance of online learning conversations for maintaining the local relevance of replies may not only be related to students' limited understanding of the subject matter, but also to some of its basic communicative constraints. As stated by Janssen, Erkens, Jaspers, and Broeken (2006), computer-mediated discussions have a low "media richness" which may cause students to have interpretation difficulties and hinder them in conducting constructive discussions. Not only do online discussions generally lack non-verbal information, but their delayed nature and the relatively high investments for writing messages (compared to synchronous or F-2-F communication) also present a limited degree of interactivity or "turn-taking" (Smith, 2003). As mentioned by Levinson (1983), interactional coherence in F-2-F communication is actively constructed by participants across turns. Consequently, the limited number of turns in online discussion affects this continuous alignment of ideas, which may account for the natural tendency of electronic discussions to diverge (Andriessen & Sandberg, 1999; Mäkitalo, Häkkinen, Salo, & Järvalä, 2001).

In addition, the limited degree of turn-taking in online discussion means that participants cannot rely on repairing any misalignments after they arrive, as in F-2-F communication, but have to put more emphasis on preventing any misalignment in advance. This normally works by a process known as "audience design" (see Clark, 1992), by means of which a person ensures the relevance of a message for the receiver. However, the "many-to-many"-character of online discussion means that participants may design their messages to be relevant in relation to the general topic of conversation, requiring a less strict alignment than when designing for relevance in relation to specific individual ideas. Fay, Garrod, and Carletta (2000) have already demonstrated this effect of many-to-many communication on audience design in a F-2-F situation. Communication in small groups showed mainly a bilateral process among pairs of communicators who were primarily sensitive to their current conversational partner, whereas communication in large groups consisted of a more unilateral process of broadcasting information to the group

at large. In addition, as computer-mediated communication (CMC) generally offers users more social freedom to react associatively and *"(…) the standards of local relevance are weakened, relative to spoken conversations in which logical adjacency is more reliably preserved"* (Herring, 1999, p. 11), it becomes clear that online discussion holds quite some constraints for producing locally relevant replies.

Increasing local relevance by providing an evaluation function

Having identified the local relevance of replies as a crucial but problematic element in creating constructive online theoretical learning conversations, we may ask ourselves how to facilitate or stimulate the creation of relevant replies. A known approach to improve the quality of students' online interaction is to develop tools that provide the participants with information about their collaboration, described by Soller, Martinez, Jermann, and Mühlenbrock (2005) as mirroring and awareness tools. Even without offering any concrete advise, simply making certain characteristics of students' collaboration explicit and possibly accompanying this with some form of ideal configuration can be a powerful means of influencing students' actions and intentions.

In our case, providing additional feedback on the relevance of students' replies seems a suitable approach, as the decreased level of turn-taking and the fewer channels of communication in online asynchronous discussions provide users with relatively little information on the knowledge and perspectives of other users (see also Nückles & Stürz, 2006) and on the degree to which they experience certain replies as relevant. In addition, as providing feedback on the relevance of a reply attributes a certain value to it, this can also function as a reward system, which can effectively increase the quantity and quality of students' contributions, as shown by Cress, Barquero, Schwan, and Hesse (in press) and by Hummel, Burgos, Tattersall, Brouns, Kurvers, and Koper (2005). Instead of trying to computationally generate information about the relevance of replies, we will use the input of users themselves. Compared to traditional mirroring systems, this has the additional benefit that it actively involves students in thinking about the relevance of replies, which might help to make them more aware of its importance for their collaborative learning process. Subsequently, this awareness might lead students to try to design their messages to have more implications for specific existing ideas and thus increase the local relevance of their replies.

In order to investigate these potential effects of feedback on the relevance of student' replies, we developed a system for anchored discussion (see Fig. 1) with an optional evaluation function. The system displays both students' subject matter (in the right-hand side of the screen) and students' threaded learning conversations (on the left). The highlighted and numbered passages in the document function as anchors for specific threads and form the link between discussion and content. As is visible in Fig. 1, the evaluation function that was integrated offers students a range of 1 to 5 to indicate the relevance of a reply by selecting a smiley of their choice. Because we were unsure in what form the evaluation function would be most effective, two different versions were developed. The first version reserves the evaluation option for the receiver of a reply, automatically popping up with every reply that this particular user receives, and allowing users to provide an additional argumentation for their evaluation. We have developed this version because it seems natural to ask the receiver of a reply to evaluate its relevance, as it is a direct response to his or her message. However, as this function limits the evaluation opportunity to a limited number of people, it may not have the same effect as when the opportunity to evaluate a certain reply is open to all participants. Therefore, the second version we developed offers all participants the option to evaluate any reply. As the layout of the discussion area does not provide enough space to open up the evaluation option automatically, this second version presents it after clicking the appropriate evaluation link. In addition, as a single message can now be evaluated multiple times, technical limitations dictate that in this version no additional argumentations can be provided.

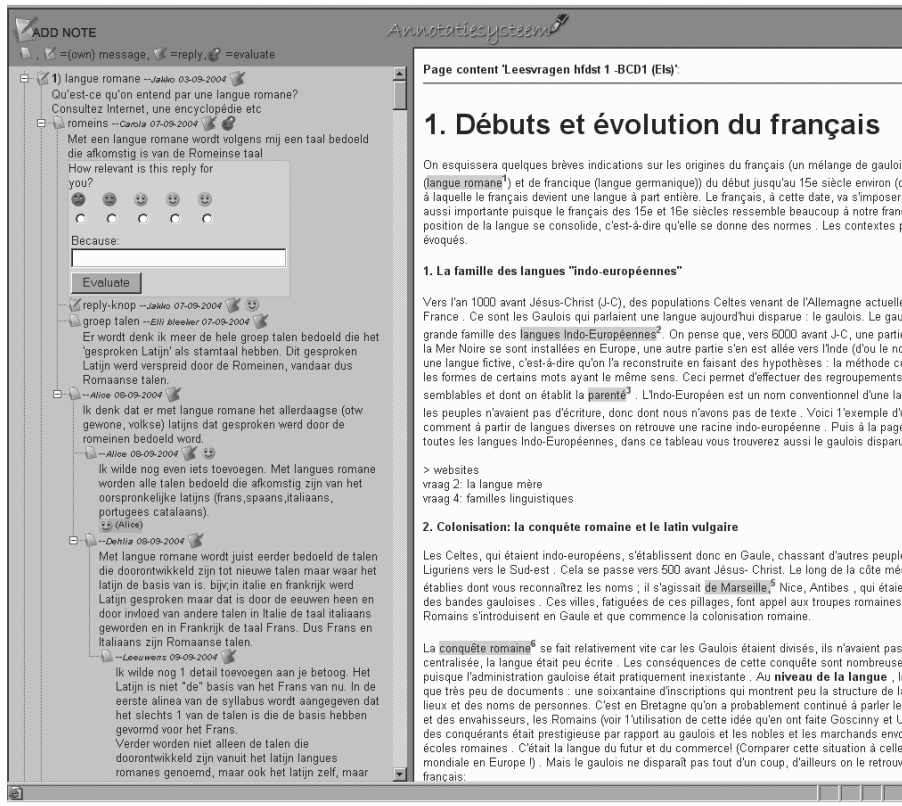

Figure 1. Partial print screen of the developed tool for anchored discussion (available at www.annotatiesysteem.nl), displaying the first version of the evaluation function: "the receiver evaluates".

Method

We have set up a comparative study with three types of anchored discussion: two versions with an additional evaluation function and one regular version without an evaluation function, to function as a control condition. Our hypothesis is that replies in the anchored discussion with an evaluation function will be more locally relevant than in the discussion groups without an evaluation function. Because we do not yet know which specific design of the evaluation function can be expected to influence the local relevance of replies the most effectively, we will test two evaluation functions that differ on how and to whom the function is accessible.

Data

This study was conducted in a Dutch university course titled "Introduction to French Linguistics". For 15 weeks the students participated in weekly seminars for which the total of 38 students was randomly divided into 4 separate groups. During the course the students had to read a syllabus (in French) based on a weekly planning. To deepen students processing of the syllabus, detailed reading questions were provided. Instead of answering these questions individually, as had been done in previous years, this study provided online discussion of the reading questions for the syllabus to stimulate a more collaborative processing of the material. Several chapters of the syllabus that were known to be difficult were selected by the teacher for online collaboration during fixed periods within the course accumulating to a total discussion period of six weeks. Two experimental groups used the described system for anchored discussion with an evaluation function, and two control groups used the same system without an evaluation function. All four groups operated independently (both on- and offline) and the online discussions were monitored by the teacher, but minimally moderated. Within the experimental condition the two different versions of the evaluation function were tested in subsequent discussion periods, forming two different phases within the study (see Table 1).

Table 1
Evaluation function per study phase

	Study Phase	
Condition	Phase 1	Phase 2
Experimental (student groups 3 and 4)	Evaluation function 1: "the receiver evaluates"	Evaluation function 2: "everyone evaluates"
Control (student groups 1 and 2)	No evaluation function	No evaluation function

Prior to the start of the discussions, students in all groups received the same document explaining how to get started with the tool and describing the goal of the discussions. No separate instruction was given for the evaluation function, which students were left to discover themselves. Students had to provide initial reactions to the reading questions and had to try to

collaboratively improve them. If their participation was sufficient, students could earn the privilege of dropping a question on the course's final exam.

Procedure

The main research data consist of the online collaboration protocols, which will be investigated in different steps. First, we will perform a selection of the data including only student-generated replies, excluding teacher-generated messages, and the first message of every thread (which will not be scored itself, but which will be used in determining the relevance of its replies). In addition, we will select the replies that are aimed at developing students' understanding of the meaning of the syllabus, as we have identified relevance of replies to be an important issue within online conversations that are already meaning-oriented. Only these will be analyzed on their local relevance. For this selection, the variable "message subject" from the coding scheme of Van der Pol, Admiraal, and Simons (2006) will be used, for which the authors already established a sufficient interrater reliability (Cohen's κ =.79). After this first selection of data, the remaining replies will be scored on their local relevance using a coding scheme, which will be constructed in the next section. Finally, the resulting local relevance scores in the different conditions will be compared using multilevel analyses.

Measuring the local relevance of replies

As no existing schemes for measuring the local relevance of replies were found, a new coding scheme has been developed. This new coding scheme is based both on Sperber and Wilson's (2004) description of relevance, as mentioned in the introduction, and on Veerman's (2000) coding scheme for the "constructive activities" in students' online theoretical learning conversations, as this also looks specifically at the relation between newly added information and earlier expressed ideas.

In our coding scheme we want to focus primarily on the relation between a reply and existing ideas as this seems to be a central aspect of the notion of relevance and only ideas that are related to and significant for existing ideas can be said to hold potential implications for answering questions, either complementing existing knowledge, or adjusting mistaken impressions. Secondly, we will investigate whether a reply contains any new ideas or merely states (dis)agreement with ideas that already exist. While messages that do relate

but do not provide any new ideas can be useful for the knowledge construction process in allowing for certain conclusions to be drawn or doubts to be settled, they are still limited in their ability to develop students' knowledge and deepen their understanding of the subject matter. In contrast, if a reply that connects to previous ideas does contain new information, it can be expected to hold stronger implications for the existing ideas it addresses. Thirdly, we divide these replies with related and new ideas into replies that either complement or questions existing ideas. As learning is usually associated with conceptual change (Roschelle, 1992) and the pursuit of newness (Paavola, Lipponen, & Hakkarainen, 2004), we will consider a negative evaluation (that helps to revise certain existing ideas) as being more locally relevant than a positive evaluation (that helps the receiver in confirming the ideas he or she already had).

In sum, by hierarchically applying these three criteria, the framework for a new coding scheme emerges (see Table 2). The four resulting categories can be used to order the different elements from Sperber and Wilson's (2004) description into a hierarchical coding scheme for the local relevance of replies.

Table 2

Hierarchical framework with four different degrees of local relevance

Reply categories	Potential cognitive effect
1: *Not relating* to existing ideas.	
2: Relating to existing ideas, but *not adding* new ideas.	Confirming a suspicion, settling a doubt
3: Relating to existing ideas and adding new ideas that *complement* existing ideas.	Answering a question, improving knowledge on a certain topic
4: Relating to existing ideas and adding new ideas that *adjust* existing ideas.	Correcting a mistaken impression

Next, a preliminary investigation of part of the data will be used to get a better view on how students shaped their collaboration in reaction to the task and to specify the four categories accordingly in order to make them sufficiently discriminative and reliably applicable. For this we will take the messages in the first thread of the first discussion round in all four groups. We see that in group 1 (control condition), the discussion of the first reading

question *("Qu'est-ce qu'on entend par une langue romane? Consultez Internet, une encyclopédie etc.")*, starts with an initial answer from student A[2]:

> answer – Student A (Sept. 6, 2004)
> I had some time today and have an answer to question 1:
> Romane Se dit des langues derivees (branch) du latin populaire
> Dictionary [..here we find the copied entry from a dictionary..]

To which, amongst others, two students respond by expressing their agreement:

> correct! – Student B (Sept. 6, 2004)
> I think that this is a good answer!!Seeing that you took it out of the dictionary I don't want to argue with you, haha!
>
> – Student C (Sept. 7, 2004)
> Cette réponse est parfait!

In group 3 (experimental condition), we also find replies that aim to complement a provided answer:

> – Student D (Sept. 8, 2004)
> I just wanted to add something. Langues romane refers to all languages that stem from the original Latin (French, Spanish, Italian, Portugese, Catalan).
>
> – Student E (Sept. 9, 2004)
> I wanted to add 1 detail to your argument. Latin is not "the" basis of contemporary French. In the first paragraph of the syllabus is indicated that it is only one of the languages that have formed the basis for the French language.
> Furthermore does one not only call the languages that are developed from Latin langues romanes, but also Latin itself, but that is silly.

[2] In these messages the original Dutch text has been translated into English (French remaining French), and in some cases partially summarized (indicated in brackets).

Looking more closely at the second message, we see that it not only adds information, but also refines a previous idea. This way of partially adjusting a previous answer is also visible in this message from group 4 (experimental condition):

> – Student F (Sept. 7, 2004)
> The Roman language was spoken by both the upper and lower classes and invaders also copied the Roman language, by means of which it was continued, and thus it was not simply a continuance of the Latin that was only spoken by the common people.

From this preliminary view on the collected data, we see that students' messages primarily consist of answers to the reading questions, in which they try to complement and correct each other where possible, to collaboratively construct an answer that is as complete and accurate as possible. This picture of students' collaboration seems to fit the framework for the local relevance replies that we created (see Table 2).

This insight in the nature of the collected data also allows us to refine the coding scheme further, by establishing some more concrete rules for the coding process (see Table 3). First, as messages sometimes proved to connect more to messages that are higher up in the tread than to the one that they technically react to, we will look at all messages in the thread to assess whether a connection exists to any earlier idea or statement. Thus, "no connection"-replies (Category 1) are often replies that ignore all earlier posted messages and directly address the original reading question. Second, to determine whether a message relates any new ideas to the idea it connects to, we will specifically be looking for things such as additional examples or argumentations. A reply that merely contains a participants' (dis)agreement, check, or bare follow-up question, will, although it does present some new *information* (for instance whether others agree or not), not be coded as containing any new *ideas* (and will be scored in Category 2). A question is regarded as "bare" if it does not contain a suggestion for a possible answer, in which case it will be regarded as introducing a new idea (however carefully it may be presented). Third, to determine whether a reply (which adds new ideas that connect with existing ideas) conflicts with questions or adjusts the idea it relates to (Category 4) instead of merely complementing it (Category 3), one can ask the question

whether - from a semantic point of view - both ideas could be true at the same time or not.

As the coding scheme is hierarchical in nature each level also contains the criteria of the preceding stages and where multiple categories apply, the highest one will be scored. Regarding its inter-coder reliability, letting two coders apply the coding scheme independently to 32 replies, after a limited training on the instrument as well as the contents of the discussion, resulted in a Cohen's κ of .74.

Table 3
Coding scheme for the local relevance of replies

Category	*Description*	*Example*
1: No connection	Is a direct reaction to the task, ignoring all preceding messages.	
2: Bare reply	Connects to an existing idea, without adding any new ideas.	- expressing (dis)agreement without new arguments or elaboration - a bare follow-up question
3: Addition	A reaction with new ideas that complement the ideas in an earlier message.	- providing additional examples - answering a question - confirming an earlier answer with new arguments or elaboration
4: Adjustment	A reaction with new ideas that adjust or question the ideas in an earlier message.	- expressing disagreement with a supporting argument

Statistical Analyses

In order to investigate whether the developed evaluation functions increase the local relevance of replies we will compare the local relevance scores in both experimental conditions with the scores in the control condition in the respective phase. Because messages are nested within students, who are nested within groups, we cannot assume the scores of individual messages to be independent of each other (e.g. messages of the same student may be more similar than messages from different students), and a multilevel approach will be taken. The model we will construct for each phase will test whether the independent dichotomous variable "condition" (with or without an evaluation

function) can explain a significant amount of the variance in local relevance scores. As the latter is measured on an ordinal scale we performed an ordinal multilevel analysis with a logit-link function and second-order penalized quasi-likelihood (PQL) estimation using the HLM program (for an extensive treatment of the methodology for ordinal multivel analysis, see Raudenbush & Bryk, 2002).

Investigation of the data shows that in both phases different messages from the same student are not totally independent. Taking "student" as a second level significantly improves the fit of the model. Besides the student level, the data show no significant amount of variance on a discussion level, nor on the level of individual threads. This investigation also shows that the additional variables we have information on ("discussion round", "reading question", "position of message in thread", and the number of given and/ or received evaluations per student), do not explain a significant amount of the variance in local relevance scores. Hence, the final model that will be used in both phases to analyze the variance in local relevance scores, is a 2-level variance only model with "condition" as explaining variable. The tests will be done one-directional with $\alpha = 5\ \%$.

Results

After describing students' participation in the online learning conversations in general we will quantitatively compare the local relevance scores (our dependent variable) in the experimental and control conditions. These scores result from coding the messages according to the developed coding scheme and will function as dependent variable in the analyses. Finally, we will give a view on how students used the evaluation functions by looking at the provided argumentations.

General participation

After coding the student-generated replies on their subject for the first selection of data, we see that the majority of students' replies is aimed at developing a better understanding of the syllabus. A total of 793 replies were posted in the two phases of the study, generated by 28 students (which gives an average of almost 5 messages per student per week). This substantial amount of messages (with an average of 4.7 messages per student per week), suggests that the online learning conversations were successfully implemented in the course.

From these 793 replies only 87 were excluded as not being meaning-oriented, of which 12 were coded as opinion-oriented and 75 as non-task related. The opinion-oriented replies all responded to an opinion-oriented reading question and the non-task related replies mostly consisted of regulative messages that commented on reasons for not (yet) participating in the discussion, the difficulty of the reading questions, typing errors, or websites that did not function correctly.

Comparing the local relevance scores between conditions

We will now compare the local relevance scores between the experimental and control conditions for each of the two phases of the study. It should be noted that these scores cannot be compared across the two phases because the discussions concerned different chapters of the syllabus, varying in difficulty and likely to produce different relevant scores. A more difficult chapter, for instance, leaves more room for students to complement and adjust each others' ideas than an easier chapter, where a satisfying conclusion may sooner be reached.

Phase 1: "The receiver evaluates"

The first phase of the study, where students in the experimental groups used the version of the evaluation function where only the receiver of a reply can evaluate, yielded a total of 307 student-generated meaning-oriented replies. Table 4 displays the local relevance scores these replies received after coding. In the experimental condition, 37 replies (or 22.3 %) were evaluated.

Table 4

Relative number of replies scored per category of local relevance in Phase 1

	Group	
Category	Control (no evaluation function)	Experimental ("the receiver evaluates")
1) No connection	29,1 %	9,6 %
2) Bare reply	33,3 %	18,7 %
3) Addition	32,6 %	55,4 %
4) Adjustment	5,0 %	16,3 %

Constructing an empty two-level model of these message scores (see Table 5) shows that a significant amount of the total variance can be attributed to the student level. Using the variance component of .92 we can calculate that 22 % of the total variance in the message scores lies at a student level (.92/{.92+3.29}=.22, calculated cf. Hox, 2002). Adding the independent variable "condition" to the model (see Table 6) shows that condition explains a significant amount of variance in the local relevance scores. When comparing the variance component in both models we see that it drops from .92 to .35 and hence the percentage of variance on a student level that is explained by condition is 62 % ({.92-.35}/.92=.62, calculated cf. Bryk & Raudenbush, 1992). The percentage of variance on the level of individual messages explained by condition is 14 % (.62*.22=.14). Note that the estimates cannot be interpreted as the estimates in normal multilevel analysis due to the nature of the link function. We can interpret the direction of the effect, but this is opposite to what we find in the model (with negative coefficients indicating positive relationships). This reversal is caused by the fact that the thresholds for the different categories have been estimated using the highest category as baseline.

Table 5

Empty multilevel model of local relevance scores for Phase 1

Fixed effect	coefficient	(SE)		p
Intercept	-1.46	(.27)		.000
	Variance component			
Random effect		*df*	χ^2	*p*
student-level error	0.92	21	84.60	.000

Table 6

Multilevel model with "condition" explaining local relevance scores for Phase 1

Fixed effect	coefficient	(SE)		p
Intercept	-0.78	(.26)		.008
Condition	-1.53	(.37)		.001
	Variance component			
Random effect		*df*	χ^2	*p*
student-level error	0.35	20	40.67	.004

Phase 2: "Everyone evaluates"

Phase 2 of the study, where the experimental condition tested the evaluation function which offered all the students the possibility to evaluate each other's replies, yielded a total of 411 student-generated meaning-oriented replies. Table 7 displays the local relevance scores these replies received after coding. In this phase of the study only 1 reply (or 0.7 %) was evaluated.

Table 7

Relative number of replies per category of local relevance in Phase 2

	Group	
Category	Control	Experimental ("everyone evaluates")
1: No connection	11,5 %	13,4 %
2: Bare reply	37,4 %	24,8 %
3: Addition	37,8 %	40,9 %
4: Adjustment	13,4 %	20,8 %

Constructing an empty two-level model of these message scores (see Table 8) shows that 7 % of the total variance can be attributed to the student level (.26/ {.26 + 3.29} = .07, calculated cf. Hox, 2002). Adding the independent variable "condition" to the model shows that in the second phase condition does not explain a significant amount of variance in the local relevance scores (see Table 9).

Table 8

Empty multilevel model of local relevance scores for Phase 2

Fixed effect	coefficient	(SE)		p
Intercept	-2.01	(.19)		.000
	Variance component			
Random effect		*df*	χ^2	*p*
student-level error	0.26	21	52.33	.000

Table 9

Multilevel model with "condition" explaining local relevance scores for Phase 2

Fixed effect	coefficient	(SE)		p
Intercept	-1.89	(.22)		.000
Condition	-0.33	(.30)		.286
	Variance component			
Random effect		*df*	χ^2	*p*
student-level error	0.25	20	48.86	.001

Students' use of the evaluation function

Of the 37 evaluations that were given in the first phase of the study 9 evaluations included an additional argumentation (provided by 5 different students). To provide a better view on how these students used the evaluation function we have collected these argumentations in Table 10. These data seem to indicate that students evaluate the relevance of a message mainly by its content and by the question whether it helps them to deepen their understanding of the subject matter. This way of interpreting the relevance of a reply as its "usefulness for their learning process" largely overlaps with the framework we constructed for measuring the local relevance of replies. However, the argumentations in Table 10 show that for students the accuracy of the provided information is also an important aspect of its relevance.

Table 10

Provided argumentations for students' evaluations of the degree of relevance of a received reply

Evaluation	Argumentation ("Because:..")
1	-
2	-
3	it is about right
4	it makes the commonalities clear instead of the differences
4	This is helpful, because I now understand the text better
4	Good point, that is indeed a bit contradicting of Malherbe

4	I think the previous answer was wrong, so I think this is a useful reaction. There was probably a mistake made between the names of departments and provinces
5	I think that is a better way of putting it
5	it is good that she poses a new question so others can respond to it again
5	that is indeed a logical summary of what is answered above
5	it indeed were only the Franks who merged with the Gallo-roman people

Conclusion & discussion

In response to the original research question, we can conclude that the presence of an evaluation function in a tool for anchored discussion can indeed increase the local relevance of students' replies. However, this effect was only found for the "receiver evaluates"-version of the two evaluation functions that were tested in this study. In the next paragraphs we will address to which aspect of (the use of) the two evaluation versions this difference in effectiveness can be attributed, why both evaluation versions differed on this aspect, and how this difference may have influenced the local relevance of students' replies.

First, when trying to identify what caused the two experimental conditions to differ in their effectiveness in increasing the local relevance of replies, we have to take into account that the two versions differed on two aspects. The "everyone evaluates"-version both lacked the direct visual presence of the evaluation function and the ability to produce a substantial amount of evaluations. Taking these two aspects together, we can say at a more general level that that students need to be confronted with and engaged in the evaluation process to a certain degree for an evaluation function to become effective. It also seems that this confrontation and engagement needs to be continuous, as the effect found in the first phase of the study did not transcend into the second phase and hence the evaluation process did not seem to present a lasting learning effect.

Why was the "everybody evaluates"-evaluation function not capable to sufficiently confront students with and engage them in the evaluation process? In response to this question, we can identify an important usability issue. While the evaluation option automatically popped-up in the first version, the second

version required an extra mouse-click for students to open it. In addition, by leaving this option open for all the students, the second version did not personally address a selection of students in their role of receiver in asking them to evaluate a certain reply, as the first version did. We suspect both aspects to have contributed to the virtual absence of evaluations in the second version. Thus, although both versions of the evaluation function made it possible to evaluate the replies of others, only the first version really afforded students to do so (defining affordance as not only making certain behavior possible, but also stimulating or inviting it). These results underline the importance of simple and personal interface design when it comes to getting online participants to respond to information requests.

How did the presence of and students' engagement with the evaluation function in the first part of the study significantly increase the local relevance of replies? Our findings, that neither the amount of evaluations per thread, nor the amount of given or received evaluations per student explained a significant amount of variation in the local relevance scores of replies, suggests that the influence of the evaluation function was broader than just the particular locations or persons that were concerned. Rather, its effect seems to have been more general, transcending the local instances where evaluations were given due to its visual presence and use in general. As such, it may have effectively introduced the concept of relevance as a new criterion for students' online collaboration, making them more aware of the importance of relevant replies.

Practical implications

This study has demonstrated that using an evaluation function for the relevance of replies can be an effective way of increasing the local relevance of replies in students' online learning conversations. It should be noted that the local relevance of replies, as operationalized in this study, might not always be the only or most important variable to focus on when trying to create successful online learning conversations. As students in this study were engaged in a rather fixed domain of which they already possessed a relatively high degree of understanding and for which the task was also structured pretty strongly, we found it appropriate to look directly and only at the potential implications of ideas for other ideas, inspired by Sperber and Wilson's (2004) account of relevance. For example, in situations with a less structured domain where students possess only a very limited level of understanding of the material, it

may be required to put more emphasis on the form of communication. In these situations, it becomes more important to establish an open or soft form of collaboration for which a "confirmation" may for instance be sometimes more valuable than an "adjustment". Similarly, in a more creative brainstorm situation, the primary requirement may not lie in the local relevance of replies, but in creating a more divergent discussion that introduces as many new and unrelated subjects as possible.

Limitations and future research

One of the limitations in this study is formed by a partial misalignment between the nature of the experimental evaluation functions and the way in which we measured its effects. Where the evaluation functions asked users to provide their personal perception of the relevance of replies, we measured its effects in terms of a more objective operationalisation of the local relevance of replies. Although the results indicate that there was at least a significant overlap of the two notions, there are also certain differences that can be identified. As indicated by the students in the argumentations for their evaluations, they assess the relevance of replies on more dimensions than the ones that were present in our coding scheme for local relevance, one of which seems to be perceived correctness of a reply. This difference between students' account of relevance and our operationalisation of local relevance seems comparable to the distinction that Scardamalia and Bereiter (2003) make between a "belief" and a "design" mode of working with knowledge. According to the authors, in a belief mode we respond to ideas by agreeing or disagreeing and asking for evidence (for and against), while in a design mode we concentrate on the usefulness of ideas, how to improve upon them, and how to build knowledge. In this respect, while our coding scheme for the local relevance focuses mainly on a design mode, students seem to take both a belief and a design mode when assessing the relevance of replies.

While this partial misalignment proved not to present a problem in this study, it does give rise to some suggestions for the future. In the development of new research several possibilities exist to bring the two notions even more in line. On the one hand, students could be stimulated to focus more on a design mode when assessing the relevance of replies, for instance by integrating explicit criteria for evaluation in the evaluation function. On the other hand, it would also be possible to adapt our measurement criteria to match those of the

students more closely, including more subjective elements such as the comprehensibility and (perceived) accurateness of a reply. This could be the start of a further refinement and operationalisation of the notions of relevance and coherence in online learning conversations, of which the presented coding scheme can be viewed as a first step. We would be particularly interested in finding a way to identify those replies that answer a particular question or fit a certain information gap in just the right way to create a leap in clarity and understanding within the diffuse mix of students' existing viewpoints, hunches, and ideas.

References

Andriessen, J., & Sandberg, J. (1999). Where is education heading and how about AI. *International Journal of Artificial Intelligence in Education, 10,* 130-150.

Bellamy, R. (1997). *Support for learning conversations.* Abstract retrieved November 11, 2002, from http://kmi.open.ac.uk/seminars/abstract.cfm?id=10.

Bernheim Brush, A. J., Bargeron, D., Grudin, J., Borning, A., and Gupta, A. (2002). Supporting interaction outside of class: Anchored discussion vs. discussion boards. In G. Stahl (Ed.), *Computer support for collaborative learning: Foundations for a CSCL community. Proceedings of CSCL 2002* (pp. 425-434). Hillsdale, NY: Lawrence Erlbaum Associates.

Bryk, A. S., & Raudenbush, S. W. (1992). *Hierarchical linear models.* Newbury Park, CA: Sage Publications.

Clark, H. H. (1992). *Arenas of language use.* Chicago: University of Chicago Press.

Cress, U., Barqeuro, B., Schwan, S., & Hesse, F.W. (in press). Improving quality and quantity of contributions: Two models for promoting knowledge exchange with shared databases. *Computers & Education.*

Dietz-Uhler, B., & Bishop-Clark, C. (2001). The use of computer-mediated communication to enhance subsequent face-to-face discussions. *Computers in Human Behaviour, 17,* 269–283.

Fay N., Garrod, S. C., & Carletta J. (2000). Group discussion as interactive dialogue or as serial monologue: The influence of group size. *Psychological Science, 11*(6), 481–486.

Herring, S. C. (1999). Interactional coherence in CMC. *Journal of Computer-Mediated Communication, 4.*

Hoadley, C. M., & Enyedy, N. (1999). Between information and communication: middle spaces in computer media for learning In C. M. Hoadley and J. Roschelle (Eds.), CSCL'99 *Proceedings of Computer Support for Collaborative Learning,* pp. 242-251. Mahwah, NJ: Lawrence Erlbaum Associates.

Hox J. J. (2002). *Multilevel analysis. Techniques and applications.* Mahwah, NJ: Lawrence Erlbaum Associates.

Hsi, S. H. (1997). *Facilitating knowledge integration in science through electronic discussion: The multimedia forum kiosk.* Unpublished doctoral dissertation, University of California, Berkeley.

Hummel, H. G. K., Burgos, D., Tattersall, C., Brouns, F., Kurvers, H., & Koper, R. (2005). Encouraging contributions in learning networks using incentive mechanisms. *Journal of Computer Assisted Learning, 21*(5), 355-365.

Janssen, J., Erkens, G., Jaspers, J., & Broeken, M. (2006). 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.

Lapadat, J. C. (2002). Written interaction: A key component in online learning. *Journal of Computer Mediated Communication, 7*(4).

Levinson, S. C. (1983). *Pragmatics.* Cambridge, UK: Cambridge University Press.

Mäkitalo, K., Salo, P., Häkkinen, P., & Järvelä, S. (2001, August). *Analysing the mechanisms of a common ground in Web-based interaction.* Paper presented at the JURE pre-conference of the 9th European Conference for Research on Learning and Instruction, Fribourg, Switzerland.

Nückles, M., & Stürz, A. (2006). The assessment tool: A method to support asynchronous communication between computer experts and laypersons. *Computers in Human Behavior, 22*(5), 917-940.

Paavola, S., Lipponen, L., & Hakkarainen, K. (2004). Models of innovative knowledge communities and three metaphors of learning. *Review of Educational Research, 74*(4), 557-576.

Raudenbush, S. W., & Bryk, A. S. (2002). *Hierarchical linear models: Applications and data analysis methods* (2nd ed.). Thousand Oaks, CA: Sage Publications.

Reyes, P., & Tchounikine, P. (2003). Supporting emergence of threaded learning conversations through augmenting interactional and sequential coherence. In B. Wasson, S. Ludvigsen, & U. Hoppe (Eds.), *Designing for*

change in networked learning environments: Proceedings of the international conference on computer support for collaborative learning (pp. 83-92). Dordrecht, The Netherlands: Kluwer Academic Publishers.

Roschelle, J. (1992). Learning by collaborating: Convergent conceptual change. *Journal of the Learning Sciences, 2*(3), 235-276.

Scardamalia, M., & Bereiter, C. (2003). Knowledge building. In J. W. Guthrie (Ed.), *Encyclopedia of education* (2nd ed.). New York: Macmillan Reference.

Smith, B. (2003). The use of communication strategies in computer-mediated communication. *System, 31,* 29-53.

Soller, A., Martinez, A., Jermann, P., & Mühlenbrock, M. (2005). From mirroring to guiding: A review of state of the art technology for supporting collaborative learning. *International Journal of Artificial Intelligence in Education, 15,* 261-290.

Sperber, D., & Wilson, D. (2004). Relevance theory. In: G. Ward and L. Horn (Eds.), *Handbook of pragmatics* (pp. 607-632). Oxford: Blackwell Publishing.

Van der Meij, H., De Vries, K., Boersma, B., Pieters, P., & Wegerif, R. (2005). An examination of international coherence in email use in elementary school. *Computers in Human Behaviour, 21*(3), 417-439.

Van der Pol, J. (2002, January). Identifying and modeling variables in complex CSCL-situations. Case study: The use of asynchronous electronic discussions. Paper presented at the presented at the Conference on Computer Support for Collaborative Learning (CSCL2002), Boulder, CO

Van der Pol, J. Admiraal, W. F., & Simons, P. R. J. (2006). The affordance of anchored discussion for the collaborative processing of academic texts. *International journal of Computer Supported Collaborative Learning, 1*(3), 339-357.

Veerman, A. L. (2000). *Computer-supported collaborative learning through argumentation*. Unpublished doctoral dissertation, Utrecht University, The Netherlands.

Webb, N. M., & Mastergeorge, A. M. (2003). Promoting effective helping behavior in peer-directed groups. *International Journal of Educational Research, 39*, 73-97.

Chapter 5

The nature, reception, and use of online peer feedback in higher education*

In higher education, the use of peer feedback currently forms a significant part of the pedagogical practice. One of the reasons for this is the increasing attention for the development of complex competencies that ask for more, and more differentiated feedback to support the learning process of students. A second reason is that the assessment of products of peers and providing peer feedback resemble professional practice. Due to a change in society's learning goals towards lifelong learning, providing and receiving feedback from work colleagues is a common learning activity in many professional practices. Finally, a more pragmatic -but not less legitimate- reason to consider the use of peer feedback is that teachers often have little time available for the individual coaching of students.

In recent years, the process of providing peer feedback is increasingly facilitated by using electronic learning environments, such as Blackboard or WebCT (both found on www.blackboard.com), which can simplify its logistics considerably. Exchanging peer feedback online also makes it easier for teachers

---

* Van der Pol, J., Admiraal, W. F., Van den Berg, B. A. M, & Simons, P. R. J. (2007). *The nature, reception, and use of online peer feedback in higher education.* Manuscript submitted for publication.

to monitor the process and to intervene, if necessary. More specialized software has been developed that not only facilitates the logistics of peer feedback, but also supports the pedagogical aspects through allowing different degrees of interactivity and different ways of relating feedback to the product. Bhalero and Ward (2001) showed that web-based peer feedback can support the pragmatic and efficient implementation of peer feedback. Web-based peer feedback seems to be suited to implement easily accessible and interactive forms of peer feedback (cf., Gehringer, 2001; Trahasch, 2004).

In this study, we examine the implementation of online peer feedback in two different educational practices in higher education. Hereby, we will focus on the nature of feedback, its reception, and its effects.

Research on peer feedback in higher education

The concept of peer feedback is used for both formative and summative feedback of students. The latter is often referred to as *peer assessment*. Research on peer assessment addresses, for example, the educational design of peer assessment (e.g., Van den Berg, Admiraal, & Pilot, 2006) and the reliability and validity of students' markings and their appraisal (e.g., Bhalerao & Ward, 2001; Sitthiworachart & Joy, 2004). In this study however, we will focus on a formative use that we associate particularly with the notion of *peer feedback*. Falchikov (1986) describes peer feedback as a method where students engage in reflective criticism of the products of other students and provide them with feedback, using previously identified criteria. This peer feedback may be a single activity, or involve a series of instances in which students supply feedback on increasingly polished versions of a piece of written work.

Insight into the effects of peer feedback, both online and offline, seems to be difficult to reach. Topping, Smith, Swanson, and Elliot (2000) state that many of the studies on peer assessment are "merely" descriptive. In addition, these studies often neglect the relation between peer feedback and a change in the products on which the feedback is focused. In recent research on peer feedback, there seems to be a shift in attention from the effects of receiving feedback towards the effect of *producing* it (see for example, Nicol & Macfarlane-Dick, 2006). Apart from a study by Orsmond, Merry, and Reiling (1996), none of the studies mentioned in the review of Dochy, Segers, and Sluijsmans (1999) report learning effects of *receiving* peer feedback. Although producing peer feedback may indeed be responsible for producing important learning results, we think that in order for students' investments in the process

of peer feedback to be worthwhile, it is also crucial to optimize the effects for the receiver. Peer feedback resembles other forms of collaborative learning, offering potential to develop new knowledge and understanding. In order to reach this potential a successful *uptake* of the feedback seems to be essential (Shekary & Tahririan, 2006). Uptake not only refers to an accurate understanding of the provided feedback, but also to the way it is used in changing written work. This other key factor of feedback seems to have drifted to the background in recent research on peer feedback. Topping et al. (2000) mention two studies that do examine the effects of offline peer feedback on the improvement of students' written products (Chaudron, 1983; Jacobs & Zhang, 1989), but both studies only relate these effects to the origin of the feedback (peer or teacher generated), and not to its characteristics. Despite the existence of some more recent studies that deduct on a more theoretical level which kind of feedback would be most beneficial for learning (e.g., Van den Berg, Admiraal, & Pilot, 2006), the relation between the nature of peer feedback and its successful uptake in practice still demands further research. The effectiveness of peer feedback, in terms of a successful uptake, hinges on the quality of the feedback that students provide. To account for the time and effort that students invest in peer feedback, it is important to look critically at the quality of the provided feedback and to investigate how this is received and used by the receiver in the revision of their products. As providing feedback to one's peers (without possessing a high level of expertise) is known to be difficult and time demanding (Dochy et al., 1999; Topping et al., 2000), a successful uptake of peer feedback is by no means guaranteed. Finally, insight into how peer feedback is received and used is not only important with regard to students' learning process, but also with regard to their motivation to keep engaged in the time and effort demanding process of providing peer feedback.

The aim of this article is to examine the relation between the nature of online peer feedback, the reception of the feedback, and its use in the revision of written texts. We conducted two separate studies to investigate these relations in two different educational contexts. In the second study we also examined differences between two tools that were used to facilitate the process of peer feedback. More specific, our research questions are:

1. In what way is the nature of peer feedback, in terms of its functions and aspects, related to a) the reception of the feedback, in terms of its evaluation by the receiver and his or her agreement with it, and b) the use of the feedback in the revision of text?

2. In what way is the reception of the feedback, in terms of its evaluation by the receiver and his or her agreement with it, related to the use of the feedback in the revision of text?
3. In what ways do two tools for peer feedback differ with respect to, functions and aspects of peer feedback, the evaluation of and the agreement on this feedback, and the use of this feedback for revision of written texts?

Next, the two studies will be presented separately, followed by a general discussion.

STUDY 1

Method

Data collection

The data for the first study have been collected during a 6 months period at the Health Care Education study at the HAN University in the Netherlands. The 27 participating students worked individually on several assignments including internship reports, essays, and reflection reports. These documents were part of the students' portfolio demonstrating their mastery of a range of competencies. The peer feedback process was organized in groups of four to ten students and the feedback was aimed at each other's portfolio products. There were no structured procedures of how, how much, when, or where to provide peer feedback. A consequence was that students did not receive feedback on all their products and did not produce a revised version of all products. In this study we limit our analysis to the products on which students received feedback and which have been revised. This resulted in a sample of 392 feedback comments on 78 products created in four groups of students.

An electronic learning environment the Virtual Learning Community (VLC) was used in order to facilitate the process of providing peer feedback (see Figure 1). This VLC showed students' documents on screen along with the provided feedback (displayed in the lower part of the screen) but lacked the possibility to link feedback to a specific selection in the document. Neither did this VLC enable the opportunity to respond to each other's feedback.

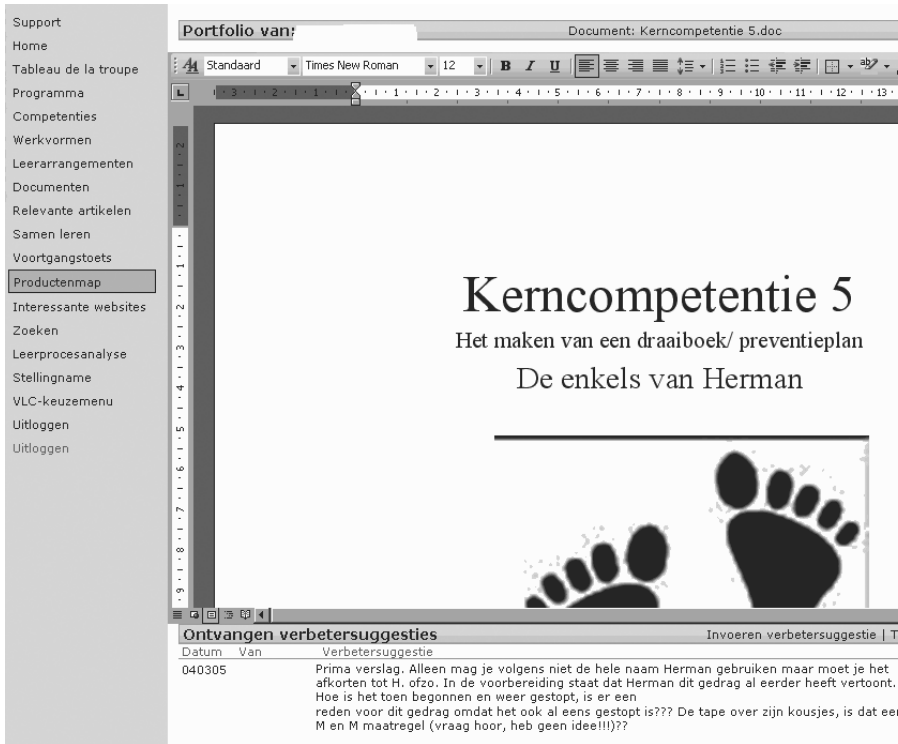

Figure 1. Print screen of students' products along with the provided feedback in the VLC.

Measures

In order to code all data on the three variables, *nature of feedback*, *reception of feedback,* and *revision of texts*, we used event sampling. If the topic of the communication had changed, a new coding unit started. Two researchers were trained for approximately two hours, after which independent coding of 21 messages (12 % of the data) resulted in an agreement of 98.3% from the perspective of both coders, which is well above the threshold of 80 % (Riffe, Lacy, & Fico, 1998). Distinguishing topics is something that is already done very much by students themselves by using keywords or layout features to signal a division of separate points. Both coders divided the 21 messages into 58 feedback units (topics).

The *nature of feedback* was measured in terms of feedback aspects and functions. We used a coding scheme developed by, Van den Berg, Admiraal, and Pilot (2006), who reported an inter-rater reliability of Cohen's κ of .93 for the feedback aspects and .85 for the feedback functions. In order to avoid to

repeat the process of establishing in inter-coder reliability for this study, the first author of Van den Berg et al. coded all data on feedback functions and aspects, following the same procedure as reported in Van den Berg et al. The feedback *functions* refer to analysis of issues, evaluation of parts of the text, explanation of the subject, and suggestions for revision. *Analysis* includes questions and comments aimed at understanding the text. *Evaluation* refers to all explicit and implicit quality statements. Arguments, which elaborate on the evaluation, belong to the category of *explanation.* Measures, which are suggested in order to improve the written text, belong to the category of *revision*. The feedback *aspects* refer to the subject of the feedback, distinguishing between content, structure, and style of the students' written work. *Content* includes feedback on the relevance of information, the clarity of the text, the argumentation, and the explanation of concepts. With *structure* we mean the inner consistency of a text, for example the relation between the main problem and the specified research questions, or between the argumentation and the conclusion. *Style* refers to the form of the text, such as language, grammar, spelling, and layout. Feedback simultaneously aimed at more than one aspect has been excluded from further analysis.

The *reception of feedback* was measured in terms of the "importance" of a comment as perceived by the receiver (on a 4-point scale with 1="not important" and 4="very important") and the receivers' agreement on the feedback which is provided (coded by the researchers as 0= "do not agree", 1="partly agree", and 2="completely agree"). The latter is coded by the researcher with an inter-rater reliability of Cohen's κ of .74. This reliability was established by comparing the independent scores of two researchers on 35 feedback comments on 8 products. In the VLC, the collected evaluations of the importance of the feedback were not displayed

In order to measure the *use of feedback* for revision of the text, all feedback comments were first sorted by the document they concerned. Second, anti-plagiarism software was used to systematically identify all changes in the original and revised versions of the documents. Then, feedback was scored as being processed and incorporated in the revised version of the text in some way (score "1"), or as not leading to a change in the document at all (score "0"). The former also includes cases where something was changed in another way than suggested in the feedback. The inter-rater reliability, in terms on Cohen's κ, is .77, based on 35 scores of two researchers referring to 8 student products.

Analysis

Data on the three variables is available on the level of topic within a message with feedback. We call this the feedback unit. Analyses of the relation between the three variables, *nature of feedback*, *reception of feedback,* and *use of feedback* were performed on the level of this feedback unit. The scores in the feedback units were analyzed using multilevel ordinal regression tests with a logit link function with use of feedback as the dependent variable and with the feedback functions and feedback aspects, respectively, as independent variables. Pearson correlations have been used to examine the relation between the reception of feedback and the revision of texts, and between the nature of feedback and the reception of feedback.

Results

In the VLC, 392 feedback units were produced by the four groups of students. In Table 1 on the next page, we show the proportion of feedback units with the particular feedback function and feedback aspect, the proportion of feedback units which are evaluated on importance and on agreement as well as the mean of the evaluations on both aspects, and the proportion of feedback units which has resulted in revision of the texts.

Table 1

Descriptive statistics for nature, reception, and use of feedback (N=392)

Variable	%	total
Nature of feedback		
Feedback functions		
Analysis	40	157
Explanation	30	118
Evaluation	61	237
Revision	52	202
Feedback aspects		
Content	71	279
Structure	3	12
Style	25	96

	mean	*SD*	*%*	*total*
Reception of feedback				
Importance of feedback as evaluated by the receiver (1=low, 4 =high)	2.50	.88	33	131
Agreement of the receiver with the received feedback as coded by the researcher (0=low, 2=high)	.64	.63	4	14

	%	*total*
Use of feedback		
Feedback comments which led to a revision in the text	49	192

Note. The sum of proportion per variable can be more than 100% as more than one feedback function or aspect can be assigned to one feedback unit.

A remarkable finding is that students received a response on 4% of the feedback comments, although the system did not provide the opportunity to respond to the received feedback. These 14 responses were created by students who found an alternative way to respond to the provided feedback, by including their reactions in the revised versions of their products.

Relation between the nature of feedback and revision of written work

The results with respect to the nature of feedback (function and aspect) are summarized in Table 2. With respect to the feedback function, we found only one significant relation with the revisions. The more feedback includes recommendations for revision, the more it results in the revision of texts. With respect to the feedback aspect, we found both content and style to be significantly related to revision. The more students focused in their feedback on the content and the style of the written work of their peers, the more their peers revised it.

Table 2

Ordinal regression with feedback functions and aspects as independent variables and the use of feedback as dependent variable

Variable	*N*	*B*	*SE*	*df*	*p*
Feedback functions					
Analysis	157	.38	.25	1	n.s.
Explanation	118	.44	.27	1	n.s.
Evaluation	237	.06	.27	1	n.s.
Revision	202	.72	.23	1	.001
Feedback aspects					
Content	279	.95	.36	1	.009
Structure	12	-.82	.66	1	n.s.
Style	96	1.12	.38	1	.003

Note. n.s. means not significant.

Relation between the reception of feedback and the use of feedback

The results with respect to the relation between the reception of feedback and the revision of the texts are summarized in Table 3. We only found one significant (positive) correlation between the evaluation of importance and the revision of texts: The more feedback is evaluated as important, the more feedback has been used to change the written work. It must be noted that the absence of a significant relation of agreement, with both evaluation and use for revision, may result from the small number of feedback comments where the agreement of the receiver could be coded (due to constraints of the tool).

Table 3

Pearson's correlations between evaluation, agreement, and use of feedback

	Evaluation of importance	Agreement receiver	Used for revision of text
Evaluation of importance		n.s. n=10	.50* n=126
Agreement receiver			n.s. n=13
Used for revision of texts			

Note. n.s. means not significant. * p≤ .05

Relation between the nature of feedback and the reception of feedback

In the correlations between the nature of feedback (its functions and aspects) and the reception of feedback (evaluation of importance and agreement of the receiver) we did not find any significant results.

Summary of the main findings

In sum, we found significant relations between, on the one hand, feedback with the function of suggesting revisions and feedback on the aspects content and style of a written text, and, on the other hand, whether or not the parts of the corresponding parts of the text have been revised. Moreover, the more feedback has been evaluated as important, the more students used it to revise their texts.

STUDY 2

Method

Data collection

The data for the second study have been gathered in a 3 months Educational Science course at Utrecht University. In this course 38 students, in groups of three to four, had to collaboratively create a set of course materials for high-school students. These course materials consisted of a theoretical chapter, a chapter with assignments, and a report accounting for the educational choices that were made. Students chose their own subjects, as well as their target group. The teacher provided some guidelines on how to design educational materials

and a set of criteria on which they would eventually be evaluated. Students of each group provided feedback on the product of one other group. The teacher provided the goals and the criteria for the feedback with different roles for students in commenting on each other's product. The feedback was provided on a concept version of the materials during a fixed 1 week period. After this, the students had one week to revise their chapters and hand them in to receive their final grades.

In this second study, we also included the use of two different tools. Besides students' regular online learning environment (Blackboard), we used a system for anchored discussion called "the Annotation system". When used for the collaborative processing of texts, anchored discussion has been found to produce to-the-point feedback in a highly task-oriented learning environment (Van der Pol, Admiraal, & Simons, 2006). In accordance, Trahasch (2004) already suggested to incorporate the possibility of annotation in tools for online peer-feedback, to afford more specific feedback. The importance of detailed feedback for learning has been underlined by Gibbs and Simpson (2004) and Webb and Mastergeorge (2003). In addition, Fiehn (2003) found shared annotations, with students anchoring their evaluations in the artifacts created by their peers, to be particularly suitable for peer assessment scenarios. Thus, students were placed in either the discussion board of Blackboard to write their feedback and upload their documents (see Figure 2), or in the Annotation system developed by Van der Pol, Admiraal, and Simons (2006, see Figure 3). The main differences between these tools are that the Annotation system displays both the document and discussion on screen and allows students to anchor their comments to specific selections of the document. Blackboard does not, presenting a regular forum discussion with the possibility to attach products to individual reactions.

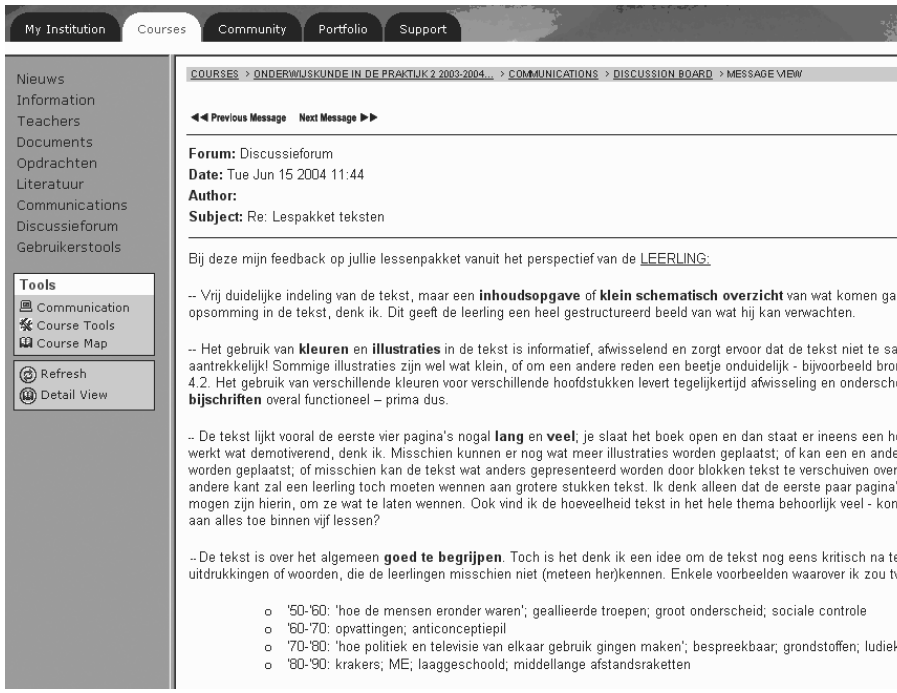

Figure 2. Print screen of students' feedback in the Blackboard discussion area.

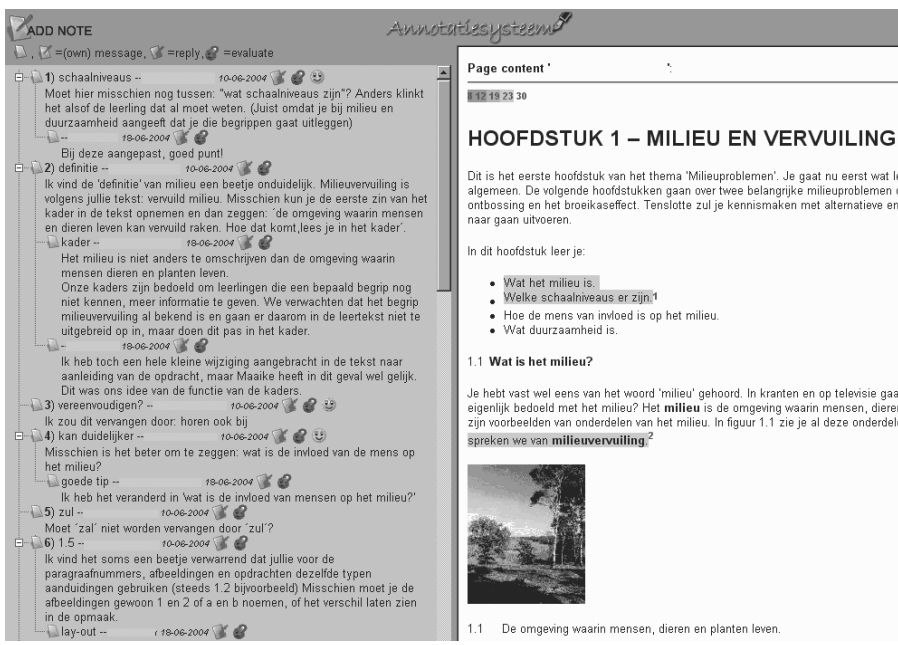

Figure 3. Print screen of students' feedback in the Annotation system.

Measures

In order to code all data on the three variables, *nature of feedback*, *reception of feedback,* and *revision of texts*, we used the same procedures as described in Study 1. Event sampling has been used in order to determine the coding unit. This means that if the topic of the communication changed, a new coding unit started. The *nature of feedback* was measured in terms of the four feedback functions (analysis, evaluation, explanation, and revision) and three feedback aspects (content, structure, and style of the students' writing). The *reception of feedback* was measured in terms of the "usefulness" of a comment as perceived by the receiver (on a 5-point scale, ranging from 1= low, to 5= high) and the receivers' agreement with the provided feedback expressed in a reaction (coded by the researcher as 0= "do not agree", 1= "partly agree", and 2= "completely agree"). Evaluating the usefulness of feedback was done both online and offline for the groups in the Annotation system, whereas for the groups using Blackboard these evaluations were only collected on paper. In both cases these evaluations were collected after the exchange of peer feedback and the revision of texts. The evaluation of the feedback in the Annotation system was directly visible online for all students, whereas the results of the evaluations in Blackboard collected offline were not available for other students. In order to measure the use of the feedback in the revision of text, again the same procedure has been followed as described in Study 1, with a score "1" for a feedback unit leading to a revision in the text in some way and "0" for feedback not leading to a change in the document at all.

Analysis

As in Study 1, data on the three variables is available on the level of topic within a message with feedback. We call this the feedback unit. Analyses of the relation between the three variables, *nature of feedback*, *reception of feedback,* and *use of feedback,* were performed on the level of feedback unit. The scores were analyzed using multilevel ordinal regression tests with a logit link function, with use of feedback as the dependent variable and with the feedback functions and feedback aspects, respectively, as independent variables. Because we will see an influence of "tool" on the presence of several feedback functions, we included it as a covariate in this regression analysis. Pearson correlations have been used to examine the relation between the reception of feedback and the revision of texts, and between the nature of feedback and the reception of feedback. Because the evaluation of usefulness in the Annotation system is linked to

entire messages, we disaggregate this evaluation score in the case of messages that contain multiple topics, attributing the same score for the several feedback units within the message. This way, a total of 160 original evaluation scores resulted in 192 disaggregated scores. This means that the statistical analyses using the evaluation scores from the Annotation system may overestimate the results. χ^2-tests and Mann-Whitney U-tests were performed in order to examine the relation between tool (Annotation system or Blackboard) and nature of feedback, reception of feedback, and revisions of texts.

Results

In both tools (Annotation system and Blackboard) 335 feedback units were produced by the six groups of students (four groups in the Annotation system and two in Blackboard). In Table 4 we show the proportion of feedback units with the particular feedback function and feedback aspect, the proportion of feedback units which are evaluated on usefulness and on agreement, and the proportion of feedback units which has resulted in revision of the texts.

Table 4

Descriptive statistics for nature of feedback, reception of feedback, and use of feedback (N=335)

Variable	%	total
Nature of feedback		
Feedback functions		
Analysis	26	87
Explanation	42	140
Evaluation	70	235
Revision	50	167
Feedback aspects		
Content	50	168
Structure	6	19
Style	44	147

	mean	*SD*	*%*	*total*
Reception of feedback				
Importance of feedback as evaluated by the receiver (1=low, 5 =high)	3.56	1.30	70	234
Agreement of the receiver with the received feedback as coded by the researcher (0=low, 2=high)	1.39	.85	34	115

	%	*total*
Use of feedback		
Feedback comments which led to a revision	55	184

Note. The sum of proportion per variable can be more than 100% as more than one feedback function or aspect can be assigned to one feedback unit.

Table 4 shows an obvious difference with the results found in Study 1. This time, students received an evaluation on about 70% of the feedback comments and a reaction on about 34%. The latter is a direct result of the more interactive nature of the tools used in this study that, in contrast with the VLC in Study 1, allowed students to react on the feedback they received.

Relation between the nature of feedback and use of feedback

The results with respect to the nature of feedback (function and aspect) are summarized in Table 5. With respect to the feedback function, we found three significant relations with the revision of the texts: The more students include an analysis of an issue, an evaluation, or a recommendation for revision in their feedback, the more students changed their written work. With respect to the feedback aspect, we found two significant relations: The more students focus in their feedback on the content and the style of the written work of their peers, the more their peers revised it.

Table 5

Ordinal regression with feedback functions and aspects as independent variables, tool as covariate, and the use of feedback as dependent variable

Variable	*N*	*B*	*SE*	*df*	*p*
Feedback functions					
Analysis	87	1.18	.33	1	≤.001
Explanation	140	.02	.28	1	n.s.
Evaluation	235	.75	.35	1	.032
Revision	167	1.20	.25	1	≤.001
Feedback aspects					
Content	168	.96	.34	1	.005
Structure	19	-.19	.51	1	n.s.
Style	147	.70	.34	1	.042

Note. n.s. means not significant.

Relation between the reception of feedback and the revision of texts

The results with respect to the relation between the reception of feedback and the use of feedback are summarized in Table 6. We found two significant (positive) correlations. The first correlation refers to the relation between the evaluation of usefulness and agreement: The more feedback has been perceived as useful, the more the receiver agrees with the provider of feedback. The second correlation refers to the relation between the level of agreement and the revision of texts: The more a student agrees with the provider of feedback, the more feedback has been used to change the written work.

Table 6

Pearson's correlations between evaluation, agreement, and use of feedback

	Evaluation of usefulness	Agreement receiver	Used for revision of text
Evaluation of usefulness		.32* n=77	n.s. n=234
Agreement receiver			.33* n=115
Used for revision of texts			

Note. n.s. means not significant. * $p \leq 0.05$

Relation between the nature of feedback and the reception of feedback

In the correlations between the nature of feedback (its functions and aspects) and the reception of feedback (evaluation of usefulness and agreement by the receiver), we did not find any significant results.

Differences between tools

A total of 335 feedback units were produced (200 by the four groups of students in the Annotation system and 135 by the two groups of students in Blackboard). An interesting detail, when comparing the use of both tools, is that one student in Blackboard (with the regular, non-anchored discussion) did spontaneously anchor his feedback (consisting of 8 feedback units) in the corresponding text by using the review function of word and attaching this reviewed document in the discussion forum. The results of χ^2-tests and nonparametric t-tests (Mann-Whitney U) that check for differences across tools are summarized in Table 7.

Table 7

Difference between the Annotation system and Blackboard

Variable	Annotation system (N=200)	Blackboard (N=135)	χ^2	df	p
	%				
Nature of Feedback					
Feedback functions					
Analysis	27	24	n.s.		
Explanation	41	44	n.s.		
Evaluation	64	80	10.5	1	.001
Revision	55	42	5.3	1	.014
Feedback aspects					
Content	51	49	n.s.		
Structure	5	7	n.s.		
Style	43	46	n.s.		
	mean	*mean*	*Z*		
Reception of feedback					
Usefulness of feedback as evaluated by the receiver (1=low, 5 =high)	3.7 n=192	2.9 n=42	-3.36		<.001
Agreement with feedback as observed by the researcher (0=low, 2=high)	1.43 n=51	1.36 n=64	n.s.		
	%				
Use of feedback					
Feedback comments which led to a revision	56	54	n.s.		

Note. The sum of proportion per variable can be more than 100% as more than one feedback function or aspect can be assigned to one feedback unit. n.s.= not significant

In Table 7 we see three significant differences between the Annotation system and Blackboard. First, the proportion of feedback units which included evaluative remarks of students was lower in the Annotation system than in Blackboard (χ^2= 10.5; df=1; p=.001). Second, the proportion of feedback units with suggestions for revisions was higher in the Annotation system than in Blackboard (χ^2= 5.3; df=1; p=.014). Third, we see a significant difference in the average scores on the scale of usefulness. In the Annotation system, students evaluated the received feedback as more useful than in Blackboard (*Z*= -3.36; p≤.001). This difference however, might be caused by the fact that the online evaluations in the Annotation system were visible to the other students, which was not the case in Blackboard. Fourth, we see that in the Annotation system, a larger amount of feedback units was evaluated**,** which can also be traced back to the different ways of collecting students' evaluations.

Summary of the main findings

In sum, we found significant relations between feedback with the function of analysis, evaluation, and proposing revisions and feedback on both content and style of a text, on the one hand, and whether or not a text has been revised, on the other hand. Moreover, the more students agreed with the received feedback, the more the feedback has been evaluated as useful; and the more students agreed with the provided feedback, the more they revised the corresponding parts of their texts. Finally, in the Annotation system students show less evaluative feedback and more feedback with suggestions for revisions than in Blackboard.

GENERAL DISCUSSION

Comparison of results

A similar result in both studies, with regard to the functions of feedback, is that students' feedback, in which they suggest concrete revisions, is related to revision of the corresponding parts of students' texts. Understandably, these concrete suggestions for revisions may be the most applied because they give the receivers the best and most direct lead for a potential change in their text. Only in the course on Education Science (the second study) other types of feedback were also related to changes in texts. Here, feedback functions with an analysis of an issue or the evaluation of parts of the text were also related to the revision of texts.

With regard to the aspects of feedback, both studies established a significant positive relationship of content and style with the revision of texts: The more feedback on content or on style, the more this feedback leads to revisions in the text. In both studies we found only a small amount of feedback comments on the structure of students' texts, similar to a finding by Van den Berg, Admiraal, and Pilot (2006). Their explanation for this finding is that feedback on structure is difficult for students to formulate properly, and also difficult to apply in revising a text. Van den Berg et al. also suggest that the task of providing feedback on the structure of a text may be better suited for the teacher.

With regard to the relation between the reception of feedback and its use for the revision of texts, we found different results in Study 1 and 2. In Study 1, the more feedback has been evaluated as important, the more it was used for revisions of the text. In Study 2, the more a receiver agrees with the provider of feedback, the more feedback has been used in the revisions of a text. Thus, the way feedback is received seems to play an important role in the use of feedback, but the exact nature of this relation may depend on the particular task and context.

In neither of the two studies any of the functions or aspects of students' feedback comments correlated significantly with the way it was received (in terms of the receivers' evaluation of importance or usefulness and the agreement with the provider of feedback). This means that for students the importance of feedback does not depend on the nature of feedback (as we have operationalized it in this study) and that there may be other variables that determine its perceived importance.

Tools

In Study 2 we found interesting differences between the Annotation system and Blackboard. The former seems to elicit less evaluative remarks of students and more suggestions for revisions than to the Blackboard environment. This could be a relevant result as students generally appreciate a stance that can be characterized as explorative and collaborative instead of evaluative and authorative (c.f., Lockhart & Ng, 1995). The higher amount of evaluations in Blackboard is also in line with the outcomes of a previous study on the collaborative processing of texts which states that students' interaction was more opinion-oriented in Blackboard than in an annotation system (Van der Pol, Admiraal, & Simons, 2006). In addition, as students in both the VLC and in Blackboard expressed a preference for the possibility of interactivity and annotation by circumventing certain medium-related constraints, we recommend further investigation of using anchored discussion (which does offer both possibilities) for the facilitation of online peer feedback.

Future research

Finally, it is important to note that we can identify a correspondence between feedback and revisions, but we cannot be sure that the revision is a *consequence* of the feedback. It might be possible that the authors of a text were already planning to make certain revisions, regardless of the feedback they received. Self-report questionnaires or a more controlled research design as in quasi-experimental studies might be useful to examine these relationships in the future. More possibilities for research that continues on this first step towards relating the nature of feedback to its reception and its effects can be found in including measures of accurateness, both in measuring the quality of feedback, as in measuring the quality improvement of students' products.

References

Bhalerao, A., & Ward, A. (2001). Towards electronically assisted peer assessment: A case study. *ALT-J, 9*(1), 26-37.

Chaudron, C. (1983, March). Evaluating writing: Effects of feedback on revision. Paper presented at the 17th Annual Teachers of English to Speakers of Other Languages TESOL Convention, Toronto, Ont.

Dochy, F., Segers, M., & Sluijsmans, D. (1999). The use of self-, peer- and co-assessment in higher education: A review. *Studies in Higher Education, 24*, 331-350.

Falchikov, N. (1986). Product comparisons and process benefits of collaborative self and peer group assessments. *Assessment and Evaluation in Higher Education, 11*(2), 146-166.

Fiehn, T., Lauer, T., Lienhard, J., Ottmann, T., Trahasch, S., & Zupancic, B. (2003). From lecture recording towards personalized collaborative learning. In B. Wasson, S. Ludigsen, & and U. Hoppe (Eds.), *Designing for change in networked learning environment. Proceedings of the International Conference on Computer Support for Collaborative Learning* (pp. 471-475). London: Kluwer Academic Publishers.

Flower, L., Hayes, J. R., Carey, L., Schriver, K., & Stratman, J. (1986). Detection, diagnosis, and the strategies of revision. *College Composition and Communication, 37*, 16-55.

Gehringer, E. (2001). Electronic peer review and peer grading in computer-science courses. In H. Walker (Ed.), *Proceedings of the 32nd ACM Special Interest Group on Computer Science Education (SIGCSE) Technical Symposium on Computer Science Education*, (pp. 139 – 143). New York: ACM Press.

Gibbs, G., & Simpson, C. (2004). Conditions under which assessment supports students' learning. *Learning and Teaching in Higher Education, 1*, 3-31.

Jacobs, G., & Zhang, S. (1989, March). *Peer feedback in second language writing instruction: Boon or bane?* Paper presented at the Annual Meeting of the American Educational Research Association, San Francisco, CA.

Lockhart, C., & Ng, P. (1995). Analyzing talk in ESL peer response groups: Stances, functions and content. *Language Learning, 45*, 605-655.

Nicol, D. J., & Macfarlance-Dick, D. (2006). Formative assessment and self-regulated learning: A model and seven principles of good feedback practice. *Studies in Higher Education, 31*, 199-218.

Orsmond, P., Merry, S., & Reiling, K. (1996) The importance of marking criteria in the use of peer assessment. *Assessment and Evaluation in Higher Education, 21*, 239-249.

Riffe, D., Lacy, S., & Fico, F. (1998). *Analyzing media messages: Using quantitative content analysis in research*. Mahwah, NJ: Lawrence Erlbaum Associates.

Shekary, M., & Tahririan, M. H. (2006). Negotiation of meaning and noticing in text-based online chat. *The Modern Language Journal, 90*(4), 557-573.

Sitthiworachart, J., & Joy, M. (2004, July). *Using web-based peer assessment in fostering deep learning in computer programming*. Paper presented at the International Conference on Education and Information Systems: Technologies and Applications , Orlando, FL.

Topping, K., Smith, E., Swanson, I., & Elliot, A. (2000). Formative peer assessment of academic writing between postgraduate students. *Assessment and Evaluation in Higher Education, 25*(2), 149-169.

Trahasch, S. (2004). Towards a flexible peer assessment system. In Y. Akpinar (Ed.), *Proceedings of 5th International Conference on Information Technology Based Higher Education and Training* (pp. 16-20). Istanbul: Institute of Electrical and Electronics Engineers (IEEE).

Van den Berg, B. A. M., Admiraal, W. F., & Pilot, A. (2006). Design principles and outcomes of peer assessment in higher education. *Studies in Higher Education, 31*(3), 341-356.

Van der Pol, J., Admiraal, W. F., & Simons, P. R. J. (2006). The affordance of anchored discussion for the collaborative processing of academic texts. *International Journal of Computer Supported Collaborative Learning, 1*(3), 339-357.

Webb, N. M., & Mastergeorge, A. M. (2003). Promoting effective helping behavior in peer-directed groups. *International Journal of Educational Research, 39*, 73-97.

Chapter 6

Reflection

Similarly to our aim of engaging students in reflection during their collaborative process of meaning-making in order to deepen their learning process and relating their newly developed understanding to their personal practice, it is also important for us to reflect on the strengths and weaknesses of the research we have presented. Looking back at the several studies that have been conducted, we can reflect on the used concepts, the followed methodology, the empirical results, the implications for practice, and on the tools that have been used; each of which may lead to recommendations for further research. In this final chapter we will make a start with this process of reflecting, rethinking, and elaborating on the presented research, which we hope will not end here.

Reflection on theoretical concepts

Resulting from on our primary concern with creating and facilitating successful online learning conversations in educational practice, we have also opted to guide the selection of relevant research and useful theoretical notions primarily by the phenomenon under investigation rather than from a predetermined theoretical framework. A drawback of this approach is that it offers a broad range of disciplines from which the study of online learning conversations can be approached. In choosing primarily to focus on the process of interaction and collaboration, we largely ignored other possible approaches and levels, such as the specific research fields on individual reading or writing. However, even within this primary focus on theories of interaction, still many different approaches are possible, each illuminating a different aspect of the subject. Here we have tried to focus our theoretical approach by drawing mainly from the fields of communication and of collaborative learning, largely ignoring

other possible approaches such as those of knowledge management, rhetoric, individual processes of reflection, social dynamics, collaborative work, or group cognition.

In spite of these restrictions, the remaining diversity of related and relevant studies within the fields of communication and collaborative learning have still resulted somewhat in an abundance of concepts that can be found throughout this dissertation. Most of this abundance is caused by the fact that we have used concepts from both theoretical and practical approaches. Whereas the more abstract concepts from communication theories have been mainly used to explore the difficulties relating to online learning conversations, we have based our concrete operationalizations and analyses largely on more practical concepts from interaction theories and interaction analysis. Thus, for instance, the more theoretical notions of *semantic* and *pragmatic grounding* from the theoretical parts of Chapters 1 and 2 have developed into the more specific and concrete operationalizations of *meaning-oriented interaction* and *regulative activity* in Chapter 3. In addition, two different levels of reasoning were followed to describe the possible benefits of anchoring students' online learning conversations in their study material. Where in Chapter 2 the main reasoning to implement a system for anchored discussion followed the more theoretical line of exploring the effect of distance on the context of online learning situations, the reasoning in Chapter 3 followed a more practical line in emphasizing the importance of meaning-oriented and efficient communication for students' collaboration. On the positive side, we have demonstrated that it is possible and fruitful to strengthen the link between communication theories and the field of CSCL, which could be even further elaborated in the future. Hereby, an interesting line of research would lie in further developing specific theory for group communication and for communication online. Generally, we think that valuable progress can be made by combining the insights of different research fields in approaching a single phenomenon.

Fortunately, there are also concepts that can be found throughout this dissertation. *Affordance*, defined as the functional design of a tool that makes possible, suggests, or even stimulates a certain use, can be identified as a central concept. We have demonstrated that the affordance of tools can be used to facilitate students' collaborative processes in an open learning environment and, thus, offers a good alternative for more direct forms of training or instruction. We therefore propose to continue research on the notion of affordance in the

future and to investigate how existing tools can be further specialized, based on a thorough analysis of the particular task. Furthermore, we hope that we have illustrated the unique nature of *online learning conversations* and propose that the specific requirements and characteristics of different forms of online learning conversations will be the subject of further research.

Regarding the meaning-oriented online learning conversations that played a central role in this research, a fruitful connection in future research may also be made to the concept of *trialogic learning* (see Paavola, Lipponen, & Hakkarainen, 2004). These authors compare a "knowledge creation", or trialogical, metaphor of learning to a monological and dialogical one, and use the trialogical metaphor to indicate a collaborative creation of knowledge and the collaborative development of shared (conceptual) artifacts. Although not exactly similar, the meaning-oriented learning conversations that form the focus of this dissertation can also be seen as presenting a trialogic relation between various peers and the study material they are collaboratively attempting to re-create and appropriate.

Finally, we think the notions of online *distance* and *conversational context* present interesting issues for future research. We have seen evidence for the idea that physical distance can also influence conceptual distance. Similar to telephone conversations, one can be "drawn" into the conversation to a greater or lesser extent. Use a cell-phone in the presence of other people, or maybe even in busy traffic, and one is almost guaranteed to be less engaged in the conversation and missing more of its meaning than when calling from an empty room, where one can more easily zoom into, or loose oneself in the conversation. On a conceptual level, this could mean that the shared conversational context can vary in strength, presence and focus, and is created by the different participants who may meet each other "somewhere halfway". This dynamic and slippery nature of the context in which an online conversation is held underlines the idea that making certain aspects of this context more salient may also influence the range of meaning that is being triggered.

Reflection on methodology

Regarding the validity of the methodology followed, we have primarily focused on establishing sufficient levels of internal validity. First, we have created consistent and inter-subjective coding schemes with according levels of inter-coder reliability. This was achieved by developing specialized schemes based on the characteristics, needs, and data in our specific research situations. A second form of internal validity regards the question whether an instrument used also measures what it is expected to measure. This aspect regards the meaning that can be attributed to the results that are obtained. In Chapter 3, where the first quantitative data were introduced, the interpretation of our measurements are still pretty straightforward, as the dichotomous variables concerned were still neutral in indicating only whether certain basic activities, such as "statements" or "critical reactions", were either present in students' conversation or not. Chapter 4, however, presents a *hierarchical* coding scheme, which required a more thorough consideration to ensure it could be meaningfully interpreted. For this reason, a considerable amount of attention was paid to the construction of the coding scheme, which, in the end, was based on Sperber and Wilson's (1995) theoretical notion of relevance, Veerman's (2000) existing coding schemes on constructive activities, and a preliminary view of the data at hand.

Regarding the external validity, an advantage of the followed methodological approach is that the data were collected in real and complex situations and can thus more easily be generalized to other educational practices. This will be of particular concern to university teachers. The possibility for generalization of the obtained results is mainly a matter of identifying the characteristics of the courses in which our studies were done and comparing these to other courses. These characteristics include for example the task, goal, guidance, group size, and level of expertise of the participating students. If these characteristics are sufficiently similar, we may reasonably expect the results that we obtained to be applicable. To ensure sufficient similarity, we suggest to restrict generalizations of our results primarily to other courses in higher education in the field of the social sciences where it is important that these courses also consist of a blended format with integrated on- and offline activities. However, future research may also explore the usefulness of anchored discussion for secondary education, where the

collaborative processing of texts may closely resemble the processes that were studied in this dissertation.

Regarding the nature of the collected data, our primary concern for reliability has also affected the scope of phenomena that we were able to study. As a result of our primary use of students' externalized dialogue, many other implicit processes have remained invisible. This approach has, for example, not provided much direct information on the psychological and social mechanisms underlying the externalized collaborative process. In addition, to ensure sufficient construct validity our analyses focused on concrete operationalizations that partly differed from the more broad theoretical notions from which they were derived. Thus, further research on these broader concepts such as distance, common ground, co-intentionality, or relevance is also warranted if we want to be able to shed more light on the theoretical assumptions on which the presented facilitation attempts were based.

Several other suggestions for future research can also be made. First, one could set out to collect more qualitative data, which could provide additional insight into the reasons why certain facilitation attempts did or did not work as expected. This data collection could even comprise students' *offline* interaction, as students learning conversations are not bound to their online learning environment alone. It is even one of the main goals of using asynchronous discussion to expand students' engagement with the subject matter into their everyday life. Second, it would also be possible to conduct quantitative research in more controlled conditions, in order to isolate the effect of different variables. As indicated by Salomon (1988), it can be a valuable approach to first try to find something that works (as we have done in this dissertation), and afterwards try to dismantle the complex phenomenon into its underlying mechanisms. This process of dismantling could focus on identifying the specific actions and interactions that form the essence of students' individual and collaborative learning processes. This would not necessarily lead to laboratory type research, as online learning environments would also present good opportunities to manipulate parts of students' online interaction.

Although gathering data on learning effects proved to be difficult in the first two studies in this dissertation, the use of peer feedback in the third one has created a practical opportunity to collect more information on these effects. As we consider this to be only a first step in the right direction, we would be interested in applying a similar approach to the collaborative processing of

abstract texts. It would, for instance, be possible to measure students' comprehension of a text before and after conducting a series of learning conversions, preferably after both shorter and longer periods of time. We think that having more information on students' individual and collective learning results will also help us to identify the most effective ways of facilitating different kinds of online learning conversations. Ideally, we would be able to establish such an optimal allocation of students' efforts and increase the quality of their learning conversations to such an extent, that for students the gains of participating in online learning conversations outweigh the costs and offering external motivation for students' participation will no longer be required.

Reflection on empirical results

In the introduction, we identified several problems related to fulfilling the learning potential of conducting online learning conversations in higher education. In short, we identified two major obstacles. The first concerned the fact that conversation for learning presents students with the challenging task of semantic grounding and a collaborative deepening of understanding. The second concerned the additional difficulties that come into play when conducting these conversations online. In response to these obstacles, the first two studies in this dissertation demonstrated ways in which students' collaboration could be facilitated successfully, whereas the third has made a start to measure the connection between the nature of students' online collaboration and its learning effects.

In the first study, students' collaboration in the system for anchored discussion has showed both an increased link between discussion and study material and an increased communicative efficiency. In short, we linked these results to an apparent inherent suitability of anchored discussion for discussing the meaning of texts. Its functional design seems to naturally present students with an answer to the question "What are we going to do here?" and on the level of individual threads, "What exactly am I talking about?". These effects are similar to the ones created by Suthers and Xu's (2002) system for artifact-mediated communication, which, as expressed by Cherubini (2007), presents messages which are "overwhelmingly" on task, and presents focused discussions with little topic drift within particular threads. In addition, we have seen that opinion-oriented messages (which were found more often in the regular discussion tool) contain more argumentations and confirmations. We

take this result as an indication for the idea that different kinds of argumentation may also differ in the amount of *effort* they require and that it may be easier for students to provide arguments for a certain opinion they have on a text than for a certain interpretation of its meaning. As we also expect that these different forms of argumentation contribute to students' learning process in a different degree, this brings us back to the concept of an *optimal collaborative load* (see Dillenbourg and Bétrancourt, 2006). Here we propose to view the word "optimal" as indicating an optimal direction of students' efforts, more than an optimal quantity. In the context of the first two studies, the optimal division can then be found in a maximal orientation of students' argumentation and conversation to processing the meaning of texts.

The notion of optimal collaborative load as being a matter of division rather than an absolute amount presumes a more or less fixed amount of effort that is available for a certain collaborative process which can be allocated in varying ways. This notion seems to concur with our finding that in discussions with a high percentage of meaning-oriented collaboration, only a very small percentage of messages contained social or regulative comments. We propose that this can be interpreted in line with the functional nature of grounding, meaning that students will engage less in social and regulating activities if there is less need to. With regard to the anchored discussion, this would mean that it presents users with a greater amount of social and regulative common ground. As a result, maintaining social relationships in order to ensure the continuance of the conversation may be less required. Many authors emphasize the importance of social and regulative activities in students' learning conversations, reasoning that they are important because we often find them. However (as indicated in Chapter 2), we want to suggest that a high degree of social and regulative activity can also be seen as reflecting participants' lack of shared collaboration goals and as using important resources that cannot be directed towards processing the meaning of the subject matter.

In the second study (using a further developed system for anchored discussion) we demonstrated de effectiveness of a personalized evaluation function on the relevance of replies. Compared to the first study, where the affordance of anchored discussion seems to direct users to processing the meaning of the text, we have seen that the effect of the evaluation function depends on its active use. In this way, the results of the first study could be labelled as true affordance effects, in the sense of *inviting a certain use because the*

tool is fit for that use to a high extent, whereas in the second study the effect on students' collaboration seems to have been a bit more direct, by increasing their awareness of the importance of providing relevant replies.

The third and final study did not attempt to facilitate students' interaction, but made a first step towards measuring the individual learning effects of students' online collaboration by studying the use of online peer feedback. Amongst others, it demonstrated the importance of offering concrete suggestions for a successful uptake of the feedback. This seems to confirm the suitability of Sperber and Wilson's (1995) notion of relevance for analyzing online learning conversations. Just like a relevant utterance, a concrete suggestion offers the receiver new implications for existing ideas, without the costs of deriving these implications being too high. The third study also found significant relations between elements of both the nature and the reception of feedback on the one hand, and the use of this feedback by the receiver on the other hand. In the future, we would be interested in taking this effect-oriented approach to the study of online learning conversations and in further developing ways to determine their effectiveness, in order to investigate how to optimally facilitate them.

Reflection on practical implications

In the previous section we have seen several empirical results from the three conducted studies, contributing valuable information for educational practice. The main implication is that using (specific forms of) anchored discussion can help teachers to create meaning-oriented and constructive online learning conversations. Therefore, to complement the regular forum discussion that may be more suitable for opinion-oriented discussion, we think it would be useful for anchored discussion functionality to become a standard feature of electronic learning environments. In a broader sense, the results implicate that "affordance works", and that it may also be used in other ways and educational tools to partly reduce the need for explicit training or instruction.

In response to the constraints for creating successful online learning conversations that we emphasized in our introduction (to demonstrate the need for research and facilitation), we also want to emphasize the possibilities. In the course of this dissertation, we have seen that it is possible to facilitate the paradoxical process of collaboratively developing new understanding that surpasses that of the individual participants. In order to illustrate this cause for

optimism, we have included two small examples (see Figure 1 and 2 below) from our first study that show students in the process of collaboratively developing meaning and understanding, using only their limited available resources and rational thought.

[7][text = 'dyad'] What is meant by the word 'dyads'. I find it in the article in many other occasions too, but the dictionary doesn't know it.--*Tamara added* October 1, 2002

my dictionary says: pair--*marieke added* October 1, 2002

But then I still don't get the way it is used. They are talking often about parents and dyads, not? so with dyads they cannot mean two parents. Then what do they mean with it? --*Tamara added* October 2, 2002

I think they mean with dyads they mean the pair parent-child. If you take this meaning, it does make sense in the text --*meijke added* October 6, 2002

Figure 1. Thread from the annotation system in study 1.

[4][text] I am not sure what is meant by "nuclear family".
--*sstudent A added* September 29, 2002

I thought maybe a household, the other what we call "the whole family" and then the society? But I have my doubts about this.. – *Student B added* September 29, 2002

I think your idea about the 'nuclear family' is not entirely correct. Nuclear can in Dutch be seen as part of the word nuclear energy. So atomic energy. If you see nuclear as 'core'. You can see the nuclear family as the core family. With this I am thinking about direct family; parents, brothers, sisters and maybe even grandparents etc. –*Student C added* September 30, 2002

I agree with student C, especially because apart from the 'nuclear family' the text speaks about: 'an extended family and the culture'...in other words, other family and culture... Then it almost has to be the immediate family, right? –*Student D added* September 30, 2002

Figure 2. Thread from the annotation system in study 1.

Reflection on tools

In this dissertation, we have shown that the development and implementation of a specialized tool for CSCL can facilitate specific aspects of students' online learning conversations. We believe that this task-specific nature of the developed tool also made it intuitive to use. In regard to this aspect, it deserves to be mentioned that in our studies, both students and teachers were able to use the tools for anchored discussion with only a very limited amount of instruction on how to use them. We also see many possible ways for further development and use of anchored discussion in educational practice. Brainstorming on the essence of anchored conversations, the thought comes to mind that ideas are always connected to other ideas. Therefore it generally seems a productive thought to integrate students' conversations more with the objects of conversation (also in the form of images or video). It would also be natural to approach this integration from both ways: Not only as providing conversations with context, but also as providing material with conversation. The second approach for instance is being followed by Churchill, Trevor, Bly, Nelson, and Cubranic (2000) whose system for "anchored conversations" presents chat windows that are inserted into a document and accompany it whenever it is opened and wherever it goes.

As demonstrated by George and Labas (2007) with their "contextual forums" system, students' discussions can also be integrated with multiple files at the same time, ranging from documents to other content, such as descriptions of learning activities. It is even possible to anchor conversations in video material, as demonstrated by Hwang, Wang, and Sharples (2007), which opens up possibilities for collaborative reflection on videotaped presentations, or personal video material in portfolio's. Like Churchill et al., George and Labas also implemented an intuitive way of accessing these contextualized conversations, making them automatically accompany the document when it is opened. Furthermore, it seems a natural thing to include instructions for students in the contextualized mix of conversation and content. While in our studies the students' task was usually presented in a separate message, the specific task may also be more incorporated in a system's design. Therefore, the new version of the Annotation system also includes the possibility for students or teachers to provide several subheadings in the discussion area, to focus students' conversations on particular aspects.

From a less technical and more pedagogical viewpoint, it would be worthwhile to investigate the possibility of converting students' efforts in conducting online learning conversations not only into individual learning gains, but also into the creation of some form of recyclable product, that presents additional returns of the conversation. This recycling of students' conversations could, for instance, be found in summarizing or filtering the most important elements of the conversation, but also in reusing parts of a discussion for one's personal portfolio, or in processing them into a personal essay that may be required for the course.

Compared to other forms of online collaborative annotation, like shared bookmarking in Connotea or Del.icio.us, or online peer feedback in Google Docs or Turnitin, anchored discussion presents more opportunity for interaction, which we would recommend to be included in these systems as well. The other way around, existing forms of contextualized discussion, like Kükäkükä or Ubiquitous 3DE, do not yet allow users to anchor notes to a particular self- picked place in the document, which we would also recommend them to include. Finally, the WebAnn system that does posses both these functionalities is not available through a browser without the installation of a plug-in or a server. Here we would also recommend the use of an easily accessible tool like the Annotation system that is purely web-based, so it can be integrated in any online learning environment by inserting a single link.

Finally, anchored discussion has shown to be a versatile tool with many possible uses that concern the discussion of online materials. Besides for students' collaborative processing of academic texts, we have seen its suitability for interactive peer feedback. In addition, as shown by Buckingham Shum and Sumner (2001), journal reviewers may also benefit from the possibility of conducting context-based conversations. Besides for official reviewing, anchored discussion may also present a fruitful possibility for researchers to discuss texts asynchronously and online. The Annotation system developed in the course of this dissertation can be tested and -under certain terms and conditions- freely used, and is available at www.annotatiesysteem.nl. Here, an annotated version of this dissertation can also be found on which we invite our readers to leave their impressions and comments, in order to continue the conversation on how to facilitate students' online learning conversations.

References

Buckingham Shum, S., & Sumner, T. (2001). JIME: An interactive journal for interactive media. *Learned Publishing, 14*(4), 273-285.

Cherubini, M. (2007). *Blog entry Kükäkükä: An online environment for artifact-centered discourse.* Retrieved, April 2, 2007, from http://www.i-cherubini.it/mauro/blog/2007/01/10/kukakuka-an-online-environment-for-artifact-centered-discourse.

Churchill, E. F., Trevor, J., Bly, S., Nelson, L., & Cubranic, D. (2000). Anchored conversations: Chatting in the context of a document. *Proceedings of the SIGCHI conference on Human Factors in Computing Systems* (pp. 454-461). New York: Association for Computing Machinery.

Dillenbourg, P., & Bétrancourt, M. (2006). Collaboration load. In J. Elen, & R. E. Clark (Eds.), *Handling complexity in learning environments: Research and theory*. Amsterdam: Elsevier.

George, S., & Labas, H. (2007). E-learning standards as a basis for contextual forums design. *Computers in Human Behavior.* doi:10.1016/j.chb.2007.01.006.

Paavola, S., Lipponen, L., & Hakkarainen, K. (2004). Models of innovative knowledge communities and three metaphors of learning. *Review of Educational Research 74*(4), 557-576.

Salomon, G. (1988, April). *AI in reverse: Computer tools that become cognitive.* Invited address at the annual meeting of the American Educational Research Association, New Orleans, LA.

Sperber, D., & Wilson, D. (1995). *Relevance: Communication & cognition.* Oxford, UK: Blackwell Publishing.

Suthers, D., & Xu, J. (2002, May). *Kükäkükä: An online environment for artifact-centered discourse.* Paper presented at the Eleventh World Wide Web Conference, Honolulu, Hawaii.

Veerman, A. L. (2000). *Computer-supported collaborative learning through argumentation.* Utrecht, The Netherlands: Utrecht University.

Hwang, W. Y, Wang, C. Y., and Sharples, M. (2007). A study of multimedia annotation of web-based materials. *Computers & Education, 48*, 680- 699.

Samenvatting

Dit onderzoek naar het faciliteren van online discussies in het hoger onderwijs richt zich specifiek op het gezamenlijk verwerken van literatuur en het geven van online peer feedback. Voor deze taken bieden vooral asynchrone discussiefora goede mogelijkheden om het samenwerkend leren van studenten te bevorderen. De mogelijkheid elkaars vragen te laten bezinken en de tijd te nemen voor een reactie stimuleert reflectie en het opnieuw activeren en onder de loep te nemen van de bestudeerde stof, opdat dit niet beperkt blijft tot de beschikbare contacturen. Ook bieden online discussiefora een open leeromgeving waar deelnemers hun interpretaties van de stof aan elkaar kunnen voorleggen, waarbij het onderbouwen van en het doorbouwen op deze interpretaties kan leiden tot het beter 'eigen' maken van de stof. Tenslotte kunnen docenten op basis van de discussies beter bepalen welke betekenis studenten precies toekennen aan bepaalde stof.
Aangezien samenwerkend leren echter een complex proces is waarbij moet worden voldaan aan vele randvoorwaarden, is succes geenszins vanzelfsprekend en is het in de praktijk vaak moeilijk om dit potentieel ook ten volle te benutten.

In de introductie (Hoofdstuk 1) van deze dissertatie worden twee belangrijke problemen aan de orde gesteld die het voeren van gezamenlijke online leerconversaties tussen studenten kunnen bemoeilijken. Allereerst stellen conversaties voor leren vaak meer eisen aan de inhoud en de kwaliteit ervan dan conversaties zoals die gevoerd worden in het dagelijks leven. Een belangrijk verschil is het nagestreefde doel, aangezien voor samenwerkend leren het creëren van gedeeld begrip (het 'semantisch grounden') centraal staat, wat meer eisen stelt en vaak lastiger is te bereiken dan alleen het waarborgen van de voortgang van de conversatie (het 'pragmatisch en sociaal grounden'), waarop in alledaagse conversaties veelal meer de nadruk ligt. Ten tweede brengt het online en asynchroon communiceren ook enkele beperkingen met zich mee, die veroorzaakt worden door de afwezigheid van veel non-verbale communicatie en de verminderde mate van *turn-taking* en interactiviteit.

In de eerste twee studies in deze dissertatie wordt onderzocht hoe de online leerconversaties van studenten die gericht zijn op het gezamenlijk verwerken van literatuur, kunnen worden gefaciliteerd. De eerste studie is vooral gericht op de efficiëntie van de communicatie en het verhogen van de mate waarin de conversatie van studenten is gericht op het verwerken van de betekenis van een tekst. Om dit te bereiken wordt een zogenaamd annotatiesysteem ingezet dat de conversatie van studenten 'verankert' in de te bespreken tekst. Deze vorm van ondersteuning wordt vergeleken met een standaard discussieforum van Blackboard. De tweede studie richt zich vooral op het verhogen van de relevantie van de reacties van studenten op elkaars berichten. Hiertoe is een evaluatiefunctie toegevoegd aan het annotatiesysteem, waarmee studenten de relevantie kunnen aangeven van de reacties die ze van elkaar ontvangen. De resultaten van eerste studie (gerapporteerd in Hoofdstukken 2 en 3) laten zien dat vergeleken met het standaard discussieforum het annotatiesysteem een meer taakgerichte, efficiënte en constructieve communicatie tussen studenten creëert. Dit maakt het uitstekend geschikt voor de eerste fases van de gezamenlijke verwerking van literatuur, waarbij de nadruk ligt op het achterhalen van betekenis en het vergroten van begrip. In de tweede studie (beschreven in Hoofdstuk 4) is gebleken dat de kwaliteit van deze communicatie tussen studenten in een annotatiesysteem verder verhoogd kan worden door de aanwezigheid van een evaluatiefunctie waarmee studenten de relevantie van ontvangen reacties kunnen aangeven. Om de relevantie van reacties te verhogen, moet deze evaluatiefunctie echter wel voldoende zichtbaar zijn en gebruikt worden.

De derde studie (in Hoofdstuk 5) gaat verder in op de kwaliteit van de reacties die studenten elkaar geven. Daarin kijken wij niet alleen naar de eigenschappen van het gegeven commentaar, maar proberen we deze ook te relateren aan hoe de feedback wordt opgevat en verwerkt door de ontvanger. Om dat laatste mogelijk te maken is voor een andere context gekozen dan in studies 1 en 2 (waar het ging om het gezamenlijk verwerken van teksten), namelijk het online geven van peer feedback. Door de veranderingen bij te houden in de schrijfproducten van studenten is onderzocht wanneer de peer feedback is gebruikt voor het reviseren van de eigen tekst. De resultaten van deze studie laten zien dat de link tussen de aard van feedback en het toepassen ervan vooral zit in de aanwezigheid van concrete suggesties voor revisie. Zoals in studie 1,

blijken ook hier de berichten van studenten in Blackboard meer evaluatief en in het Annotatiesysteem meer constructief van aard (door het geven van concrete suggesties).

De algemene conclusie die getrokken kan worden (zie Hoofdstuk 6) is dat het functionele ontwerp van een systeem de samenwerkingsprocessen tussen studenten kan ondersteunen. Voor docenten betekent dit dat het gebruik van een annotatiesysteem (al dan niet met evaluatiefunctie) een betere keuze kan zijn dan een algemeen discussieforum, vooral wanneer het gaat om de betekenisgerichte en constructieve bespreking van (moeilijke) teksten.

Curriculum Vitae

Jakko van der Pol was born on April 6th 1976 in Utrecht, The Netherlands and completed his secondary education in 1994 at the Sint-Vitus College in Bussum. After studying Natural Sciences and Innovation Management for a year, he started the study of Educational Sciences at Utrecht University, and graduated in 2001. During his masters program, he specialized in learning with new media and performed an internship at the Dutch television station Teleac/NOT. After completing his studies, he worked as a programmer and lecturer at Utrecht University and the University of Amsterdam. Since 2002, Jakko has been working as a PhD candidate with a focus on CSCL at the IVLOS Institute of Education at Utrecht University. In the course of his research project, he coordinated the development of the Annotation system. He has also worked for four months at the centre for research and support of training and its technologies (CRAFT), led by Pierre Dillenbourg at the EPFL in Lausanne. Currently, Jakko works as an educational developer at VU University Amsterdam.

Publications

Journal articles

Van der Pol, J., Admiraal, W. F., Van den Berg, B. A. M., & Simons, P. R. J. (2007). *The nature, reception, and use of online peer feedback in higher education.* Manuscript submitted for publication.

Van der Pol, J., Admiraal, W. F., & Simons, P. R. J. (2007). *Peer evaluation in online anchored discussion for an increased local relevance of replies.* Manuscript submitted for publication.

Van der Pol, J., Admiraal, W. F., & Simons, P. R. J. (2006). The affordance of anchored discussion for the collaborative processing of academic texts. *International Journal of Computer Supported Collaborative Learning, 1*(3), 339-357.

Van der Pol, J., Admiraal, W. F., & Simons, P. R. J. (2006). Context enhancement for co-intentionality and co-reference. *Journal of Artificial Intelligence & Society, 20*(3), 301-313.

Van der Pol, J., & Admiraal, W. F. (2003). Het succesvol inzetten van asynchrone elektronsiche discussies [Succesfully implementing asynchronous online discussions]. *Onderzoek van Onderwijs*, *32*, 26-31.

Paper presentations

Cherubini, M., Van der Pol, J., & Dillenbourg, P. (2005, July). *Grounding is not shared understanding: Distinguishing grounding at an utterance and knowledge level.* Paper presented at CONTEXT'05, the Fifth International and Interdisciplinary Conference on Modeling and Using Context, Paris.

Van der Pol, J., & Admiraal, W. F. (2005, May) Het ontwikkelen van een gedetailleerd codeerschema voor het analyseren van samenwerkend leren in asynchrone discussies [Developing a detailed coding scheme for analysing collaborative learning in asynchronous discussions]. In M. Valcke, K. de Cock, D. Gombeir, & R. Vanderlinde (Eds.), *Meten en onderwijskundig onderzoek. Onderwijs Research Dagen 2005* (p. 449-451). Gent, Belgium: Universiteit Gent/Vakgroep Onderwijs.

Admiraal, W. F., Akkerman, S., Groep, J. van de, Laat, M. de, Van der Pol, J., & Simons, P. R. J. (2003, August). How people in virtual groups (fail to) interact. In L. Mason, S. Andreuzza, B. Arfè, & L. Del Favero (Eds.),

Improving learning, fostering the will to learn (pp. 238-239). Padova, Italy: University of Padova/Department of Developmental and Socialisation Psychology.

Akkerman, S., Beers, P. J., Van den Bossche, P., Van der Pol, J., & Mulder, I. (2003, May). Het bereiken van een gemeenschappelijke kennisbasis in teams: een zoektocht naar een gemeenschappelijk kader [Reaching shared knowledge in teams: a search for common ground]. In W. Jochems (Ed.), *Grenzeloos leren, Proceedings 30e Onderwijs Research Dagen* (p. 52). Heerlen: Open Universiteit Nederland.

Van der Pol, J., Admiraal, W. F., & Simons, P. R. J. (2003, June). Grounding in electronic discussions: Standard (threaded) versus anchored discussion. In B. Wasson, S. Ludvigsen, & U. Hoppe (Eds.), *Designing for Change in Networked Learning Environments* (pp. 77-81). Dordrecht, The Netherlands: Kluwer Academic Publishers.

Van der Pol, J. (2002, May). Tijdsbesteding en grounding in elektronische a-synchrone discussies [Time investments and grounding in electronic asynchrounous discussions]. In F. Deams, R. Rymenans, & G. Rogiest (Eds.), *Onderwijsonderzoek in Nederland en Vlaanderen* (pp. 65-66). Antwerpen, Belgium: Universiteit van Antwerpen.

Van der Pol, J. (2002, January). *Identifying and modeling variables in complex CSCL-situations. Case study: The use of asynchronous electronic discussions*. Paper presented at the Conference on computer support for collaborative learning (CSCL2002), Boulder, CO.

Poster presentations

Van der Pol, J. (2006, October). *Making dialogic learning trialogic*. Poster session presented at the Interaction et Pensée: Perspectives Dialogiques Conference, Lausanne, Switzerland.

Van der Pol, J. (2004, January). *Adding documents to document-centered discussion*. Poster session presented at the Vereniging voor Onderwijs Research Theme Conference, Utrecht, The Netherlands.