Computers & Education 56 (2011) 379–387

Contents lists available at ScienceDirect

Computers & Education

journal homepage: www.elsevier.com/locate/compedu

Effects of face-to-face versus chat communication on performance in a collaborative inquiry modeling task

Patrick H.M. Sins [a,*], Elwin R. Savelsbergh [b], Wouter R. van Joolingen [c], Bernadette H.A.M. van Hout-Wolters [d]

[a] Faculty of Social and Behavioral Sciences, Research Centre Learning in Interaction, Utrecht University, P.O. Box 80140, 3508 TC Utrecht, The Netherlands
[b] Freudenthal Institute for Science and Mathematics Education, University of Utrecht, P.O. Box 80000, 3508 TA Utrecht, The Netherlands
[c] Faculty of Behavioral Sciences, University of Twente, P.O. Box 217, 7500 AE Enschede, The Netherlands
[d] Graduate School of Teaching and Learning, University of Amsterdam, Spinozastraat 55, Amsterdam, The Netherlands

A R T I C L E I N F O

Article history:
Received 9 March 2010
Received in revised form
26 August 2010
Accepted 26 August 2010

Keywords:
Computer-mediated communication
Cooperative/collaborative learning
Interactive learning environments
Secondary education
Simulations

A B S T R A C T

In many contemporary collaborative inquiry learning environments, chat is being used as a means for communication. Still, it remains an open issue whether chat communication is an appropriate means to support the deep reasoning process students need to perform in such environments. Purpose of the present study was to compare the impact of chat versus face-to-face communication on performance within a collaborative computer-supported modeling task. 44 Students from 11th-grade pre-university education, working in dyads, were observed during modeling. Dyads communicated either face-to-face or through a chat tool. Students' reasoning during modeling was assessed by analyzing verbal protocols. In addition, we assessed the quality of student-built models. Results show that while model quality scores did not differ across both conditions, students communicating through chat compressed their interactions resulting in less time spent on surface reasoning, whereas students who communicated face-to-face spent significantly more time on surface reasoning.

1. Introduction

Current views on science education emphasize the importance of getting students engaged in scientific inquiry activities to improve their understanding of both science content and scientific practices. These conceptions of learning can be implemented by having students 'act like scientists' in a collaborative inquiry environment (Reid, Zhang, & Chen, 2003; van Joolingen & de Jong, 1997). In order to attain the intended learning outcomes in such environments students need to engage in deep reasoning which involves active learning processes, such as relating ideas, looking for patterns and principles, and attempting to integrate new information with prior knowledge and experience (Entwistle, 1979, 1988, 2001; Marton & Säljö, 1976, 1997; Ramsden, 1992). Moreover, to allow such processing, students will need to create and maintain common ground in a comparatively ill-defined situation (Clark & Brennan, 1991). A successful implementation of collaborative inquiry, therefore requires intensive high quality discourse among students.

While such discourse may be hard to achieve in many classroom settings (Barron, 2003; Lemke, 1990; Manlove, Lazonder, & de Jong, 2006), it may seem even more difficult to realize in distance learning environments. Yet, in many cases, remote collaboration is required, and consequently many computer-supported collaborative inquiry learning environments provide a chat tool to support students' online communication (e.g., Co-Lab: van Joolingen, de Jong, Lazonder, Savelsbergh, & Manlove, 2005; ModelingSpace: Avouris, Dimitracopoulou, & Komis, 2003; Belvedere; Suthers, 2006; and VCRI: Janssen, Erkens, Kanselaar, & Jaspers, 2007). Although most students are rather fluent with chat tools, the communication channel remains relatively narrow (Baltes, Dickson, Sherman, Bauer, & LaGanke, 2002; Ingram, Hathorn, & Evans, 2000). This means that these tools restrict the exchange of auditory, visual and nonverbal communication cues which normally help groups to regulate interaction, express information and monitor feedback from others (Straus, 1997).

* Corresponding author. Tel.: +31 302534926; fax: +31 302532352.
E-mail addresses: Patrick.Sins@gmail.com (P.H.M. Sins), E.R.Savelsbergh@uu.nl (E.R. Savelsbergh), W.R.VanJoolingen@utwente.nl (W.R. van Joolingen), B.H.A.M.vanHout-Wolters@uva.nl (B.H.A.M. van Hout-Wolters).

0360-1315/$ – see front matter
doi:10.1016/j.compedu.2010.08.022

The aim of the current research is to find out whether chat comprises an appropriate communication channel in a collaborative inquiry learning environment, or whether face-to-face collaboration should be preferred to enable deep reasoning during collaborative inquiry. The study is performed in the context of a collaborative inquiry task involving dynamic systems modeling, implemented in the Co-Lab learning environment (van Joolingen et al., 2005).

1.1. Reasoning processes in collaborative inquiry modeling of dynamic systems

In collaborative inquiry modeling environments, such as Co-Lab and ModelingSpace, students study the behavior of dynamic systems. Traditionally, in secondary science education students gain little insight into these systems, because of the advanced mathematics needed to compute the behavior of these systems. Computer modeling environments may help to overcome these problems, by providing a graphical network presentation of the model, and by taking over the tedious task of solving differential equations, allowing learners to experiment with, revise and evaluate their models (cf., Hogan & Thomas, 2001; Penner, 2001; Schwarz & White, 2005).

The degree to which students obtain a learning benefit from such computer-based modeling, highly depends on the nature of their reasoning. A review conducted by Covington (2000) supports this contention, showing that deep reasoning will create the optimal conditions for achievement in a variety of subject-matter areas, including science, whereas surface processing is linked to poor achievement.

In a previous study, we operationalized deep and surface reasoning processes in the context of scientific inquiry modeling (Sins, Savelsbergh, & van Joolingen, 2005). Therefore, we conducted a detailed analysis of verbal protocols to identify processes secondary students engage in during a collaborative inquiry modeling task. Based on discourse analysis, we identified five types of reasoning processes: *Analyzing, Inductive Reasoning, Quantifying, Explaining and Evaluating.* When students are *analyzing*, they decompose the system they are modeling and identify important model elements (i.e. quantities or relations between quantities) to be implemented in their model. *Inductive reasoning* occurs when students elaborate on how they think elements in their model interact and on how the model should behave. Students can make their ideas about model elements and relations more precise by specifying quantities in the model with numerical values, relations are worked out in equations (i.e. *quantifying*). *Explaining* involves students clarifying to their partners how elements in their model are related to each other. Finally, when *evaluating* their model, students judge whether their model is consistent with their own beliefs, with the data obtained from experiments and/or with descriptions of the system being modeled. Moreover, we found that most of these reasoning processes could take place at different levels of quality. Specifically, a process could occur with or without substantiation. Substantiation could involve reference to experiential knowledge (i.e. knowledge from everyday experience), physics knowledge (i.e. use of terminology, concepts, and formulas common in mathematics), mathematics knowledge (i.e. use of terminology, concepts and formulas common in physics), or to model components. Reasoning processes that are substantiated by referring to knowledge, are taken as indicative of deep reasoning. *Explaining* and *Inductive Reasoning* with substantiation involve deep reasoning by definition and includes students either referring to knowledge or to model components. By contrast, surface reasoning involves unelaborated processes where students do not refer to knowledge.

In a previous study, where 60 secondary students (30 dyads) were communicating through chat in the Co-Lab environment, we could confirm the validity of this schema, as we found that the time spent on deep processing was a significant predictor of achievement ($r = 0.46$, $p < 0.01$), whereas the time spent on surface processing had no additional effect (Sins, van Joolingen, Savelsbergh, & Van Hout-Wolters, 2008).

1.2. Fostering deep reasoning: frameworks to predict the effects of different communication media

An important consideration in the design of computer-supported learning environments aimed at fostering students' collaborative inquiry modeling, is whether chat is an appropriate means of communication in order for students to engage in deep reasoning; or whether it is preferable to perform collaborative modeling face-to-face. As we will argue in this section, one could draw on different theoretical frameworks to argue in support of or against using computer-mediated chat in these environments.

One of the most widely used frameworks to predict the effect of different communication media on performance for different task types, is the task-media fit hypothesis (McGrath & Hollingshead, 1993). This theory states that the effectiveness of a communication mode for a given task depends on the degree of fit between the richness of information that can be transmitted via that system's technology and the information richness requirements of that task. The richness of a medium depends on the following attributes: (1) immediacy of feedback, (2) the number of cues and channels available, (3) language variety and (4) the degree to which intent is focused on the receiver. By these criteria, face-to-face communication is considered the richest medium in that people can rapidly convey content and meaning both verbally and nonverbally, whereas communication through chat is low in media richness since merely textual information is conveyed.

McGrath and Hollingshead (1993) propose four general task categories, namely *generative*; *intellective*; *decision-making*; and *negotiation* tasks, with the latter type requiring the highest information richness. Moreover, Rana, Turoff, and Hiltz (1997) identify task complexity[1] as an important additional task dimension to be taken into consideration. Finally, as Zigurs and Buckland (1998) argue, defining a task should not only include *what* must be accomplished in order to meet stated goals, but also *how* those goals are attained by collaborating partners. Several studies have corroborated the predictions of the task-media fit hypothesis regarding the performance of chat versus face-to-face communication for different task types (e.g., Daly, 1993; Hollingshead, McGrath, & O'Connor, 1993; Mennecke, Valacich, & Wheeler, 2000; Straus & McGrath, 1994; Van der Meijden & Veenman, 2005). Collaborative inquiry modeling has essential characteristics of a *negotiation* task since it requires students to sustain common ground, to agree on a shared approach and to resolve conflicting viewpoints. Therefore, the task-media fit hypothesis would predict that a rich communication mode is required.

By contrast, from an entirely different viewpoint it has been argued that students may be well able to adapt their communication strategies to the constraints of chat and take benefit of its affordances (Condon & Cech, 1996a; 1996b). For instance, since written messages have more permanence than spoken utterances, students communicating through chat have more time to process incoming messages. This

[1] Task complexity is defined here as the subjective mental load a person experiences during a particular task.

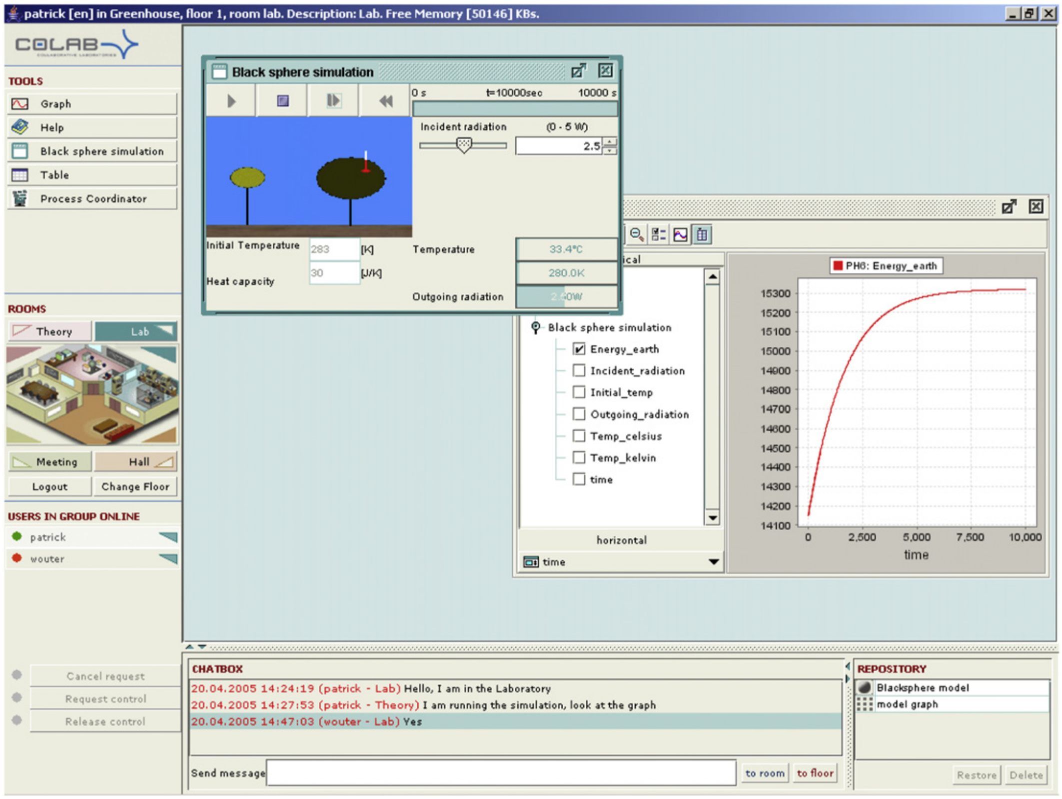

Fig. 1. Screenshot of the simulation of the heating of a black sphere. Results of the experiments are provided in a graph or table.

would mean that they are able to interpret complex instructions more easily than in spoken interaction. In addition, computer-mediated chat allows students to revise messages before they send them which would facilitate the composition of more precise and concise instructions. Finally, chat may give students more time to plan clearer, less ambiguous instructions and messages than are usually found in spoken dialogues, where there is a constant pressure to keep the interaction going.

Condon and Cech (1996a, 1996b) claim that the constraints of computer-mediated chat stimulate students to increase the efficiency of their interaction by compressing it (i.e. compression effect). The compression effect has been corroborated in a survey of Newlands, Anderson, and Mullin (2003) where it is shown that as people adapt to chat they will produce less linguistic output and at the same time change the way in which they collaborate to achieve mutual understanding. Also, Barile and Durso (2002), Basque and Pudelko (2004), and Dennis (2003) found that the use of a chat tool proved to be as effective as face-to-face communication in collaborative writing tasks. Furthermore, Ruberg, Moore, and Taylor (1996) found that synchronous computer-mediated communication leads to more experimentation, sharing of ideas, increased and more distributed participation compared to face-to-face communication. Finally, Condon and Cech (1996a, 1996b) and Jonassen and Kwon (2001) show that groups who use computer-mediated chat compensate for the communication constraints by being more concise in their interactions. They found that students using a chat tool eliminate unnecessary elaborations and repetitions, and seek to increase the efficiency of their communication because of the slower pace of interaction. These students were found to be more task-oriented and compressed their communication resulting in more efficient interactions. With respect to computer-based modeling, this framework expects that students compensate for restrictions of computer-mediated chat by being more selective and more thorough in their interactions than students in the face-to-face condition which would be reflected in more deep reasoning.

1.3. Research question

Based on these theoretical perspectives, no conclusive evidence is provided for or against chat as a communication mode for collaborative inquiry modeling tasks. Consequently, this study will address the following research question: *What is the effect of chat versus face-to-face communication on students' reasoning during modeling and on the quality of their model?*

2. Method

2.1. Participants

This study involved 44 students (aged 16–18 years) from 11th-grade pre-university education, with a major in science. Students from 6 classes from 6 different schools located in Amsterdam, The Netherlands participated in this study.

Based on recommendations from the literature to compose mixed-ability groups without the within group difference becoming too large (e.g. Webb, 1989; Webb & Farivar, 1994); for each participating class students were paired in mixed-ability dyads, based on their grades in physics and mathematics. The selection procedure was such that the dyads were composed of either a low and a middle graded or a middle and a high graded student. All participants were familiar with using chat software. Dyads were randomly assigned to one of two conditions: a synchronous computer-mediated communication condition ($n = 11$) and a face-to-face condition ($n = 11$). To avoid interference between students in different conditions, sessions with chat or face-to-face groups were organized on separate occasions. In addition, each session involved a maximum of 4 dyads working on the modeling task. Students' teachers were not present during these sessions; only technical assistance was provided to students by the first author. Students were awarded € 20 for their participation.

2.2. Material

Student dyads performed the modeling task within the Co-Lab environment (van Joolingen et al., 2005). Their task was to revise a simple pre-built model of the global mean temperature on earth (see Appendix A for the assignment). Students' task was to extend and adapt the initial model, such that the outcomes would be in accordance with the experimental tests they could do on a simulation (see Fig. 1 for a screenshot of the simulation).

Students constructed their models in the model editor tool of Co-Lab. Successful completion of the task would require the identification of the variables heat capacity and temperature, of the relation between these variables, and of a feedback loop which runs from energy earth to temperature (see Fig. 2 for the correct model structure).

Pairs of students in the face-to-face condition sat behind a single screen together, their Co-Lab environment had no chat tool. Students in the chat condition worked on individual computers, and they could only communicate through the Co-Lab chat tool.

2.3. Procedure

The full experiment took two sessions, each of about two and a half hours on two separate days. In the first session, students were first presented with a short plenary introduction to the Co-Lab environment. Then, in both conditions, students were provided with an instruction manual to get acquainted with the system dynamics modeling in Co-Lab. In this instruction manual, students were presented with an example model of a water tank. The water tank model, its elements (i.e. variables and relationships between variables) and how it can be built in Co-Lab were explained and students could execute the model in Co-Lab. The instruction took about one and a half hour. Subsequently, students were asked to revise the water tank model themselves for the remaining time. In the second session, dyads were presented with the black sphere modeling task. Students in both conditions worked for 2 h on the modeling task.

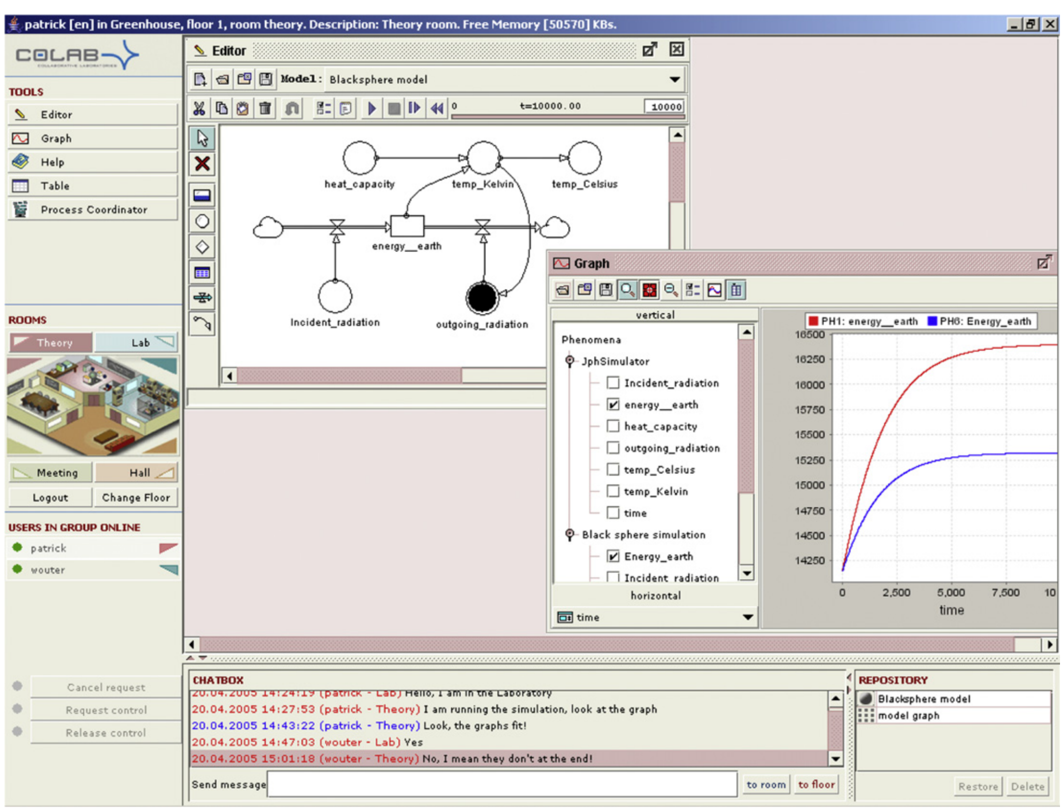

Fig. 2. Screenshot of the model editor in Co-Lab. In this particular example an accurate model diagram is provided of the black sphere simulation. This model shows that the energy of the earth is influenced by the incident radiation from the sun (i.e. energy inflow) and the outgoing radiation (i.e. energy outflow). The outflow is influenced by the temperature of the earth. Finally the temperature is influenced by the energy of the earth and the heat capacity of the earth.

2.4. Analysis

To obtain verbal protocols, voice recordings of dyads working in the face-to-face condition were transcribed, whereas protocols of the chat messages were obtained from the software logs.

Reasoning processes like analyzing or explaining may involve several turns by both partners in a dyad (cf. Naidu & Järvelä, 2006). Therefore, the unit of analysis is the 'process-episode' level (cf. Strijbos, Martens, Prins, & Jochems, 2006; de Wever, Schellens, Valcke, & van Keer, 2006), an episode being a period of coherent continuous talk on a single issue, rather than single utterances (Chi, 1997). Each episode was scored on two dimensions[2]: a) type of reasoning processes; and b) type of reference (see Appendix B). Reasoning episodes in which students are elaborating on the modeling task and connect to their knowledge, which can be either gained from the task at hand or prior knowledge, were designated as deep reasoning. Episodes in which students posit unelaborated statements without reference to available knowledge were labeled as surface reasoning (see Table 1 for definitions and examples). The remaining episodes were scored indifferent, and excluded for the purpose of this analysis. Interrater reliability was determined by calculating Cohen's kappa, which was considered to be acceptable ($n = 176$, Cohen's kappa $= 0.65$) (Heuvelmans & Sanders, 1993).

Because the length of episodes differed from just one or two utterances to several minutes, we calculated proportions of time dyads spent on each reasoning process. This means that we determined the total time dyads spent on a particular reasoning process in minutes and divided this figure by the total amount of time students worked on the modeling task (i.e. 120 min). All further analyses are based on these proportions, unless specified otherwise.

The quality of the students' models was assessed using a strict scoring procedure: points were awarded for each correct variable name and specification; for correct links between variables; and for correct specifications for these relations. Points were subtracted for each incorrect relation.

We chose the Mann–Whitney *U* test to analyze between-group differences, because Shapiro–Wilk tests revealed that the distributions of some of the variables did not satisfy normality assumption. In addition, data showed outliers in these variables making the Mann–Whitney *U* test more robust compared to employing *t*-test (Fay & Proschan, 2010).

Table 1
Definitions and examples of deep and surface reasoning processes.

	Definition	Deep reasoning	Surface reasoning
Analyzing	Interpreting or identifying model elements	S1: Is temperature in Fahrenheit or in Celsius? S2: We should use Kelvin [*reference to physics knowledge*] S1: Yes, the scale of Kelvin	S1: But what is this variable? S2: Let's see, which variables are related to the outflow of energy? S1: Outflow hmm
Inductive reasoning	Qualitative elaboration on model elements	S1: If the earth receives more heath from the sun, it gives off more heath, so it becomes warmer [*reference to experiential knowledge*] S2: If the earth. Oh, yes more radiation S1: Yes S2: Yes, we have a nice hypothesis S1: and the earth will give off more radiation S2: If the earth emits radiation more easily, than it will become much less warm on earth [*reference to experiential knowledge*]	–
Quantifying	Specification of numerical value(s) or equation(s)	S1: Yes, shall we specify the starting value of this variable? S2: That is the default energy of the earth according to our model S1: Yes, I think that we have to specify a value in joule [*reference to physics knowledge*]	S1: Here, I think you should leave it at 50 S2: Outgoing radiation is zero S1: I think you should just leave it at 50 and they are not related to each other
Explaining	Explication of how model elements work or providing rationales why elements were included	S1: What does heath capacity stand for? S2: The capacity of heath S1: Yes, what does that mean? S2: It's the energy that is required to make a particular object warmer with 1° [*reference to physics knowledge*]	–
Evaluating	Judgment of model elements	S1: The model draws a straight line S2: Yes, it should behave according to an exponential function [*reference to mathematics knowledge*] S1: I think it has something to do with the loop between energy earth and the outgoing radiation	S1: Wait a minute, the graph is going straight now S2: What happens? S1: It stays exactly the same S2: It makes no sense, our model is totally wrong

[2] In addition to these two codes, the analysis scheme of Sins et al. (2005) also includes the category: 'topic focus of students' reasoning'. We did not include this code in the present analyses, since reasoning processes coupled with the type of reference students make during process-episodes provide sufficient information concerning students' level of reasoning.

To investigate the relations between reasoning processes and model quality, we calculated non-parametric Spearman rank correlations between model quality scores and the proportion of time students spent on deep and surface reasoning processes. Between-condition differences for these correlations were analyzed more closely by conducting the Fisher Z-test.

3. Results

3.1. Reasoning processes

The total number of process-episodes was strongly different between the two conditions, with dyads in the face-to-face condition having significantly more episodes ($M = 85.64$, SD $= 16.78$) compared to dyads in the chat condition ($M = 20.82$, SD $= 9.78$; $U = 0.0$, $p < 0.001$). Table 2 shows the mean percentages of time dyads spent on the deep and surface processes. As can be seen from the table, the variation between different dyads was rather large. Overall, dyads communicating face-to-face spent significantly more time on surface processes ($M = 25.82\%$, SD $= 8.32$) compared to dyads communicating through chat ($M = 9.76\%$, SD $= 8.32$; $U = 11.0$, $p < 0.001$). No significant difference could be found between the two conditions with respect to deep processes. When inspecting the differences between the two conditions concerning the proportion of time spent on deep and surface processes more closely, several differences were found. Students in the chat condition spent significantly more time than students in the face-to-face condition on inductive reasoning with reference to model components ($U = 31.0$, $p = 0.03$). Within the surface processes category, it turns out that face-to-face dyads spent significantly more time than the chat dyads on: evaluating with no reference to knowledge ($U = 22.0$, $p = 0.01$), quantifying with no reference to knowledge ($U = 18.0$, $p = 0.005$), and analyzing with no reference to knowledge ($U = 25.5$, $p = 0.02$).

For the remaining time, dyads in the face-to-face condition spent significantly more time on off-task communication ($M = 8.14\%$, SD $= 5.82$) compared to the chat group ($M = 4.25\%$, SD $= 7.28$; $U = 28.0$, $p = 0.03$). No significant between-group differences were found on time spent on reading the learning material in Co-Lab ($U = 34.0$, $p = 0.09$) and on other reasoning processes that had not been specified as either deep or surface reasoning ($U = 50.0$, $p = 0.52$).

3.2. Model quality

Model quality scores did not differ significantly between the chat dyads ($M = 6.97$, SD $= 1.41$) and the face-to-face dyads ($M = 7.02$, SD $= 1.74$; $U = 58$, $p = 0.90$).

3.3. Relation between reasoning processes and model quality

In the chat condition we found no significant correlation between total number of process-episodes and model quality score ($r = -0.01$, $p = 0.97$). However, in the face-to-face condition we found a significant negative relation between the number of process-episodes and the quality of students' models ($r = -0.67$, $p = 0.02$). This finding implies that for students in the face-to-face condition, a higher number of brief process-episodes goes together with models of low quality.

Given our previous findings (Sins et al., 2008) we expected a positive correlation between proportion of time spent on deep reasoning and model quality scores, and no effect of the proportion of time spent on surface reasoning. Table 3 shows the correlations between the proportion of time spent on deep and surface reasoning processes at the one hand and students' model quality scores at the other. For the chat condition, the correlation found is well in line with those previous findings, although it is not significant in this study ($r = 0.31$, $p = 0.31$). For the face-to-face condition, correlations are also not significant, but the trends in the data suggest that the relations may be different from those in the chat condition. As shown in Table 3, the data suggest that there is no relation between the proportion of time spent on deep reasoning and model quality score ($r = 0.01$, $p = 0.97$), while there might be a negative correlation between the proportion of time spent on surface reasoning and model quality ($r = -0.43$, $p = 0.17$). Although not significant, this trend in the data is different compared to the chat condition.

Within the chat condition, there is a significant positive correlation between the proportion of time spent on quantifying with reference to knowledge (i.e. deep process) and model quality score ($r = 0.63$, $p = 0.04$). One significant negative correlation is found within the face-to-face group between the proportion of time spent on quantifying with no reference to knowledge (i.e. surface process) and model quality score ($r = -0.76$, $p = 0.007$). The Fisher Z-test reveals that there is a significant difference between the face-to-face and the chat condition with respect to the correlations found between these two variables ($Z' = 2.28$, $p = 0.02$).

Table 2
Mean percentages of time spent on deep and surface processes for the chat condition and the face-to-face condition (standard deviations in parentheses).

	Chat	Face-to-face	U	p
Deep processes	15.45 (23.16)	12.12 (9.73)	55.0	0.72
Evaluating and reference to knowledge	0.62 (1.10)	1.10 (1.20)	37.0	0.11
Explaining and reference to knowledge	0.23 (0.48)	0.00 (0.00)	49.5	0.15
Quantifying and reference to knowledge	5.44 (6.78)	4.60 (5.44)	56.5	0.79
Inductive reasoning and reference to knowledge	0.64 (1.45)	2.49 (3.65)	38.5	0.094
Inductive reasoning and reference to components	7.29 (24.18)	2.21 (3.70)	31.0*	0.03
Analyzing and reference to knowledge	1.20 (1.80)	1.72 (2.24)	49.0	0.42
Surface processes	9.76 (8.40)	25.82 (8.32)	11.0**	0.001
Evaluating and no reference to knowledge	3.50 (6.60)	9.71 (3.94)	22.0**	0.01
Quantifying and no reference to knowledge	3.10 (5.83)	8.00 (4.96)	18.0**	0.005
Analyzing and no reference to knowledge	3.17 (3.79)	8.11 (4.42)	25.5*	0.02

$*p < 0.05$; $**p < 0.01$.

Table 3
Correlations between proportion of time spent on deep and surface reasoning processes at the one hand and model quality scores at the other together with the results of a Fisher *Z*-test for differences between the chat condition and the face-to-face condition.

	Chat	Face-to-face	Z'	*p*
Deep processes	0.31	0.01	0.04	0.97
Evaluating and reference to knowledge	−0.27	−0.32	0.11	0.92
Explaining and reference to knowledge	0.36	0.00	0.75	0.45
Quantifying and reference to knowledge	0.63*	0.48	0.45	0.65
Inductive reasoning and reference to knowledge	0.15	0.34	−0.40	0.69
Inductive reasoning and reference to components	0.49	−0.32	1.74	0.08
Analyzing and reference to knowledge	0.50	−0.12	1.35	0.18
Surface processes	−0.09	−0.43	0.74	0.46
Evaluating and no reference to knowledge	−0.34	−0.26	−0.16	0.87
Quantifying and no reference to knowledge	0.15	−0.76**	2.28*	0.02
Analyzing and no reference to knowledge	−0.08	−0.26	0.37	0.71

$*p < 0.05$; $**p < 0.01$.

4. Discussion

In this study we compared the impact of chat versus face-to-face communication on students' reasoning and performance during collaborative inquiry modeling. Based on our findings, there is no evidence against the use of computer-mediated chat as a communication mode during inquiry modeling, as both the amount of deep reasoning and the resulting model quality scores were the same across both conditions.

In general, our findings are in agreement with the adaptive use of the communication constraints of chat tools as advocated by the compression hypothesis (Condon & Cech, 1996a, 1996b; Jonassen & Kwon, 2001). First, it turned out that dyads in the chat condition spent substantially less time on surface reasoning and off-task communication compared to the face-to-face dyads, whereas the time spent on deep reasoning remained the same. Second, in the face-to-face condition, the total number of reasoning episodes was significantly higher; with the total time on task remaining equal, this implies a more fragmented discourse. Moreover, there was a negative correlation between the number of episodes and model quality scores. In the chat condition there was no such effect. Third, zooming in on a specific type of reasoning process, we found that dyads in the chat condition spent a significantly more time on inductive reasoning with explicit reference to model components. This provides a clear illustration of how students in the chat condition compensate for not being able to point to the model element they are talking about (cf. Suthers, Hundhausen, & Girardeau, 2003).

By contrast, our findings seem to not only disagree with the expectations that would be derived from the task-media fit hypothesis, but also with empirical findings where face-to-face dyads outperform chat groups (e.g. Carey & Kacmar, 1997; Hollingshead et al., 1993; Marttunen & Laurinen, 2009; Mennecke et al., 2000; Straus, 1997; Van der Meijden & Veenman, 2005). We identified two factors that may help explain this discrepancy in findings. First, the groups in the present study consisted of dyads rather than the larger groups used in other studies (cf. Baltes et al., 2002). It may be that the limitations of chat impede communication to a larger extent in larger groups, because it becomes more difficult for students to keep track of the discussion and of each other when communicating through chat. Second, the students in our study had a great deal of experience with using chat as a communication mode. If students are less skilled in typing and using chat software they may be less concise and less task-focused than the dyads in our study, since they first have to learn to use the chat tool.

Given a lack of power, we could not directly corroborate the relation between deep reasoning and model quality in this study. However, for the chat condition findings are in line with those in our previous study (Sins et al., 2008). For instance, we found a significant positive correlation between time spent on quantifying with reference to knowledge (i.e. deep process) for this group. Although in the face-to-face condition correlations for deep reasoning were not significant either, the trend seems to suggest that in this condition more time spent on surface reasoning consistently relates to dyads scoring lower on model quality. For example, correlation analysis within this group revealed that dyads who spent more time quantifying with no reference to knowledge, arrived at models of lower quality. This finding is consistent with the correlation between time spent on quantifying with no reference to knowledge and model quality scores ($r = -0.51$, $p < 0.05$) reported in Sins et al. (2005). In addition, this result is in line with findings of Löhner, van Joolingen, and Savelsbergh (2003). They found that trial-and-error behavior of students, who communicated face-to-face during an inquiry modeling task, was significantly negatively related with their model quality scores. An implication is, that uncomplicated communication has its drawbacks for students' performance in inquiry modeling tasks. In agreement with Liu, Liao, and Pratt (2009) we underline the importance of considering the ways in which users adjust to constraints of computer-mediated communication in order to determine its effects on reasoning and achievement during computer-based modeling.

Although the collaborative computer-supported inquiry modeling task we used in our study was classified as a complex task, students in the chat condition compensated for this by being more selective and more thorough in their interactions than students in the face-to-face condition resulting in more efficient modeling in the former group (cf. Fidas, Komis, Tzanavaris, & Avouris, 2005). However, an important consideration to take into account is that the reported findings are applicable only within the context of the present study and of its participants. It remains an open issue in how far these findings can be generalized to other settings.

All in all, results in the chat and face-to-face groups were equally good. However, inevitably, compared to face-to-face communication, chat slows down the working process and in general chat groups will take longer than face-to-face groups to complete a given learning task (Baltes et al., 2002; Bordia, 1997; Chen, Chen, & Tsai, 2009). Therefore, we speculate that, given sufficient time, the chat group may outperform the face-to-face group. Whether this hypothesis will hold true for this task, and for other complex collaborative inquiry tasks, needs to be tested in further research.

Acknowledgements

We are grateful to the teachers and their students for their participation in the study reported in this paper. We also would like to thank Iris Dicke and Johan van Strien for their assistance in transcribing and coding the data.

Appendix A. Task Co-Lab black sphere

There has been a lot of publicity about the earth's changing climate. Scientists all around the world are trying to understand what is going on, in order to predict what will happen next, or maybe more importantly, to give advice on what to do about it. The earth's climate is a very complex system, however, and even with all those scientists working on it, uncertainties remain. In such a situation scientists usually begin by making all kinds of simplifications. They first try to understand this simplified system, for instance by making a computer model. Then they use the computer model of the simplified system to make predictions about the real earth. Then they compare their predictions to reality, and consider what refinements are most needed.

In this module, you'll take a similar approach. We have made a very simplified small scale version of the earth and the sun: In our laboratory we have ignored the differences between oceans, forests and deserts. All that remains of the earth is a black sphere and at some distance you'll find a strong light, which takes the function of the sun. Not too similar to the world we live in, you'll say, and you are right. Nevertheless, you can investigate how the earth's temperature responds to changes of solar activity, and what the effects will be if the earth's surface changes color, for instance because it becomes covered with ice. Once you have got a computer model to make proper predictions about this simplified situation, you'll have discovered the basic model structure that underlies even the most advanced climate models today.

To summarize, your goal in this module will be to build a model that can predict the temperature of a black sphere (the barren earth) after being exposed to a source of light (the sun) for some while. To assist you, we provide you with an initial but still incomplete model. This model shows that the energy content of the earth is influenced by an energy inflow (incident radiation from the sun) and an energy outflow (outgoing radiation). Your goal is to extend this model in such a way in that it will provide a provide a good match with the data you obtained from the simulation of the black sphere. In order to fulfill this goal, you'll need to find out first which factors play a role, and how they depend on each other.

Appendix B. Coding scheme for cognitive processes employed during modeling

Type of reasoning process	
Analyze	Interpreting or identifying model elements
Inductive reasoning	Qualitative elaboration on model elements
Quantify	Specification of numerical value(s) or equation(s)
Explain	Explication of how model elements work or providing rationales why elements were included
Evaluate	Judgment of model elements
Other processes	
Read	Students read or paraphrase
Off-task (no further coding)	Students talk about topics unrelated to the inquiry modeling task
Type of reference	
None	No reference to model components or knowledge
Physics knowledge	Use of terminology, concepts (i.e. units, quantities), formula's common in physics
Mathematics knowledge	Use of terminology, concepts, formula's common in mathematics
Experiential knowledge	Knowledge from everyday experience
Correspondence model graph and data	Students refer to (the extent of) correspondence between model output and experimental data
Data from simulation	Experimental data from the simulation (i.e. table/graph)
Html-documents	Information about the black sphere problem provided in the html-documents

References

Avouris, N. M., Dimitracopoulou, A., & Komis, V. (2003). On analysis of collaborative problem solving: an object-oriented approach. *Computers in Human Behavior, 19*(2), 147–167.
Baltes, B. B., Dickson, M. W., Sherman, M. P., Bauer, C. C., & LaGanke, J. S. (2002). Computer-mediated communication and group decision making: a meta-analysis. *Organizational Behavior and Human Decision Processes, 87*(1), 156–179.
Barile, A. L., & Durso, F. T. (2002). Computer-mediated communication in collaborative writing. *Computers in Human Behavior, 18*, 173–190.
Barron, B. (2003). When smart groups fail. *The Journal of the Learning Sciences, 12*(3), 307–359.
Basque, J., & Pudelko, B. (14–17 September, 2004). The effect of collaborative knowledge modeling at distance on performance and on learning. In *Paper presented at the first international conference on concept mapping*. Pamplona, Spain.
Bordia, P. (1997). Face-to-face versus computer-mediated communication: a synthesis of the experimental literature. *The Journal of Business Communication, 34*(1), 99–120.
Carey, J. M., & Kacmar, C. J. (1997). The impact of communication mode and task complexity on small group performance and member satisfaction. *Computers in Human Behavior, 13*(1), 23–49.
Chen, Y., Chen, N., & Tsai, C. (2009). The use of online synchronous discussion for web-based professional development for teachers. *Computers & Education, 53*, 115–1166.
Chi, M. T. H. (1997). Quantifying qualitative analyses of verbal data: a practical guide. *The Journal of the Learning Sciences, 6*(3), 271–315.
Clark, H. H., & Brennan, S. E. (1991). Grounding in communication. In L. B. Resnick, J. Levine, & S. D. Behrend (Eds.), *Perspectives on socially shared cognition* (pp. 127–149). Washington, DC: American Psychological Association.
Condon, S. L., & Cech, C. G. (1996a). Functional comparison of face-to-face and computer-mediated decision-making interactions. In S. C. Herring (Ed.), *Computer-mediated communication: Linguistic, social, and cross-cultural perspectives* (pp. 65–80). Philadelphia: John Benjamin.
Condon, S. L., & Cech, C. G. (1996b). Discourse management in face-to-face and computer-mediated decision-making interactions. *Electronic Journal of Communication/La Revue Electronique de Communication, 6*(3). Retrieved August 3, 2009, from. http://www.cios.org/EJCPUBLIC/006/3/006314.HTML.

Covington, M. V. (2000). Goal theory, motivation, and school achievement: an integrative review. *Annual Review of Psychology, 51*, 171–200.

Daly, B. L. (1993). The influence of face-to-face versus computer-mediated communication channels on collective induction. *Accounting, Management, and Information Technologies, 3*(1), 1–22.

de Wever, B., Schellens, T., Valcke, M., & van Keer, H. (2006). Content analysis schemes to analyze transcripts of online asynchronous discussion groups: a review. *Computers & Education, 46*(1), 6–28.

Dennis, J. K. (2003). Problem-based learning in online vs. face-to-face environments. *Education for Health, 16*(2), 198–209.

Entwistle, N. (1979). Stages, levels, styles, or strategies: dilemmas in the description of thinking. *Educational Review, 31*, 123–132.

Entwistle, N. (1988). *Styles of learning and teaching*. London: David Fulton.

Entwistle, N. (2001). Styles of learning and approaches to studying in higher education. *Learning Styles in Higher Education, 30*(5/6), 593–602.

Fidas, C., Komis, V., Tzanavaris, S., & Avouris, N. (2005). Heterogeneity of learning material in synchronous computer-supported collaborative modelling. *Computers & Education, 44*, 135–154.

Fay, M. P., & Proschan, M. A. (2010). Wilcoxon–Mann–Whitney or *t*-test? On assumptions for hypothesis tests and multiple interpretations of decision rules. *Statistics Surveys, 4*, 1–39.

Heuvelmans, A. P. J. M., & Sanders, P. F. (1993). Beoordelaarsoverstemming. [Interjudgement reliability measurement]. In T. J. H. M. Eggen, & P. F. Sanders (Eds.), *Psychometrie in de praktijk*. [*Psychometrics in practice*] (pp. 443–469). Arnhem, The Netherlands: CITO.

Hogan, K., & Thomas, D. (2001). Reasoning comparisons of students' systems modeling in ecology. *Journal of Science Education and Technology, 10*(4), 319–345.

Hollingshead, A. B., McGrath, J. E., & O'Connor, K. M. (1993). Group task performance and communication technology: a longitudinal study of computer-mediated versus face-to-face work groups. *Small Group Research, 24*(3), 307–333.

Ingram, A. L., Hathorn, L. G., & Evans, A. (2000). Beyond chat on the Internet. *Computers & Education, 35*, 21–35.

Janssen, J., Erkens, G., Kanselaar, G., & Jaspers, J. (2007). Visualization of participation: does it contribute to successful computer-supported collaborative learning? *Computers & Education, 49*(4), 1037–1065.

Jonassen, D. H., & Kwon, H. I. (2001). Communication patterns in computer-mediated versus face-to-face group problem solving. *Educational Technology Research and Development, 49*(1), 35–51.

Lemke, J. (1990). *Talking science: Language, learning, and values*. Norwood, NJ: Ablex/Elsevier.

Liu, S., Liao, H., & Pratt, J. A. (2009). Impact of media richness and flow on e-learning technology acceptance. *Computers & Education, 52*, 599–607.

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

Manlove, S., Lazonder, A. W., & de Jong, T. (2006). Regulative support for collaborative scientific inquiry learning. *Journal of Computer Assisted Learning, 22*(2), 87–98.

Marton, F., & Säljö, R. (1976). On qualitative differences in learning. I–outcome and process. *British Journal of Educational Psychology, 46*, 4–11.

Marton, F., & Säljö, R. (1997). Approaches to learning. In F. Marton, D. J. Hounsell, & N. J. Entwistle (Eds.), *The experience of learning* (2nd ed.). (pp. 39–58) Edinburgh: Scottish Academic Press.

Marttunen, M. J., & Laurinen, L. I. (2009). Secondary school students' collaboration during dyadic debates face-to-face and through computer chat. *Computers in Human Behavior, 25*, 961–969.

McGrath, J. E., & Hollingshead, A. B. (1993). Putting the "group" back in group support systems: some theoretical issues about dynamic processes in groups with technological enhancements. In L. M. Jessup (Ed.), *Group support systems: New perspectives* (pp. 78–96). New York: MacMillan Publishing.

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

Naidu, S., & Järvelä, S. (2006). Analyzing CMC content for what? *Computers & Education, 46*(1), 96–103.

Newlands, A., Anderson, A., & Mullin, J. (2003). Adapting communicative strategies to computer-mediated communication: an analysis of task performance and dialogue structure. *Applied Cognitive Psychology, 17*, 325–348.

Penner, D. E. (2001). Cognition, computers, and synthetic science: building knowledge and meaning through modeling. In W.G.Secada. (Ed.), *Review of research in education* (pp. 1–35). Washington DC: American Educational Research Association.

Ramsden, P. (1992). *Learning to teach in higher education*. London: Routledge.

Rana, A. R., Turoff, M., & Hiltz, S. R. (1997). Task and technology interaction (TTI): a theory of technological support for group tasks. In*Proceedings of the thirtieth annual Hawaii international conference on system sciences, Vol. 2* (pp. 66–76). Washington DC: IEEE Computer Society.

Reid, D. J., Zhang, J., & Chen, Q. (2003). Supporting scientific discovery learning in a simulation environment. *Journal of Computer Assisted Learning, 19*(1), 9–20.

Ruberg, L. F., Moore, D. M., & Taylor, C. D. (1996). Student participation, interaction, and regulation in a computer-mediated communication environment: a qualitative study. *Journal of Educational Computing Research, 14*(3), 243–268.

Schwarz, C. V., & White, B. Y. (2005). Metamodeling knowledge: developing students' understanding of scientific modeling. *Cognition and Instruction, 23*(2), 165–205.

Sins, P. H. M., Savelsbergh, E. R., & van Joolingen, W. R. (2005). The difficult process of scientific modeling: an analysis of novices' reasoning during computer-based modeling. *International Journal of Science Education, 27*(14), 1695–1721.

Sins, P. H. M., van Joolingen, W. R., Savelsbergh, E. R., & Van Hout-Wolters, B. H. A. M. (2008). Motivation and performance within a collaborative computer-based modeling task: relations between students' achievement goal orientation, self-efficacy, cognitive processing, and achievement. *Contemporary Educational Psychology, 33*(1), 58–77.

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

Straus, S. G., & McGrath, J. E. (1994). Does the medium matter? The interaction of task type and technology on group performance and member reactions. *Journal of Applied Psychology, 79*(1), 87–97.

Strijbos, J., Martens, R. L., Prins, F. J., & Jochems, W. M. G. (2006). Content analysis: what are they talking about? *Computers & Education, 46*(1), 29–48.

Suthers, D. D. (2006). A qualitative analysis of collaborative knowledge construction through shared representations. *Research and Practice in Technology Enhanced Learning, 1*(2), 1–28.

Suthers, D. D., Hundhausen, C. D., & Girardeau, L. E. (2003). Comparing the roles of representations in face-to-face and online computer supported collaborative learning. *Computers & Education, 41*, 335–351.

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

van Joolingen, W. R., & de Jong, T. (1997). An extended dual search space model of scientific discovery learning. *Instructional Science, 25*(5), 307–346.

van Joolingen, W. R., de Jong, T., Lazonder, A. W., Savelsbergh, E. R., & Manlove, S. (2005). Co-Lab: research and development of an online learning environment for collaborative scientific discovery learning. *Computers in Human Behavior, 21*, 671–688.

Webb, N. M. (1989). Peer interaction and learning in small groups. *International Journal of Educational Research, 13*, 21–39.

Webb, N. M., & Farivar, S. (1994). Promoting helping behavior in cooperative small groups in middle school mathematics. *American Educational Research Journal, 31*(2), 369–395.

Zigurs, I., & Buckland, B. K. (1998). A theory of task/technology fit and group support systems effectiveness. *Management Information Systems Quartely, 22*(3), 313–334.