

Facilitating argumentative knowledge construction with computer-supported collaboration scripts

Karsten Stegmann · Armin Weinberger · Frank Fischer

Received: 19 July 2006 / Accepted: 13 September 2007 /
Published online: 3 November 2007

© International Society of the Learning Sciences, Inc.; Springer Science + Business Media, LLC 2007

Abstract Online discussions provide opportunities for learners to engage in argumentative debate, but learners rarely formulate well-grounded arguments or benefit individually from participating in online discussions. Learners often do not explicitly warrant their arguments and fail to construct counterarguments (incomplete formal argumentation structure), which is hypothesized to impede individual knowledge acquisition. Computer-supported scripts have been found to support learners during online discussions. Such scripts can support specific discourse activities, such as the construction of single arguments, by supporting learners in explicitly warranting their claims or in constructing specific argumentation sequences, e.g., argument-counterargument sequences, during online discussions. Participation in argumentative discourse is seen to promote both knowledge on argumentation and domain-specific knowledge. However, there have been few empirical investigations regarding the extent to which computer-supported collaboration scripts can foster the formal quality of argumentation and thereby facilitate the individual acquisition of knowledge. One hundred and twenty (120) students of Educational Science participated in the study with a 2×2-factorial design (with vs. without script for the construction of single arguments and with vs. without script for the construction of argumentation sequences) and were randomly divided into groups of three. Results indicated that the collaboration scripts could improve the formal quality of single arguments and the formal quality of argumentation sequences in online discussions. Scripts also facilitated the acquisition of knowledge on argumentation, without affecting the acquisition of domain-specific knowledge.

Keywords Computer-supported collaboration scripts ·
Argumentative knowledge construction · Online discussion

An important goal of university education is for students to learn how to apply specific theoretical concepts in order to develop warranted arguments and counterarguments. In

K. Stegmann (✉) · A. Weinberger · F. Fischer
LMU München, Leopoldstrasse 13, 80802 Munich, Germany
e-mail: karsten.stegmann@psy.lmu.de

addition, it has been assumed that such argumentative processes might be positively related to the acquisition of domain-specific knowledge (Baker 2003; Kuhn and Goh 2005). Computer-supported collaborative learning in asynchronous settings might provide students with the appropriate conditions for acquiring knowledge on argumentation as well as domain-specific knowledge, as learners have the time they need to engage in high-quality argumentative processes (Kuhn and Goh 2005; Marttunen 1992). However, learners rarely use this time advantage, unless they are provided additional support for constructing better arguments and argumentation sequences. One approach for fostering the quality of argumentation involves the use of computer-supported collaboration scripts (Kollar et al. 2005). The quality of argumentation, in turn, can be hypothesized as positively related to the individuals' acquisition of knowledge.

Although plausible, most of these assumptions have not yet been subject to systematic empirical investigation. Therefore, in this empirical study, we investigate how different computer-supported collaboration scripts can improve specific argumentative processes. In addition, we examine the effects of these collaboration scripts on online discussions.

Computer-supported collaboration scripts

Collaboration scripts are instructional plans that specify and sequence collaborative learning activities. When needed, these scripts assign various activities to the different learners (Kobbe et al. 2007). Collaboration scripts typically focus on those activities that are associated with deeper cognitive elaboration and thereby facilitate knowledge acquisition, but that learners seldom perform spontaneously (King 2007). The script approach is not limited to specific activities, but may focus on those activities regarded as beneficial for learning (Kobbe et al. 2007; Kollar et al. 2006) and/or aim to reduce extraneous activities like off-topic talk (Baker and Lund 1997). Hence, each collaboration script is explicitly or implicitly based on an approach to collaborative learning that specifies crucial cognitive and social processes as well as associated activities, learning goals, and the relationship between these elements.

Collaboration scripts differ in the degree to which the sequences of activities are supposed to be internalized by the individual. According to Pea (2004), structuring approaches may aid learning in at least two different ways. Approaches that focus on the internalization of the scripted activities can be termed *scaffolding approaches to scripting*. Examples for these approaches are Scripted Cooperation (O'Donnell and Dansereau 1992), Ask to Think–Tel Why (King 1997), Structured Academic Controversy (Johnson and Johnson 1994), and Reciprocal Teaching (Palincsar and Brown 1984). Other approaches consider collaboration scripts as external aids for better understanding complicated domain concepts or processes. These external aids are not meant to be internalized by the individual. This second group of approaches can be labeled as *distributed intelligence approaches to scripting* (Pea 2004). In the case of scripts for argumentation that aim to facilitate both knowledge on argumentation and domain knowledge, the *scaffolding approach* seems to be appropriate. Collaboration scripts should provide learners with a scaffold to enable them to participate in high-quality argumentation far beyond their current level of competence and construct knowledge on argumentation that is distributed by the script.

Different *computer-supported collaboration scripts* have been developed and explored that support online discussions by integrating scaffolds into the communication interface (Baker and Lund 1997; Hron et al. 1997; Jermann and Dillenbourg 2003; Kollar et al. 2005). The typical implementations are prompts or sentence starters (Nussbaum et al. 2002,

April; Weinberger 2003), buttons that open text-boxes for specific speech acts (Baker and Lund 1997; Hron et al. 1997), or input text fields (Kollar et al. 2005). These elements are typically accompanied by instructional hints that specify the activity to be performed. However, scripts also constrain collaboration. It has been suggested that if scripts are too restrictive, they might reduce cognitive processing and motivation due to over-scripting (Cohen 1994; Dillenbourg 2002). To what extent over-scripting occurs may heavily depend on the individual's prior knowledge. It has been assumed that learners have knowledge on how to act in collaborative learning, i.e., they have an internal script (Kollar et al. 2006) that might guide and constrain individual activities. The instructional support provided can be regarded as an external script (Kollar et al. 2006). Kollar et al. (2005) described this interplay between external (i.e., instructional) collaboration script and the learner's internal (i.e., cognitive) script for the task as a "person-plus-script system," as described by Perkins (1993). On the one hand, if the internal script is weak and the external script provides little structure, the person-plus-script system may not have enough regulatory "knowledge" to handle the complex task at hand. On the other hand, if the internal script and the external script are both well developed, learners may be frustrated by being restricted by the external script. Therefore, it has been argued that the external collaboration script should complement the learner's internal script for the task as much as possible to facilitate activities that foster individual knowledge acquisition (Carmien et al. 2007).

Scripting argumentative knowledge construction

Which specific argumentative activities have to be addressed by a collaboration script to improve online discussions and to foster knowledge construction? In this section, we describe two perspectives on argumentation in discussions and identify components and processes that can be facilitated using collaboration scripts.

Perspectives on argumentation

The formal quality of argumentation can be described by at least two dimensions, i.e., the construction of single arguments and the construction of argumentation sequences (Weinberger and Fischer 2006). Focusing on the construction of single arguments emphasizes the individual aspects of argumentation, such as the explicit occurrence of a reason (van Eemeren 2003; Voss et al. 1983). Focusing on the construction of argumentation sequences places more emphasis on mutual reference during argumentation, such as arguments that counter the arguments of a learning partner (Jermann and Dillenbourg 2003; Resnick et al. 1993). These aspects can be regarded as complementary perspectives on argumentation (Kuhn et al. 1997).

In research on collaborative argumentation, Toulmin's (1958) argumentation model is often cited (Clark et al. 2007; Kuhn 1991; Leitão 2000; Means and Voss 1996). His model has been regarded as an alternative to the approach of formal logic. In formal logic, real-life argumentation is always incomplete to some degree and no valid inferences can be drawn. For example, the argument "George will succeed in his professional life, because he is a good school student" must be prefaced by the premise that "Students who are good in school have good chances in professional life." However, in real-life argumentation, these elements are often implicit. Against the background of formal logic, all arguments can be evaluated by universal norms. In Toulmin's (1958) model, single arguments can be described in a universal, formal way, but the quality of an argument can be only evaluated

with respect to the nature of the problem context. Therefore, he distinguishes between several components, some of which are optional. Due to the number of interrelated and partly optional components, the model is quite complex and is typically not applied fully in everyday argumentation. Doubts have been expressed with respect to its appropriateness as a prescriptive model for learners (see Voss and van Dyke 2001). It has been argued that a less complex model might be more beneficial for learners. Therefore, we propose a simplified core model that comprises the components of claim, grounds (which support the claim) and qualifications (which limit the validity of the claim). In this simplified model, a completely explicit argument consists of a claim supported by grounds and limited by qualifications. The *claim* expresses the position on an argument, e.g., “I’m against it” or (in the context of attribution theory) “The father is attributing internally stable.” The elements that support the claim in Toulmin’s (1958) model, which we subsume under the term grounds, are *data*, *warrant* and *backing*. *Data* involves factual information, such as an observation, which supports the acceptance of the claim. In learning settings, this information is provided within the description of a problem case, e.g., “The father said that his daughter has the ability to succeed in her exam.” A *warrant* justifies the inference between data and claim. Warrants are usually theoretical laws, rules, or definitions, e.g., a warrant derived from attribution theory may be “Ascribing success to ability is an internal attribution.” *Backing* is evidence such as statistics or expert opinions that is in line with the warrant, e.g. “Studies have shown that ascribing causes for success in exams has a significant effect on future motivation to learn.” Qualifiers and their associated rebuttals are the elements in Toulmin’s (1958) model that limit the validity of a claim. We use the term *qualifications* for these elements, because they qualify the relationship between claim and warrant. Toulmin’s (1958) *qualifier* reflects uncertainty with regard to the validity of the claim and is usually expressed using modal adverbs such as “perhaps” or “probably.” While the qualifier only expresses a potential limitation, the so-called *rebuttal* describes the state of affairs when the claim is invalid, e.g., “[...] provided that the father tells the truth.”

Construction of argumentation sequences captures the dynamic of argumentative dialogue consisting of arguments, counterarguments, and integrations. The ideal pattern proposed by Leitão (2000) is designed to organize turns in a way that promotes the construction of valid knowledge in collaboration. In this dynamic sequence, learning partners first try to justify their (initial) position by constructing *arguments*. Then, *counterarguments* challenge this position and may lead to a reconsideration of the initial argument. The counterargument is not necessarily the opposite of the initial argument, but calls it into question. As a minimum, a counterargument makes the acceptability of the initial position less certain. Finally, learners construct replies and may possibly synthesize their initial positions in an *integration* or decide which alternatives fit the best.

However, these models for formally describing argumentation do not answer the question of what the formal quality of argumentation means for knowledge construction. On one hand, it may not be beneficial in every context to formulate complete arguments that include all components. For example, in a debate (i.e., a conversation with the goal of convincing the audience), constructing all single arguments with “probably” as qualifier and constructing rebuttals for each claim may not be appropriate for achieving the intended goal of convincing the audience. On the other hand, for acquiring well-structured knowledge in a certain area, it may be beneficial for a group of learners to identify the limits of claims as well as possible counterarguments. Hence, we argue that the criteria used for evaluating the formal structure of argumentation must be defined with consideration given to type of dialogue at hand. Walton and Krabbe (1995) proposed a model of ideal dialogue types that were further developed by Keefer et al. (2000). The framework of

Keefer and colleagues differentiates between four dialogue types depending on the initial situation, the dialogue's main goal, the methods or means of the dialogue, and the participants' goals. The four dialogue types are (a) the critical discussion, (b) the explanatory inquiry, (c) the eristic discussion, and (d) the consensus dialogue. Walton and Krabbe (1995) and Keefer et al. (2000) emphasize that these types are ideal descriptions, but typically overlap with one another. For example, knowledge acquisition is the main goal of an explanatory inquiry. However, during an explanatory inquiry, two learners can defend different opinions as they would in a critical discussion. Hence, a way to include contextual aspects when defining the quality of argumentation is to consider the dialogue type. Based on the framework of Keefer et al. (2000), the construction of knowledge through collaborative argumentation can be classified as explanatory inquiry. In this dialogue type, a higher degree of formal structure of single arguments is supposed to be related to deep cognitive elaboration. Providing grounds (Baker 2003) and qualifying the relationship between claim and grounds by considering possible alternative viewpoints or explanations (Kuhn et al. 1997; Spiro and Jehng 1990) is assumed to contribute to the active construction of a knowledge representation and to be related to deeper cognitive elaboration. For example, Baker (2003) identifies analogies between supporting claims and self-explaining (Chi et al. 1989). In this line of thinking, the formal structure of a single argument can be regarded as an indicator of self-explaining. Furthermore, a high degree of formal structure in a single argument can be seen as an elaborated representation of the situation model (Kintsch 1991) of the individual learner that bases a larger number of claims on non-superficial inferences. First, claims that are supported by grounds can be regarded as self-explanations that indicate a high degree of formal structure. Second, qualifications of the relationship between claim and ground indicate the consideration of alternative (self-) explanations that indicate a high degree of formal structure. In turn, this deeper cognitive elaboration might foster knowledge construction.

What is the formal quality of argumentation with respect to argumentation sequences? It has been hypothesized that the construction of complete argumentation sequences, and thereby the construction of this kind of sophisticated knowledge representations, is related to deep cognitive processes (Leitão 2000). Andriessen et al. (2003a) have argued that comparing and contrasting different positions during an argument may induce socio-cognitive conflicts *sensu* Doise and Mugny (1984). Socio-cognitive conflicts occur when learners discuss divergent or incompatible views. While counterarguments may raise socio-cognitive conflicts, integrations may initiate the process of resolving the socio-cognitive conflicts (Nastasi and Clements 1992).

However, Leitão (2000) raises the point that counterarguments may not always be the opposite of an initial argument and, therefore, may not always increase socio-cognitive conflicts. They may “merely” present a different perspective on the same issue and thereby broaden the understanding of the concepts being learned. Considering counterarguments may lead the learner to elaborate different perspectives, compare different possible solutions, and decide which solution is most likely. Therefore, according to Leitão (2000), argumentation sequences that foster knowledge construction can be described according to this sequence: argument, counterargument, and integration.

Having defined what we mean by the formal quality of argumentation in the context of construction of knowledge through collaborative argumentation, we now turn to the question of *learning outcomes*: What can be learned by participating in collaborative argumentation? Argumentative knowledge construction has been seen to foster both the acquisition of knowledge on argumentation (Kuhn 1991) as well as domain-specific knowledge. In reviewing the existing literature (Andriessen et al. 2003b; Astleitner 2002;

Baker 2003; Clark et al. 2007; Dillenbourg 2004; Hanley 1995; King 1997; Koschmann 2003; Kuhn 1991; Leitão 2000; Marttunen 1994; Pithers 2000; Stein and Miller 1996), the following assumption seems rather straightforward: since learners engaged in collaborative argumentation are elaborating the domain content of their discussion more deeply, they should thus acquire more and better organized domain knowledge (Jermann and Dillenbourg 2003). Beyond acquiring domain-specific knowledge through argumentation, Kuhn's work (Kuhn 1991; Kuhn and Goh 2005) shows evidence that learners engaged in argumentative knowledge construction actually acquire domain-general *knowledge on argumentation*.

Although knowledge on argumentation starts to develop from an early age (Stein and Bernas 1999), studies have shown that adults' knowledge on argumentation is often applied in a suboptimal manner. Adults rarely base their claims on grounds (Kuhn et al. 1997) and rarely consider counterarguments (Leitão 2000). However, studies also show that learners who are engaged in argumentative discourse acquire knowledge on argumentation (Kuhn et al. 1997). The acquisition of knowledge on argumentation in online discussions may be further promoted by fostering the formal quality of argumentation in online discussions. Having introduced argumentative knowledge construction as an approach to collaborative learning, we now turn our attention to how the processes of collaborative argumentation can be facilitated with a computer-supported script.

Argumentative computer-supported collaboration scripts

To date, several empirical studies have provided evidence that computer-supported collaboration scripts are able to foster specific processes and outcomes of argumentative knowledge construction (Jermann and Dillenbourg 2003; Kollar et al. 2005; Weinberger 2003). For example, scripts that were developed to foster the quality of argumentation sequences enhanced both the quality of argumentation sequences and domain-specific knowledge acquisition (Jermann and Dillenbourg 2003). For a more detailed description of the script approach see Kobbe et al. (2007). However, only the higher quality of social interaction served to improve domain-specific knowledge acquisition. Weinberger (2003) developed two scripts for the facilitation of both specific patterns of social interaction and specific epistemic activities. These scripts were able to enhance the specific qualities of online discussions for which they were designed, i.e., a social script facilitated specific social interactions and an epistemic script facilitated specific epistemic activities. The script supporting specific epistemic activities allowed the task to be solved with a lesser degree of cognitive engagement. Hence, the higher quality of online discussion is not a sufficient condition for facilitating knowledge acquisition. To some degree, the script must also make collaboration more extensive with respect to the activities that are seen to foster knowledge acquisition.

Collaborative learners may be supported in constructing a single argument according to the simplified Toulmin (1958) model and in the construction of argumentation sequences according to the Leitão (2000) cycle of knowledge construction. Scripts for the construction of single arguments may encourage the use of grounds (data, warrant, and/or backing) or supporting a claim using some qualification (qualifier and/or rebuttal) to indicate the consideration of alternative explanations. Each component of the complete simplified core structure of a single argument might be supported by specific structures implemented in the user-interface. Scripting the construction of argumentation sequences provides a socio-cognitive structure (together with the associated hints and prompts) that supports the construction of counterarguments and integrations and thus fosters extended argumentation sequences. The script therefore should both help differentiate the components of a complete argumentation sequence (argument, counterargument, and integration) as well as suggest

specific sequences for these components. Computer-supported scripts should facilitate the differentiation and sequencing of argumentation sequence components within the computer interface through a dedicated socio-cognitive structure and the associated hints and prompts.

Research questions

To date, research has not focused systematically on supporting both of the aspects of argumentation (structure of single arguments and argumentation sequences) with appropriate support measures. It is unclear to what extent the formal quality of argumentation can be enhanced using collaboration scripts. Furthermore, there has been little experimental research on the assumption that the enhanced formal quality of collaborative argumentation has a positive effect on individual knowledge acquisition. We have formulated two research questions to address these issues:

1. To what extent is the formal quality of argumentation in online discussion affected by a script for the construction of single arguments, a script for the construction of argumentation sequences, and their combination?

We expect that a script for the construction of single arguments should improve the formal quality of single arguments. Moreover, we expect that a script for the construction of argumentation sequences should improve the formal quality of argumentation sequences. The scripts are designed to affect different dimensions of argumentation and hence, we expect additive effects for both components (i.e., no interaction of the two factors).

2. To what extent is individual knowledge acquisition affected by a script for the construction of single arguments, a script for the construction of argumentation sequences, and their combination?

Regarding the individual's acquisition of knowledge on argumentation, we expect that the support from dedicated collaboration scripts should firstly facilitate the acquisition of knowledge on argumentation, as the necessary information about arguments and argumentation is accessible through the scripts' representations. The script for the construction of single arguments should foster knowledge on single arguments while the script for the construction of argumentation sequences should foster knowledge on argumentation sequences. Both collaboration scripts should facilitate the acquisition of domain-specific knowledge by encouraging learners to participate actively in argumentation. Better counterarguments and argumentation sequences with well-formed arguments may facilitate domain-specific knowledge more than the added effects of counterarguments with bare claims plus grounded and qualified arguments that are not related to other arguments. Therefore, the combined use of both scripts may lead to a positive interaction effect, i.e., learners supported by both scripts may acquire more domain-specific knowledge as compared to the added effects of both scripts used in isolation.

Method

Participants and design

One hundred twenty (120) students of Educational Science at the University of Munich participated in this study. The mean age of the participants was $M = 23.08$ ($SD = 4.05$) years,

of which 101 (84.2%) were female, while 19 (15.8%) were male. This represents the regular proportion between female and male students of Educational Science at the University of Munich. With respect to gender composition, 25 (62.5%) groups consisted only of female participants, 15 (37.5%) groups had members of both genders, and no groups consisted only of male participants. Participation was required in order for freshmen to receive a course credit at the end of the term for a mandatory introduction course. In this way, the experimental learning environment was included as part of the regular curriculum. However, the learning outcomes of the experimental session did not count towards the students' overall grade. The experimental session involved a motivational theory (see below) usually covered in the introductory course and took the place of a three-hour lecture session plus seminar. We independently varied (1) *the script for the construction of single arguments* (without vs. with) and (2) *the script for the construction of argumentation sequences* (without vs. with). The participants were assigned to groups of three and each group was randomly assigned to one of the four experimental conditions. Thus, each cell of the 2×2 -design contained ten groups of three. The gender composition within the four experimental conditions was distributed similarly. The pure female groups made up 50 to 70%. The groups with mixed gender made up 50 to 30%. The three students working together electronically were each assigned to one of three different laboratory rooms.

Unit of analysis and statistical tests

Learners within one group of three cannot be seen as acting independently. As a result, the measures observed for these learners also cannot be regarded as independent. Therefore, we used the values of one individual per group. Because individual knowledge acquisition is a main point of interest within this study, we decided to use the individual learners as the unit of analysis. Hence, we randomly selected one learner from each group of three. Therefore, in the analysis, each of the 40 groups is represented by one member. The mean age of the selected participants was $M = 23.18$ ($SD = 3.69$) years. Thirty-one (77.5%) were female, while nine (22.5%) were male.

Material

The subject of the learning environment was Weiner's (1985) attribution theory and its application in education. This theory is able to explain learners' motivation on the basis of the kinds of causes to which they attribute their own success or failure. The theory defines several concepts for explaining the relationship between attribution and learning motivation, most importantly the concept of locus of control and the concept of stability. Locus of control distinguishes between the internal and external factors people regard as the causes of success or failure. Stability comprises the stable and variable factors people deem responsible for their success or failure. For the attribution of success and failure, different combinations of these features are functional or dysfunctional with respect to learning motivation. The students read a three-page description of this theory.

In the collaborative learning phase, three problem cases from practical contexts were used as a basis for online discussions. Each problem case was realistic, complex and allowed learners to construct different arguments based on the attribution theory. The case "Math" describes the attributions of a student with respect to his poor performance in mathematics. In the case "Class reunion" a math tutor talks about how he tries to help female students deal with success and failure in assignments. The case "Asia" describes differences in school performance between Asian and American/European students that

were explained by the attribution theory. This kind of task is popular in virtual learning scenarios where learners are asked to acquire knowledge on how to apply a specific theory to problems (see Nistor 2003).

Collaborative learning task

The group's task was to analyze the three problem cases in an 80-minute collaborative learning phase and to determine a joint solution for each problem case. The three students in each group were separately placed in one of three laboratory rooms. An asynchronous, text-based discussion board was used for collaboration.

This discussion board allowed for the exchange of text messages that resembled emails. Learners could either start a new topic by posting a new message or reply to messages that had been posted previously. Each message consisted of a subject line, author information, date, time, and the message body. While the learning environment set author, date, and time automatically, the learners had to enter the subject line and the body of the message. Each of the three cases was discussed on a separate discussion board and learners could switch between these boards at any time during the collaborative learning phase.

Implementation of the scripted construction of arguments

The discussion board allowed for the implementation of different types of computer-supported collaboration scripts.

1. The *control group* received no additional support in solving the three problem cases.
2. The *script for the construction of single arguments* was implemented within the interface of the discussion board by a set of input text boxes (see Fig. 1). In accordance with our simplified version of Toulmin's (1958) model, the script consisted of input text boxes for a claim, grounds and qualifications. Each text box of the interface was to be filled out by the learners to construct a completely explicit argument. Subsequently, the learners were asked to add the argument to the message body by clicking on a command button. By clicking the command button, the system compiled the contents of the three input text boxes into a pre-specified textual structure for the individual messages. The three input text boxes were then cleared. Hence, the learners could either construct the next single argument or contribute their message to the online discussion. However, learners were not limited to using the three input text boxes for constructing single arguments. Questions, comments or expressions of emotion could also be directly written into the main input text box, without using the script for the construction of single arguments.
3. The *script for the construction of argumentation sequences* aimed to facilitate specific argumentation sequences of argument-counterargument integration (following Leitão 2000). Hence, the subject of the posted message was automatically pre-set, depending on its position in the cascading discussion thread. The first message of a discussion thread was labeled "Argumentation." The answer to an argument was automatically labeled as "Counterargumentation" and a reply to a counterargument was labeled "Integration." The next message was again labeled counterargument, then integration and so on. In this way, there was a default path through the discussion according to Leitão's model (see Fig. 2). Each message could contain several single arguments. When necessary, learners were also able to change the subject of their message.

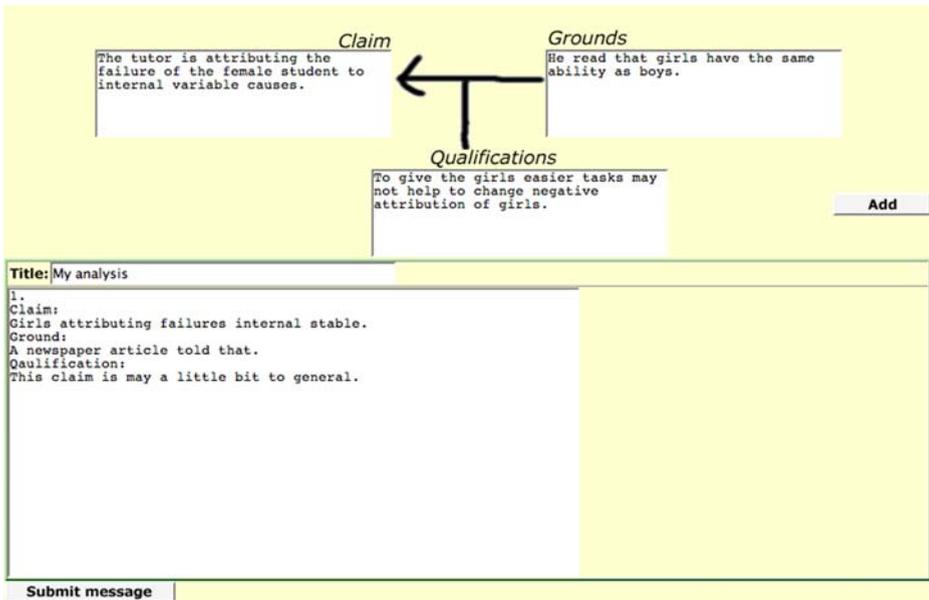
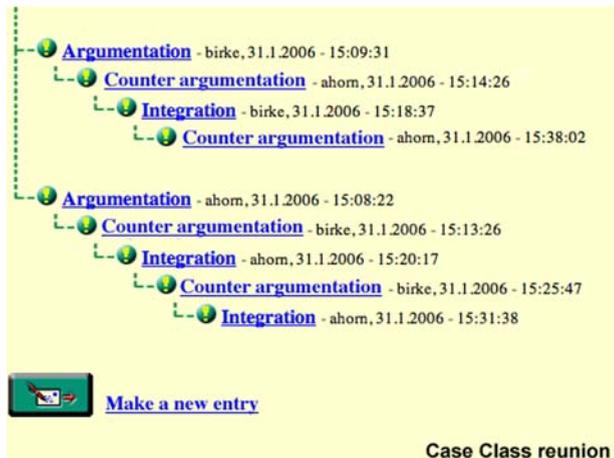


Fig. 1 The interface of the script for the construction of arguments. This extension was placed between the description of the cases and the regular user interface. It comprises input text fields for claims, grounds and qualifications. With a click on the add button, the argument was pasted to the input text field of the regular interface and the input text fields of the extension were cleared

- 4. In the *combined condition*, learners were supported with both scripts during collaboration. The interface contained the three input fields for argument construction (see Fig. 1). Each message, independent of its position in the sequences of argumentation, was supported by these additional fields. The subjects of the messages were pre-set automatically by the script for the construction of argumentation sequences. Hence, the script provoked a sequence of argument, counterargument, and integration (see Fig. 2).

Fig. 2 The goal structure of the script for the construction of argumentation sequences. The script is suggesting a sequence of argumentation, counter argumentation, and integration



Procedure

First, a pre-test on prior domain-specific knowledge and on prior knowledge on argumentation was administered along with a questionnaire on general variables (e.g., gender, age; 15 min.). Subsequently, the participants were given 15 min. to individually read the three-page description of the attribution theory. Learners were then introduced to the learning environment (20 min.). In the next phase, the learners collaborated for 80 min. in groups of three with the task of composing joint analyses of the cases. Finally, the students took individual post-tests on domain-specific knowledge related to the attribution theory and a test of knowledge on argumentation (about 20 min.). As a minimum, students were asked to fill out a questionnaire consisting of several control variables (15 min.). Time on task was held constant in all four conditions.

Data sources and instruments

Assessing the formal quality of online collaborative argumentation

The individuals' contributions to the online discussion were used as the data source for assessing the quality of collaborative argumentation. Discourse corpora was segmented and then analyzed with a coding scheme developed by Weinberger and Fischer (2006). First, trained coders segmented the discourse corpora into propositional units. The segmentation was based on propositional units that could be evaluated as true or false. For example, the sentence "The teacher is attributing external and variable" was segmented into "The teacher is attributing external" and "[The teacher is attributing] variable." The coders achieved an agreement of 83% during the training with respect to the segmentation of the discourse corpora. To reduce the amount of data, the online discussion of only one of the three cases was analyzed. About 5% of the discourse data (about 500 out of 10,000 segments) was used for training the six coders. The coders were unaware of participant characteristics but were aware of the experimental condition. The coders were aware of the treatment condition of the online discussion as they included the prompts of the script. For a detailed description of this coding scheme see Weinberger and Fischer (2006).

Process analysis

With regard to the *argumentation in online discussion*, the propositional units were then coded with respect to the quality of single arguments and with respect to the quality of argumentation sequences. (1) *The formal quality of single arguments* has been defined as share of segments that were coded as claims with grounds and/or qualifications. The coders distinguished between bare claims, claims with grounds, claims with qualifications, and claims with grounds and qualifications. *Bare claims* are neither supported by grounds, nor restricted by a qualifier, e.g., "The teacher is attributing variable." *Grounds* are reasons given in support of a claim. These can be data, i.e., information from the case description, or warrants, i.e., concepts and explanations from the attribution theory. In the context of this study, learners may support claims with case information or concepts from the respective attribution theory. For instance, the claim "The teacher is attributing to variable causes." is based on the data (given as case information) and the claim "The teacher ascribes Michael's failure to laziness" is based on the warrant "Seeing the cause in laziness is a variable attribution." *Qualifications* limit the validity of the claim by introducing alternative

viewpoints and explanations, e.g., by adding “to the extent that the student tells the truth in the school counseling session” to the argument.

Assessing (2) *the formal quality of argumentation sequences* was performed with the help of sequence analyses. Therefore, the first step was to identify counter arguments and integrations in the contributions to the online discussion. The coders distinguished between arguments, counterarguments and integrations. Arguments initiate a new argumentation sequence and consist of claims that have not been discussed before. Counterarguments are arguments that attack other arguments and are identified on the basis of differences between claims: If a claim opposes or attacks a preceding claim, the later claim is coded as a counterargument. For instance, the argument “The teacher is supporting Michael by ascribing his failure to laziness” can be countered by the argument “But Michael obviously does not believe what his teacher is telling him and is blaming his lack of talent instead.” Integrations resolve the conflict or tension between arguments and counterarguments on a higher level. This means that arguments and counterarguments are integrated to define a perspective in which the main claims can be sustained in a logically consistent and coherent way. These integrations stand in contrast to more additive methods of synthesis or consensus building, where the opposing arguments are simply collated such that the resulting statement contains contradictions and inconsistencies. Note that learners are not limited to writing counterarguments and integrations that address the arguments of their learning partners, but may also construct counterarguments or integrations for their own arguments. For example, learner A could state the argument “Michael will not improve in math!” Learner B contradicts this argument with the counterargument “But the teacher provides a positive attributional pattern to Michael!” As reply, learner A may state the integration “Michael’s current attributional pattern will lead to a failure in the next math exam, but if he adopts the teacher’s pattern, he may succeed in math.” Six trained coders evaluated the quality of single arguments and argumentation sequences independently with a sufficiently high reliability (Cohen’s between $\kappa = .50$ and $\kappa = .69$, median $\kappa = .61$). To analyze the sequences on the level of the messages exchanged, we aggregated the assigned codes on the message level. Each message was classified as argument, counterargument, or integration based on the codes assigned to the propositional segments within the message. A message was classified as argument when no propositional segment within the message was coded as counterargument or integration. A counterargument message contained at least one propositional segment that was coded as counterargument, but no segment that was coded as integration. Respectively, messages with at least one propositional segment classified as integration were coded as integration. The next step was performed using the software tool MEPA, developed by Erkens (1998). With this tool, we computed the probability of transitions between the aforementioned message types (argument, counterargument, or integration) for each group of three. With respect to the quality of argumentation sequences, we considered the transitions from argument to counterargument, counterargument to integration, and integration to counterargument. In addition, we analyzed the percentage of non-argumentative messages, arguments, counterarguments, and integrations.

Knowledge tests

Both the *pre-test* and the *post-test* of domain-specific knowledge involved the analysis of an authentic problem case using Weiner’s attribution theory. The *post-test for knowledge on argumentation* also included the transfer of knowledge on argumentation to another domain.

Knowledge pre-tests

Prior knowledge on argumentation was tested using the case problem “Choosing a Major,” which had to be analyzed by the participants individually. The case “Choosing a Major” is about the influence of parents’ attributional pattern on their daughter’s choice of subject for her university education. Participants were asked to write their analysis on a sheet of paper. To determine prior knowledge on argumentation, the analyses were segmented into propositional units and coded according to the procedure suggested by Weinberger and Fischer (2006) and described above (“Assessing the quality of online collaborative argumentation” and “Process analysis” sections). After segmentation, the segments were coded with respect to the two aspects of knowledge on argumentation: Knowledge on the construction of single arguments and knowledge on argumentation sequences. As an indicator of (1) *prior knowledge on the construction of single arguments*, we used the number of arguments with grounds and/or qualifications, e.g., “The father is attributing internal, because he says that his daughter is gifted.” As an indicator of (2) *prior knowledge on the construction of argumentation sequences*, we used the number of counterarguments and integrations. For example, the learner contradicts his own claim “The father has a positive attributional pattern” with “Despite the positive influence of her father, the student has a negative attributional pattern.” Segmented individual case analyses (already used to assess the prior knowledge on argumentation, see above) were the basis for assessing (3) *domain-specific prior knowledge*. Two experts identified 18 different appropriate propositional units, i.e., instances of applying concepts from the attribution theory in the context of analyzing the case. These propositional units were then used as a reference for evaluating the learners’ case analyses. The number of appropriate propositional units externalized in the case analysis was used as an indicator of prior domain-specific knowledge. For instance, within the utterance “The parents attribute their failures in math to stable causes,” case information (the parents failed in math and blamed their abilities) was related to theoretical concepts (attribution to stable causes). The median of Cohen’s Kappa values for categorizing domain-specific prior knowledge was sufficiently high with $\kappa = .72$. Due to the fact that the propositional units were scaled dichotomously (externalized vs. not externalized), Guttman’s split-half reliability test was performed. The reliability was sufficiently good ($r = .62$).

Knowledge post-tests

In the test of (1) *knowledge on the construction of single arguments*, participants were asked to recall components of single arguments (claim, ground, and qualification). In addition, participants were asked to formulate single arguments about “smoking.” They were asked to construct completely explicit arguments that contained all the components of the simplified Toulmin model. The participants were to provide a claim (e.g., “Smoking causes cancer”), a ground (“People who smoke are more likely to get cancer”), and a qualification (“Not all people who smoke get cancer”). Students were given one point for each of the three types of components that were appropriately contained in their responses. Hence, the test scores could range from 0 to 6 points. In the test of (2) *knowledge on the construction of argumentation sequences*, participants were asked to recall components of argumentation sequences (argument, counterargument, and integration) in a free format paper-and-pencil test. One point was assigned for each of the three components. Furthermore, participants were asked to formulate an argumentation sequence about “smoking.” The arguments that learners constructed were analyzed with respect to the

components of an argumentation sequence (argument, counterargument, and integration). Learners were to provide an argumentation (e.g., “Smoking causes health problems”), contradict the argumentation with a counterargument (e.g., “Not everybody who smokes develops health problems”), and construct an integration (e.g., “Not everybody who smokes will develop health problems, but the likelihood of developing health problems will be much greater for people who smoke than for those who do not”). One point was given for each of the three correctly applied components. Hence, the range of this test was also 0 to 6 points. Two trained coders evaluated the tests independently with sufficiently high reliability (Cohen’s $\kappa = .83$). The Guttman split-half coefficient indicated sufficiently high levels of internal test consistency for knowledge on the construction of single arguments (Guttman split-half $r = .89$) and for the knowledge on the construction of argumentation sequences (Guttman split-half $r = .79$).

In order to measure (3) *the acquisition of domain-specific knowledge*, participants had to individually analyse another case, namely the “Text Analysis” case, with the help of the attribution theory. In the case “Text Analysis,” a student talks about the reasons for failing a text analysis exam. Experts identified 32 different propositional units related to applications of theoretical concepts from the theory. These propositional units were coded using the coding rules described for the domain-specific *prior* knowledge test above. The reliability of this test was sufficiently high (Guttman split-half $r = .60$). Due to the fact that we used different cases in pre-test and post-test, we used the share of these propositional units instead of the number of explicit applications. Furthermore, we used the similarity between the students’ analyses and the expert analyses as an indicator of domain-specific knowledge. The increase in similarity between the learner’s solution and the expert solution from pre-test to post-test was used as an indicator of the acquisition of domain-specific knowledge.

Treatment check

As a treatment check, we computed (1) the proportion of learners’ follow-up activities intended by the script (e.g., the input field “grounds” of the script for the construction of single arguments contained a warrant, backing and/or data), (2) the proportion of prompts that were followed by a non-intended reaction (e.g., the input field “grounds” of the script for the construction of single arguments contained a qualification), and (3) the share of unreciprocated prompts, i.e., the input field “grounds” of the script for the construction of single arguments remained empty. Overall, 655 prompts were presented to the learners in the three experimental conditions with script support. The share of reactions on these prompts as foreseen by the different scripts was $M = 68.13\%$ ($SD = 26.87$). The share of unforeseen reactions was $M = 7.58\%$ ($SD = 14.36$) and the share of prompts that caused no reaction was $M = 24.29\%$ ($SD = 22.42$). Hence, the learners reacted as intended to the majority of prompts to an extent comparable to other computer-supported collaboration scripts (see Weinberger et al. 2007).

Results

Randomization check and control variables

To ensure that randomization was successful, we examined the prior knowledge on argumentation and domain-specific knowledge with regard to differences between (1) experimental groups and (2) unselected vs. selected individuals within the groups of three.

These tests were conducted on a 20% level to avoid type II errors. No significant differences were found between experimental groups (prior knowledge on the construction of single arguments: $F(1, 36) < 1.00, p > .2$; prior knowledge on the construction of argumentation sequences: $F(1, 36) < 1.00, p > .2$; domain-specific prior knowledge: $F(1, 36) < 1.00, p > .2$; also none of the pair wise comparisons were significant) or between selected vs. unselected individuals (prior knowledge on the construction of single arguments: $t(118) = -.24, p > .2$; prior knowledge on the construction of argumentation sequences: $t(118) = -1.16, p > .2$; domain-specific prior knowledge: $t(60.56) = -1.18, p > .2$).

Furthermore, we monitored whether the different experimental conditions had an effect on the learner's participation in argumentation, i.e., the amount of words and messages posted to the discussion board. The script for the construction of argumentation sequences affected the amount of words contributed to the discourse, but this effect was only of small size ($F(1, 116) = 4.43, p < .05, \eta^2 = .04$). In online discussions supported with the script for the construction of argumentation sequences, learners contributed about 100 words per learner more to the discourse in comparison with the control group. The script for the construction of single arguments had no effect on the amount of words contributed to the discourse ($F(1, 116) = 1.35, n.s.$). No interaction effect was found ($F(1, 116) = .31, n.s.$). With respect to the amount of posted messages, the script for the construction of single arguments had a large negative effect ($F(1, 116) = 26.52, p < .05, \eta^2 = .19$). While learners without this script posted about seven messages, learners with this script only contributed about four messages to the discourse. The script for the construction of argumentation sequences had no effect on the amount of posted messages ($F(1, 116) = .07, n.s.$) and no interaction effect of the scripts could be observed ($F(1, 116) = .07, n.s.$).

Hence, the script for the construction of single arguments reduced the number of posted messages, but the number of words per message increased, i.e., learners posted longer messages and learner's overall participation was not affected. The script for the construction of argumentation sequences increased the activity of the learners as indicated by the amount of words posted during the discussion. However, the effect was small. Nevertheless, this effect has to be taken into account when interpreting the other effects of the script on the outcomes of argumentative knowledge construction.

RQ1: Effects of the scripts on collaborative argumentation in online discussions

We examined the effects of the two computer-supported collaboration scripts and their combination on the processes of collaborative argumentation. In particular, we analyzed the effects of the scripts on the formal quality of single arguments and the formal quality of argumentation sequences.

The script for the construction of single arguments substantially and strongly affects the formal quality of single arguments (see Table 1; $F(3, 33) = 12.83, p < .05, \eta^2 = .61$). The script for the construction of argumentation sequences had no effect on the formal quality of single arguments ($F(3, 33) = 1.88, n.s.$). No interaction effect was found ($F(3, 33) < 1.00, n.s.$). Figure 3 illustrates the effect of the script for the construction of single arguments. The single argument on the left came from an online discussion without support of the script, while the argument on the right was taken from a discussion supported by the script for the construction of single arguments. Both single arguments are part of a longer message and include the same claim, which is mainly "The parents are attributing to talent." This is a correct inference between Weiner's (1985) attribution theory and the statement of the parents in the problem case "math," which was one of the aforementioned problem cases analyzed by the learners. The example for an argument constructed without the support of the script for the

Table 1 Share of arguments in discourse by degree of formal structure of argumentation sequences and experimental condition: means (*M*) and standard deviations (*SD*)

		Control condition	Script for the construction of single arguments	Script for the construction of argumentation sequences	Combined condition
Bare claims	<i>M</i> (%)	70.75	41.60	73.09	53.75
	<i>SD</i>	16.17	12.37	22.98	14.54
Supported claims	<i>M</i> (%)	11.39	35.17	15.35	29.05
	<i>SD</i>	10.04	13.54	19.68	11.57
Qualified claims	<i>M</i> (%)	2.06	11.72	2.71	13.62
	<i>SD</i>	3.34	9.47	5.08	8.53
Supported and qualified claims	<i>M</i> (%)	2.03	4.51	0.56	0.48
	<i>SD</i>	4.71	5.86	1.76	1.51

construction of single arguments shows a bare claim without any supporting grounds or limiting qualifications. In contrast, the argument of the scripted condition contains the three components claim, grounds, and qualifications. Beneath the claim, it contains a ground as well as a qualification. As the ground, the learner provides the father’s statement, which counts as data supporting the claim. The claim is qualified by an assumption that is based on the naïve prior knowledge of the learner.

In online discussions supported with the script for the construction of single arguments, nearly 25% less bare claims were formulated in comparison to discussions that were not supported by this script ($F(1, 36) = 20.39, p < .05, \eta^2 = .36$). Instead, the share of supported

**Example of Formal Quality of Single Arguments
With and Without the Script for the Construction of Single Arguments**

<i>Script for the construction of single arguments</i>	
Without script	With script
<p>[The] mother ascribes the bad performance [in math] to absence of talent.</p>	<p>Claim: [The] self-attribution of the Peters family is that they have no talent for math</p> <p>Ground(s): Statement [of the] father [that] the Peters were always [known as] the scare of math teachers</p> <p>Qualification(s): It is not realistic that in a family not at least one member has a basic talent [for math]</p>

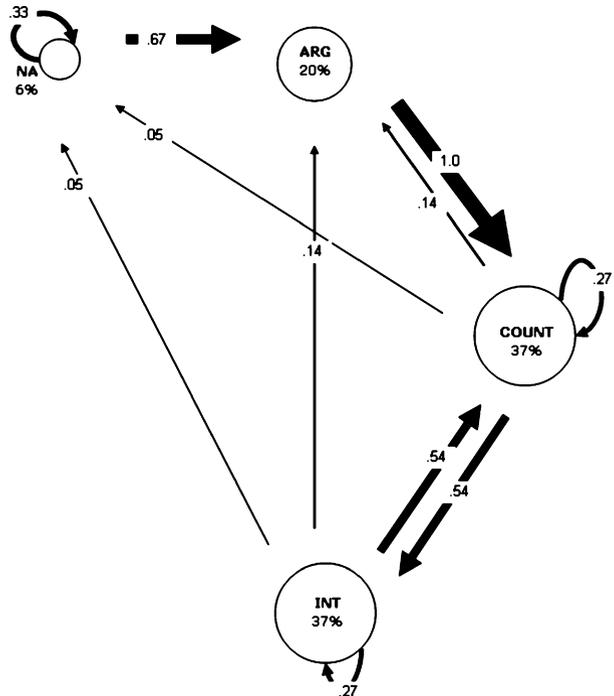
Fig. 3 Example of formal quality of single arguments with and without the script for the construction of single arguments. Two single arguments taken from two different online discussions of the sample examined. Both arguments have a similar claim (“The parents are attributing on talent.”), but the claim is supported by data from the problem case and qualified by (naive) prior knowledge only in the condition with script for the construction of single arguments

claims ($F(1, 36) = 17.44, p < .05, \eta^2 = .33$) as well as the share of qualified claims ($F(1, 36) = 21.23, p < .05, \eta^2 = .37$) were positively affected by the script for the construction of single arguments (see Table 1). This supports our expectation that the script for the construction of single arguments specifically facilitates the formal quality of single arguments.

With respect to the effects of the argumentative collaboration scripts on the formal quality of the argumentation sequences, we analyzed both the proportion of the specific message types (argument, counterargument, integration, and non-argumentative message) as well as the probabilities of the transitions from argument to counterargument, counterargument to integration, and integration to counterargument. To assess the effects of the scripts, we constructed an ideal model of argumentation sequences. If a script affects the probability of a transition positively, this could be interpreted as higher formal quality in the case of the transition from argument to counterargument. However, a higher probability of the transition from argument to a non-argumentative contribution is viewed as having a negative effect on the formal quality of argumentation sequences. Figure 4 is an illustration of a constructed ideal argumentation sequence in an explanatory inquiry dialogue. The basis of this model was a constructed dialogue between three learners in which all learners follow the script for the construction of argumentation sequences. Furthermore, we considered the fact that the learners may need to coordinate with respect to their learning task. The resulting proportions of contribution types and probabilities of transitions between them should therefore be taken as indications, not as absolute values.

The circles in the figure represent the four different types of contributions to an argumentative discourse with respect to the argumentation sequence: (1) argument (ARG), (2) counterargument (COUNT), (3) integration (INT), and (4) non-argumentative contributions (NA), as, for example, the coordination of the collaboration. The size of the circles is directly related to the share of the specific type of contribution in the overall

Fig. 4 Illustration of the shares of the four different message types (size of circles) and the probabilities of transitions between them (thickness of arrows) of an ideal structure of an argumentation sequence

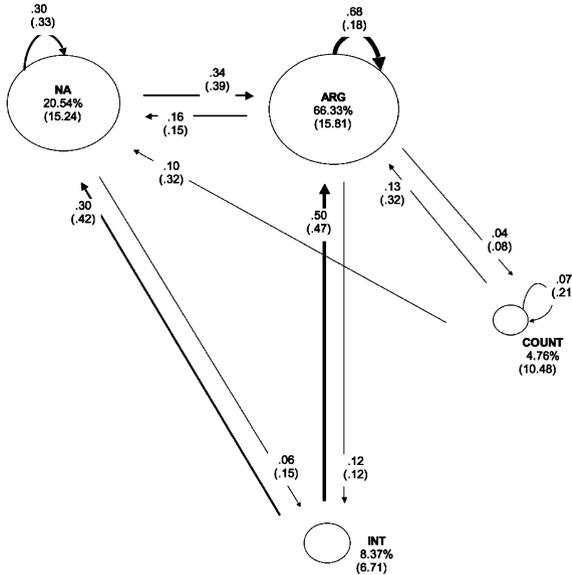


discussion. The arrows between the types of contributions represent the probability that a specific type of contribution follows the type of contribution where the arrow starts. The thickness of arrows is proportional to the probability of this transition. In the ideal argumentative discourse with respect to the structure of argumentation sequences, the argumentation either starts with a first argument (ARG) or a message for the coordination of collaboration (e.g., “What next?”; NA). In our ideal model, non-argumentative contribution (NA) would have a share of about 6% of all contributions. If the first message is related to coordination, there is a probability of $p = .33$ that the response is also a message for coordination (e.g., “Let’s start at the first of the three problem cases.”) or a probability of $p = .67$ that the response is an argument (ARG). This means that, for example, in a discourse that consists of 100 messages, six of these messages would be non-argumentative. Two of them (about 33%) would be answered with non-argumentative messages and four (about 67%) would be followed by an argument. Furthermore, in our ideal model, each ($p = 1.0$) argument (ARG) would be answered by a counterargument (COUNT). There is a probability of $p = .27$ that a counterargument would be followed by another counterargument (e.g., totally disagreeing with the claim of the previous counterargument) or by an integration ($p = .54$; e.g., integrating parts of the counterargument into the new claim). There is a probability of $p = .54$ that an integration would be answered with a counterargument and a probability of $p = .27$ that it would be answered by another integration. The share of counterarguments and integrations should be about 37% each. This would correspond to a discourse in which each initial argument would, on average, be followed by about four counterarguments and four integrations (in a sequence in order of the before mentioned probabilities). If the argumentation sequence on a particular initial argument should end, there is a probability of $p = .05$ that a non-argumentative message (NA) may follow a counterargument (COUNT) or integration (INT). There is a probability of $p = .14$ that a new initial argument (ARG) will follow a counterargument (COUNT) or an integration (INT).

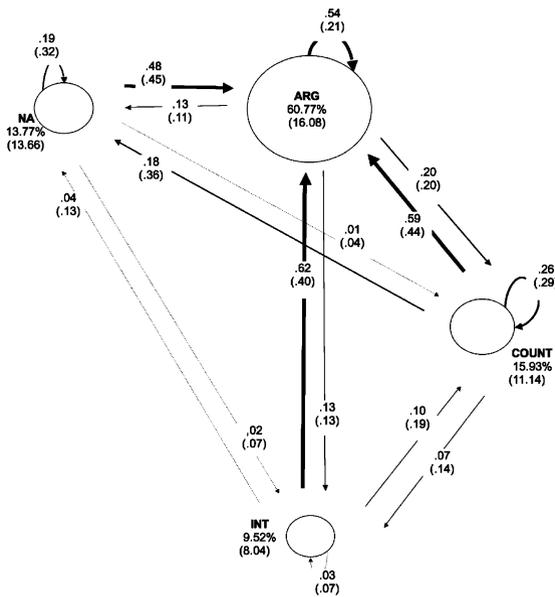
We used MANOVA to analyze the proportions of the specific message types and the transitions between them. The script for the construction of argumentation sequences strongly affected the *formal quality of argumentation sequences* with respect to the message types posted ($F(4, 33) = 3.95, p < .05, \eta^2 = .32$; see Figs. 5 and 6). The script for the construction of single arguments had no effect ($F(4, 33) < 1.00, n.s.$) and no interaction effect ($F(4, 33) < 1.00, n.s.$) was found. Analysis of script effects on the level of the different message types showed that the script for the construction of argumentation sequences affected the proportion of non-argumentative messages ($F(1, 36) = 4.48, p < .05, \eta^2 = .11$) and the proportion of counterarguments ($F(1, 36) = 14.88, p < .05, \eta^2 = .29$). The learners supported by this script produced about 10% non-argumentative messages, while learners without this support produced about 20% of contributions of this type. With respect to counterarguments, the share was about three times higher with script for the construction of argumentation sequences (about 18%) than without this script (about 6%).

In a MANOVA with respect to the probabilities of the transitions from argument to counterargument, counterargument to integration, and integration to counterargument, the script for the construction of argumentation sequences had a strong effect ($F(3, 34) = 3.85, p < .05, \eta^2 = .32$; see Figs. 5 and 6). The script for the construction of single arguments had no effect ($F(3, 34) < 1.00, n.s.$) on these transitions and no interaction effect ($F(3, 34) = 1.10, n.s.$) was found. The univariate ANOVA of the transitions from argument to counterargument, counterargument to integration, and integration to counterargument showed that only the script for the construction of argumentation sequences positively affected the transition from argument to counterargument ($F(1, 36) = 5.39, p < .05, \eta^2 = .13$). Learners supported with the

Effects of the scripts on the formal quality of argumentation sequences (part I)



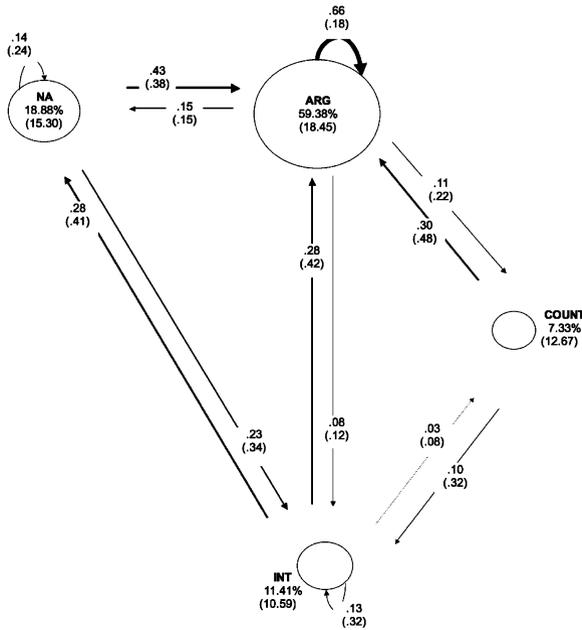
a Control condition



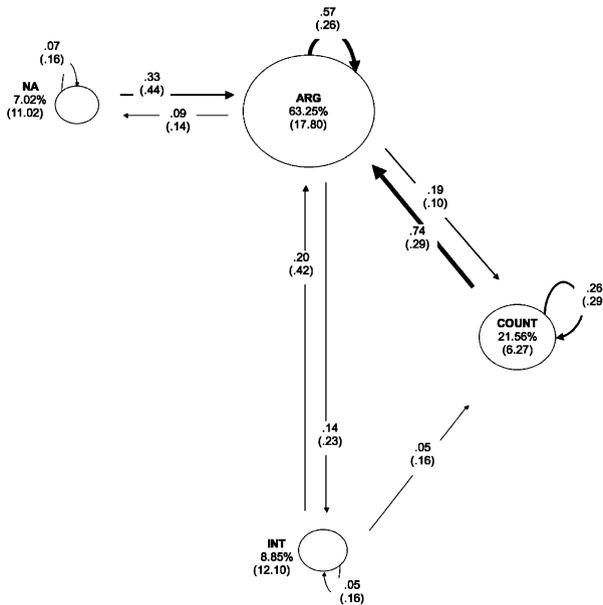
b Script for the construction of argumentation sequences

Fig. 5 Illustration of the shares of the four different message types (*size of circles*) and the probabilities of transitions between them (*thickness of arrows*) by experimental groups (part I): (a) control group and (b) script for the construction of argumentation sequences

Effects of the scripts on the formal quality of argumentation sequences (part II)



c Script for the construction of single arguments



d Combined condition

Fig. 6 Illustration of the shares of the four different message types (*size of circles*) and the probabilities of transitions between them (*thickness of arrows*) by experimental groups (part II): (c) script for the construction of single arguments and (d) combined condition

script for the construction on argumentation sequences were about two times more likely to answer an argument with a counterargument than learners without the support of this script.

In summary, when comparing these outcomes with the ideal argumentation sequences (see Fig. 4), we find that the script for the construction of argumentation sequences specifically facilitates the proportion of counterarguments and the transition from argument to counterargument in the direction of the ideal model. Overall, the formal quality of argumentation sequences increased.

RQ2: Effects of scripts on outcomes of argumentative knowledge construction

Next, we examined the effects of the two scripts on the *outcomes* of argumentative knowledge construction, i.e., acquisition of knowledge on the construction of single arguments, the acquisition of knowledge on construction of argumentation sequences, as well as the acquisition of domain-specific knowledge.

In the individual post-test, learners who had been supported by the script for the construction of single arguments achieved about double the score in the knowledge test on the construction of single arguments than learners without the support of this script (see Table 2). The script for the construction of single arguments had a significant and large effect on the *acquisition of knowledge on the construction of single arguments* ($F(1, 36) = 10.00, p < .05, \eta^2 = .22$), i.e., participants who had learned in these conditions during the discussion performed better in the individual post-test. The script for construction of argumentation sequences ($F(1, 36) < 1.00, n.s.$) had no effect and no interaction effect could be found ($F(1, 36) < 1.00, n.s.$).

The findings on learning outcomes support our expectation that the script for the construction of single arguments specifically facilitates the acquisition of knowledge on the construction of single arguments.

Learners supported with the script for the construction of argumentation sequences had substantially higher scores in the knowledge test with respect to the construction of argumentation sequences test than learners in the other conditions. This effect was significant and large ($F(1, 36) = 17.02, p < .05, \eta^2 = .32$). The script for the construction of single arguments ($F(1, 36) < 1.00, n.s.$) had no effect and no interaction effect of the scripts could be found ($F(1, 36) < 1.00, n.s.$).

Table 2 Outcomes of argumentative knowledge construction in the experimental conditions: means (*M*) and standard deviations (*SD*)

	Control condition	Script for the construction of single arguments	Script for the construction of argumentation sequences	Combined condition
Acquisition of knowledge on construction of single arguments	<i>M</i> 2.10 <i>SD</i> 1.52	4.50 2.27	3.20 1.69	4.50 1.84
Acquisition of knowledge on construction of argumentation sequences	<i>M</i> 1.60 <i>SD</i> 2.01	2.00 4.50	4.30 2.11	4.50 2.17

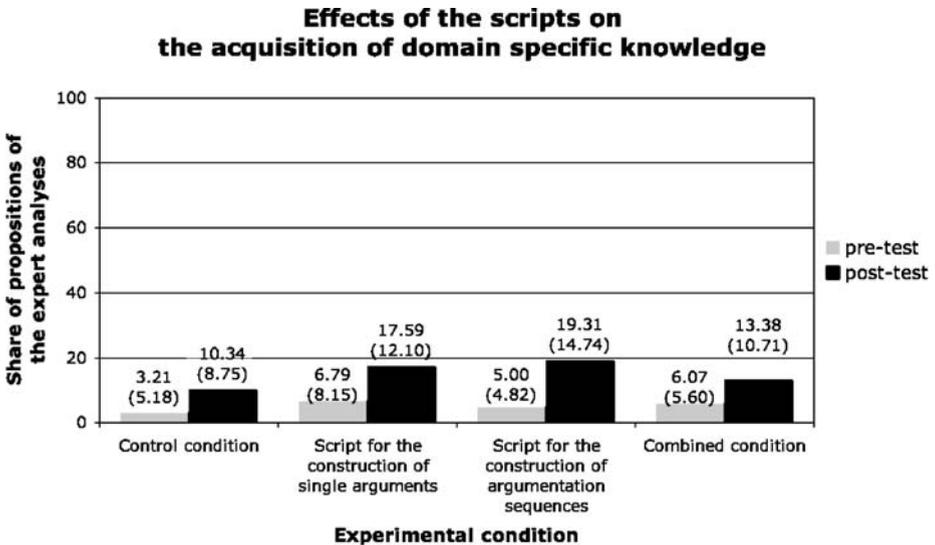


Fig. 7 Effect of the scripts on the acquisition of domain-specific knowledge: similarity of pre-test and post-test to the expert analysis

These findings support our expectation that the script for the construction of argumentation sequences specifically facilitates the acquisition of knowledge on the construction of argumentation sequences.

With respect to the acquisition of domain-specific knowledge, we examined the pre-test to post-test gain with the help of the general linear model for repeated measures. The analyses show that all participants significantly improved from pre-test to post-test ($F(1, 36) = 22.90$, $p < .05$, $\eta^2 = .39$). The similarity of analyses of learners with the expert analysis was $M = 5.27\%$ ($SD = 6.00$) for the pre-test case. With respect to the post-test, the similarity was about ten percentage points higher ($M = 15.27\%$; $SD = 11.86$). The average gain between pre-test and post-test was 7.13 percentage points in the control condition, 10.80 percentage points for learners supported only by the script for the construction of single arguments, and 14.31 percentage points for learners supported only by the script for the construction of argumentation sequences. The average gain between pre-test and post-test for learners in the combined condition was similar to the control condition (7.37 percentage points). However, neither the script for the construction of single arguments ($F(1, 36) < 1.00$, *n.s.*), nor the script for the construction of argumentation sequences affected ($F(1, 36) < 1.00$, *n.s.*) domain-specific knowledge acquisition (see Fig. 7). No interaction effect could be found ($F(1, 36) = 1.64$, *n.s.*). Hence, the findings of this study did not support our expectations with regard to the effects of the scripts on domain-specific knowledge acquisition.

Discussion

We expected that the script for the construction of single arguments would facilitate the quality of single arguments and that the script for the construction of argumentation sequences would facilitate the quality of argumentation sequences for learners during online

discussions (RQ1). Both scripts effectively facilitated the specific discourse processes of argumentative knowledge construction as intended. The proportion of single arguments with a higher formal quality increased for learners who were supported with the script for the construction of single arguments. The script for the construction of argumentation sequences facilitated the formal quality of argumentation sequences as indicated by the increased proportion of counterarguments and transitions from arguments to counterarguments. The scripts improved online discussions in the positive direction of ideal models of argumentation, but the higher formal quality of argumentation was still far from the ideal. For example, even with support of the scripts, every second argument was a bare claim; the share of counterarguments and integrations did not even approach the level of the ideal model, and the important transitions from counterargument to integration and vice versa were still unlikely during the 80-min collaborative learning phase. Nevertheless, the argumentative collaboration scripts seem to be appropriate for facilitating argumentative knowledge construction. One might argue that the process-related findings are merely a kind of (expensive) treatment check. We argue, however, that the specific facilitation of processes in online discussions through scripts is not self-evident. Scripts are not strict rules that must be followed in a specific way. Instead, learners can decide to ignore the script or use it in a different way. This view is further emphasized by our finding that about 40% of the prompts in this study were ignored or used in a manner different than intended.

With respect to the second research question, mixed results were found. As expected, both scripts successfully facilitated the acquisition of knowledge on argumentation. Learners specifically acquired more knowledge on single arguments when supported by the script for the construction of single arguments and more knowledge on argumentation sequences when supported with the script for the construction of argumentation sequences as compared to learners without script support. The scripts had additive effects with regard to knowledge on argumentation, i.e., they did not interact negatively, and can thus be used in combination to foster the two facets of knowledge on argumentation. However, the scripts did not affect domain-specific knowledge acquisition. There may be a trade-off between the acquisition of knowledge on argumentation and domain-specific knowledge acquisition. The learners' tendency to focus efforts more on the knowledge on argumentation might be due to the time constraints set by this study. The learners may have elaborated individual concepts more extensively, but also expended effort constructing formally complete single argument and argumentation sequences. This may explain why learners in the scripted experimental groups acquired as much domain-specific knowledge as the unscripted learners, but additional knowledge on argumentation. It is of note that the slightly higher participation of learners who were supported by the script for the construction of argumentation sequences did not lead to higher domain-specific knowledge acquisition. Hence, this study was inconclusive with respect to the facilitation of domain-specific knowledge acquisition through argumentative scripts. There is empirical evidence that the scripts would also affect domain knowledge acquisition if longer time frames were considered (Kuhn and Goh 2005). Another explanation could be that external argumentative collaboration scripts have less impact on deep cognitive elaboration than internal scripts. The findings of Kollar et al. (2005) indicate that learners may have to internalize some argumentative knowledge before this knowledge can effectively facilitate the acquisition of domain-specific knowledge. But the insufficient compliance to the script may also lead to a relatively poor formal quality of the online discussion. A higher compliance with the script (in conjunction with a higher formal quality of argumentation) may lead to higher domain-specific knowledge acquisition, even in a shorter time frame.

However, further restrictions on the learners have to be implemented carefully, especially in light of “over-scripting” concerns (Dillenbourg 2002) and the expected decrease of motivation.

Some limitations of the study and its findings must also be considered. First, the intervention of this study is of a short duration. The extent to which the findings would apply to more long-term interventions is unclear. One might expect that after some initial coordination losses, unstructured groups might improve their self-regulated collaboration once they have interacted several times. Although this seems to be a broadly shared notion in CSCL research, review articles on collaborative learning are not encouraging with respect to unstructured collaboration—neither for short term nor for long term collaborations (Cohen 1994; Webb and Palincsar 1996). One might further expect a change in the way that instructional support, which does not alter over a specific period of time, affects collaboration once longer durations are considered. This might be especially true for collaboration scripts. Since the static structure of the collaboration script is being internalized by the learners, the external script may become a disturbing factor (Kollar et al. 2005), like training wheels hinder further learning when they are not removed in time (see Carroll and Carrithers 1984). Thus, we argue that longer durations of script interventions do not call for less structured instructional approaches, but for more flexible and dynamic support for learners with increasing self-regulation competencies (Donmez et al. 2005; Pea 2004).

Moreover, more long-term interventions highlight another “hot issue” in CSCL research that has not been addressed in this study: the lack of scientific knowledge on strategies for integrating computer-supported small group learning into the overall classroom or seminar lesson. This issue is related to the question about the role the teacher can play in a computer-supported classroom or seminar and how scaffolding on different levels can be “orchestrated” (Fischer and Dillenbourg 2006, April) to exploit potential synergies between them (Tabak 2004).

A further limitation of this study is related to the aspect of the intervention’s duration. One might object that the measures defined as indicators of individual knowledge acquisition are also rather short-term measures. There is research indicating that better elaboration of the learning material during the learning phase might have even more pronounced benefits when longer retention intervals are considered (Dochy et al. 2003; Howe 2005). The present study only involves immediate measures of knowledge acquisition. One might speculate that learners who collaborated with the support of the scripts retained better domain knowledge as compared to unstructured groups. This may have been due to their substantially better argumentation (and therefore probably improved cognitive elaboration) during the collaboration phase.

What conclusions can be drawn for the application of collaboration scripts? The learning environments realized in this study show some effective and less effective ways for improving lessons and lectures by using the Internet for more short-term case-based discussions. Online discussions have a large potential for improving the quality of argumentation in the context of knowledge construction and learning. In line with other studies (Kuhn 1991; Kuhn et al. 1997), the analysis of argumentation in unstructured online discussions showed that learners rarely base their claims on grounds and rarely construct counterarguments. Computer-supported scripts provide a way to improve argumentation processes without extensive prior training. In that respect, the findings of the present study are in line with a series of other studies, which indicate that scripts can facilitate specific processes and outcomes in a highly targeted way (see King 2007; Rosenshine and Meister 1994, 1996; Weinberger et al. 2005). With these scripts, learners acquire additional domain-general skills without imparting the development of domain expertise.

Acknowledgement The study has been funded by the DFG. We would like to thank Christof Wecker for commenting on an earlier version of this article.

References

- Andriessen, J., Baker, M., & Suthers, D. (2003a). 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* (Vol. 1, pp. 1–25). Dordrecht: Kluwer.
- Andriessen, J., Erkens, G., van de Laak, C., Peters, N., & Coirier, P. (2003b). Argumentation as negotiation in electronic collaborative writing. In J. Andriessen, M. Baker, & D. Suthers (Eds.), *Arguing to learn: Confronting cognitions in computer-supported collaborative learning environments* (Vol. 1, pp. 1–25). Dordrecht: Kluwer.
- Astleitner, H. (2002). Teaching critical thinking online. *Journal of Instructional Psychology*, 29, 53–76.
- Baker, M. (2003). Computer-mediated argumentative interactions for the co-elaboration of scientific notions. In J. Andriessen, M. Baker, & D. Suthers (Eds.), *Arguing to learn: Confronting cognitions in computer-supported collaborative learning environments* (Vol. 1, pp. 1–25). Dordrecht: Kluwer.
- Baker, M., & Lund, K. (1997). Promoting reflective interactions in a CSCL environment. *Journal of Computer Assisted Learning*, 13(3), 175–193.
- Carmien, S., Kollar, I., Fischer, G., & Fischer, F. (2007). The interplay of internal and external scripts—a distributed cognition perspective. In F. Fischer, H. Mandl, J. Haake, & I. Kollar (Eds.), *Scripting computer-supported collaborative learning: Cognitive, computational, and educational perspectives* (pp. 303–326). New York: Springer.
- Carroll, J. M., & Carrithers, C. (1984). Blocking learner error states in a training-wheel system. *Human Factors*, 26(4), 377–389.
- Chi, M. T., Bassok, M., Lewis, M., Reimann, P., & Glaser, R. (1989). Self-explanations: How students study and use examples in learning to solve problems. *Cognitive Science*, 13(2), 145–182.
- Clark, D. B., Sampson, V., Weinberger, A., & Erkens, G. (2007). Analytic frameworks for assessing dialogic argumentation in online learning environments. *Educational Psychology Review*, 19(3), 343–374.
- Cohen, E. G. (1994). Restructuring the classroom: Conditions for productive small groups. *Review of Educational Research*, 64, 1–35.
- Dillenbourg, P. (2002). Over-scripting CSCL. In P. A. Kirschner (Ed.), *Three worlds of CSCL: Can we support CSCL* (pp. 61–91). Heerlen: Open University of the Netherlands.
- Dillenbourg, P. (2004). “Split Where Interaction Should Happen,” a model for designing CSCL scripts. In P. Gerjets, P. A. Kirschner, J. Elen, & R. Joiner (Eds.), *Instructional design for effective and enjoyable computer-supported learning* (pp. i–ii, CD-ROM). Tübinge: Knowledge Media Research Center.
- Dochy, F. J. R. C., Segers, M., & van den Bossche, P. (2003). Effects of problem-based learning: A meta-analysis. *Learning and Instruction*, 13(5), 533–568.
- Doise, W., & Mugny, G. (1984). *The social development of the intellect*. Oxford: Pergamon.
- Donmez, P., Rosé, C. P., Stegmann, K., Weinberger, A., & Fischer, F. (2005). Supporting CSCL with automatic corpus analysis technology. In T. Koschmann, D. Suthers, & T.-W. Chan (Eds.), *Proceedings of the International Conference on Computer Supported Collaborative Learning—CSCL 2005* (pp. 125–134). Taipei, TW: Erlbaum.
- Erkens, G. (1998). Multiple episode protocol analysis (MEPA 3.0): Department of Educational Sciences. The Netherlands: Utrecht University.
- Fischer, F., & Dillenbourg, P. (2006, April). *Challenges of orchestrating computer-supported collaborative learning*. Paper presented at the 87th Annual Meeting of the American Education Research Association (AERA).
- Hanley, G. L. (1995). Teaching critical thinking: Focusing on metacognitive skills and problem solving. *Teaching of Psychology*, 22(1), 68–72.
- Howe, C. (2005). *Group work and conceptual growth in science: Taking account of post-group effects*. Paper presented at the 11th Biennial Conference of EARLI.
- Hron, A., Hesse, F.-W., Reinhard, P., & Picard, E. (1997). Strukturierte Kooperation beim computer-unterstützten kollaborativen Lernen. *Unterrichtswissenschaft*, 25(1), 56–69.
- Jermann, P., & Dillenbourg, P. (2003). Elaborating new arguments through a CSCL script. In P. Dillenbourg (Ed.), *Learning to argue* (Vol. 1, pp. 205–226). Dordrecht, The Netherlands: Kluwer.
- Johnson, D. W., & Johnson, R. T. (1994). Constructive conflict in schools. *Journal of Social Issues*, 50(1), 117–137.
- Keefe, M. W., Zeitz, C. M., & Resnick, L. B. (2000). Judging the quality of peer-led student dialogues. *Cognition and Instruction*, 18(1), 53–81.

- King, A. (1997). Ask to think—Tel why: A model of transactive peer tutoring for scaffolding higher level complex learning. *Educational Psychologist*, 32(4), 221–226.
- King, A. (2007). Scripting collaborative learning processes: A cognitive perspective. In F. Fischer, I. Kollar, H. Mandl, & J. M. Haake (Eds.), *Scripting computer-supported collaborative learning: Cognitive, computational, and educational perspectives* (pp. 13–37). New York: Springer.
- Kintsch, W. (1991). The role of knowledge in discourse comprehension: A construction–integration model. In G. Denhière & J.-P. Rossi (Eds.), *Text and text processing* (pp. 107–153). Amsterdam: Elsevier.
- Kobbe, L., Weinberger, A., Dillenbourg, P., Harrer, A., Hämäläinen, R., & Fischer, F. (2007). Specifying computer-supported collaboration scripts. *International Journal of Computer-Supported Collaborative Learning*, 2(2–3), 211–224.
- Kollar, I., Fischer, F., & Hesse, F. W. (2006). Collaboration scripts—a conceptual analysis. *Educational Psychology Review*, 18(2), 159–185.
- Kollar, I., Fischer, F., & Slotta, J. D. (2005). Internal and external collaboration scripts in webbased science learning at schools. In T. Koschmann, D. Suthers, & T. W. Chan (Eds.), *Computer supported collaborative learning 2005: The next 10 Years* (pp. 331–340). Mahwah, NJ: Erlbaum.
- Koschmann, T. (2003). CSCL, argumentation and deweyan inquiry: Argumentation is learning. In J. Andriessen, M. Baker, & D. Suthers (Eds.), *Arguing to learn: Confronting cognitions in computer-supported collaborative learning environments* (Vol. 1, pp. 1–25). Dordrecht: Kluwer.
- Kuhn, D. (1991). *The skills of argument*. Cambridge: Cambridge University Press.
- Kuhn, D., & Goh, W. W. L. (2005). Arguing on the computer. In T. Koschmann, D. Suthers, & T. W. Chan (Eds.), *Computer supported collaborative learning 2005: The next 10 years* (pp. 125–134). Mahwah, NJ: Erlbaum.
- Kuhn, D., Shaw, V., & Felton, M. (1997). Effects of dyadic interaction on argumentive reasoning. *Cognition and Instruction*, 15(3), 287–315.
- Leitão, S. (2000). The potential of argument in knowledge building. *Human Development*, 43, 332–360.
- Marttunen, M. (1992). Commenting on written arguments as a part of argumentation skills: Comparison between students engaged in traditional vs on-line study. *Scandinavian Journal of Educational Research*, 36(4), 289–302.
- Marttunen, M. (1994). Assessing argumentation skills among Finnish university students. *Learning and Instruction*, 4(2), 175–191.
- Means, M. L., & Voss, J. F. (1996). Who reasons well? Two studies of informal reasoning among children of different grade, ability and knowledge levels. *Cognition and Instruction*, 14, 139–178.
- Nastasi, B. K., & Clements, D. H. (1992). Social-cognitive behaviors and higher-order thinking in educational computer environments. *Learning and Instruction*, 2, 215–238.
- Nistor, N. (2003). Koalah: A problem-based virtual seminar on the WWW. In R. C. Geibert & S. Hunter-Harvey (Eds.), *Web-wise learning: Wisdom from the field* (pp. 337–362). Philadelphia, PA: Xlibris.
- Nussbaum, E. M., Hartley, K., Sinatra, G. M., Reynolds, R. E., & Bendixen, L. D. (2002, April). *Enhancing the quality of on-line discussions*. Paper presented at the Annual meeting of the American Educational Research Association.
- O'Donnell, A. M., & Dansereau, D. F. (1992). Scripted cooperation in student dyads: A method for analyzing and enhancing academic learning and performance. In R. Hertz-Lazarowitz & N. Miller (Eds.), *Interaction in cooperative groups. The theoretical anatomy of group learning* (pp. 120–141). Cambridge: University Press.
- Palincsar, A. S., & Brown, A. L. (1984). Reciprocal teaching of comprehension-fostering and comprehension-monitoring activities. *Cognition & Instruction*, 1(2), 117–175.
- Pea, R. D. (2004). The social and technological dimensions of “scaffolding” and related theoretical concepts for learning, education and human activity. *The Journal of the Learning Sciences*, 13(3), 423–451.
- Perkins, D. N. (1993). Person-plus: A distributed view of thinking and learning. In G. Salomon (Ed.), *Distributed cognition: Psychological and educational considerations* (pp. 88–110). Cambridge: University Press.
- Pithers, R. T. (2000). Critical thinking in education: A review. *Educational Researcher*, 42(3), 237–249.
- Resnick, L. B., Salomon, M., Zeitz, C., Wathen, S. H., & Holowchak, M. (1993). Reasoning in conversation. *Cognition and Instruction*, 11, 347–364.
- Rosenshine, B., & Meister, C. (1994). Reciprocal teaching: A review of the research. *Review of Educational Research*, 64, 479–630.
- Rosenshine, B., & Meister, C. (1996). Teaching students to generate questions: A review of the intervention studies. *Review of Educational Research*, 66(2), 181–221.
- Spiro, R. J., & Jehng, J. C. (1990). Cognitive flexibility and hypertext: Theory and technology for the nonlinear and multidimensional traversal of complex subject matter. In D. Nix & R. J. Spiro (Eds.), *Cognition, education, and multimedia: Exploring ideas in high technology* (pp. 163–205). Hillsdale: Erlbaum.

- Stein, N. L., & Bernas, R. S. (1999). The early emergence of argumentative knowledge and skills. In J. Andriessen & P. Coirier (Eds.), *Foundations of argumentative text processing* (pp. 97–116). Amsterdam: Amsterdam University Press.
- Stein, N. L., & Miller, C. A. (1996). A theory of argumentive understanding: Relationship among position preference, judgements of goodness, memory and reasoning. *Argumentation*, 7, 183–204.
- Tabak, I. (2004). Synergy: A complement to emerging patterns of distributed scaffolding. *Journal of the Learning Sciences*, 13(3), 305–355.
- Toulmin, S. (1958). *The uses of argument*. Cambridge: Cambridge University Press.
- van Eemeren, F. H. (2003). A glance behind the scenes: The state of the art in the study of argumentation. *Studies in communication Sciences*, 3(1), 1–23.
- Voss, J. F., Tyler, S. W., & Yengo, L. A. (1983). Individual differences in the solving of social science problems. In R. F. Dillon & R. R. Schmeck (Eds.), *Individual differences in cognition* (pp. 205–232). New York: Academic.
- Voss, J. F., & van Dyke, J. A. (2001). Narrative structure, information certainty, emotional content, and gender as factors in a pseudo jury decision-making task. *Discourse Processes*, 32(2&3), 215–243.
- Walton, D. N., & Krabbe, E. C. W. (1995). *Commitment in dialogue. Basic concepts of interpersonal reasoning*. Albany, NY: State University of New York Press.
- Webb, N. M., & Palincsar, A. S. (1996). Group processes in the classroom. In D. C. Berliner & R. C. Calfee (Eds.), *Handbook of educational psychology* (pp. 841–873). New York: Macmillan.
- Weinberger, A. (2003). *Scripts for Computer-Supported Collaborative Learning Effects of social and epistemic cooperation scripts on collaborative knowledge construction*. Doctoral dissertation, Ludwig-Maximilians-Universität (LMU), Munich, Germany. Available at: http://edoc.ub.uni-muenchen.de/archive/00001120/01/Weinberger_Armin.pdf.
- Weinberger, A., & Fischer, F. (2006). A framework to analyze argumentative knowledge construction in computer-supported collaborative learning. *Computers & Education*, 46(1), 71–95.
- Weinberger, A., Reiserer, M., Ertl, B., Fischer, F., & Mandl, H. (2005). Facilitating computer-supported collaborative learning with cooperation scripts. In R. Bromme, F. W. Hesse, & H. Spada (Eds.), *Barriers and Biases in network-based knowledge communication in groups* (pp. 15–37). Dordrecht: Kluwer.
- Weinberger, A., Stegmann, K., Fischer, F., & Mandl, H. (2007). Scripting argumentative knowledge construction in computer-supported learning environments. In F. Fischer, H. Mandl, J. Haake, & I. Kollar (Eds.), *Scripting computer-supported communication of knowledge—cognitive, computational and educational perspectives* (pp. 191–211). New York: Springer.
- Weiner, B. (1985). An attributional theory of achievement motivation and emotion. *Psychological Review*, 92, 548–573.