Argumentation in Web-based Collaborative Inquiry Learning: Scripts for Writing and Scripts for Talking Aren't the Same

Ingo Kollar & Frank Fischer, University of Munich, Leopoldstr. 13, D-80802 Munich, Germany
Email: ingo.kollar@psy.lmu.de, frank.fischer@psy.lmu.de
James D. Slotta, Ontario Institute for Studies in Education, Toronto, Canada
slotta@oise.utoronto.ca

Abstract: We use the script concept to describe knowledge structures that help individuals understand and act in specific contexts as well as scaffolds structuring collaborative learning. External scripts can be presented in different ways, e.g. as written text. For internal scripts, it is not clear whether they have identical effects on collaborative argumentation processes in oral vs. written discussions. We empirically investigated the effects of two differently structured external scripts on the structural quality of written and oral arguments produced in dyads with either low or high structured internal scripts. External scripts were presented in a written mode at specific instances in a web-based inquiry learning environment. Results indicate that the high structured external script strongly improved the structural quality of the written arguments, but had hardly any effects on orally produced arguments, which were instead more strongly influenced by the learners' internal scripts.

Introduction

The phenomenon of argumentation in collaborating groups has attracted a wealth of researchers in recent years, especially in science education and inquiry learning (e.g., Driver, Newton, & Osbourne, 2000; McNeill, Lizotte, Krajcik, & Marx, 2006). Inquiry learning has been described as a well-developed pedagogical approach in science education which puts learners in the role of scientists to solve science problems (e.g., Quintana et al., 2004; Slotta, 2004). Learners are usually provided with authentic science problems and empirical data or rich simulations that help them in their inquiry. During inquiry learning, students engage in a variety of knowledge-generating and knowledge-monitoring activities such as hypothesis generation, modeling, searching for empirical evidence, conducting experiments and collaborative argumentation. When inquiry learning is to be conducted collaboratively, students often require instructional support for collaborative argumentation. As one promising instructional scaffold, collaboration scripts have been described (e.g., Fischer, Kollar, Mandl & Haake, 2007). One design issue of collaboration scripts is related to how the script instructions should be presented to students – either in a written or oral mode or through some other modality. It is yet unclear whether there are differential effects of different types of presentation used to provide learners with such collaboration scripts. Further, the effectiveness of collaboration scripts is likely to interact with the learners' prior argumentation skills, which we conceptualize as "internal scripts on collaborative argumentation" (Kollar, Fischer, & Slotta, 2007). Another interesting open question with regard to these internal scripts is concerned with the extent to which they differentially influence students' collaborative oral or written argumentation. Therefore, the main research question for this study was: How do differently structured internal and external scripts interact in the process of evoking particular argumentative discourse patterns in dyadic collaborative inquiry learning?

The role of the structural quality of arguments in collaborative inquiry learning

Descriptive models about the logical and conversational structure of argumentation mainly come from psycholinguistics (e. g., Toulmin, 1958; van Eemeren, Grootendorst, & Henkemans, 1996). These models have been used by learning theorists to both analyze and facilitate collaborative argumentation in learning tasks. In related research, collaborative argumentation has been analyzed on at least two dimensions – an argument structure dimension and an argument sequence dimension (e.g., Stegmann, Weinberger, Fischer, & Mandl, 2004). In this paper, we focus on the structural aspect of arguments.

The most prominent model for analyzing the structure of arguments has been developed by Toulmin (1958). In its original version, the model distinguishes between six structural components that can be part of an argument. First, an argument usually consists of a *claim* that represents the speaker's position with respect to the topic at hand. Second, an argument may be backed up with *data* that support the claim. Third, a *warrant* may specify in what way the data support the claim. To support the warrant further, a speaker may fourth refer to some general law or principle, which is called *backing*. Fifth, a speaker may constrain the validity of his or her argument by introducing a qualifier such as "maybe" or "mostly". Finally, a speaker may be more explicit in constraining his or her argument by including a *rebuttal* that specifies concrete circumstances under which the claim is not true. In practice, it has been shown that some of the argument components are not easy to segregate.

In empirical studies which analyzed the structural quality of single arguments that were produced by learners, therefore very often a simplified Toulmin-like model has been used (e.g., Driver et al., 2000; McNeill et al., 2006; Stegmann et al., 2004). In the same way, in the empirical study we report here, we also constrain the structural model to three components: claims, data, and reasons (which comprise both warrants and backings).

The role of scripts in collaborative argumentation

Within psychology, the script concept has been used in at least two different ways. On the one hand, cognitive psychology uses the term to describe cognitive schemata that guide people in understanding and acting in particular sequential events such as going to a restaurant (e.g., Kolodner, 2007; Schank & Abelson, 1977). These scripts can be described as "internal" because they are regarded as being located inside the cognitive structure of a particular individual (see the restaurant script in Schank & Abelson, 1977). On the other hand, instructional psychology has used the term "script" to describe a specific type of scaffold, namely "collaboration scripts" which provide, distribute and sequence specific activities and roles among members of a collaborative learning group (Kollar, Fischer, & Hesse, 2006). Those scripts can be described as "external" because they usually are externally presented as a group starts working on a particular collaborative task. However, parts of an external collaboration script may be continuously internalized. In sections below, we describe how both internal and external scripts can influence collaborative argumentation, which will be described in the following.

Internal scripts for collaborative argumentation

From a cognitive perspective, scripts are internal knowledge structures that individuals develop over multiple similar experiences of events (e.g., Kolodner, 2007; Schank & Abelson, 1977). For example, most people have a grocery store script. This script puts the activities "getting a shopping cart", "choosing items", "putting items in shopping cart", "going to the cash desk" etc. in a meaningful sequence that allows us both to act properly in the grocery store and to understand stories others tell us about a grocery store event.

Let us consider how the script notion can be transferred to collaborative argumentation. Just like we go through multiple grocery store episodes over our life course, we also go through a number of situations in which collaborative argumentation is the key activity to engage in (see Reznitskaya, Anderson, McNurlen, Nguyen-Jahiel, Archodidou, & Kim, 2001). For example, collaborative argumentation is necessary when we are negotiating with our partner about who is doing the dishes, when we are discussing with colleagues how to approach a new work project, or when we are debating with an insurance company about an insurance fee. Thus, it can be assumed that students will use internal argumentation-related scripts they have developed elsewhere when they come to a collaborative inquiry learning episode (see also Sandoval & Millwood, 2005; Stein & Albro, 2001). In that respect, argumentation-related scripts can at least partially be defined as having a domain-general character. Yet, it can be assumed that individuals will vary in the degree to which their argumentation scripts comply with a Toulmin-like argumentation model. While one individual may hold a script specifying that arguments should always be backed up with data and reasons, others may have developed argumentation scripts that stipulate to produce striking arguments with plausible claims without backing them up with evidence. In compliance with Schank (1999), we assume that internal scripts for collaborative argumentation are not completely culturally shared and thus partly idiosyncratic in nature. For the sake of a first investigation, we hope to be able to identify at least two different types of learners: learners with rather high-structured internal argumentation scripts (i.e., whose internal scripts exhibit a rather high compliance with a simplified model of argument structure quality) and learners with rather low-structured internal argumentation scripts (i.e., whose internal scripts exhibit a rather weak compliance with the model of argument structure quality).

One open question is whether internal scripts on collaborative argumentation are equally important for oral and written argumentation. It is possible that individuals hold different internal scripts for situations in which oral argumentation is required and for situations in which written argumentation is required. This is especially important since existing web-based inquiry learning environments such as WISE (Slotta, 2004) sometimes require students to argue orally and sometimes in a written mode.

External scripts for collaborative argumentation

From an instructional psychology perspective, scripts are scaffolds that structure collaboration between two or more learning partners. Originally developed for face-to-face groups (e.g., O'Donnell & Dansereau, 1992), collaboration scripts have extensively been used in the context of Computer-Supported Collaborative Learning over the last years (e.g., Fischer, et al., 2007). In contrast to content-specific scaffolding techniques such as providing learners with worked-out examples (e.g., Atkinson & Renkl, 2007), collaboration scripts do not provide content-specific support, but rather directly modify the collaboration process.

In a recent literature review, Kollar et al. (2006) identified five conceptual components of collaboration scripts. First, collaboration scripts are always designed towards a specific pedagogical goal. For example, they can be designed to facilitate the acquisition of domain-specific knowledge (e.g., about Newtonian physics) or of domain-general knowledge (e.g., argumentation strategies). Second, collaboration scripts always induce certain

activities to be conducted by the learners (e.g., summarizing, explaining, constructing arguments). Third, a script always more or less explicitly puts these activities into a particular sequence (e.g., first "summarizing", then "questioning", then "arguing"). Fourth, collaboration scripts always either imply or explicitly specify and distribute specific collaboration roles among the learners (e.g., "summarizer" vs. "questioner"). And fifth, a collaboration script is always characterized by a specific type of presentation, i.e. the modality in which the script instructions are presented to the learners (e.g., in a written or a oral mode).

When reviewing empirical literature on collaboration scripts, at least two aspects are important for this paper. First, it appears that collaboration scripts can vary a lot concerning the degree of structuredness they impose on the collaborative learning process. While some scripts provide learners with rather rough guidelines on the phases of collaboration without giving specific guidance on how to act in the single phases (such as the ArgueGraph script developed by Dillenbourg & Jermann, 2007), other scripts are way more prescriptive and sometimes even ask learners to use sentence starters or text prompts to perform specific discourse activities. For example, with respect to collaborative argumentation, Stegmann et al. (2004) have developed different collaboration scripts that were used in a text-based asynchronous learning environment, in which students were supposed to analyze authentic problem cases on the basis of a psychological theory. In one script, when learners were to compose a new message, they were demanded to (a) name the claim they want to make, (b) give evidence that supports the claim and (c) give a qualifier that specifies under which circumstances the claim is valid (see Toulmin, 1958). In another script, the messages of the learners were pre-structured by a pre-defined title such as "argument", "counterargument", or "reply" (see Leitao, 2000). The empirical results indicated that both scripts alone significantly improved the dimension of argumentation they had been designed for. The first script improved the construction of formally complete arguments in the sense of the Toulmin-like argumentation model, and the second script raised the number of counterarguments and replies during discourse in comparison to an unstructured control condition. From these experimental results, it can be concluded that collaboration scripts can be designed to facilitate argumentation processes within collaborating groups. However, it is yet unclear, how much structure an argumentation-related collaboration script should provide to facilitate collaborative argumentation and whether the potentially positive effects of a high degree of structuredness can only be expected for learners with specific internal scripts.

Second, there is hardly any research concerning the effects of different types of presentation of collaboration script instructions on learning processes and outcomes. However, in literature on multimedia learning, it has been demonstrated that the modality in which instruction is presented to learners can have strong effects on outcomes such as retention and transfer performance (e.g., Ginns, 2005). Therefore, in this study we want to take a closer look at whether differently well-structured collaboration script that are designed in a written mode have differential effects on written and oral argumentation processes within collaborating dyads.

Goals of the study

The present paper presents in-depth process analyses of a quasi-experimental study by Kollar et al. (2007), which had a focus on individual knowledge acquisition through collaborative argumentation. In contrast, in this paper we focus on two process-related research questions: 1. What are the effects of two differently well-structured external scripts that give written instruction for collaborative argumentation on the structural quality of arguments produced in dyads with differently well-structured internal scripts on collaborative argumentation? 2. To what extent do differently well-structured internal and external scripts have differential effects on written and oral collaborative argumentation? Concerning the first question, we set up two competing hypotheses – an additive effects hypothesis and an interactive effects hypothesis. The *additive effects hypothesis* predicts that holding a high structured in contrast to a low structured internal script and being provided with a high structured in contrast to a low structured external script will have positive effects on the structural quality of the produced arguments). The positive effects of high structured internal and external scripts will be additive. In contrast, the *interactive effects hypothesis* predicts that the high structured external script will especially help learners with low structured internal scripts to produce more arguments with a high structural quality because only these learners require additional support. Learners with high structured internal scripts will be better off with a low structured external script because they already are capable of engaging in high-quality dialogue and argumentation. Imposing unnecessary constraints with a high structured external script on them, will at best have no effect, but may even be detrimental for these learners. With respect to the second question, we set up an undirected hypothesis, since prior research has not yet investigated whether there are differential effects of internal and external scripts on oral and written discussions: Internal and external scripts for collaborative argumentation will have differential effects on the structural quality of arguments that are produced either orally or in a written mode.

Method

Participants and design. As described in Kollar et al. (2007), participants were 90 students from grades 8 to 10 from two German secondary schools (M_{Age} = 15.3 years; SD = 0.99). They collaborated in dyads on a

biology module of the Web-based Inquiry Science Environment WISE (Slotta, 2004). 48 students were female, 42 were male. We established a 2x2-factorial design with the structuredness of the learners' internal scripts (high vs. low) and the external script (high vs. low) as independent factors. All dyads were homogeneous with respect to gender and internal scripts. Dyads were randomly assigned to either the low or the high structured external script condition. Unequal distribution over the four cells of the experimental design (see Table 1) was due to the fact that the study was conducted over two sessions, with some students missing on the second day.

Table 1: Overview over the experimental design.

		External scripts	
		Low structured	High structured
Internal scripts	Low structured	*N* = 22 (11 dyads)	*N* = 20 (10 dyads)
	High structured	*N* = 24 (12 dyads)	*N* = 24 (12 dyads)

Setting. For the purpose of this study, students were taken out of their class and seated in a separate room. During collaboration, a teacher was not present. All dyads collaborated on a German version of the WISE module "The Deformed Frogs Mystery" (Shear, Bell, & Linn, 2004). During collaboration, the two learning partners shared one laptop computer. We used screen capturing software that recorded both screen actions and talk over the course of collaboration. Parts of student dialogue were later transcribed and used for discourse analysis. Also, we recorded every written utterance that occurred over the course of collaboration (see below).

Learning environment. Dyads collaborated on the WISE "Deformed Frogs" module. In this module, students learn that huge numbers of frogs with missing or multiple limbs, shrinked heads, deformed jaws etc. have been found in the late 90's and that science has not yet found a commonly agreed explanation for this problem. The module presents dyads with two hypotheses (a parasite hypothesis and an environmental-chemical hypothesis) they are to discuss throughout their work on the module by looking at different sorts of background information (such as maps, photographs, journal articles). The module is segmented into five content-specific activities ("The Problem", "Where can these frogs be found?", "What's in the water?", "Why study frogs?", "Summary") with several steps (e.g., "Take a note", "What kinds of deformities have been found?", "Discuss the two hypotheses" etc.) attached to each activity.

Procedure. The experiment was conducted in two sessions. In the first session, learners were asked to fill in a questionnaire on demographic variables, domain-specific prior knowledge, and several individual learning prerequisites. Also, the students completed a test to assess the structuredness of their internal scripts (see below). Within the two weeks before the second session (collaboration session) took place, these tests were analyzed and each learner was identified as holding either a high or low structured internal script. Then, homogenous dyads were established. In the second session, students collaborated on the WISE module for 120 minutes followed by a post test phase in which we assessed domain-specific and domain-general knowledge (for these results, please see Kollar et al. 2007), motivation and acceptance..

Independent variables. The two versions of the *external script* were implemented at several identical points in the Deformed Frogs module. At the end of each activity, the final step asked students to discuss the two hypotheses on the background of the evidence they had gathered in the particular activity and to type their arguments into designated text boxes. In the low structured external script condition, learners were asked to orally discuss and write their arguments into a blank text box. In the high structured external script condition, dyads were also told to orally discuss and write their arguments down, but received written instruction on how the noted arguments should look like. In compliance with the simplified argument structure model introduced above, the script presented learners with nine text boxes which were pre-specified in that they put learner A into the role of an advocate for the parasite hypothesis, requiring to first provide data upon which to base her or his argument, then name a claim he or she was going to make, and then to give a reason specifying why the data would support the claim. Each component was to be typed into a separate textbox, each of which was further pre-structured with a sentence starter (e.g., "I observed that…" for the "data" component). Then, in compliance with the model of argument sequences developed by Leitao (2000), learner B was to give a counterargument (again including the three structural components "data", "claim", "reason" including the sentence starters), and finally both partners were to collaborate to find an integrative argument (again with the three components). Then learner B was to advocate the environmental-chemical hypothesis, and A had to give a counterargument. Again, the analogous supportive functions of the script were implemented. Over the course of the dyad's collaboration, roles were switched several times in order to avoid biased information processing. Also, in correspondence with claims made for example by Pea (2004), the script was continuously faded out during the course of collaboration. For example, in a later "Discuss the two hypotheses"-phase, students were no more asked to specify the structural components of their arguments, and finally were even free to write their argumentation in

a completely unstructured way. At several instances during oral discussions, learners were reminded of the argumentation rules specified in the high structured external script, but not explicitly forced to follow them.

The learners' *internal scripts* on collaborative argumentation were assessed in the first session. In this test, students were provided with a fictitious protocol of a science-related discussion between two girls. This protocol included both complete and incomplete arguments (with respect to the argument structure model) and both complete and incomplete argumentation sequences (with respect to Leitao's argument sequence model). The learners' task was to read the protocol and to find instances of good and poor argumentative moves and to give reasons for why they were good or poor. In total, participants could reach up to 20 points on this measure. In average, students reached a value of $M = 3.49$ ($SD = 2.38$), the median value was 3. The sample was split into halves, so that 42 students of the final sample were identified as holding a low structured internal script and 48 students as holding a high structured internal script (the unequal distribution was due to absenteeism in the second session). Since, according to the script concept introduced by Schank and Abelson (1977), scripts guide both our understanding and our acting in particular situations, we also validated the results of the internal scripts test by looking at the arguments that were produced during collaboration by learners with low and high structured internal scripts in the low structured external script condition. The results largely confirmed the results of the initial internal scripts test. Learners with high structured internal scripts based on the initial test produced more formally complete arguments and more counterarguments and in tendency more integrative arguments than learners who had scored low in the initial test (see Kollar et al., 2007).

Dependent variables. With respect to dependent variables, we focused on the frequencies of arguments with different levels of structural quality. Two kinds of externalizations of the dyads were taken into account. First, we analyzed arguments produced in overall discourse, without differentiating between arguments that were produced in a written or oral format. For economic reasons, we abstained from a full transcription of oral discussions, and instead transcribed ten intervals of five minutes length each of the discussions in each dyad ("time sampling"; Bakeman & Gottman, 1997). Second, to get a more detailed impression on the effects of internal and external scripts when the external script with its written instructions was salient, we looked more closely at those arguments that were typed into the text boxes in the two conditions.

To perform these analyses, we first needed to separate argumentative talk from non-argumentative talk. Argumentative talk was defined by the speaker at least implicitly or explicitly uttering a claim that was relevant to the deformed frogs problem. The units of analysis were turns. Interrater reliability was sufficient ($\kappa = .78$). For further analyses, only these discourse elements were used that had been rated as instances of argumentative talk. However, since prior research (e.g., Resnick, Salmon, Zeitz, Wathen, & Holowchak, 1993) has shown that arguments can develop over multiple turns and speakers, we first needed to adopt a new segmentation procedure to identify arguments. As a constituent for an argument, we used the implicit or explicit presence of a claim. We then looked in the surrounding discourse corpus for further argument components that obviously were related to this claim. Once a new claim was uttered, a new segment was coded. For segmentation, we used principles introduced by Strijbos, Martens, Prins and Jochems (2006) which lead to two agreement indicators of 79 % and 81 % (for a more detailed description of the method, see Strijbos et al., 2006). After segmentation, every identified argument was assessed with respect to *argument structure quality*. Three levels of argument structure quality were distinguished: (a) arguments consisting only of a claim (low quality), (b) arguments consisting of a claim and data *or* a reason (medium quality), and (c) arguments consisting of a claim, data, and a reason (high quality). Interrater reliability with respect to argument structure quality was sufficient ($\kappa = .68$).

Control variables. As control measures, we used motivation, interest, prior collaboration experience, prior experience in learning with the computer and prior domain-specific knowledge. All of these constructs were measured by a paper-pencil-questionnaire that was to be completed in the first session. In none of these variables, we found a systematic difference between the four experimental groups except for the biology grade in the preceding year. It appeared that students with high structured internal scripts outperformed students with low structured internal scripts ($F(1,87) = 7.38$; $p < .01$; $Eta^2 = .08$). Thus, there is obviously a connection between high argumentation skills and getting good grades in biology. We consider this result as a confirmation of the validity of the internal script test, since empirical research (Means & Voss, 1996) has shown that learners exhibit higher quality of argumentation when they have more domain-specific knowledge.

Statistical analyses. With respect to each discourse category, we computed ANOVA's to see whether it was affected by the different internal and external scripts. For all analyses, the alpha level was set to 5 %.

Results

Effects of internal and external scripts on argument structure quality in overall discourse

We first looked at the effects of internal and external scripts on the structural quality of all arguments, no matter if they were produced in oral or in written discourse. The mean frequencies and standard deviations of arguments with low, medium, and high structural quality are presented in table 2.

Separate ANOVA's revealed that the structuredness of the learners' internal scripts had significant effects on the production of arguments with high ($F(1,40) = 10.48$; $p < .01$; $Eta^2 = .21$) and with medium structural quality ($F(1,40) = 7.32$; $p < .01$; $Eta^2 = .16$), but not on the production of arguments with low argument structure quality ($F(1,40) < 1$; *n.s.*). I.e., learners with high structured internal scripts produced more arguments with higher levels of argument structure quality than learners with low structured internal scripts, while there were no differences with respect to the production of arguments with low structural quality.

Table 2: Mean frequencies and standard deviations of arguments on different levels of argument structure quality (in overall discourse).

	Low structured internal scripts				High structured internal scripts			
Overall argument structure quality	Low structured external script		High structured external script		Low structured external script		High structured external script	
	M	*SD*	M	*SD*	M	*SD*	M	*SD*
Low	10.09	*4.11*	7.50	*5.89*	11.55	*6.92*	9.33	*5.19*
Medium	14.36	*6.04*	15.00	*10.46*	28.27	*12.51*	15.92	*5.87*
High	5.64	*3.50*	8.30	*5.77*	13.09	*6.35*	10.92	*4.68*

For the structuredness of the external script, however, only one significant, but small main effect was found, namely for the production of argument with a medium structural quality (with the high structured external script reducing the number of such arguments, $F(1,40) = 4.57$; $p < .05$; $Eta^2 = .10$). However, we also found a small, but significant interaction effect on the same measure ($F(1,40) = 5.62$; $p < .05$; $Eta^2 = .12$), indicating that the positive effect of the low structured external script on the production of arguments with medium structural quality was especially due to the high number of such arguments produced by learners with high structured internal scripts. A similar interaction effect for arguments with high structural quality failed to reach statistical significance ($F(1,40) < 2.40$; *n.s.*). No further effects were significant ($F(1,40) < 2.00$; *n.s.*). To sum up, in overall discourse, mainly the learners' internal scripts were influential. The external script had at best a rather weak impact on the production of arguments with medium structural quality.

Effects of internal and external scripts on argument structure quality in written discourse

To answer the second research question (whether differently well-structured internal and external scripts have differential effects on the structural quality of arguments that were either produced in an oral or a written mode), we conducted analogous analyses with respect to the arguments that were produced in a written mode (for the descriptive results, see table 3).

Separate ANOVA's on the single argument structure quality categories revealed a large and significant main effect of the external script on the production of arguments with a high structural quality ($F(1,41) = 23.86$; $p < .01$; $Eta^2 = .37$). A marginally significant main effect for the external script was also found for the production of arguments with medium structural quality ($F(1,41) = 3.14$; $p = .08$; $Eta^2 = .07$). For arguments with low structural quality, no main effect was found for the external script ($F(1,41) = 1.68$; *n.s.*). Interestingly, the learners' internal scripts did not have effects on any of the three categories ($F(1,41) < 1$; *n.s.*).Thus, when looking only at those arguments that were produced when the external script was salient, only the external script was influential, while effects of the learners' internal scripts disappeared.

Table 3: Mean frequencies and standard deviations of arguments on different levels of argument structure quality (in written discourse).

	Low structured internal scripts				High structured internal scripts			
Written argument structure quality	Low structured external script		High structured external script		Low structured external script		High structured external script	
	M	*SD*	M	*SD*	M	*SD*	M	*SD*
Low	3.00	*2.49*	2.40	*2.50*	3.92	*2.11*	2.50	*3.17*
Medium	3.00	*1.48*	3.70	*1.83*	2.92	*2.02*	4.33	*2.46*
High	0.36	*0.50*	3.30	*3.68*	0.67	*0.89*	3.83	*2.04*

Discussion

In this paper, we presented in-depth process analyses on a data set described in Kollar et al. (2007). We investigated the interplay and differential effects of differently structured internal and external scripts on the structural quality of arguments produced by dyads in web-based inquiry learning. The first research question

was: "What are the effects of two differently well-structured external scripts that give written instruction for collaborative argumentation on the structural quality of arguments produced in dyads with differently well-structured internal scripts on collaborative argumentation?" Connected to this question there were two competing hypotheses, an additive and an interactive effects hypothesis. With respect to this question, none of the two hypotheses received full support. While interaction effects were solitary and not homogeneously distributed over the three levels of argument structure quality, we did find that holding a high as compared to a low structured internal script led to the production of arguments with a higher structural quality, when the overall discourse material was examined. This result has at least two implications. First, it can be seen as a validation of the results of the initial internal scripts test that asked students to analyze a fictitious discourse protocol. Referring back to Schank and Abelson's (1977) notion of scripts as internal knowledge structures that permit both *understanding* and *acting* in related situations, there indeed is a correspondence between these two functional components. Second, the result fits nicely with the notion of internal scripts as knowledge structures that have developed through participation in numerous related situations over one's life course and thus are relatively stable (Kolodner, 2007; Schank & Abelson, 1977). In our study, students indeed used their internal scripts that they probably have developed not only in school settings, but also in other instances, to regulate their collaborative argumentation. The high structured external script, in contrast, did not have clear effects on argument structure quality when looking at the overall discourse. Obviously, the high structured external script was not powerful enough to deeply change or even replace the learners' internal scripts by a more – from a scientific perspective – sophisticated one.

This, however, does not mean that the high structured external script was ineffective, because when looking at those phases of collaboration, in which the written external script instructions were highly salient (namely in written discourse), we found large effects compared to the low structured external script in the direction of an improvement of argumentative processes. This result fits with prior research for example by Stegmann et al. (2004) which demonstrated the power of collaboration scripts for successfully modifying collaborative argumentation processes in a highly specific manner. Thus, the second research question ("To what extent do differently well-structured internal and external scripts have differential effects on written and oral argumentation?") can clearly be answered as "Yes, there are differential effects". While the internal scripts obviously were influential for oral discussions (which took a large part in overall discourse), they were ineffective in written argumentation – and the reverse was true for the external script. Thus, for both internal and external scripts, it can be argued that scripts for talking are not simultaneously scripts for writing. Instead, internal scripts as measured by an argumentation analysis task rather seem to be a script for talking (but not for writing), while the external script with its written instructions seemed to be a script for writing (but not for talking). From a pedagogical perspective, this result points to the need of finding ways to transfer the positive effects of providing learners with a high structured external scripts to situations in which such a script is not available. Probably, more long-term effects can only be reached through a more long-term intervention. To confirm this, future studies are needed.

Coming back to the two competing hypotheses concerning the first question, it can thus be argued that there was no additive effect of high structured internal and external scripts in the sense of simultaneous effects of the two. Instead, an additive effect has occurred in a chronological way – both high structured internal scripts and high structured external scripts improved the structural quality of collaborative argumentation, but in different phases of collaboration. As a consequence, designers of web-based collaborative inquiry learning environments should present high structured external scripts in a written mode to improve written argumentation processes. However, future research should also try to find ways how to transfer the positive effects to oral discussions – or find external scripts that have the potential to improve oral discussions in a way similar to the written script in this study improved written discussions.

References

Atkinson, R. K. & Renkl, A. (2007). Interactive example-based learning environments: Using interactive elements to encourage effective processing of worked examples. *Educational Psychology Review, 19,* 375-386.

Bakeman, R. & Gottman, J. M. (1997). *Observing interactions.* New York: Cambridge University Press.

de Jong, T. & van Joolingen, W. R. (1998). Scientific discovery learning with computer simulations of conceptual domains. *Review of Educational Research, 68*(2), 179-201.

Dillenbourg, P. & Jermann, P. (2007). Designing integrative scripts. In F. Fischer, I. Kollar, H. Mandl & J. M. Haake (Eds.), *Scripting computer-supported collaborative learning: Cognitive, computational, and educational perspectives* (pp. 275-301). New York: Springer.

Driver, R., Newton, P. & Osborne, J. (2000). Establishing the norms of scientificargumentation in classrooms. *Science Education, 84*, 287–312.

Fischer, F., Kollar, I., Mandl H., & Haake, J. M. (Eds.) (2007). *Scripting Computer-Supported Collaborative Learning – cognitive, computational and educational perspectives.* New York, NY: Springer.

Ginns, P. (2005). Meta-analysis of the modality effect. *Learning & Instruction, 15*(4), 313-331.
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. (in press). Internal and external scripts in computer-supported collaborative inquiry learning. *Learning & Instruction.*
Kolodner, J. L. (2007). The roles of scripts in promoting collaborative discourse in learning by design. In F. Fischer, I. Kollar, H. Mandl & J. M. Haake (Eds.), *Scripting Computer-Supported Collaborative Learning – cognitive, computational, and educational perspectives* (pp. 237-262). New York, NY: Springer.
Leitão, S. (2000). The potential of argument in knowledge building. *Human Development, 43*, 332-360.
McNeill, K. L., Lizotte, D. J, Krajcik, J., & Marx, R. W. (2006). Supporting students' construction of scientific explanations by fading scaffolds in instructional materials. *The Journal of the Learning Sciences*, 15(2), 153-191.
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*(2), 139-178.
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). New York: Cambridge University Press.
Pea, R. (2004). The social and technological dimensions of scaffolding and related theoretical concepts for learning, education, and human activity. *Journal of the Learning Sciences, 13*(3), 423-451.
Quintana, C., Reiser, B. J., Davis, E. A., Krajcik, J., Fretz, E., Duncan, R. G., Kyza, E., Edelson, D. & Soloway, E. (2004). A scaffolding design framework for software to support science inquiry. *The Journal of the Learning Sciences, 13*(3), 337-386.
Resnick, L. B., Salmon, M., Zeitz, C. M., Wathen, S. H. & Holowchak, M. (1993). Reasoning in conversation. *Cognition and Instruction, 11*(3&4), 347-364.
Reznitskaya, A., Anderson, R. C., McNurlen, B., Nguyen-Jahiel, K., Archodidou, A. & Kim, S.-Y. (2001). Influence of oral discussion on written argument. *Discourse Processes, 32*(2&3), 155-175.
Sandoval, W. A. & Millwood, K. A. (2005). The quality of students' use of evidence in written scientific explanations. *Cognition and Instruction, 23*(1), 23-55.
Schank, R. C. (1999). *Dynamic memory revisited*. Cambridge, NY: Cambridge University Press.
Schank, R. C. & Abelson, R. P. (1977). *Scripts, plans, goals and understanding*. Hillsdale, NJ: Erlbaum.
Shear, L., Bell, P. & Linn, M. C. (2004). Partnership models: The case of the deformed frogs. In M. C. Linn, E. A. Davis & P. Bell (Eds.), *Internet environments for science education* (pp. 289-311). Mahwah, NJ: Erlbaum.
Slotta, J. D. (2004). Web-Based Inquiry Science Environment. In M. C. Linn, E. A. Davis & P. Bell (Eds.). *Internet Environments for Science Education* (pp. 203-231). Mahwah, NJ: Erlbaum.
Stegmann, K., Weinberger, A., Fischer, F. & Mandl, H. (2004). Scripting argumentation in computer-supported learning environments. In P. Gerjets, P. A. Kirschner, J. Elen, & R. Joiner (Eds.), *Instructional design for effective and enjoyable computer-supported learning. Proceedings of the first joint meeting of the EARLI SIGs Instructional Design and Learning and Instruction with Computers* (S. 320-330). Tübingen: Knowledge Media Research Center.
Stein, N. L. & Albro, E. R. (2001). The origins and nature of arguments: Studies in conflict understanding, emotion, and negotiation. *Discourse Processes, 32*(2), p. 113-133.
Strijbos, J.-W., Martens, R. L., Prins, F. J. & Jochems, W. M. G. (2006). Content analysis: What are they talking about? *Computers & Education, 46*, 29-48.
Toulmin, S. (1958). *The uses of argument*. Cambridge, UK: Cambridge University Press.
Van Eemeren, F. H., Grootendorst, R. & Henkemans, F. S. (1996). *Fundamentals of argumentation theory – a handbook of historical backgrounds and contemporary developments.* Mahwah, NJ: Erlbaum.