Vogel, Freydis and Wecker, Christof and Kollar, Ingo and Fischer, Frank (2017) Socio-cognitive scaffolding with collaboration scripts: a meta-analysis. Educational Psychology Review, 29 (3). pp. 477-511. ISSN 1040-726X

Socio-Cognitive Scaffolding with Computer-Supported Collaboration Scripts: a Meta-Analysis

Freydis Vogel[1] · Christof Wecker[2] · Ingo Kollar[3] ·
Frank Fischer[2]

Abstract Scripts for computer-supported collaborative learning (CSCL) offer socio-cognitive scaffolding for learners to engage in collaborative activities that are considered beneficial for learning. Yet, CSCL scripts are often criticized for hampering naturally emerging collaboration. Research on the effectiveness of CSCL scripts has shown divergent results. This article reports a meta-analysis about the effects of CSCL scripts on domain-specific knowledge and collaboration skills. Results indicate that CSCL scripts as a kind of socio-cognitive scaffolding can enhance learning outcomes substantially. Learning with CSCL scripts leads to a small positive effect on domain-specific knowledge ($d = 0.20$) and a large positive effect on collaboration skills ($d = 0.95$) compared to unstructured CSCL. Further analyses reveal that CSCL scripts are particularly effective for domain-specific learning when they prompt transactive activities (i.e., activities in which a learner's reasoning builds on the contribution of a learning partner) and when they are combined with additional content-specific scaffolding (worked examples, concept maps, etc.). Future research on CSCL scripts should include measures of learners' internal scripts (i.e., prior collaboration skills) and the transactivity of the actual learning process.

Keywords Collaboration scripts · Collaboration skills · Computer-supported collaborative learning · Domain-specific knowledge · Socio-cognitive scaffolding · Transactivity

Research on computer-supported collaborative learning (CSCL) deals with the question how digital technologies can be used to help groups of learners collaborate on a high level

✉ Freydis Vogel
    freydis.vogel@tum.de

[1]   Technical University of Munich, Arcisstr. 21, 80333 Munich, Germany

[2]   LMU Munich, Leopoldstr. 13, 80802 Munich, Germany

[3]   University of Augsburg, Universitätsstr. 10, 86159 Augsburg, Germany

(Koschmann 1996). A high level of collaboration is reached when groups of learners engage in certain socio-cognitive activities such as explaining (Webb et al. 2009), questioning (King 1998), or arguing (Andriessen, Baker, and Suthers 2003). Through an engagement in such activities, students are assumed to acquire both domain-specific knowledge (i.e., knowledge about the topic that is discussed within the group) and cross-domain skills such as collaboration skills or argumentation skills. In short, collaborative learning is credited for its high potential to facilitate learners' cognitive development (Mugny and Doise 1978; Schwarz and Linchevski 2007). However, as prior research on collaborative learning—both in face-to-face and in computer-mediated settings—has shown, learners often have difficulties engaging spontaneously in beneficial collaborative learning activities (Cohen 1994; Kuhn, Shaw, and Felton 1997). In this article, the term "unstructured collaboration" is used to refer to such situations in which learners are not supported with respect to their collaborative learning activities.

To overcome these problems, learners can be supported by means of socio-cognitive scaffolding that guides them through collaborative activities that enhance learning. In computer-supported learning, such scaffolding for collaboration can be provided through CSCL scripts. A CSCL script is a kind of instructional support that uses computer technology to provide learners with guidance about how to interact (Kollar, Fischer, and Hesse 2006). CSCL scripts are believed to be effective for domain-specific and domain-general learning by means of different mechanisms. A CSCL script may prompt learners directly to engage in collaborative activities that promote deep elaboration of the learning material. For instance, a CSCL script may prompt learners to build upon each others' contributions. A study by Stegmann et al. (2007) compared the effects of a CSCL script that structured online discussions in the domain of education to unstructured CSCL. Domain-specific knowledge was measured by students' performance in applying their knowledge about attribution theory in a case analysis. Results of this study showed that those students who learned collaboratively with the support of a CSCL script outperformed students who collaborated without a CSCL script with respect to their domain-specific knowledge. The CSCL script might have induced beneficial learning activities by structuring the discussion, and, thus, the CSCL script increased knowledge acquisition.

Furthermore, CSCL scripts may deliberately distribute different learning materials among the participants of a small group in order to induce knowledge interdependence. Knowledge interdependence is likely to increase interaction by interfering with the small group's tendency to converge too early (Dillenbourg and Hong 2008). Knowledge interdependence may also be induced by having interdisciplinary teams learn together on one task. The goal of scripting interdisciplinary teams can be twofold: first, increasing interaction by knowledge interdependence; second, having interdisciplinary problem solving as a genuine learning goal. For example, Rummel et al. (2009) developed a CSCL script to scaffold the collaborative work of physicians and psychotherapists on a joint diagnosis and therapy plan for a patient. This CSCL script and an advanced version of this script (enriched with instructional prompts and phases for reflective self-explanation) were compared to a control condition. The learners' collaboration skills were assessed in individual post-tests that required them to describe relevant aspects of good collaborative work in computer-supported collaboration settings. Results showed that the CSCL scripts had positive effects on learners' collaboration skills compared to the control condition (Rummel et al. 2009).

In short, a variety of CSCL scripts have been developed and analyzed in empirical studies (e.g., Demetriadis, Egerter, Hanisch, and Fischer 2011; Ertl, Fischer, and Mandl 2006;

Gelmini-Hornsby, Ainsworth, and O'Malley 2011; Gijlers and de Jong 2009; Haake and Pfister 2010; Kollar, Fischer, and Slotta 2007; Noroozi, Teasley, Biemans, Weinberger, and Mulder 2013; Rummel, Mullins, and Spada 2012; Schellens, Van Keer, De Wever, and Valcke 2007; Stegmann et al. 2007; Van Aalst and Chan 2007; Weinberger, Ertl, Fischer, and Mandl 2005). Yet, despite broad agreement concerning the high potential of CSCL scripts for enhancing collaborative learning, research about its effectiveness yielded mixed results (see Molinari, Sangin, Dillenbourg, and Nüssli 2009; Weinberger et al. 2010). Hence, even though a consensus exists that structured collaborative learning should lead to better learning outcomes than unstructured collaborative learning, little is known about the general effectiveness of learning with CSCL scripts. A comprehensive meta-analysis about the effects of CSCL scripts does not seem to have been conducted so far. Furthermore, the factors that might moderate the effectiveness of CSCL have not systematically been investigated. Meta-analyses and reviews that are concerned with a similar topic were either focused on face-to-face collaboration (Howard 1996) or CSCL merely based on argumentation (Noroozi, Weinberger, Biemans, Mulder, and Chizari 2012), or they less comprehensively review research on specific collaboration scripts in CSCL and face-to-face learning without reporting effect sizes (Kollar et al. 2006).

This article presents a meta-analysis that (a) investigates the effects of CSCL scripts on learners' domain-specific knowledge and collaboration skills and (b) examines to what extent different factors in the learning environment moderate the effectiveness of CSCL scripts. The following theoretical overview discusses potential explanatory mechanisms for effects of CSCL scripts on collaborative learning.

CSCL Scripts as Socio-Cognitive Scaffolding

Prior research suggests that CSCL scripts may enhance learners' domain-specific knowledge and collaboration skills. In this article, the term "domain-specific knowledge" is used for learning outcomes that are tied to a specific domain or subject (e.g., knowledge about the functioning of the neuron or mastery of algebraic problem solving). In contrast, the term "collaboration skills" is used for skills that can be applied in collaborative settings across different domains, such as argumentation skills (Noroozi et al. 2013) or peer assessment skills (Ulicsak 2004). As the theoretical underpinning for the present meta-analysis, we build on the *script theory of guidance* (SToG; Fischer, Kollar, Stegmann, and Wecker 2013).

The Role of Learners' Internal Scripts According to the script theory of guidance, collaboration skills can be regarded as "internal scripts" that—once they are consolidated—guide an individual's understanding and acting in similar situations. Based on Schank's theory of dynamic memory, Fischer et al. (2013) assume that internal scripts may consist of four partially hierarchically ordered components:

1. The *play* component comprises an individual's knowledge about the type of collaborative endeavor in which he or she and his or her co-learners are engaged (such as collaborative problem solving or collaborative argumentation). This component encompasses knowledge and expectations about an appropriate sequence of scenes and about their functions in a collaborative setting.

2. A *scene* component encapsulates an individual's knowledge and expectations about the specific situations that occur within a play (such as phases of individual idea generation, collaborative pooling of individually generated ideas, or the joint development of a solution for a problem). Each scene encompasses knowledge about an appropriate sequence of activities and about their functions within the scene.

3. A *scriptlet* component contains an individual's knowledge and expectations about the kinds of activities that can occur within a scene (such as considering a peer's ideas and evaluating their appropriateness for solving a problem). Each scriptlet comprises knowledge on how to perform a particular activity by executing specific (cognitive) operations.

4. A *role* component comprises knowledge and expectations regarding the question of how different activities that are expected in a particular scene are distributed among the members of the group. Hence, a role component intersects with the scene components, just as it is the case in a movie script.

According to the script theory of guidance, a learner may adapt and reconfigure internal script components in a new collaborative situation and develop new instances of an internal script (Fischer et al. 2013). For instance, a learner may already have developed an internal script about collaborative reading that incorporates the activity "asking thought-provoking questions." In a new situation in which the learner is asked to solve a problem together with a learning partner, the learner might generate a new configuration of internal script components for this collaborative problem-solving situation, which includes the activity of asking thought-provoking questions.

Often, however, learners are not able to spontaneously engage in beneficial collaborative learning activities (Cohen 1994). One reason for this may be that they lack appropriate internal scripts (Fischer et al. 2013). Therefore, "external" CSCL scripts may be employed to support students in dealing with the collaborative learning situation, which may in turn result in better individual learning outcomes compared to unstructured collaboration (King 2007). CSCL scripts are employed to induce beneficial collaborative learning activities by prompting cognitive operations necessary to perform these activities, sequencing the resulting activities, and distributing them among specific roles that are assigned to individual learners (Kobbe et al. 2007).

Corresponding to the components of the internal script, external CSCL scripts may present scaffolds on the three hierarchical script levels (play, scene, and scriptlet levels; Fischer et al. 2013). Play-level scaffolds in a CSCL script stimulate activities by specifying a sequence of scenes. Scene-level scaffolds in a CSCL script specify and potentially sequence activities that are necessary for achieving the function of the scene. Scriptlet-level scaffolds in a CSCL script specify individual operations that are necessary for contributing a specific activity. A CSCL script may also distribute the specified activities among roles that are assigned to individual learners during the collaboration (Fischer et al. 2013).

Supporting Learning with CSCL Scripts The basic assumption for how CSCL scripts support learning is that they are designed in a way that guides students in performing meaningful and beneficial collaborative learning activities, which may result in more positive learning outcomes, both with respect to domain-specific knowledge and collaboration skills, as compared to unstructured collaboration (King 2007). Recently, diverse types of CSCL scripts have been developed to foster learning in various domains, such as computer science (Demetriadis et al. 2011; Haake and Pfister 2010), economics (Huang, Wu, and Chen 2012),

life sciences (Noroozi et al. 2013), educational science (Stegmann et al. 2007), medicine, and psychology (Rummel et al. 2009). These CSCL scripts are designed to implement distinct collaborative learning scenarios that are assumed to enhance domain-specific learning by fostering sets of collaborative learning activities that evoke a deep elaboration of the learning material at hand (King 2007).

In general, studies about collaborative learning have shown that there is a range of collaborative learning activities that lead to domain-specific individual learning. Prototypical examples of beneficial collaborative learning activities are reciprocal questioning and explaining, creating and sharing external representations of knowledge, or engaging in discursive learning activities (King 1992; Kwon and Cifuentes 2009; Webb et al. 2009). Webb et al. (2009), for instance, reported in their study about elementary school students' algebraic problem solving that the collaborative activity of giving explanations during small group learning dialogue is positively related to domain-specific learning outcomes. Teasley (1997) pointed out the importance of other-oriented transactive activities. The most important characteristic of these transactive activities is that they take the learning partners' contributions into account (e.g., by criticizing, refining, or extending these contributions).

Transactive activities also play a pivotal role in Chi's (2009) proposed classification of collaborative learning activities. In this classification, types of learning activities are arranged in a hierarchical order from "passive" on the lowest level, "active" and "constructive" on intermediate levels, to "transactive" (or "interactive") on the highest level. Each level incorporates the characteristics of the activities on the level(s) below. "Passive" means that a learner passively receives information (e.g. watching a video). "Active" applies when an action is performed physically using the learning material (e.g. pointing to words at a text). "Constructive" applies to activities that are characterized by the production of knowledge beyond the information the learner can extract directly from the learning material (e.g., summarizing a text in one's own words). "Transactive" applies to activities by which learners build on their learning partners' contributions (e.g., by extending, refining, or revising them; Chi 2009; Teasley 1997).

During collaboration, a mixture of all types of activities is likely to occur. But, according to Chi (2009), there are qualitative differences between the learning mechanisms that accompany each type of activity. The cognitive operations stimulated during the performance of constructive and transactive activities are assumed to be more beneficial for knowledge acquisition than those stimulated when collaborating mainly on the two lower levels. During both the constructive and transactive activities, learners must cognitively elaborate learning material that is presented to them either by the learning environment or by their learning partners. In contrast, collaboration on the passive or active levels does not involve the cognitive elaboration of the learning material. Furthermore, transactive activities that explicitly refer to the contributions made by the learning partner may provide additional benefits. First, the learning partner may serve as an additional source of knowledge and perspectives (Clark and Sampson 2007) and mutual scaffolding (De Wever, van Keer, Schellens, and Valcke 2010; Ismail and Alexander 2005) and thus help the learners to further develop knowledge and skills. Second, transactive activities during collaborative learning may lead to socio-cognitive conflict that can trigger cognitive development and domain-specific knowledge acquisition by learners' efforts to resolve the conflict (Mugny and Doise 1978; Schwarz and Linchevski 2007).

Consequently, transactive—and, to a lesser degree, constructive—activities have a higher chance to promote learning than collaboration on the active or even passive levels (Chi 2009). By supporting such higher-level collaborative activities as peer feedback and assessment

(Ulicsak 2004) or social-discursive argumentation (Noroozi et al. 2013), CSCL scripts may contribute to deeper elaboration of the learning material and to improved domain-specific learning (King 2007). For example, scripting a social-discursive argumentation about the application of an educational theory to cases facilitated learners' engagement in formulating arguments based on this theory. Through the engagement in formulating arguments, a deeper level of elaboration of this theory was reached, accompanied by the acquisition of more domain-specific knowledge with respect to the theory (Weinberger et al. 2010). However, studies report rather heterogeneous results regarding the effects of learning with CSCL scripts compared to unstructured CSCL on domain-specific knowledge.

Although fostering the acquisition of domain-specific knowledge is the first priority of many educators, CSCL scripts may also promote learners' development of collaboration skills. Collaboration skills typically targeted by CSCL scripts include argumentation skills (Noroozi et al. 2012), peer feedback skills (Demetriadis et al. 2011), and skills for productive interdisciplinary collaboration (Rummel and Spada 2005; Rummel et al. 2009). By repeatedly engaging in a collaboration with the support of a CSCL script, learners will acquire skills in performing productive collaborative activities (Anderson 1992). Through repeated participation in the collaborative situation—and possibly the gradual removal of the CSCL script ("fading"; Wecker and Fischer 2011)—the learners' internal script can be modified by iteratively employing a new configuration of internal script components (Fischer et al. 2013; Schank 1999).

Potential Moderators of the Effectiveness of CSCL Scripts

The variety of results with respect to the effects of CSCL scripts on the acquisition of domain-specific knowledge and collaboration skills leads to an important question: What is it that makes certain CSCL scripts effective and other CSCL scripts less effective or even ineffective? A closer look at some of the CSCL scripts that are described in the literature indicates that their design differs substantially from each other with respect to at least three dimensions. First, the CSCL scripts vary extensively with respect to the collaborative activities they prompt. Here, we focus on *transactivity* because transactive activities are assumed to be most beneficial for collaborative learning (Chi 2009; Chi and Wylie 2014; Teasley 1997). Second, CSCL scripts differ in terms of how much structure they induce. While some CSCL scripts provide less structuring by only being focused on the play level, others provide rather deeply structured support on the scene level or, even more deeply, on the scriptlet level (Fischer et al. 2013). Third, CSCL scripts vary in the extent to which they are accompanied by *additional content-related support*. While some CSCL scripts only structure the collaboration, others are coupled with domain-specific support such as worked examples or concept maps. The present meta-analysis covers these three different factors as potential moderators of the effectiveness of CSCL scripts.

Transactivity A first possible moderator influencing the effectiveness of CSCL scripts is related to the type of activity that the CSCL scripts are designed to stimulate. Chi (2009; Chi and Wylie 2014) recently synthesized findings from empirical studies that dealt with the relation between different learning activities and learning outcomes and identified interactive activities as being most beneficial for learning (as compared to collaboration on the constructive, active, or passive levels). In the script theory of guidance, Fischer et al. (2013) postulate

that "the more a given CSCL practice requires the transactive application of knowledge, the better this knowledge is learned through participation in this CSCL practice" (p. 60). The theoretical assumption is that the contributions of the learning partner might offer additional resources and stimulate cognitive elaboration by, for instance, contributing another perspective or asking thought-provoking questions (Chi 2009; Teasley 1997). Thus, transactive activities can enhance the development of knowledge about content, which the learners transactively elaborate on (Fischer et al. 2013).

A first rough inspection of results of empirical studies about learning with CSCL scripts that are designed to stimulate transactivity also supports these assumptions. For instance, a study by Noroozi et al. (2013) analyzed the effects of two different CSCL scripts for transactive knowledge sharing and transactive discussion and compared them to the effects of unstructured collaboration. The results indicated that learning with CSCL scripts for transactive knowledge sharing and transactive discussion positively affected both the learning activities and domain-specific knowledge acquisition compared to learning collaboratively without any support by a script.

For CSCL scripts that are not designed to stimulate transactivity, some studies did not reveal positive effects of CSCL scripts on individual learning as compared to a control condition. For example, a study by Mäkitalo, Weinberger, Häkkinen, Järvelä, and Fischer (2005) used a script that scaffolded the application of a theory in a computer-supported collaborative problem-solving task. More specifically, the script presented learners with a pre-specified structure for their messages to their learning partners. Contrary to what was hypothesized, learners in the control condition who collaborated without a script outperformed learners who were supported by the script in applying the theory in an individual post-test case analysis. This study supports the assumption that CSCL scripts that are not designed to stimulate transactivity might not be optimally effective for learning. Experimental studies that directly test the relevance of transactivity for the effectiveness of CSCL scripts are rare, however, and too specific for a general conclusion (e.g., Noroozi et al. 2013). Therefore, this meta-analysis tries to merge prior findings to arrive at more robust conclusions with respect to whether CSCL scripts that provoke learners' engagement in transactive activities are more effective than CSCL scripts that do not.

Script Level CSCL scripts are often criticized for providing too much structure and being too rigid. This criticism is based on the expectation that a high degree of structure and rigidity in these scripts might have a negative effect on learners' motivation and self-regulation and thus on their engagement in "naturally occurring" collaboration. This drawback, in turn, might potentially lead to reduced or even negative effects of CSCL scripts on learning outcomes, sometimes called "over-scripting" (Dillenbourg 2002). However, for learners with collaboration skills that are not yet fully developed in all of their components, CSCL scripts that provide less structure might not be effective for learning because they do not explicitly specify the activities that are functional for learning (Rienties et al. 2012; Stegmann, Mu, Gehlen-Baum, and Fischer 2011). To what extent a CSCL script structures the collaboration can be approximately operationalized by the script level that the CSCL script mainly focuses. A CSCL script that mainly focuses the play level has a lower degree of structure than a CSCL script that mainly focuses the scene level, which in turn has a lower degree of structure than a CSCL script that mainly focuses the scriptlet level.

To compare the effects of different script levels on learning activities and outcomes requires studies in which external scripts for each script level are developed for the collaborative

situation that should be supported (e.g., a script for peer feedback). So far, we are not aware of any studies on CSCL scripts that experimentally manipulated the focus on one of the three script levels for analyzing their effectiveness on learning activities and outcomes. Nevertheless, it seems possible to characterize CSCL scripts from the literature with respect to whether they provide a lower degree of structure, i.e., they primarily target the play level, or whether they provide a higher degree of structure by primarily targeting the scene level or the scriptlet level. To be sure, there may be instances in which a script that, for example, mainly focuses on the scriptlet level also strongly emphasizes the scene level (even more as some other scripts that mainly focus this level), and vice versa. However, these are atypical and probably rare cases. Given the current degree of differentiation and level of reporting in the literature synthesized in this meta-analysis, the proposed distinction is the best approximation to the script level currently available and enables us to provide at least some preliminary insights concerning the question whether the amount of structure of the script moderates the effect of CSCL scripts on learning outcomes.

Additional Content-Related Support By definition, CSCL scripts scaffold collaborative learning by prompting rather formal (content-independent) aspects of collaborative learning activities, such as question asking, argumentation, or explaining (Kollar et al. 2006). Yet, to engage in high-quality collaborative learning activities (e.g., social-discursive argumentation), learners might also need scaffolds that support them in their processing of the domain-specific content (Sadler 2004). For the purposes of this meta-analysis, we differentiate between three levels at which content-related support may be provided:

At the highest level, CSCL scripts may be combined with domain-specific scaffolds that are designed to support the processing of content-related information in problem-solving tasks, such as a content scheme (Ertl, Kopp, and Mandl 2008) or content-specific scaffolds for hypothesis generation in simulation-based inquiry learning (Gijlers and de Jong 2009). At the intermediate level, content-related information may only be made available, without offering learners guidance that might facilitate its processing during problem solving. This is the case when, for instance, learners are provided with reading assignments about background information (Rummel and Spada 2005). Here, domain-specific content serves as a knowledge resource that may or may not be accessed and used for problem solving. At the lowest level, no additional content-related support may be offered at all while students are learning with a CSCL script (Haake and Pfister 2010).

Ideally, scaffolding targeted at different aspects, such as a collaboration script supporting learners' activities and content-related scaffolding supporting the application of domain-specific knowledge, amplifies each other's effectiveness; that is, they constitute what has been called "synergistic scaffolding" (Tabak 2004). Only a few studies directly compared the effect of the combination of CSCL scripts and additional content-related support in a two-by-two factorial design (e.g., Ertl et al. 2008). However, a meta-analytic moderator analysis can at least indicate to what extent the effects of the CSCL scripts synthesized in this meta-analysis might be affected by the degree of additionally presented content-related support.

Further Potential Moderators Studies on CSCL differ in various other features. For instance, they differ with respect to the level on which outcomes are measured. While some

studies measure outcomes on the level of individual learners, other studies measure the outcome on the level of small groups of two or three persons. Other varying features of the studies are overall learning time, the educational background of the participants, the subject discipline of the study, the country the study was conducted, and the time-space modes of collaboration. The meta-analysis also explores these variables as potential moderators for the effects of CSCL scripts on domain-specific learning outcomes and collaboration skills. Yet, given the lack of theoretical assumptions concerning the effects of these moderators, the respective analyses have a somewhat exploratory character and are intended to stimulate further theory development and empirical research.

Research Questions

The present meta-analysis was conducted to investigate the following four research questions and associated hypotheses:

Research question 1. What is the effect of collaborative learning with a CSCL script compared to collaborative learning without a CSCL script on the acquisition of (a) domain-specific knowledge and (b) collaboration skills? We hypothesized positive effects of collaborative learning on both domain-specific knowledge and collaboration skills with a CSCL script compared to learning without a CSCL script.

Research question 2. To what extent do the effects of CSCL scripts on (a) domain-specific learning outcomes and (b) collaboration skills depend on the prompting of transactive activities by the CSCL scripts? As transactive activities are supposed to benefit learning outcomes more than other types of activities, for instance collaboration on the active or constructive levels (Chi 2009), we expected CSCL scripts that prompt an engagement in transactive activities to have stronger effects on domain-specific learning and the development of collaboration skills than CSCL scripts that do not prompt an engagement in transactive activities.

Research question 3. To what extent do the effects of CSCL scripts on (a) domain-specific learning outcomes and (b) collaboration skills depend on the script level that is mainly addressed? If scripts with a high degree of structure (i.e., mainly targeting the scriptlet level) would be less helpful than scripts with a lower degree of structure (i.e., mainly targeting the play level or the scene level), this would be in line with the over-scripting hypothesis that scripts may decrease motivation and engagement in collaboration (Dillenbourg 2002).

Research question 4. To what extent do the effects of CSCL scripts on (a) domain-specific learning outcomes and (b) collaboration skills depend on (the level of) additional content-related support? We expected that higher-level additional content-related support (i.e., the provision of content-specific scaffolds such as worked examples or content schemes) would be associated with stronger effects of CSCL scripts on domain-specific knowledge and collaboration skills than additional content-related support at lower levels (i.e., content information merely as a knowledge resource or no content presentation at all).

Method

Criteria for Inclusion

For inclusion in this meta-analysis, studies had to meet the following requirements:

Independent Variable We included only studies that compared learners who had learned with a CSCL script to learners who also engaged in CSCL but were not externally supported by a CSCL script. For this meta-analysis, learners are considered to be supported by a CSCL script if they are assigned roles or if at least one learning partner is prompted for at least one activity in which the learner addresses one or several learning partner(s). Examples include reading a text to the learning partner, explaining something to a learning partner, asking questions, or contributing an argument to a discussion. Typically, however, CSCL scripts prompt collaborators to engage in a sequence of activities, which often are compatible with each other. Support can be implemented either in computer-mediated communication or in face-to-face interaction that is supported by technology. In the latter case, the computer is used as a tool to prompt the activities specified by the script. We deliberately excluded comparisons of CSCL scripts to other types of instructional support (e.g., modeling, Rummel et al. 2009). First, these comparisons address a different research question that is not the focus of this meta-analysis (e.g., "What is the effect of collaborative learning with a CSCL script compared to collaborative learning with modeling on the acquisition of domain-specific knowledge and collaboration skills?"). Yet, these kinds of research questions are addressed very rarely in primary research and thus a meta-analytic synthesis would not be feasible. Second, meta-analyses typically focus not only on general effects but also on the role of potential moderator variables. By including moderator analyses, we still find out more about the conditions of effective learning with CSCL scripts beyond the results the primary studies can provide.

Dependent Variables We included only studies that report at least one objective quantitative outcome measure of domain-specific knowledge or collaboration skills. For domain-specific knowledge, a measure was required to assess the learner's knowledge about content that was relevant to the domain for which the collaborative learning environment was designed. For collaboration skills, a measure was required to assess skills that are necessary to engage appropriately in a collaborative situation. Measures for domain-specific learning outcomes and/or collaboration skills were obtained either on the individual level or on the group level.

Experimental or Quasi-Experimental Design The meta-analysis included only studies that manipulated the availability of the CSCL script in an experimental or quasi-experimental design. A study follows an experimental or quasi-experimental manipulation of a CSCL script if in at least one experimental condition learners are supported by means of a CSCL script and in at least one control condition learners collaborate without script support.

Availability of Data A requirement for inclusion in the meta-analysis is that studies had to report sufficient quantitative information for calculating an effect size (e.g., F value along with denominator degrees of freedom, p value, effect size, or mean, standard deviation, and number of participants) regarding at least one dependent variable for the comparison of the experimental and the control condition.

Source In order to secure high-quality studies in the meta-analysis, only studies published in peer-reviewed journals were considered eligible. Furthermore, studies were included only if the report was written in English.

Selection Procedure

We searched two different bibliographic databases (ERIC and ISI Web of Science) in July 2013 by using the following search string: "(scaffold* OR script*) AND (learn* OR know*) AND (collaborat* OR cooperat*) AND (computer* OR CSCL OR techno*)." The search was limited to articles in peer-reviewed academic journals. The two database searches yielded 248 articles.

The articles obtained from the bibliographic databases were evaluated for their relevance in two steps. First, their titles and abstracts were screened to distinguish between potentially relevant and clearly non-relevant articles. Second, the full texts of the potentially relevant articles were examined and coded for relevance according to the criteria for inclusion.

Assessment of Relevance Based on Title and Abstract A coding scheme was developed to classify the relevance of the articles reflecting the first criterion for inclusion (i.e., the independent variable). Two coders (two of the authors of this paper) practiced coding the articles by applying the coding scheme. After training, the coders achieved perfect inter-rater reliability based on 19 independently double-coded articles (Cohen's $\kappa = 1.00$). The coding of the remaining articles resulted in a first selection of potentially relevant publications comprising 175 articles. The full texts of these 175 relevant articles were then acquired as pdf documents for the second stage of the evaluation for relevance.

Assessment of Relevance Based on Full Texts of Articles Three coders (three of the authors of this paper) examined the full texts of all articles to evaluate whether the articles met all of the criteria for inclusion. If an article failed to meet one of the criteria, it was omitted from the sample of articles included in the meta-analysis. The coders evaluated the following criteria regarding the content: (a) whether the article reported a study about learning with CSCL scripts and (b) whether at least one measure for domain-specific knowledge or collaboration skills was reported. For the articles that met the first two criteria, the coders evaluated the formal criteria: (c) whether the CSCL script was manipulated as an independent factor in an experimental or quasi-experimental study, (d) whether sufficient quantitative data to estimate an effect size were reported for at least one measure of domain-specific learning outcome or collaboration skills, and (e) whether the study was reported in English in a peer-reviewed journal.

Beyond that, the coders checked whether the study had been reported more comprehensively (e.g., with a broader set of dependent variables or more detailed information about experimental manipulations) in another article. If another article reported information that the respective publication did not contain, the information from the two publications was merged at a later stage when coding the study features.

After training, each coder evaluated the whole sample of 175 articles. The agreement concerning the decision of whether to include individual articles was acceptable (Fleiss' $\kappa = .722$). Each of the 26 individual articles for which there was any disagreement was discussed until a consensus about inclusion was reached. This set of assessments of relevance led to the exclusion of 153 articles. A total of 22 articles met all criteria for inclusion. Some of

the accepted articles reported more than one experiment, and some of the experiments contained more than one comparison between an experimental condition and a corresponding control condition; therefore, the number of comparisons included in the meta-analysis is larger than the number of publications included.

Description of the Articles Included in the Meta-Analysis

The 22 articles (published between 2005 and 2012) that were finally included comprised 24 studies containing 34 comparisons, from which 45 effect sizes were estimated. Overall, the studies included involved 2825 participants ($M_N = 101$; $SD_N = 99.85$), whose mean age was 20 years on average ($M_{age} = 20.17$; $SD_{age} = 4.85$). In 21 of the articles, personal computers were used as electronic devices, and in only one article mobile devices were used (Huang, Wu, and Chen 2012). The average time that the participants worked with support from the treatment was about 4 h ($M_{duration} = 239.08$ min; $SD_{duration} = 341.45$). The shortest overall learning time supported by the treatment in a study was 12 min (Molinari et al. 2009), and the longest was about 24 h (Haake and Pfister 2010). The studies included represent a certain bandwidth of further characteristics, such as the educational background of the participants or the subject discipline in which the study was conducted.

Coding of Study Features

In a further step, we developed a coding scheme for categorizing the CSCL scripts regarding the three main moderators (transactivity, script level, and additional content-related support) and four exploratory moderator variables (educational background of participants, subject discipline, country in which the study was conducted, time-space modes for collaboration). Again, the three coders first practiced the application of the coding scheme to assure inter-coder agreement. Specifically, after training, the coders categorized all treatments described in seven of the articles that were initially found to be relevant. They achieved sufficient agreement for each of the three coded moderators (see the descriptions below). Because all three moderators reflected an implicit ordering, we used the intraclass correlation coefficient (ICC) as a measure of inter-rater agreement. More specifically, the two-way mixed-model, absolute agreement ICC for single measures was used. Also, for the exploratory moderator variables, sufficient agreement was achieved. Fleiss' Kappa was used to determine the inter-coder reliability of these characteristics because they were coded on a nominal scale. The remaining articles were distributed equally among the three coders and coded individually. The moderator variable levels of all studies are presented in the appendix.

Transactivity We distinguished between *CSCL scripts that prompt transactive activities* and *CSCL scripts that do not prompt transactive activities*. According to Chi (2009), an activity is transactive if it builds upon or directly refers to a contribution of a learning partner, such as responding to a learning partner's request, or criticizing a learning partner's argument. In contrast, CSCL scripts that do not prompt transactive activities ask learners only to contribute activities without direct uptake of a partner's contribution, such as reading something to the learning partner or writing an argument. Agreement among the coders for this variable was satisfactory ($ICC = .82$).

Script Level CSCL scripts may contain scaffolds on the play level, the scene level, and the scriptlet level. Typically, one of the three levels is the main focus of a CSCL script. We defined the moderator variable *script level* of the CSCL script in a study as the script level with the highest proportion of activities that were prompted by the CSCL script. To clearly distinguish among CSCL scripts that offer scaffolds addressing the three script levels, we used a working definition for each of these three levels based on the script theory of guidance (Fischer et al. 2013):

> *Play-level* scaffolds prompt two or more learners to engage in collaborative activities in a sequence of scenes that are necessary for executing the "play" (e.g., discussing an issue together, finding a joint solution for a problem). The only individual activities that are prompted by play-level scaffolds are activities that (a) are to be performed individually between the collaborative activities and (b) do not differentiate among the learners of a learning group (e.g., two learners first read the same text individually, then both learners discuss the text).
> *Scene-level* scaffolds prompt learners to perform individual activities that are necessary to collaboratively engage in the scenes that are scaffolded on the play level. The activities on the scene level are usually different for different learners of a learning group (e.g., for a discussion, one learner might be requested to contribute arguments while the other learners may be requested to listen to the arguments and to think about potential points of criticism).
> *Scriptlet-level* scaffolds prompt learners to conduct individual (typically mental) operations that are needed to perform the activities required on the scene level. For example, in order to formulate an argument, a learner could be requested to first think about a claim and then search for appropriate evidence that supports that claim. The inter-rater agreement for the script level moderator was $ICC = .61$ and can be considered satisfactory (Orwin 1994, p. 152).

Additional Content-Related Support All studies that met the inclusion criteria were further coded with respect to the degree of additional content-related support. The code "*No content-related support*" was applied to studies in which learners received no domain-specific content in the learning environments. The code "*Content-related information only*" was given to studies in which the learning environment contained domain-specific content, such as background readings, summaries of a theory, or dictionaries, without any guidance on how to apply this information. The code "*Content-related scaffolds*" was applied to a study when learners received support on how to strategically work with the domain-related material (e.g., worked examples, content scheme, scaffolds for hypothesis generation in a simulation for inquiry learning). The inter-rater agreement was satisfactory ($ICC = .61$).

Calculation of Effect Sizes

The chosen index of effect size was the unbiased estimate of the standardized mean difference between experimental and control conditions with standard deviations pooled across groups (Hedges 1981; Hedges and Olkin 1985). Following Cohen's (1988) suggestion, effect sizes of $d = 0.20$–0.49 are considered as small effects, effect sizes of $d = 0.50$–0.79 as medium effects, and effect sizes of $d = 0.80$ and higher as large effects (Cohen 1988).

If a study reported more than one measure for domain-specific knowledge or collaboration skills, we averaged the effect sizes for each of these types of outcomes to arrive at a single effect size per outcome type for each study. None of the studies reported the correlations between different dependent variables. Therefore, the corresponding within-study effect variance was calculated by assuming an inter-correlation of 0.5 (see Borenstein, Hedges, Higgins, and Rothstein 2009, p. 228). The effect sizes from the individual primary studies are listed in the Appendix.

Statistical Analysis

Synthesis of Effects The random-effects model with inverse-variance weighting was used for synthesizing the effect sizes because study features and the kinds of measures vary substantially among studies. Hence, a unique common effect size for all studies cannot be assumed (Borenstein et al. 2009, pp. 69ff).

To compare effect sizes of CSCL scripts between moderator levels, we calculated the mean effect sizes and their confidence intervals for each subgroup assuming random effects within moderator levels. We used separate estimates of the between-study variance. However, if one of the moderator levels contained five or fewer studies, the variance estimates were pooled (Borenstein et al. 2009).

The criterion of statistical significance was set to 5 % for all statistical tests. Post hoc power analyses were conducted in order to assess the sensitivity of the present meta-analysis to detect existing main and moderator effects. The results of these analyses are reported along with the results from the corresponding statistical tests.

Assessment of Publication Bias We generated funnel plots to assess the extent of publication bias (see section on sensitivity analysis). A funnel plot is a scatter plot of the effect sizes against their respective standard errors. In the case of a publication bias, the plot is skewed and shows irregular densities, especially for higher standard errors (Begg 1994; Greenhouse and Iyengar 1994).

We additionally employed the trim-and-fill algorithm (Duval and Tweedie 2000a, 2000b) to estimate the number of effects that may be missing in the literature due to publication bias and to assess to what extent the findings might differ if such effect sizes were observed and reported. We used the estimator L for the number of suppressed studies (Duval and Tweedie 2000a, p. 92).

Results

Mean Effect of CSCL Scripts on Domain-Specific Knowledge and Collaboration Skills

(RQ 1a) To estimate the mean effect of learning with CSCL scripts on domain-specific knowledge, we synthesized 33 single effect sizes derived from 21 articles. As expected, the amount of between-study variability in effect sizes cannot be explained by within-study sampling error alone, $Q(df = 32) = 137.88$, $p = .02$, which validates the use of the random-effects model in estimating the effect of CSCL scripts on domain-specific knowledge. As hypothesized, a statistically significant positive, although small, average effect was found (see Table 1). Post hoc power analysis revealed that the statistical power of this test was 87 %.

Table 1 Overall mean effect size and mean effect sizes within moderator levels for domain-specific knowledge along with their confidence intervals and inferential statistical information

	k	d_{Mean}	$CI_{90\%}$	z	p
Overall effect	33	0.20*	[0.04; 0.36]	2.08	.019
Transactivity					
Transactive activities not prompted	6	0.05	[−0.46; 0.56]	0.17	.432
Transactive activities prompted	27	0.22*	[0.05; 0.39]	2.08	.019
Script level					
Play level scaffolds	4	−0.06	[−0.50; 0.38]	−0.22	.587
Scene level scaffolds	15	0.28*	[0.05; 0.52]	1.97	.024
Scriptlet level scaffolds	14	0.19	[−0.06; 0.45]	1.25	.105
Degree of additional content-related support					
No content-related support	7	0.18	[−0.31; 0.66]	0.60	.275
Content-related information only	16	0.14	[−0.06; 0.35]	1.16	.122
Content-related scaffolds	10	0.30*	[0.09; 0.51]	2.34	.009

Note. *$p < .05$, one-tailed

(RQ 1b) To estimate the average effect of learning with CSCL scripts on collaboration skills, we synthesized 12 single effect sizes taken from eight articles. Again, the amount of between-study variability in effect sizes cannot be explained by within-study sampling error alone, $Q(df = 11) = 42.30$, $p < .01$, which validates the use of the random-effects model. The synthesis yielded a statistically significant large positive effect (see Table 2), which supports the hypothesis that collaborative learning with a CSCL script has a beneficial impact on the development of collaboration skills compared to collaborative learning without a CSCL script. Post hoc power analysis showed that the statistical power of this test was 99 %. The effect of learning with CSCL scripts on collaboration skills was significantly greater than the effect of learning with CSCL scripts on domain-specific knowledge ($z_{Diff} = 3.24$; $p < .01$, two-tailed).

The sample included two outliers with respect to effects on collaboration skills. When the two outliers were excluded, the synthesis still yielded a statistically significant positive effect of medium size (see Table 2). The post hoc power of this test was 97 %.

Moderator Analyses

The substantial variation among the individual studies' effects for both domain-specific knowledge and collaboration skills warrants the consideration of moderators to investigate how CSCL scripts should be designed to help students reach high levels of knowledge and skill acquisition. The estimated mean effects for the levels of the moderator variables are reported in Tables 1 and 2.

The Role of Transactivity In this moderator analysis, the effects of CSCL scripts that prompt transactive activities were compared to the effects of CSCL scripts that do not prompt transactive activities.

(RQ 2a) With respect to effects on domain-specific knowledge, the difference between the two subgroups was not statistically significant, $Q(df = 1) = 0.28$, $p = .60$. Post hoc power analysis indicated a statistical power of 8.2 %. (RQ 2b) Concerning the effects on collaboration

Table 2 Overall mean effect size and mean effect sizes within moderator levels for collaboration skills along with their confidence intervals and inferential statistical information

	k	d_{Mean}	$CI_{90\%}$	z	p
Overall effect	12	0.95*	[0.63; 1.27]	4.95	.001
Overall effect[a]	10	0.65*	[0.49; 0.81]	6.54	.001
Transactivity					
Transactive activities not prompted	3	2.15*	[1.69; 2.61]	7.65	.001
Transactive activities not prompted[a,b]	1	–	–	–	–
Transactive activities prompted	9	0.62*	[0.42; 0.81]	5.24	.001
Script level					
Play level scaffolds	2	0.62	[−0.11; 1.36]	1.39	.082
Scene level scaffolds	5	0.63*	[0.16; 1.09]	2.23	.013
Scriptlet level scaffolds	5	1.47*	[0.97; 1.97]	4.83	.001
Scriptlet level scaffolds[a]	3	1.22*	[0.65; 1.78]	3.56	.001
Degree of additional content-related support					
No content-related support[b]	1	–	–	–	–
Content-related information only	8	1.13*	[0.69; 1.57]	4.25	.001
Content-related information only[a]	6	0.62*	[0.39; 0.84]	4.56	.001
Content-related scaffolds	3	0.70*	[0.03; 1.36]	1.71	.043

*$p < .05$, one-tailed

[a] Estimated without outliers

[b] Not estimated due to small number of studies

skills, the difference between the two subgroups was also not statistically significant in the expected direction, $Q(df=1)=22.37$, $p<.001$. Post hoc power was 99.7 %.

The Role of the Script Level The second moderator analysis concerned the question of whether CSCL scripts that target a certain script level (e.g., the play level) are more or less effective than scripts that target other script levels (e.g., the scene or the scriptlet level).

(RQ 3a) The differences among the effects of the three script level subgroups on domain-specific knowledge were not statistically significant, $Q(df=2)=1.27$, $p=.53$. Post hoc power analysis showed that the statistical power of this test was 15.7 %. (RQ 3b) Regarding the effects on collaboration skills, the differences among the three script level subgroups were again not statistically significant, $Q(df=2)=.78$, $p=.09$. Post hoc power analysis indicated a statistical power of 11.3 %. Deleting the two outliers mentioned above resulted in a small change for the subgroup with scriptlet level scaffolds (see Table 2). The differences between the three script level subgroups were not statistically significant, $Q(df=2)=2.11$, $p=.35$. The post hoc power of this test was 23.6 %.

The Role of Additional Content-Related Support In the third moderator analysis, we compared three levels of content-related support.

(RQ 4a) With respect to domain-specific knowledge, the differences among the three subgroups were not statistically significant, $Q(df=2)=0.77$, $p=.68$. Post hoc power analysis showed that the statistical power of this test was 11.3 %. (RQ 4b) With respect to collaboration

skills, no mean effect could be estimated for CSCL scripts without content-related support due to an insufficient number of studies in this subgroup ($k = 1$). The difference between the two remaining subgroups was not statistically significant, $Q(df = 1) = 0.78$, $p = .38$. The post hoc power was 14.3 %. Deleting the two outliers mentioned above yielded a lower effect size for the subgroup with content-related information only (see Table 2). Still, there was no significant difference between the two subgroups, $Q(df = 1) = 0.19$, $p = .67$. The post hoc power of this test was 7.2 %.

Further Moderators

Concerning the influence of overall learning time, two meta-regressions were run with the effect sizes for domain-specific knowledge and collaboration skills, respectively, as the criterion variables and the natural logarithm of the duration of the study as the only predictor. For these meta-regressions, the between-study variance component was estimated using the approximated method of moments as described by Raudenbush (1994, pp. 310 −311). For domain-specific knowledge, longer studies tended to have slightly lower effect sizes, $b = -0.14$; $SE = 0.11$; $CI_{90\%} = [-0.31; 0.04]$; $\beta = -0.20$. However, this tendency was not statistically significant, $p = .21$, two-tailed. For collaboration skills, also no statistically significant dependency on study duration was found, $b = -0.09$; $SE = 0.45$; $CI_{90\%} = [-0.88; 0.70]$; $\beta = -0.06$; $p = .85$, two-tailed. In a further moderator analysis, we analyzed the effects of CSCL scripts as a function of two levels of measurement. On the first level, the outcome was measured on the individual level, while in the second level the outcome was measured on the small group level. For domain-specific learning outcomes, in studies that measured the outcomes on the individual level, the effect of CSCL scripts was higher ($d = 0.31$) than in studies that measured the outcomes on the group level ($d = -0.02$), but the difference was not statistically significant, $Q(df = 1) = 2.52$, $p = .11$. Also, for learning collaboration skills the effect of CSCL scripts was higher in studies that measured the outcome on the individual level ($d = 0.72$) compared to studies that measured the outcome on the group level ($d = 0.62$), but the difference was not statistically significant, $Q(df = 1) = 0.29$, $p = .59$.

Also, for each of the further characteristics (educational background of participants, subject discipline, country in which the study was conducted, time-space modes for collaboration), we conducted a moderator analysis to check if the effects of CSCL scripts are influenced by these characteristics. The results showed that only the subject discipline was a significant moderator for the effectiveness of CSCL scripts on collaboration skills. Other characteristics did not turn out to be statistically significant moderators (see Table 3).

Sensitivity Analyses

To assess the potential impact of publication bias on the findings of this meta-analysis, we generated funnel plots for the two types of outcome measures. The funnel plot for domain-specific knowledge shows different densities of comparisons for different levels of the standard error of the effect sizes (see Fig. 1). Yet, the funnel plot is rather symmetric. Only for very low standard errors does the plot seem to be skewed toward larger effect sizes, whereas the opposite holds for large standard errors. Overall, the funnel plot does not indicate publication bias with respect to the effects of CSCL scripts on domain-specific knowledge.

Table 3 Characteristics of the 22 articles finally included in the meta-analysis and moderator-analysis based on the characteristics for domain-specific knowledge and collaboration skills

Coding dimensions (Fleiss'Kappa) and codes/no. of articles	Domain-specific knowledge					Collaboration skills				
	k	d	Q	df	p	k	d	Q	df	p
Educational background of participants (1.00)			0.23	1	.63			0.23	1	.63
School level (elementary school, high school)/6	8	0.26*				3	0.63**			
University-, college-level/16	25	0.28				7	0.65**			
Subject discipline (.68)			2.43	3	.49			9.85	3	.02
Natural sciences (e.g., biology, physics)/3	5	0.35*				2	0.64**			
Social sciences (e.g., education)/9	15	0.10				3	0.97**			
Technical disciplines (e.g., engineering, computer science)/3	3	−0.04				1	0.81**			
Various disciplines/other/no discipline indicated/7	10	0.23				4	0.43**			
Country in which the study was conducted (.73)			4.68	3	.20			0.01	1	.90
Asia (Hong Kong, Taiwan)/2	3	0.72**				0				
Europe excluding Germany (e.g., Belgium, Netherlands, Switzerland, UK)/7	9	0.18				1	0.61*			
Germany/12	20	0.11				9	0.65**			
USA/1	1	0.32				0				
Time-space modes for collaboration (1.00)			0.42	2	.81			0.02	1	.90
Face to face only/2	2	0.27**				0				
Distance only/14	22	0.20*				5	0.69**			
Face to face and distance/6	9	0.24				5	0.63**			

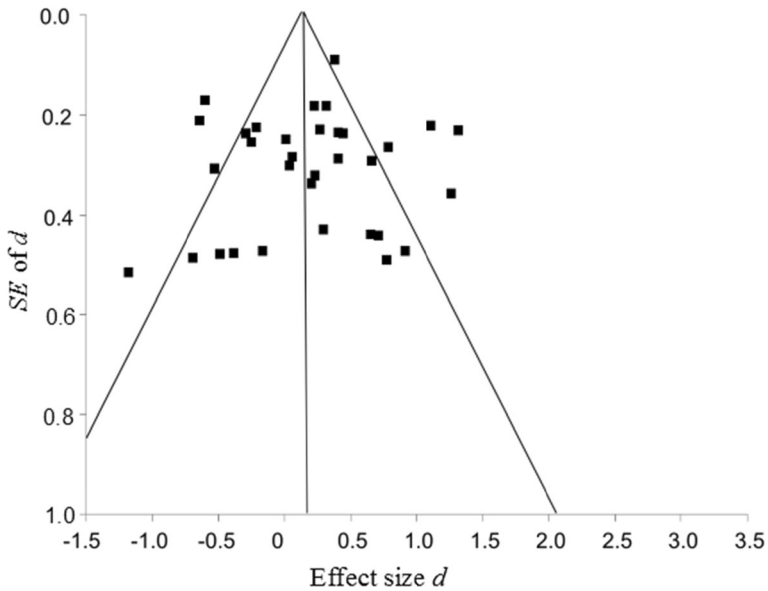Note. *$p < .05$, one-tailed; **$p < .01$, one-tailed

Fig. 1 Funnel plot for the effects of collaborative learning with CSCL scripts on domain-specific learning outcomes

With respect to collaboration skills, larger effect sizes were reported only in studies with larger standard errors. The opposite was true for smaller effect sizes. The smaller effect sizes for collaboration skills all stem from the studies with smaller standard errors (see Fig. 2). This funnel plot exhibits a skewed pattern. Potentially, comparisons with lower effect sizes from
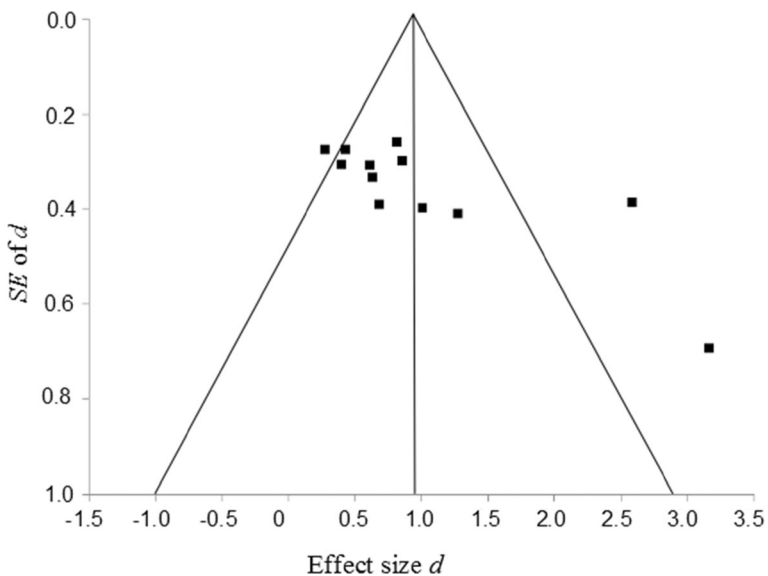

Fig. 2 Funnel plot for the effects of collaborative learning with CSCL scripts on the development of collaboration skills

studies with large standard errors (smaller sample sizes) might be missing due to publication bias.

To cross-validate the conclusions from this graphical analysis, the trim-and-fill method was employed for the effects of CSCL scripts on both domain-specific knowledge and collaboration skills (Duval and Tweedie 2000a, 2000b). With respect to the effects on domain-specific knowledge, the "trim" part of this procedure estimated the number of potentially suppressed comparisons to be 0 (i.e., 0 %). Therefore, the "fill" part of the procedure did not lead to any change of the estimated mean effect of $d = 0.20$ ($CI_{90\%}$ [0.04; 0.36], $p = .02$). These analyses can be viewed as evidence against a strong publication bias with respect to the effects of CSCL scripts on domain-specific knowledge.

With respect to collaboration skills, the number of potentially missing comparisons was estimated to be 1 (i.e., 8.33 %). The estimate of the mean effect after interpolating the potentially suppressed comparisons was $d = 0.85$ ($CI_{90\%}$ [0.51; 1.19], $p < .01$). Again, these analyses testify against a strong publication bias with respect to the effects of CSCL scripts on collaboration skills.

Discussion

General Findings

This meta-analysis was conducted to answer the question to which extent socio-cognitive scaffolding with CSCL scripts affects domain-specific knowledge as well as collaboration skills in CSCL environments compared to unstructured collaboration. Generally, the results of this meta-analysis support the hypothesis that collaborative learning with CSCL scripts has positive effects on the acquisition of domain-specific knowledge and collaboration skills compared to collaborative learning without CSCL scripts.

The results indicate more specifically that, on average, CSCL scripts have a small but statistically significant positive effect on learners' domain-specific knowledge. The mechanism is likely to be that CSCL scripts stimulate collaborative activities (e.g., explanation to a learning partner, peer feedback) that go hand in hand with high-level cognitive activities, which are beneficial for domain learning (Dillenbourg and Hong 2008; King 2007). Moreover, CSCL scripts have a large positive effect on collaboration skills. The likely mechanism here is that repeated participation in collaborative situations and the practice of collaborative activities leads to an internalization of the knowledge underlying performance in such situations. The script theory of guidance of Fischer et al. (2013) explains this effect as a consequence of the modification of learners' internal collaboration script components (Fischer et al. 2013).

These findings lead to the conclusion that CSCL scripts should be designed in such a way that they facilitate collaborative activities that are functional in a given type of collaborative situation. These activities, enacted in an increasingly self-directed way, may likewise become part of a learner's repertoire. Thus, the benefits of learning with CSCL scripts go beyond the mere support within specific collaborative learning settings designed to improve domain-specific knowledge acquisition. Supporting learners' collaborative activities is of even higher instrumentality with respect to the development of collaboration skills (see Demetriadis et al. 2011; Rummel et al. 2009).

Moderators of the Effectiveness of Learning with CSCL Scripts

There was, however, a substantial amount of heterogeneity among the individual effect sizes included in the meta-analysis. In addition to validating the choice of random-effects models for the synthesis of the individual effect sizes, this finding also warranted further exploration of potential moderators of the effectiveness of CSCL scripts. Nevertheless, the moderators did not yield any significant differences in the expected directions. Further implications are discussed in the following subsections.

Transactivity Although transactivity was not a statistically significant moderator in the expected direction, the results show a positive effect of CSCL scripts that prompt transactive activities on domain-specific knowledge, whereas no such effect could be established for CSCL scripts that do not prompt transactive activities (see Table 2). These results are in line with our hypothesis and with theoretical assumptions advocated in the research literature about transactive (Teasley 1997) or "interactive" (Chi 2009) activities as the most beneficial kind of activities in collaborative learning, but in lack of a statistically significant moderator effect they do not constitute strong evidence for these assumptions. Hence, more research is needed to directly test the validity of these assumptions for CSCL scripts.

The results also revealed a positive effect of medium size for CSCL scripts that prompt transactive activities on collaboration skills. However, the average effect size for CSCL scripts that do not prompt transactive activities was even larger (see Table 3). The so-called ICAP (Interactive, Constructive, Active, Passive) hypothesis claims that transactive activities are more beneficial for learning in general than are other kinds of activities in collaborative learning without differentiating between various learning outcomes (Chi 2009). In contrast, the present finding highlights that what is most effective with respect to one learning outcome is not necessarily most effective with respect to another learning outcome. A rather straightforward conclusion would be that the ICAP hypothesis does not hold as generally as originally stated by Chi (2009). Instead, the beneficial effects of transactive activities might be restricted to domain-specific knowledge. In fact, most of the studies included in Chi's original analysis focused on domain-specific knowledge and not on the development of collaboration skills. The development of collaboration skills might depend more on the amount of practice of the corresponding activities than on the transactivity of the interaction.

Script Level Also, the moderator *script level* was not statistically significant. With respect to domain-specific knowledge, only the subgroup with CSCL scripts mainly targeting the scene level had a statistically significant positive effect compared to unstructured CSCL, whereas CSCL scripts targeting mainly the play level or the scriptlet level did not have statistically significant effects (see Table 2). From a methodological point of view, the effects might not have become statistically significant due to unreliable estimators and a too small number of available effect sizes. As already noted, also the classification of CSCL scripts according to the script level can be problematic because even a script with a focus on one of the levels may provide more support on one of the other levels than some CSCL scripts

that have their main focus on this other level. To even complicate things, according to the SToG, the optimal script level to support a learner is also highly dependent on each learner's prior knowledge and skills (Fischer et al. 2013). Therefore, the optimal script level will be different for different learners. In absence of detailed information about the learners' prior knowledge and skills, the present operationalization was the closest approximation to the optimal script level possible given the available information from the studies included. Future research should focus specifically on the match of the script level to the prior internal scripts of the learners.

With respect to collaboration skills, the results further showed that CSCL scripts on each script level had consistently positive effects (see Table 2). Interestingly, there was no indication of an over-scripting effect (Dillenbourg 2002) with respect to collaboration skills; CSCL scripts that target the scriptlet level even had a statistically significant large positive effect on collaboration skills. One—of several possible—explanations for this finding might be that many learners in the studies included in the present meta-analysis may have had very low prior collaboration skills, which would make them good candidates for scriplet level CSCL scripts. Rather poor collaboration skills have indeed been reported in a number of studies on CSCL scripts as a justification for the introduction of CSCL scripts in the first place (e.g., Kollar et al. 2007; Stegmann, Weinberger, and Fischer 2007; Rummel, Spada, and Hauser 2009).

Content-Related Support There were also no statistically significant differences in effectiveness among the three levels of content-related support. However, the descriptive results are at least compatible with the hypothesis that additionally provided content-related support might alter the effects of CSCL scripts on domain-specific learning outcomes (Ertl et al. 2008): A significant mean effect of CSCL scripts on domain-specific knowledge was found only in the subgroup of studies in which learners were also supported with content-related scaffolds (see Table 2). In contrast, the estimated mean effects for studies in which learners collaborated without any additional content-related support or with content-related information only fell below the benchmark of a small effect and were not statistically significant (Cohen 1988). With additional content-related scaffolds, learners could acquire a certain amount of domain-specific knowledge. This might have increased the chance that learners elaborate on the learning content using the strategy prompted by the CSCL script, therefore acquiring even more domain-specific knowledge. As indicated before, further research is needed to substantiate this explanation for the observed descriptive pattern of results.

The hypothesis that additional content-related support would also have a positive impact on the effects of CSCL scripts on the development of collaboration skills (Sadler 2004) could not be fully tested in this meta-analysis. The effects of CSCL scripts were not found to differ significantly between the subgroup with content-related information only and the subgroup with content-related scaffolds. However, due to the small number of studies in the subgroup without content-related support, no mean effect could be estimated for this group of studies. Therefore, the results with respect to the role of the degree of domain-specific support for the development of collaboration skills remain inconclusive at this point.

Implications for Theory and Research

Overall, the results of this meta-analysis are compatible with the core theoretical assumptions of this paper (Chi 2009; Dillenbourg 2002; Fischer et al. 2013). For example, the finding that CSCL scripts that prompt transactive activities have a positive significant effect on domain-specific knowledge acquisition is in line with the *transactivity principle* in the script theory of guidance. This principle states that "the more a given CSCL practice requires the transactive application of knowledge, the better this knowledge is learned through participation in this CSCL practice" (Fischer et al. 2013, p. 58). However, learning with CSCL scripts that prompt transactive activities was not substantially effective for learning collaboration skills. Therefore, we would suggest a revision of the transactivity principle by differentiating between different types of learning outcomes.

Further, the findings regarding the script level as a moderator of the effectiveness of CSCL scripts for domain-specific learning do neither confirm nor contradict the *optimal external scripting level principle* from the script theory of guidance (Fischer et al. 2013, p. 62). This principle states that learners benefit most from collaborative learning when the CSCL script addresses a script level that is one level beyond the script level for which learners have already developed internal script components for the scripted collaboration (Fischer et al. 2013). Assuming that the scripts in the studies included were developed for learners with a medium level of prior collaboration skills (i.e., internal script components at least developed at the scriptlet level), the largest effect for CSCL scripts targeting on the scene level would be reasonable. Yet, this assumption cannot be validated because so few of the studies reported the learners' internal script levels. In light of the statistically non-significant moderator analysis, however, more research on these issues is needed to empirically investigate these issues discussed in the literature on CSCL scripts.

Additional content-related support might play a certain role, at least for the effectiveness of CSCL scripts on domain-specific knowledge. Therefore, perhaps the effect of additional content-related support should be included as a factor in theoretical assumptions about the effectiveness of learning with CSCL scripts. However, in the current meta-analysis, the level of additional content-related support varied only between and not within the individual studies (except for the study by Ertl et al. 2008). As a consequence, more empirical research would be needed to investigate the combined effect of CSCL scripts and content-related support in two-by-two factorial design studies.

Limitations

Some limitations of this meta-analysis need to be taken into account. The funnel plots suggest that publication bias might be a problem with respect to the effects on collaboration skills. The inclusion of unpublished studies might lead to a different estimate of the mean effect on collaboration skills (Begg 1994). However, the trim-and-fill method did not support the hypothesis that a substantial publication bias has occurred.

Furthermore, CSCL research involves many qualitative studies and case analyses that could not be included in this meta-analysis, even though they may yield

important insights for reaching a deeper understanding of the mechanisms underlying CSCL scripts (see, e.g., Hämäläinen and Arvaja 2009). In addition, a considerable number of quantitative studies on CSCL scripts could not be included in this meta-analysis because they did not manipulate CSCL scripts experimentally (e.g., Hummel et al. 2011) or because they did not report outcomes on one of the two target measures used in this meta-analysis (e.g., Yeh, Lo, and Huang 2011).

One specific shortcoming of the present meta-analysis has to be taken into account when considering the role of the transactivity induced by the CSCL script: The analysis of the role of transactivity for the effects of CSCL scripts capitalized on whether the CSCL scripts do or do not prompt transactive activities. Given the variables included and reported in the primary studies, it was not possible to take into account to what extent learners actually engaged in transactive activities during their collaboration. Learners may not necessarily perform the activities that the instructions request of them appropriately (Kirschner, Strijbos, Kreijns, and Beers 2004). Nevertheless, studies have shown that CSCL scripts can be highly effective in changing the collaboration process in the intended way (Stegmann et al. 2007; Weinberger et al. 2010). Therefore, for this meta-analysis, the information regarding whether transactive activities were prompted by the CSCL script is the best estimator for the transactive activities the learners actually were using while learning collaboratively.

Concerning the moderator *script level*, we already emphasized the problems (1) that, given the amount of missing information about the precise numbers of scaffolds on the different script levels, we could only roughly classify the CSCL scripts according to their main focus and (2) that, according to the SToG (Fischer et al. 2013), the optimal script level is also a function of each learner's prior knowledge and skill. To gain a deeper understanding of the optimal degree of support, future studies need to address this issue directly and provide detailed descriptions of the CSCL scripts employed and the relation of support to the learners' prior knowledge and skills. This would enable the authors of future meta-analyses to try a different approach.

Regarding the methodological decisions made for this meta-analysis, further limitations have to be taken into account. The database search was conducted with a narrow set of search terms that led to only 248 articles. A less narrow set of search terms might lead to a higher number of articles, which might include more studies about structured CSCL. Also, by the rather narrow criteria for inclusion, studies that were published in conference proceedings or in languages other than English were excluded from the meta-analysis. Nevertheless, the narrow criteria for inclusion were chosen deliberately to assure the quality of the included studies and to avoid problems that may result from journal publications that are based on previously published conference papers, which may make it difficult to assess whether two publications report data from the same or different samples.

In contrast, this meta-analysis is very inclusive regarding the amount of scaffolding a CSCL script must minimally provide to be included in this study. The amount of scaffolding might influence the effectiveness of a CSCL script. Yet, in many of the primary studies the descriptions of the CSCL scripts were too vague. Thus, it is very difficult to code a moderator variable based on these descriptions. For most studies, we do not know which scaffolds were provided how often, repeated how often, and so on.

As reported, the statistical power for detecting main effects on domain-specific knowledge acquisition and collaboration skills was very high. However, for the moderator analyses statistical power was rather insufficient due to the number of available studies on the effects of CSCL scripts. Thus, more research is needed to obtain a more reliable picture of the effects of the different moderators we investigated.

Conclusions

The research summarized in this meta-analysis provides evidence for the effectiveness of CSCL scripts as a form of socio-cognitive scaffolding with respect to both domain-specific knowledge ($d = 0.20$) and collaboration skills ($d = 0.65$), with more pronounced effects for the latter. This meta-analysis shows some initial support for the role the investigated moderators play for the effectiveness of CSCL scripts. Future research should focus on experimental manipulations of the moderator variables. In particular, because the optimal script level is likely to be highly dependent on the learners' prior internal scripts (Fischer et al. 2013), future research should track the effectiveness of CSCL scripts at different script levels in relation to the learners' prior internal scripts (Kollar et al. 2007; Rienties et al. 2012). Furthermore, the studies included in this meta-analysis only rarely report data on the actual degree of transactivity of the learners' collaborative activities (Jeong and Joung 2007; Noroozi et al. 2013). To advance research about CSCL scripts, transactivity should routinely be included as a standard process variable, measured by means of a common procedure, and reported in publications.

With respect to practice, facilitating learners with socio-cognitive scaffolding by CSCL scripts that support their (computer-supported) collaboration activities increases the likelihood of domain-specific learning. Designing CSCL scripts in a way that facilitates the use of transactive activities seems to be promising for the acquisition of domain-specific knowledge. Yet, maybe due to the low statistical power, a significant moderator effect of transactivity could not be found. Also, the combination of CSCL scripts and additional content-related support might increase the effectiveness with respect to domain-specific knowledge, although currently there is only limited evidence in this direction. Furthermore, a CSCL script might be effective for domain-specific knowledge if it targets the scene level, although also in this case the available evidence is still weak. In contrast, when the only instructional aim is to foster learners' collaboration skills, the scriptlet level might be the most promising target for a CSCL script. In particular, if the only instructional goal is to foster collaboration skills, the use of CSCL scripts can be highly recommended to educators.

Compliance with Ethical Standards

Conflict of Interest The authors declare that they have no conflict of interest.

Appendix

Table 4 Studies and corresponding values of variables and effect sizes

Study comparison	Moderators			Subjects		Unit	Dependent variable	Combined effect size			Single effect size
	Transactivity	Script level of scaffolds	Content-specific support	n control	n exp.			d	V_d	90 % CI	d_{single}
Demetriadis, Egerter, Hanisch, and Fischer (2011)	Prompted	Scene level	Content scaffold	29	34	Individual	Domain-specific knowledge	0.01	0.06	[−0.40; 0.43]	
							Collaboration skills	0.81	0.07	[0.38; 1.24]	
Ertl, Fischer, and Mandl (2006)											
With conceptual support	Prompted	Scriptlet level	Content scaffold	10	10	Individual	Domain-specific knowledge	0.21	0.11	[−0.35; 0.76]	0.54
											−0.13
Without conceptual support	Prompted	Scene level	Content information only	12	11	Individual	Domain-specific knowledge	0.42	0.08	[−0.06; 0.89]	0.61
											0.22
Ertl, Kopp, and Mandl (2008)											
With content scheme	Prompted	Scriptlet level	Content scaffold	13	13	Small group	Domain-specific knowledge	0.41	0.06	[0.02; 0.80]	0.75
											0.07
											0.67
											0.92
											−0.14
											−0.15
											0.75

Table 4 (continued)

Study comparison	Moderators			Subjects		Unit	Dependent variable	Combined effect size			Single effect size
	Transactivity	Script level of scaffolds	Content-specific support	n control	n exp.			d	V_d	90 % CI	d_{single}
Without content scheme	Prompted	play level	Content information only	14	13	Small group	Domain-specific knowledge	−0.21	0.05	[−0.58; 0.16]	−0.29 / 0.03 / −0.31 / 0.30 / −0.77 / −0.37 / −0.06
Gelmini-Hornsby, Ainsworth, and O'Malley (2011)	Prompted	Play level	No content support	9	9	Small group	Collaboration skills	0.61	0.09	[0.10; 1.12]	1.10 / 0.45 / 0.00 / 0.90
Gijlers and de Jong (2009)											
Shared scratchpad vs. control	Prompted	Scene level	Content scaffold	22	22	Individual	Domain-specific knowledge	−0.29	0.06	[−0.68; 0.10]	−0.66 / 0.08
Shared table vs. control	Prompted	Scene level	Content scaffold	22	22	Individual	Domain-specific knowledge	0.45	0.06	[0.05; 0.83]	0.33 / 0.56
Haake and Pfister (2010)	Prompted	Play level	No content support	21	21	Small group	Domain-specific knowledge	−0.52	0.09	[−1.03; −0.02]	
Huang, Wu, and Chen (2012)	Prompted	Scriptlet level	Content scaffold	30	30	Individual	Domain-specific knowledge	0.79	0.07	[0.35; 1.23]	
Kollar, Fischer, and Slotta (2007) *Low prior knowledge*	Prompted	Scriptlet level	Content scaffold	20	22	Individual	Domain-specific knowledge	0.04	0.09	[−0.46; 0.54]	

Table 4 (continued)

Study comparison	Moderators			Subjects		Unit	Dependent variable	Combined effect size			Single effect size
	Transactivity	Script level of scaffolds	Content-specific support	n control	n exp.			d	V_d	90 % CI	d_{single}
High prior knowledge							Collaboration skills	0.41	0.09	[−0.10; 0.91]	
	Prompted	Scriptlet level	Content scaffold	22	26	Individual	Domain-specific knowledge	0.06	0.08	[−0.41; 0.53]	
							Collaboration skills	0.86	0.09	[0.37; 1.35]	
Molinari, Sangin, Dillenbourg, and Nüssli (2009)	Not prompted	Scriptlet level	Content information only	32	28	Individual	Domain-specific knowledge	−0.25	0.07	[−0.67; 0.18]	
Noroozi, Teasley, Biemans, Weinberger, and Mulder (2013)											
Combined transactive discussion and memory script	Prompted	Scene level	No content support	30	30	Individual[a]	Domain-specific knowledge	−0.64	0.04	[−0.99; −0.29]	−0.57 / −0.55 / −0.80
Transactive discussion script	Prompted	Scene level	No content support	30	30	Individual[a]	Domain-specific knowledge	1.10	0.05	[0.74; 1.48]	1.19 / 1.20 / 0.92
Transactive memory script	Prompted	Scene level	No content support	30	30	Individual[a]	Domain-specific knowledge	1.31	0.05	[−0.88; −0.31]	1.54 / 1.60 / 0.80
Rienties, Giesbers, Tempelaar, Lygo–Baker, Segers, and Gijselaers (2012)	Prompted	Scriptlet level	No content support	82	61	Individual	Domain-specific knowledge	−0.59	0.03	[−0.88; −0.31]	
Rummel and Spada (2005)	Prompted	Play level	Content information only	9	9	Small group	Domain-specific knowledge	0.25	0.10	[−0.30; 0.77]	1.42 / −1.66 / 0.98

Table 4 (continued)

Study comparison	Moderators			Subjects		Unit	Dependent variable	Combined effect size			Single effect size
	Transactivity	Script level of scaffolds	Content-specific support	n control	n exp.			d	V_d	90 % CI	d_{single}
Rummel, Mullins, and Spada (2012)	Prompted	Scene level	No content support	18	18	Individual	Collaboration skills	0.63	0.11	[0.08; 1.18]	−1.95
											−0.98
											−0.27
											0.50
											2.00
											2.64
				39	38	Individual	Domain-specific knowledge	0.32	0.03	[0.01; 0.62]	
Rummel, Spada, and Hauser (2009)											
Script plus elaboration support	Prompted	Scene level	Content information only	8	8	Small group	Domain-specific knowledge	−0.16	0.22	[−0.94; 0.62]	
				8	8	Small[b] group	Collaboration skills	0.28	0.08	[−0.18; 0.73]	−0.36
											0.42
											−0.27
											−0.66
											0.90
											0.24
											−0.24
											−0.16
											1.01
											0.62
											1.55
Script only	Prompted	Scene level	Content information only	8	8	Small group	Domain-specific knowledge	−0.48	0.23	[−1.28; 0.31]	–

Table 4 (continued)

Study comparison	Moderators			Subjects		Unit	Dependent variable	Combined effect size			Single effect size
	Transactivity	Script level of scaffolds	Content-specific support	n control	n exp.			d	V_d	90 % CI	d_{single}
							Collaboration skills	0.43	0.08	[−0.02; 0.89]	−0.25
											0.95
											−0.25
											−0.25
											0.83
											0.39
											0.26
											0.00
											1.06
											0.66
											1.34
Schellens, Van Keer, De Wever, and Valcke (2007)	Prompted	Scene level	Content information only	223	286	Individual	Domain-specific knowledge	0.38	0.01	[0.23; 0.53]	
Stegmann, Wecker, Weinberger, and Fischer (2012)	Prompted	Scene level	Content information only	24	24	Individual	Domain-specific knowledge	0.66	0.09	[0.18; 1.15]	
Stegmann, Weinberger, and Fischer (2007)							Collaboration skills	2.59	0.15	[1.95; 3.23]	
Combined script for single arguments and argumentation sequences	Prompted	Scene level	Content information only	10	10	Individual	Domain-specific knowledge	0.30	0.19	[−0.41; 1.01]	
							Collaboration skills	0.99	0.16	[0.35; 1.67]	1.29
											0.68

Table 4 (continued)

Study comparison	Moderators			Subjects		Unit	Dependent variable	Combined effect size			Single effect size
	Transactivity	Script level of scaffolds	Content-specific support	n control	n exp.			d	V_d	90 % CI	d_{single}
Script for argumentation sequences	Prompted	Scene level	Content information only	10	10	Individual	Domain-specific knowledge	0.71	0.20	[−0.02; 1.44]	0.11
							Collaboration skills	0.67	0.15	[0.04; 1.33]	1.23
Script for single arguments	Prompted	Scriptlet level	Content information only	10	10	Individual	Domain-specific knowledge	0.66	0.19	[−0.07; 1.38]	1.37
							collaboration skills	1.39	0.17	[0.60;1.95]	1.40
Ulicsak (2004)	Prompted	Play level	No content support	25	26	Individual	Domain-specific knowledge	0.23	0.03	[−0.08; 0.53]	0.13
											0.21
											0.41
											0.16
Van Aalst and Chan (2007)											
Knowledge building principles	Prompted	Scriptlet level	No content support	9	14	Individual	Domain-specific knowledge	1.27	0.13	[0.67; 1.85]	1.31
											1.22
Principles support	Prompted	Scriptlet level	No content support	23	24	Individual	Domain-specific knowledge	0.27	0.05	[−0.11; 0.65]	0.68
											−0.14
Weinberger, Ertl, Fischer, and Mandl (2005)											
Epistemic script	Not prompted	Scene level	Content information only	8	8	Small group	Domain-specific knowledge	−0.69	0.24	[−1.49; 0.12]	

Table 4 (continued)

Study comparison	Moderators			Subjects		Unit	Dependent variable	Combined effect size			Single effect size
	Transactivity	Script level of scaffolds	Content-specific support	n control	n exp.			d	V_d	90 % CI	d_{single}
Combined epistemic and social script	Prompted	Scriptlet level	Content information only	8	8	Small group	Domain-specific knowledge	−0.38	0.23	[−1.16; 0.41]	
social script	Prompted	Scriptlet level	Content information only	8	8	Small group	Domain-specific knowledge	0.78	0.24	[−0.04; 1.59]	
Weinberger, Stegmann, and Fischer (2010)	Not prompted	Scriptlet level	Content information only	9	9	Individual	Domain-specific knowledge	0.91	0.22	[0.13; 1.70]	
	Not prompted	Scriptlet level	Content information only	9	9	Individual	Collaboration skills	3.17	0.48	[2.02; 4.31]	

[a] The data for domain-specific knowledge were derived from individual outcomes that were included in a value for shared knowledge

[b] The data for domain-specific knowledge were derived from small group outcomes and individual outcomes averaged for the small groups

References

References marked with an asterisk indicate studies included in the meta-analysis

Anderson, J. R. (1992). Automaticity and the ACT theory. *The American Journal of Psychology, 105*(2), 165–180. doi:10.2307/1423026.

Andriessen, J., Baker, M., & Suthers, D. (2003). Argumentation, computer support, and the educational context of confronting cognitions. In J. Andriessen, M. Baker, & D. Suthers (Eds.), *Arguing to learn: Confronting cognitions in computer-supported collaborative learning environments* (pp. 1–25). Dordrecht: Kluwer.

Begg, C. B. (1994). Publication bias. In H. M. Cooper & L. V. Hedges (Eds.), *The handbook of research synthesis* (pp. 399–409). New York, NY: Sage.

Borenstein, M., Hedges, L. V., Higgins, J. P. T., & Rothstein, H. R. (2009). *Introduction to meta-analysis*. Chichester, UK: Wiley.

Chi, M. T. H. (2009). Active-constructive-interactive: A conceptual framework for differentiating learning activities. *Topics in Cognitive Science, 1*(1), 73–105. doi:10.1111/j.1756-8765.2008.01005.x.

Chi, M. T. H., & Wylie, R. (2014). The ICAP framework: Linking cognitive engagement to active learning outcomes. *Educational Psychologist, 49*(4), 219–243. doi:10.1080/00461520.2014.965823.

Clark, D. B., & Sampson, V. D. (2007). Personally-seeded discussions to scaffold online argumentation. *International Journal of Science Education, 29*(3), 253–277. doi:10.1080/09500690600560944.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Erlbaum.

Cohen, G. C. (1994). Restructuring the classroom: Conditions for productive small groups. *Review of Educational Research, 64*(1), 1–35. doi:10.3102/00346543064001001.

De Wever, B., Van Keer, H., Schellens, T., & Valcke, M. (2010). Structuring asynchronous discussion groups: Comparing scripting by assigning roles with regulation by cross-age peer tutors. *Learning and Instruction, 20*(5), 349–360. doi:10.1016/j.learninstruc.2009.03.001.

*Demetriadis, S., Egerter, T., Hanisch, F., & Fischer, F. (2011). Peer review-based scripted collaboration to support domain-specific and domain-general knowledge acquisition in computer science. Computer Science Education, 21(1), 29–56. doi:10.1080/08993408.2010.539069

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

Dillenbourg, P., & Hong, F. (2008). The mechanics of CSCL macro scripts. *International Journal of Computer-Supported Collaborative Learning, 3*(1), 5–23. doi:10.1007/s11412-007-9033-1.

Duval, S., & Tweedie, R. (2000a). A nonparametric "trim and fill" method of accounting for publication bias in metaanalysis. *Journal of the American Statistical Association, 95*(449), 89–98. doi:10.1080/01621459.2000.1047390.

Duval, S., & Tweedie, R. (2000b). Trim and fill: A simple funnel-plot-based method of testing and adjusting for publication bias in meta-analysis. *Biometrics, 56*(2), 455–463. doi:10.1111/j.0006-341X.2000.00455.x.

*Ertl, B., Fischer, F., & Mandl, H. (2006). Conceptual and socio-cognitive support for collaborative learning in videoconferencing environments. Computers & Education, 47(3), 298–315. doi:10.1016/j.compedu.2004.11.001

*Ertl, B., Kopp, B., & Mandl, H. (2008). Supporting learning using external representations. Computers & Education, 51(4), 1599–1608. doi:10.1016/j.compedu.2008.03.001

Fischer, F., Kollar, I., Stegmann, K., & Wecker, C. (2013). Toward a script theory of guidance in computersupported collaborative learning. *Educational Psychologist, 48*(1), 56–66. doi:10.1080/00461520.2012.748005.

*Gelmini-Hornsby, G., Ainsworth, S., & O'Malley, C. (2011). Guided reciprocal questioning to support children's collaborative storytelling. International Journal of Computer-Supported Collaborative Learning, 6(4), 577–600. doi:10.1007/s11412-011-9129-5

*Gijlers, H., & de Jong, T. (2009). Sharing and confronting propositions in collaborative inquiry learning. Cognition and Instruction, 27(3), 239–268. doi:10.1080/07370000903014352

Greenhouse, J. B., & Iyengar, S. (1994). Sensitivity analysis and diagnostics. In H. M. Cooper & L. V. Hedges (Eds.), *The handbook of research synthesis* (pp. 383–398). New York, NY: Sage.

*Haake, J. M., & Pfister, H.-R. (2010). Scripting a distance-learning university course: Do students benefit from net-based scripted collaboration? International Journal of Computer-Supported Collaborative Learning, 5(2), 191– 210. doi:10.1007/s11412-010-9083-7

Hämäläinen, R., & Arvaja, M. (2009). Scripted collaboration and group-based variations in a higher education CSCL context. *Scandinavian Journal of Educational Research, 53*(1), 1–16. doi:10.1080/00313830802628281.

Hedges, L. (1981). Distribution theory for Glass's estimator of effect size and related estimators. *Journal of Educational Statistics, 6*(2), 107–128. doi:10.2307/1164588.

Hedges, L., & Olkin, I. (1985). *Statistical methods for meta-analysis*. Orlando, FL: Academic Press.

Howard, B. C. (1996, February). A meta-analysis of scripted cooperative learning. Paper presented at the Annual Meeting of the Eastern Educational Research Association, Boston, MA. Retrieved from http://files.eric.ed.gov/fulltext/ED404353.pdf

*Huang, H.-W., Wu, C.-W., & Chen, N.-S. (2012). The effectiveness of using procedural scaffoldings in a paperplus- smartphone collaborative learning context. Computers & Education, 59(2), 250–259. doi:10.1016/j.compedu.2012.01.015

Hummel, H. G. K., van Houcke, J., Nadolski, R. J., van der Hiele, T., Kurvers, H., & Löhr, A. (2011). Scripted collaboration in serious gaming for complex learning: Effects of multiple perspectives when acquiring water management skills. *British Journal of Educational Technology, 42*(6), 1029–1041. doi:10.1111/j.1467-8535.2010.01122.x.

Ismail, H. N., & Alexander, J. M. (2005). Learning within scripted and nonscripted peer-tutoring sessions: The Malaysian context. *The Journal of Educational Research, 99*(2), 67–77. doi:10.3200/JOER.99.2.67-77.

Jeong, A., & Joung, S. (2007). Scaffolding collaborative argumentation in asynchronous discussions with message constraints and message labels. *Computers & Education, 48*(3), 427–445. doi:10.1016/j.compedu.2005.02.002.

King, A. (1992). Facilitating elaborative learning through guided student-generated questioning. *Educational Psychologist, 27*(1), 111–126. doi:10.1207/s15326985ep2701_8.

King, A. (1998). Transactive peer tutoring: Distributing cognition and metacognition. *Educational Psychology Review, 10*(1), 57–74. doi:10.1023/A:1022858115001.

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.

Kirschner, P. A., Strijbos, J. W., Kreijns, K., & Beers, P. (2004). Designing electronic collaborative learning environments. *Educational Technology Research & Development, 52*(3), 47–66. doi:10.1007/BF02504675.

Kobbe, L., Weinberger, A., Dillenbourg, P., Harrer, A., Hämäläinen, R., Häkkinen, P., & Fischer, F. (2007). Specifying computer-supported collaboration scripts. *International Journal of Computer-Supported Collaborative Learning, 2*(2–3), 211–224. doi:10.1007/s11412-007-9014-4.

Kollar, I., Fischer, F., & Hesse, F. W. (2006). Collaboration scripts—a conceptual analysis. *Educational Psychology Review, 18*(2), 159–185. doi:10.1007/s10648-006-9007-2.

*Kollar, I., Fischer, F., & Slotta, J. D. (2007). Internal and external scripts in computer-supported collaborative inquiry learning. Learning and Instruction, 17(6), 708–721. doi:10.1016/j.learninstruc.2007.09.021

Koschmann, T. (1996). Paradigm shifts and instructional technology: An introduction. In T. Koschmann (Ed.), *CSCL: Theory and practice of an emerging paradigm* (pp. 1–26). Mahwah, NJ: Erlbaum.

Kuhn, D., Shaw, V., & Felton, M. (1997). Effects of dyadic interaction on argumentative reasoning. *Cognition and Instruction, 15*(3), 103–121. doi:10.1207/s1532690xci1503_1.

Kwon, S. Y., & Cifuentes, L. (2009). The comparative effect of individually-constructed vs. collaboratively-constructed computer-based concept maps. *Computers & Education, 52*(2), 365–375. doi:10.1016/j.compedu.2008.09.012.

*Mäkitalo, K., Weinberger, A., Häkkinen, P., Järvelä, S., & Fischer, F. (2005). Epistemic cooperation scripts in online learning environments: Fostering learning by reducing uncertainty in discourse? Computers in Human Behavior, 21(4), 603–622. doi:10.1016/j.chb.2004.10.033

*Molinari, G., Sangin, M., Dillenbourg, P., & Nüssli, M.-A. (2009). Knowledge interdependence with the partner, accuracy of mutual knowledge model and computer-supported collaborative learning. European Journal of Psychology of Education, 24(2), 129–144. doi:10.1007/BF03173006

Mugny, G., & Doise, W. (1978). Socio-cognitive conflict and structure of individual and collective performances. *European Journal of Social Psychology, 8*(2), 181–192. doi:10.1002/ejsp.2420080204.

Noroozi, O., Weinberger, A., Biemans, H. J. A., Mulder, M., & Chizari, M. (2012). Argumentation-based computer supported collaborative learning (ABCSCL): A systematic review and synthesis of fifteen years of research. *Educational Research Review, 7*(2), 79–106. doi:10.1016/j.edurev.2011.11.006.

*Noroozi, O., Teasley, S. D., Biemans, H. J. A., Weinberger, A., & Mulder, M. (2013). Facilitating learning in multidisciplinary groups with transactive CSCL scripts. International Journal of Computer-Supported Collaborative Learning, 8(2), 189–223. doi:10.1007/s11412-012-9162-z

Orwin, R. G. (1994). Evaluating coding decisions. In H. Cooper & L. V. Hedges (Eds.), *The handbook of research synthesis* (pp. 139–162). New York: Sage.

Raudenbush, S. W. (1994). Random effects models. In H. Cooper & L. V. Hedges (Eds.), *The handbook of research synthesis* (pp. 301–321). New York: Sage.

*Rienties, B., Giesbers, B., Tempelaar, D., Lygo-Baker, S., Segers, M., & Gijselaers, W. (2012). The role of scaffolding and motivation in CSCL. Computers & Education, 59(3), 893–906. doi:10.1016/j.compedu. 2012.04.010

*Rummel, N., & Spada, H. (2005). Learning to collaborate: An instructional approach to promoting collaborative problem solving in computer-mediated settings. The Journal of the Learning Sciences, 14(2), 201–241. doi: 10.1207/s15327809jls1402_2

*Rummel, N., Spada, H., & Hauser, S. (2009). Learning to collaborate while being scripted or by observing a model. International Journal of Computer-Supported Collaborative Learning, 4(1), 69–92. doi:10.1007/ s11412-008-9054-4

*Rummel, N., Mullins, D., & Spada, H. (2012). Scripted collaborative learning with the Cognitive Tutor Algebra. International Journal of Computer-Supported Collaborative Learning, 7(2), 307–339. doi:10. 1007/s11412-012-9146-z

Sadler, T. D. (2004). Informal reasoning regarding socioscientific issues: A critical review of research. Journal of Research in Science Teaching, 41(5), 513–536. doi:10.1002/tea.20009.

Schank, R. C. (1999). Dynamic memory revisited. Cambridge: Cambridge University Press.

*Schellens, T., Van Keer, H., De Wever, B., & Valcke, M. (2007). Scripting by assigning roles: Does it improve knowledge construction in asynchronous discussion groups? International Journal of Computer-Supported Collaborative Learning, 2(2–3), 225–246. doi:10.1007/s11412-007-9016-2

Schwarz, B., & Linchevski, L. (2007). The role of task design and argumentation in cognitive development during peer interaction: The case of proportional reasoning. Learning and Instruction, 17(5), 510–531. doi: 10.1016/j.learninstruc.2007.09.009.

*Stegmann, K., Weinberger, A., & Fischer, F. (2007). Facilitating argumentative knowledge construction with computer-supported collaboration scripts. Computer-Supported Collaborative Learning, 2(4), 421–447. doi: 10.1007/s11412-007-9028-y

Stegmann, K., Mu, J., Gehlen-Baum, V., & Fischer, F. (2011). The myth of over-scripting: Can novices be supported too much? In H. Spada, G. Stahl, N. Miyake, & N. Law (Eds.), CSCL2011 conference proceedings. Vol. 1: Long papers. Connecting computer-supported collaborative learning to policy and practice (pp. 406–413). International Society of the Learning Sciences.

*Stegmann, K., Wecker, C., Weinberger, A., & Fischer, F. (2012). Collaborative argumentation and cognitive elaboration in a computer-supported collaborative learning environment. Instructional Science, 40(2), 297–323. doi:10.1007/s11251-011-9174-5

Tabak, I. (2004). Synergy: A complement to emerging patterns of distributed scaffolding. The Journal of the Learning Sciences, 13(3), 305–335. doi:10.1207/s15327809jls1303_3.

Teasley, S. D. (1997). Talking about reasoning: How important is the peer in peer collaborations? In L. B. Resnick, R. Saljo, C. Pontecorvo, & B. Burge (Eds.), Discourse, Tools, and Reasoning: Situated Cognition and Technologically Supported Environments (pp. 361–384). Berlin: Springer-Verlag.

*Ulicsak, M. H. (2004). "How did it know we weren't talking?" An investigation into the impact of self-assessments and feedback in a group activity. Journal of Computer Assisted Learning, 20(3), 205–211. doi: 10.1111/j.1365-2729.2004.00083.x

*Van Aalst, J., & Chan, C. K. K. (2007). Student-directed assessment of knowledge building using electronic portfolios. The Journal of the Learning Sciences, 16(2), 175–220. doi:10.1080/10508400701193697

Webb, N. M., Franke, M. L., De, T., Chan, A. G., Freund, D., Shein, P., & Melkonian, D. K. (2009). "Explain to your partner": Teachers' instructional practices and students' dialogue in small groups. Cambridge Journal of Education, 39(1), 49–70. doi:10.1080/03057640802701986.

Wecker, C., & Fischer, F. (2011). From guided to self-regulated performance of domain-general skills: The role of peer monitoring during the fading of instructional scripts. Learning & Instruction, 21(6), 746–756. doi:10. 1016/j.learninstruc.2011.05.001.

*Weinberger, A., Ertl, B., Fischer, F., & Mandl, H. (2005). Epistemic and social scripts in computer-supported collaborative learning. Instructional Science, 33(1), 1–30. doi:10.1007/s11251-004-2322-4

*Weinberger, A., Stegmann, K., & Fischer, F. (2010). Learning to argue online: Scripted groups surpass individuals (unscripted groups do not). Computers in Human Behavior, 26(4), 506–515. doi:10.1016/j. chb.2009.08.007

Yeh, S. W., Lo, J. J., & Huang, J. J. (2011). Scaffolding collaborative technical writing with procedural facilitation and synchronous discussion. International Journal of Computer-Supported Collaborative Learning, 6(3), 397–419. doi:10.1007/s11412-011-9117-9.