EUR Research Information Portal

'Supporting socially shared regulation during collaborative task-oriented reading'

Erasmus University Rotterdam
Making Minds Matter

'Supporting socially shared regulation during collaborative task-oriented reading'

Jolique Kielstra[1] · Inge Molenaar[1] · Roel van Steensel[2,3] · Ludo Verhoeven[1]

Abstract

This study examined how to improve students' regulation of task-oriented reading (TOR). TOR encompasses reading and information processing needed to perform a specific task. Previous studies suggest students can benefit from a collaboration script to enhance socially shared regulation of TOR. The collaboration script elicits discussions about task perception, strategy selection, and strategy reflection. This study aimed to examine the depth and socially sharedness of metacognitive regulation when working with a collaboration script among 44 prevocational secondary school students working in groups of four. In addition, we examined the consequent improvement of individual task representation, strategy selection, and strategy reflection after working with the script. The analysis of group discussions indicated that the collaboration script facilitated mainly low-level metacognitive regulation of TOR. However, after working with the script, students did improve their ability to determine a correct representation of a high-level task and to reflect on the most appropriate reading strategy for these tasks. Hence, we concluded that the 'Y-read?' collaboration script did elicit shared regulation during TOR.

Keywords Task-oriented reading · Collaboration · Shared regulated learning · External collaboration script · Secondary school

✉ Jolique Kielstra
jolique.kielstra@ru.nl

Inge Molenaar
inge.molenaar@ru.nl

Roel van Steensel
vansteensel@essb.eur.nl

Ludo Verhoeven
ludo.verhoeven@ru.nl

1   Behavioural Science Institute, Radboud University Nijmegen, 9104, 6500 HE Nijmegen, The Netherlands

2   Faculty of Humanities, Department of Language, Literature, & Communication, Free University, Amsterdam, The Netherlands

3   Department of Psychology, Erasmus School of Social and Behavioural Sciences, Education & Child Studies, Rotterdam, The Netherlands

Introduction

In secondary schools, students often engage in task-oriented reading (TOR): reading with the purpose of processing information for the execution of a specific task (Anmarkrud et al., 2013; Vidal-Abarca et al., 2010). These tasks and the textual information that is relevant to fulfil them may differ. For example, while one task may require a student to retrieve details from a text, another might require a student to integrate information from different parts of the text. Students are taught diverse reading strategies for coping with these different tasks (Okkinga et al., 2021). However, even when students have knowledge about these reading strategies, they might still have difficulty selecting and applying appropriate strategies when encountering different tasks. This appears to be particularly true for students with less advanced reading skills, who are concentrated in prevocational education in the Netherlands (De Milliano et al., 2016; Okkinga et al., 2021; Salmerón et al., 2016). Therefore, this study aims to investigate how these students can be provided with support in learning how to select and apply appropriate reading strategies when confronted with different tasks.

Task complexity and strategy use

In TOR, students must make various decisions regarding what to read and how to process the text in order to complete the task (Gil et al., 2015; Serrano et al., 2018; Vidal-Abarca et al., 2010). To do so, students have to be aware of differences in the nature of tasks. One way of categorizing tasks is by means of taxonomies that indicate task complexity. These taxonomies are generally based on two variables: (1) whether relevant information can be found on a local or global level and (2) the required level of inferencing. The combination of these variables results in task types that are sometimes categorized as 'low-level', 'intermediate-level', or 'high-level' (OECD, 2003; Rouet et al., 2001; Van Steensel et al., 2013; Vidal-Abarca et al., 1998). Low-level tasks focus on retrieving explicit details from the text with little to no inferencing. Intermediate-level tasks require students to make local inferences and usually lead to a 'targeted strategy', during which initial browsing transforms into careful reading once the location of the required information is identified. Finally, high-level tasks require global processing and usually require an 'intensive reading' strategy with extensive inferencing.

Self-regulation in TOR

To effectively apply different reading strategies, students have to be aware of task demands and the consequences these demands have for selecting an appropriate reading strategy. To make such decisions, students need to engage in self-regulated learning (De Milliano et al., 2016; Rouet et al., 2017). Self-regulated learning refers to the ability to actively steer and adapt one's learning to meet learning goals and overcome challenges (Winne & Hadwin, 1998). While self-regulation encompasses cognitive, behavioural, motivational, affective and metacognitive components, for the present study we mainly focus on students' ability to engage in metacognitive regulation to coordinate and control cognitive learning in response to task demands during TOR. In this context, it is important that students have sufficient knowledge about task demands and reading strategies to engage in metacognitive control and monitoring of their reading (Winne & Hadwin, 1998). This implies being able to select an appropriate reading strategy to fulfil the task at hand (De

Backer et al., 2016). In general, models of self-regulated learning suggest three regulation phases: the pre-task preparation phase (i.e., orientation and planning), the execution phase (i.e., monitoring), and the post-task reflection phase (i.e., evaluation), also referred to as the forethought, performance and self-reflection phases (Zimmerman, 2008). During pre-task preparation, students orient themselves to characteristics of the task and text. Accordingly, students create a representation of the task based on these characteristics (Rouet et al., 2017; Winne & Hadwin, 1998). This ideally involves predicting whether information required to fulfil the task needs to be obtained on a local or global level and assessing the level of inferencing needed to process this information (Van Steensel et al., 2013). Students then use this task representation to plan and select an appropriate reading strategy. During task execution, students monitor their progress towards successful task execution in relation to the chosen strategy (Vidal-Abarca et al., 2010). This helps them to determine whether the selected strategy is appropriate for the task or whether it needs to be adapted (Panadero, 2017). Finally, during post-task reflection, students evaluate whether their perception of the task was correct in hindsight and whether their reading strategy of choice was effective and efficient. These evaluations help students gain insight into how reading strategies can be applied in future tasks and improve their TOR (Rouet et al., 2017; Winne & Hadwin, 1998). However, as mentioned before, prevocational students experience difficulty in adapting their reading strategy to task demands and therefore are in need of support to effectively regulate TOR (De Milliano et al., 2016). In this study, we propose to let learners collaborate and also support them with a collaboration script to develop the needed skills to regulate TOR.

Collaborative learning, social and self-regulation

Collaborative learning can support the development of learners' metacognitive knowledge (Molenaar et al., 2011, 2014). When learners engage in collaborative learning, they are required to verbalize metacognitive regulation, in the context of TOR, their task perception, strategy selection and strategy reflection to group members. Through explicating and discussing, students become aware of how they develop a task perception, select a strategy, reflect on their strategy use, and receive feedback from group members on their approach. As such, students influence each other in a spiral-like fashion: each student contributes knowledge to the social system, altering the interpersonal plan during collaboration and consequently eliciting new activities from the group members. In this way, individual students appropriate metacognitive knowledge provided by the social system and, in turn, contribute with their enhanced participation in the development of the social system (Volet et al., 2009). Such articulation of metacognitive knowledge has been associated with improved socially shared regulation (De Backer et al., 2016) and, consequently, individual regulation (Molenaar et al., 2014). Socially shared regulated learning (SSRL) refers to instances in which students jointly monitor and control their learning during collaborative learning (De Backer et al., 2022; Iiskala et al., 2021; Järvelä & Hadwin, 2015). During collaboration, increased transactivity can improve students' individual knowledge through shared knowledge construction (Fisher et al., 2013; Teasley, 1997).

Transactive interaction facilitates SSRL in two ways: (1) students contribute deep-level metacognitive activities (depth), and (2) students build on and elaborate on other group members' earlier metacognitive contributions (sharedness). Generally, low-and deep-level metacognitive regulation are distinguished (De Backer et al., 2016, 2022). Low-level metacognitive regulation is characterized by examining surface-level text and task

characteristics. For example, during the pre-task preparation phase, low-level metacognitive regulation is the discussion of task demands, i.e., *"this is a simple question"*, whereas deep-level orientation discusses task demands and relates them to reading strategies, i.e., *"for a simple question we only need to search for the answer in the text"* (Butler, 2002; De Backer et al., 2016). Hence contributions of deep-level metacognitive regulation are more profound contributions about the nature of text and task in which a connection is made with the location of relevant information in the text. These deep-level contributions have the potential to foster SSRL and consequently translate into individual knowledge on how to use metacognitive regulation during TOR (De Backer et al., 2016). Previous research indicates that the depth of a student's metacognitive contributions is predictive of the success of socially shared regulation and individual regulation of TOR (Hämäläinen et al., 2008; Iiskala et al., 2021).

Following the assumption of transactivity, students have to respond to, and elaborate on, group members' contributions (Fisher et al., 2013). This sharedness of metacognitive regulation refers to the extent students engage in and relate to each other's metacognitive activities (Molenaar et al., 2014). Previous research distinguished different levels of sharedness in student metacognitive regulation: students can ignore each other's metacognitive activities, accept each other's metacognitive contributions without further demands for justification, share their existing knowledge without translating it into meaningful metacognitive activities or engage in the co-construction of metacognitive activities (Molenaar et al., 2014). When students co-construct, they elaborate on others' metacognitive contributions, explaining and questioning each other's thinking and providing feedback (Järvelä & Hadwin, 2015). In this type of interaction, students formulate contributions that individual group members could not have generated themselves (Hämäläinen et al., 2008). Particularly this latter type of interaction fosters meaningful SSRL and potentially translates into improved individual regulation of TOR (Molenaar et al., 2014). However, previous research indicates when pre-vocational secondary education students collaborate, they generally engage in no or low metacognition regulation of TOR (Okkinga et al., 2021). These students need guidance while collaborating, this study uses a script to support students in discussing task and text characteristics, and to connect these characteristics with reading strategies.

Scripting socially shared regulation of TOR

It is well known that transactivity is rare during collaborative learning, and external scripts have mitigated this problem (Fisher et al., 2013). In fact, external scripts have been found to effectively support knowledge development and collaboration (Hämäläinen et al., 2008; Radkowitsch et al., 2020). Although researchers have proposed the application of external scripts to support deep level metacognitive regulation and enhance SSRL (Järvelä & Hadwin, 2015; Miller & Hadwin, 2015), few studies have effectuated this proposal.

Järvelä and Hadwin (2015) proposed three design principles for supporting SSRL, namely: (1) increasing learner awareness of individual and group members' learning processes, (2) supporting externalization of individual and group members' learning processes and helping to share and interact, and (3) prompting acquisition and activation of regulatory processes. Within the context of the present study, we have operationalized these design principles into a macro-and micro-level collaboration script that aims to facilitate depth and sharedness of metacognitive regulation (see Appendix A for an elaborated description).

The 'Y-read?' external collaboration scripts are structured around the four elements play, scene, scriptlets, and roles from Fisher et al. (2013). The macro-level script combines the components 'play', 'scene' and 'roles' (following Fisher et al., 2013; Radkowitsch et al., 2020) and fosters collaborative pre- and post-task regulation before and after individual reading (following Järvelä et al., 2016). Play provides general task definitions detailing the main goal for collaboration. Within the context of the present study, play supports pre- and post-task socially shared regulation of TOR. A 'scene' encourages a sequence of activities within the 'play' scenario. In the present study, there are four scenes, specifically pre-task group preparation, individual task execution, individual reflection and post-task group reflection (see Fig. 2 and Appendix A). A novelty of this script is the interaction between collaborative and individual elements, i.e., the combination of individual task execution and reflection combined with collaborative pre- and post-task regulation to elicit socially shared regulation. Hence the function of the script is directed at supporting socially shared regulation, which is a rare focus in script research (Radkowitsch et al., 2020). In addition, students are assigned different 'roles' across 'scenes', such as group leader, writer, task guard or text guard, to support sharedness through information interdependency among group members. Although roles have been investigated previously (Weinberger et al., 2005), a full automation flow of role information and related learning materials to different participants on four different tablets that synchronously drive face-to-face collaboration as developed in Y-read is, to the best of our knowledge, novel. Also, roles were rotated across the four sessions to increase motivation and learning (Taylor & Baek, 2019).

Finally, the micro-level script further facilitates pre- and post-task regulation. 'Scriptlets' arrange activities within a 'scene' and thus construe knowledge about the sequencing of these activities. For example, during pre-task regulation, students were urged to create a task representation in order to select a reading strategy. This scriptlet triggers students to externalize their perceptions of the task and consequently build a shared understanding of task complexity before selecting an appropriate strategy (Järvelä & Hadwin, 2015). This pre-task planning scriptlet, and the other scriptlets, are designed to elicit deep-level meta-cognitive regulation.

Present study

In this study, we examined whether the 'Y-read?' collaboration script elicits deep and shared metacognitive regulation and helps individual students to engage in more successful individual regulation of TOR. As mentioned before, students' discussions of TOR and the regulation thereof often remain superficial, suggesting they need guidance (Okkinga et al., 2021). Collaboration scripts have the potential to elevate group discussions towards more advanced regulation of TOR. We are particularly interested in the effects of such a script on the depth and sharedness of metacognitive regulation during TOR assignments. The present study thus aimed to find an answer to the following research questions:

1. *How does a collaboration script impact the depth and socially sharedness of metacognitive regulation?*
2. *To what extent does a collaboration script consequently affect individual task representation, strategy selection and strategy reflection?*

We expected the script to help groups engage in deep and shared metacognitive regulation (Molenaar et al., 2014). To explore the first question, we first examined the depth and

sharedness of metacognitive regulation during group discussions when working with the collaboration script. We expected that role division in the macro-level script would encourage sharedness among students through the support of joint negotiation and aligning metacognitive regulation, while the micro-level script would support the depth of their regulation by offering students support discussing pre-and post-task regulation. We examined depth and sharedness by analyzing group interaction when working with the script. Finally, we explored the second question by examining whether students improved on individual metacognitive regulation (i.e., individual task representation, strategy selection and strategy reflection) by comparing the pre-and post-test scores. We expected a positive effect of socially shared regulation on the ability of students to individually regulate their learning (Molenaar et al., 2014).

Method

Design

This study followed a pre-test, post-test intervention design. The intervention consisted of a collaboration script that was used in four lessons to support students' collaboration during reading tasks they conducted in small groups. Before and after the four lesson intervention, students completed a TOR-test, during which their task representation, strategy selection and strategy reflection were assessed. Both the collaboration script and the pre- and post-tests were integrated in a newly developed digital learning environment: 'Y-read?' (see Fig. 1).

Participants

The study was conducted in a classroom setting with 44 10th Grade pre-prevocational secondary school students as part of the regular Economics program. The average age of students was 15.7 years ($SD=0.9$), and 23 identified as male. The participants were divided over two classes with the same teacher. During the intervention participants collaborated in 11 groups of four students each. Parental and student approval was required for participation in this study.
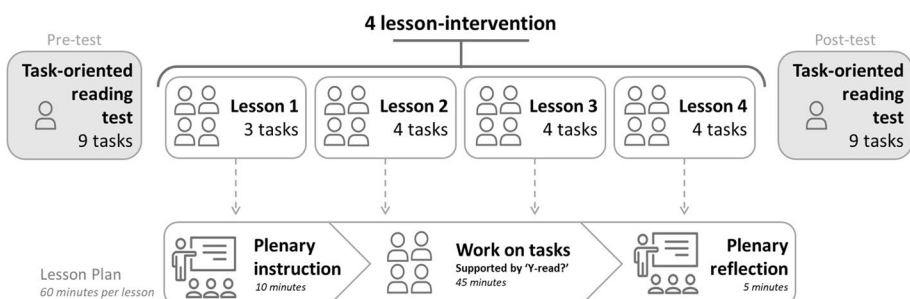
Fig. 1 Study design

Intervention

Both the TOR-test and collaboration script were integrated in the 'Y-read?' environment in such a way that the students had to adopt the script. The TOR-test was developed to measure students' task representation, strategy selection and strategy reflection before and after the intervention. This test will be described under Measures. During the 4-lesson intervention, students were placed in groups, each with their own device, while being able to discuss planning and reflection activities face-to-face. For the duration of the lesson, students were seated together, even when working individually in 'Y-read?'. When working together during planning and reflection (see Fig. 2), the micro-level script aimed to help students articulate and align their regulatory activities, such as task representation, strategy selection and strategy reflection, while the macro-level script guided the students synchronously through the pre- and post-task phase.

The aim of the collaboration script was to elicit group discussions around pre- and post-task regulation. The macro-level script guided students through the different 'scenes', respectively, collaborative pre-task group planning, followed by individual task execution and a collaborative post-task group reflection (see Fig. 2). Task representation and strategy selection were addressed in pre-task group planning, and strategy reflection was addressed in post-task group reflection. Each lesson consisted of three or four iterations of the script. After group reflection, students were redirected to the collaborative planning phase for the next task. To structure group discussions, students were automatically assigned roles, which alternated over the four lessons (see Fig. 2). The leader (A) guided discussions and
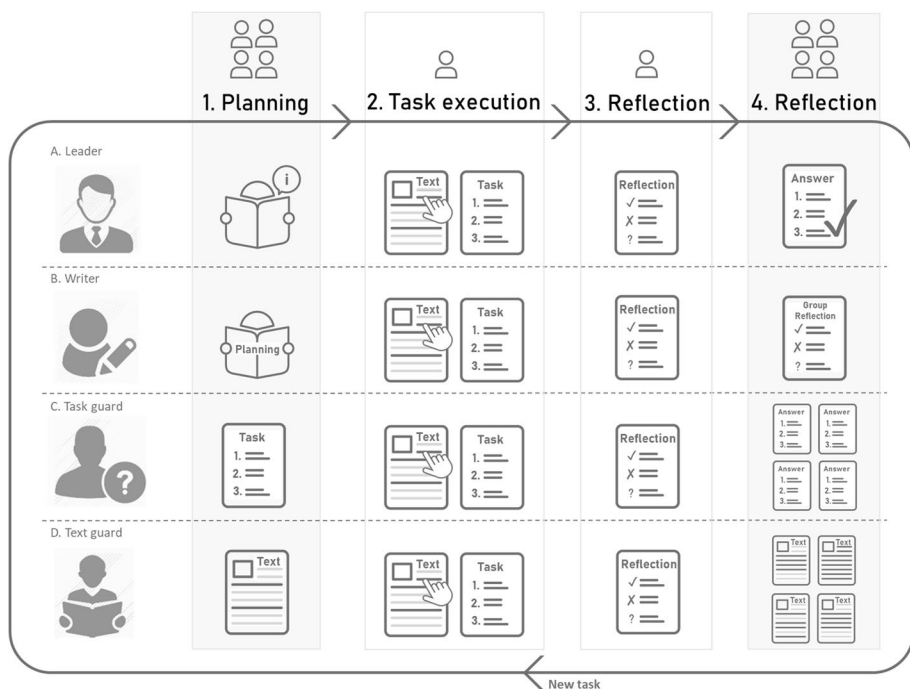

Fig. 2 Task flow of the 'Y-read?' collaboration script

had information about the task and reading strategies. The writer (B) was in charge of writing down group decisions on shared task representation, strategy selection and strategy reflection. The task guard (C) was given the task and was instructed to discuss task characteristics with group members. Finally, the text guard (D) was given the text and was instructed to discuss content characteristics with group members. By alternating the roles over lessons, students learned to approach a task from different viewpoints.

After 'Y-read?' assigned the roles at the beginning of each lesson, the pre-task group planning phase started, during which the micro-level script provided support for each group member's activities according to their role (see column "1. Planning" in Fig. 2). Distribution of information in this phase was based on the distributed information principle from Jigsaw designs and used to elicit interdependency (Strijbos & De Laat, 2010). By providing each group member with a part of the information necessary to create a plan, the collaboration script compelled students to collaborate. First, the group was guided to discuss their task representation. The leader was given questions to guide the discussion (A1 in Fig. 2), such as *"What kind of assignment is it? (For example, multiple-choice or a fill-in schedule)"* and *"What information can you already find in the text?"*. The writer was given the planning page on which students had to fill in two questions regarding their task representation (B1 in Fig. 2). These questions and their answer options were: (1) *"How difficult do you think this task will be?"* On a scale from 1 to 10; 1 being very easy, and 10 being very difficult, and (2) *"Where can you find the relevant information in the text to answer this assignment?"* On a scale from 1 to 10; 1 being in 1 place, and 10 being in many places. The task guard was given the task (C1 in Fig. 2), and the text guard was given the blurred version of the text, with only the title, paragraph headers and picture visible (D1 in Fig. 2). Second, the group was guided to discuss their strategy selection. To select a strategy, the leader was given a question to guide the discussion (A1 in Fig. 2), such as *"Which reading strategy is best for the assignment and the text?"*. In addition, the leader was given information on reading strategies. For example, the leader had information on how to apply a searched strategy, i.e. *"You look for information in the text, paying attention to specific details in the text, such as names, dates, and places. Headings above the text can help with this because headings provide information about the content of a paragraph. Often the answer is almost literally in the text."*. The writer was given the planning page (B1 in Fig. 2) on which students had to indicate which reading strategy they were going to use (i.e., searched reading, targeted reading or intensive reading). In addition, the writer had to write down why the group selected this strategy. Together with the information given to the task guard (C1 in Fig. 2) and text guard (D1 in Fig. 2), the group was able to select a strategy.

After this collaborative stage, the macro-level script gave each student access to the text and task, so that they could execute the task (see column "2. Task execution" in Fig. 2). When they completed the task, each student was asked individually which reading strategy they used and which reading strategy would have been best in hindsight for this task (see column "3. Reflection" in Fig. 2). Subsequently, they were given a waiting screen until all group members finished the task and individual reflection. Once all group members were finished, they had to reflect on the task together (see column "4. Reflection" in Fig. 2). The script supported collaboration by again providing each student with a piece of information necessary for strategy reflection. The leader was given the correct answer and additional questions to guide the discussion (A4 in Fig. 2), such as *"What answers did everyone give to the assignment? (Task guard has the answer)"* and *"What reading strategies did everyone use? (Text guard has the answer)"*. The writer was given the same questions as during planning (B4 in Fig. 2), only now with instructions to answer these questions again with

the acquired knowledge of the task. The task guard was given the task answers of each group member (C4 in Fig. 2). The text guard was able to see which strategy each group member had used, based on information students had provided during individual reflection (D4 in Fig. 2). After filling out the reflection, the script guided students back to planning to work on a new task collaboratively.

Measurements

Conversation analysis; measuring regulation of TOR

To determine how the micro-level collaboration script influenced the depth of groups' pre- and post-task shared regulation, all group conversations were recorded with voice recorders. All audio recordings of group discussions were transcribed. While it was unclear which group member was speaking on the recording, one could clearly hear when there was a change of speaker. Therefore, transcriptions were made on the utterance level. Each utterance represents an expression about a single topic by a single student.

First, each utterance during pre- and post-task regulatory discussions was coded into one of the main categories derived from the coding scheme of Molenaar et al. (2014): cognitive activities, metacognitive activities, procedural activities, relational activities, off-task activities, teacher/researcher activities and not codable activities. Cognitive activities included all utterances concerning task content and elaborations of this content (e.g., asking questions regarding domain knowledge, or elaborating on or summarizing content). Metacognitive activities included utterances concerning monitoring and controlling cognitive activities (e.g., orientation on a task and text, create a plan, or evaluate or reflect on the task and text). Procedural activities concern utterances about working in 'Y-read?' (e.g., discussing how to use 'Y-read?' or role responsibilities). Relational activities included utterances about social interaction between group members (e.g., supporting or encouraging group members to actively participate in the group discussion). Off-task activities concern utterances unrelated to the task or text (e.g., discussing last night's soccer game). Teacher/researcher activities contain contributions made by the teacher or researchers (e.g., the teacher or researcher offers support on working with 'Y-read?'). Finally, all utterances that are too short or unclear to interpret fell into the category of not codable activities. When two or more consecutive utterances fell under the same code, the same code was given for each of these utterances.

Second, each metacognitive activity was further analyzed to define whether students engaged in low- or deep-level metacognitive regulation, inspired by the coding scheme of De Backer et al. (2016). For this study, in which we aimed to measure how the collaboration script influenced pre- and post-task regulatory discussions, we focused on low- and deep-level regulation of metacognitive activities such as orientation, planning and reflection (see Table 1). Low-level orientation is solely focused on examining content characteristics and task demands, whereas deep-level orientation involves processing content characteristics and task demands by activating prior knowledge or generating hypotheses. Solely formulating a problem-solving plan without substantive arguments is considered low-level planning, whereas connecting task and/or content characteristics in creating a plan is a form of deep-level planning. Finally, low-level reflection only encompasses students evaluating and checking their learning outcomes, whereas deep-level reflection indicates reflective judgements for future tasks. We additionally registered how often students

Table 1 Low- or deep-level metacognitive regulation

Metacognitive activity		Description	Example
Orientation	Low-level	Solely focused on examining content characteristics and task demands	Student 1: *"What kind of assignment is it?"* Student 2: *"You just have to fill in three things so that is not very difficult."*
	Deep-level	Involves processing content characteristics and task demands by activating prior knowledge or generating hypotheses	Student 1: *"You just have to fill in three negative consequences of increasing premium or benefit."* Student 2: *"I think this is about state pension, the income of workers will drop as a result."*
Planning	Low-level	Solely formulating a problem-solving plan without substantive arguments	Student 1: *"I think searching for this task."*
	Deep-level	Connecting task and/or content characteristics with creating a plan	Student 1: *"I think searched reading, because when you look up those terms mentioned in the task you will find the answer."*
Reflection	Low-level	Only encompasses students evaluating and checking their learning outcomes	Student 1: *"What was the correct answer?"* Student 2: *"The first is SER, the other CBS and the other CPB."* Student 3: *"I think I was completely wrong."*
	Deep-level	Indicates reflective judgements for future tasks	Student 1: *"Where could we find the relevant information in the text?"* Student 2: *"Everywhere."* Student 3: *"I agree everywhere, so next time we should read intensively."*

literally read out loud the questions provided by the script to measure uptake of these scripted questions. Since these questions were literally read from the script, we did not regard these utterances as self-initiated metacognitive activities, and therefore we have not coded these questions under low- or deep-level activities. Instead, they were coded as scaffold. Again, in case two or more consecutive utterances fell under the same code, the same code was given for each of these utterances.

Each utterance was coded (n = 8713 utterances) and 1027 (11.78%) of these utterances were randomly selected and coded by a trained second coder to determine interrater reliability (Lombard et al., 2002). There was moderate overall interrater-reliability agreement on the main categories ($K=0.73$) with a high agreement on the main category metacognition ($K=0.80$). On the subcategories interrater-reliability indicates an almost excellent overall agreement ($K=0.88$) and a high overall agreement on the coding of low- and deep-level metacognitive activities ($K=0.84$).

Conversation analysis; measuring shared regulation

After coding all utterances into main categories and subcoding metacognitive utterances, metacognitive episodes were determined to provide insight into the sharedness of intra-group social metacognitive interaction (Molenaar et al., 2014). Group pre- and post-task discussions were divided into metacognitive episodes according to changes in topic of discussion. Metacognitive episodes comprised a sequence of utterances of which at least one is metacognitive in nature; such episodes end after the last utterance on the same topic. These metacognitive episodes were then coded per episode to distinguish four types of intra-group social metacognitive interaction: ignored, accepted, shared and co-constructed (see Table 2). An ignored social metacognitive activity occurs when a group member's metacognitive utterance is ignored, such as when a student asks what strategy to use, and another student responds with an utterance about a video on their phone. An accepted social metacognitive activity occurs when a metacognitive utterance is followed up with a cognitive activity of another group member, for example, when a student asks about task complexity and another student responds by reading the correct answer. A shared metacognitive social activity occurs when students share metacognitive activities and ideas with each other, but do not build on each other's contributions, for example, when two students share which reading strategy they plan on using. A co-constructed social metacognitive activity occurs when students build on each other's metacognitive activities, for example, when a student shares which reading strategy they deem appropriate for the task and another student responds with an argument why this strategy is appropriate for the task.

In total 938 metacognitive episodes were identified and coded by the first author of this paper. In addition, 152 (16.2%) episodes were randomly selected and coded by a trained second coder to determine interrater reliability (Lombard et al., 2002). There was good overall agreement ($K=0.80$) and moderate to excellent agreement for the types of intra-group interaction ($K_{ignored}=0,91$, $K_{accepted}=0,93$, $K_{shared}=0,71$, $K_{co-constructed}=0,84$).

Depth and sharedness of regulation across the lessons

These analyses grant insight into groups' depth and social sharedness of metacognitive regulation during TOR across the lessons, however information regarding the groups' interaction per lesson is thereby lost. For this reason, a qualitative exploratory analysis was added to the results to visualize how depth and social sharedness of metacognitive regulation

Table 2 Four types of intra-group social metacognitive interaction

Type	Description	Example
1 Ignored	Ignored social metacognitive activities occur when the group members do not relate to or engage with another group member's metacognitive activity	Student 1: *"I think we should use targeted reading."* Student2: *"You have to look in different places."* [in text] Student 1: *"Nice selfie."*
2 Accepted	Accepted social metacognitive activities occur when the group members reply to a metacognitive activity with a cognitive activity	Student1: *"It is a pretty difficult text, I think we have to read it intensively"* Student 2: *"Where can we find the relevant information?"* Student 1: *"I think on, like, seven places. Seven places is a lot."* [referring to number of parts in text]
3 Shared	Shared social metacognitive activities occur when group members exchange metacognitive activities. They share their ideas, but do not build on each other's comments	Student 1: *"I read targeted, but I should have read intensively."* Student2: *"First I read this and then I read intensively."*
4 Co-constructed	When group members do build on each other's metacognitive activities, we speak of co-constructed social metacognitive activities	Student1: *"Which strategy are you going to use to execute the task?"* Student 2: *"Searching, targeted or intensive? I think targeted."* Student 1: *"Yes, that is what I think as well."* Student 3: *"Because you have to look something up."*

evolved over the four lessons for each group. To visualize the differences in intra-group interaction and depth of metacognitive activities over the course of the four lessons the information from the conversation analyses was plotted in interaction graphs for each group (see Appendix B). The blue area graphs display the intra-group interaction of each group over the course of the four lessons, while the grey bar charts display the depth of metacognitive activities of each group over the course of the four lessons.

In addition, each groups' and students' correct answers on the planning and reflection questions were visualized to examine how groups and individuals evolved over the lessons. After each task a student was asked individually *"which reading strategy would have been best in hindsight for this task?"* (see Fig. 2, column "3, Reflection"). Subsequently, it was examined whether the chosen strategy was appropriate for the task complexity. If this was the case, the student received a point, with a maximum of 4 points per lesson. These correct answers are plotted anonymously for each student in the graph on top of the blue area graph. On top of the grey bar graphs a line graph is plotted, consisting of two lines representing the correct group answer on strategy selection and strategy reflection (see Fig. 2, column "1. Planning" and column "4. Reflection"). If the selected and/or reflected strategy was appropriate for the task complexity, the group received a point, with a maximum of 4 points per lesson.

Individual pre- and post-TOR-test

Students completed the TOR-test individually in 'Y-read?' behind a laptop. At the start of the TOR-test, students were asked questions about task representation and strategy selection on the planning page. To fill out the planning, they also had access to the task and limited information regarding the text (i.e., they could read the title and subheaders while the rest was blurred). After submitting their planning, they were directed to the text, of which only the title and headers were visible to read. To read the rest of the text, students could deblur a sentence at a time. During this execution phase, students were able to switch back and forth between text and task at any time. After reflection, students were redirected to the planning page for the next task. Before entering the test, students were given an example task to familiarize themselves with the individual 'Y-read?'-environment.

In the reading test, students completed nine tasks centered on the topic of environmental sustainability. This topic was chosen because it was not part of students' Economics program. We developed tasks that were comparable to tasks students commonly come across in school materials. These tasks included multiple-choice questions, open questions, matching tasks, and a summary in which students had to fill in missing words. These tasks were related to one text, and we designed the tasks so that they varied in complexity (i.e., low-level, intermediate-level, and high-level) based on the amount of reading and inferencing necessary to complete a task. Prior to a task, students were asked about their task representation and strategy selection. After completing a task, students were asked about their strategy reflection. The questions regarding task representation, strategy selection and strategy reflection are the same as in the collaboration script, only these questions were now answered individually. *Task representation* assessed students' ability to recognize task complexity based on two predictions. First, students were asked about the amount of reading necessary (i.e., same question as in the collaboration script). Second and different from the collaboration script, students were asked about the amount of inferencing deemed necessary. For this they were asked to finish the following statement: *'While reading the text…'* by choosing from two options (1) *'…I can probably literally find the information*

I am looking for' or (2) *'…I probably have to think about the information to make the assignment'*. Questions about *Strategy selection* and *Strategy reflection* were the same as in the collaboration script. However, strategy evaluation was not used as a measure during this study. Instead, this question was added as a self-report question to trigger students to think about their executed strategy (i.e., question about used strategy). Strategy reflection assessed students' ability to reflect on regulation of TOR.

Procedure

For the intervention, which consisted of four lessons, the teacher followed a predesigned lesson plan consisting of plenary instruction, work on tasks, and a plenary reflection (see Fig. 1 for lesson plan). Students worked in groups of four on 3 or 4 tasks while being guided by 'Y-read?'. These tasks were part of the regular Economics curriculum. Roles in the collaboration script and what was expected of students were explained during the first lesson. In addition, students were encouraged to ask questions regarding these roles and 'Y-read?' in case of unclarities. The teacher or researcher would answer these questions when related to the procedure. Questions regarding the task or text were not answered. Instead, students were encouraged to converse with group members about content-related questions.

At the start of each lesson students were given ten minutes of instruction. The first lesson involved plenary instruction on the 'Y-read?'-environment and shared collaboration, while instruction in the second lesson focused on task representation. Instruction in the third lesson focused on differences in reading strategies, and the final lesson focused on reflection of previous lessons. After instruction, students worked in groups for 45 min, supported by 'Y-read?'. Each lesson ended with five minutes of plenary reflection on the tasks students had worked on.

Analyses

To answer the first research question, we first examined the depth and social sharedness of metacognitive regulation of TOR while groups were supported by the macro-and micro-level collaboration script. For this purpose, we analyzed the depth of metacognitive regulation through the distribution of different types of metacognitive activities, focusing particularly on low- and deep-level metacognitive activities during pre- and post-task discussions. Second, we examined sharedness of regulation when working with the script by analyzing the distribution of different types of intra-group interaction during metacognitive episodes. In addition, we investigated whether sharedness of metacognitive regulation (i.e., the type of intra-group interaction) was correlated with the depth of metacognitive regulation (i.e., low- or deep-level metacognitive regulation). This requires that we calculate the correlation at utterance level (i.e., low-and high-level utterances) with metacognitive episodes (i.e., ignored, accepted, shared and co-constructed) to examine whether students showed more deep-level activities during shared or co-constructed interaction, as we would expect. However, it could also have been possible that students built on each other's low-level contributions, or that a student with a deep-level contribution was ignored, or group members solely accepted this statement as truth and did not elaborate on it any further.

For this reason, we performed a correlation to check whether our assumption (namely that deep-level utterances are associated with shared or co-constructed interactions) is correct.

Another way to gain insight into groups' depth and social sharedness of metacognitive regulation during TOR is by looking at the learning process (Malmberg et al., 2013; Paans et al., 2020). For this reason, we executed an exploratory qualitative analysis to visualize the depth and sharedness across the four lessons. In addition, groups' and students' correct answers on the planning and reflection questions were visualized to examine how groups and individuals evolved over the lessons. By visualizing these answers with depth and social sharedness of metacognitive regulation together in a graph (see Appendix B), we hoped to gain more insight in groups' collaboration as well as possible differences between groups.

For the second research question, we examined to what extent students' individual task representation, strategy selection and strategy reflection had improved after the intervention. For this purpose, a paired-samples t-test was executed to test differences between pre- and post-test. A power analysis with power set at 0.80 and an expected effect size of 0.25, and an $\alpha = 0.05$ using a linear multiple regression, indicated a minimum sample size of 42 participants needed to reach statistical significance. For inclusion in the analysis, students had to be present during at least three of the four lessons. Students from one group were absent more frequently and therefore excluded from this analysis. Also, students that were only present during one of the TOR-tests were excluded. After applying these rules, 38 students were included in the analysis.

Results

Regulation of TOR when working with a collaboration script

Table 3 presents the frequency of different activities during pre- and post-task group discussions. The table shows strong engagement in metacognitive activities, both before and after the task: in both stages around sixty percent of the utterances were metacognitive, which is consistent with the fact that only pre and post-task regulation was performed collaboratively. A smaller share of utterances involved procedural and relational activities. Cognitive activities and motivational comments were rare.

Table 3 Overview number of utterances per main category

Main categories	Pre-task		Post-task	
	N	%	N	%
Metacognition	2318	62.84	1848	59.6
Cognition	13	0.35	112	3.6
Procedural	509	13.80	413	13.3
Relational	513	13.91	480	15.5
Motivation	8	0.22	21	0.7
Off task	159	4.31	51	1.6
Teacher/Researcher	77	2.09	56	1.8
Not codable	92	2.49	120	3.9
Total	3688	100	3101	100

We then investigated the nature of the metacognitive activities. Of the metacognitive utterances during pre-task regulation, 314 (13.55%) involved students repeating scripted questions as presented to the leader and writer, 1419 (61.22%) involved orientation on task and content, and 577 (24.89%) involved planning. Few pre-task metacognitive utterances involved monitoring (N=8, 0.35%). Of the 1848 metacognitive utterances during post-task regulation, students repeated scripted questions 225 (12.18%) times, while 1622 utterances (87.77%) involved reflection, and 1 (0.05%) involved monitoring.

Subsequently, we investigated whether students used low or deep-level metacognitive activities during pre-and post-task regulation (N=3669, see Table 4). During pre-task regulation, 1909 (52.03%) utterances were coded as low-level metacognitive activities. On average per group, low-level metacognitive activities occurred during 173 (SD=68.77) instances. Group A (N=45) showed the least- and Group H (N=250) showed most low-level metacognitive activities. Deep-level metacognitive activities were rare and occurred during 87 (2.37%) utterances. On average, each group discussed deep-level metacognitive activities during 7.91 (SD=9.95) instances, with Group C discussing no deep-level activities and Group H discussing deep-level activities during 30 utterances. During post-task regulation, 1672 (45.57%) utterances were coded as low-level metacognitive activities. On average per group, low-level metacognitive activities occurred during 152 (SD=63.08) instances. Group C (N=73) showed the least- and Group H (N=267) showed the most low-level metacognitive activities. Deep-level metacognitive activities were only measured once during post-task regulation.

Shared regulation of TOR when working with a collaboration script

Table 5 shows the intra-group metacognitive interaction (sharedness) during pre-and post-task discussion. We established a total of 938 metacognitive episodes, distributed quite evenly across pre-and post-task (n=484, 52% and n=454, 48%, respectively). On average per group, 44 (SD=13.70)

Table 4 Low versus deep-level metacognitive activities for each group

	Pre-task				Post-task			
	Low		Deep		Low		Deep	
	N	%	N	%	N	%	N	%
Group A	45	33.8	2	1.5	86	64.7	0	0.0
Group B	82	49.4	4	2.4	80	48.2	0	0.0
Group C	137	65.2	0	0.0	73	34.8	0	0.0
Group D	139	56.3	1	0.4	107	43.3	0	0.0
Group E	235	50.1	24	5.1	210	44.8	0	0.0
Group F	218	54.2	3	0.7	181	45.0	0	0.0
Group G	232	60.9	10	2.6	139	36.5	0	0.0
Group H	250	45.7	30	5.5	267	48.8	0	0.0
Group I	232	59.9	6	1.6	148	38.2	1	0.3
Group J	143	39.6	1	0.3	217	60.1	0	0.0
Group K	196	53.6	6	1.6	164	44.8	0	0.0
Total	1909	52.03	87	2.37	1672	45.57	1	0.03

Table 5 Pre- and post-task intra-group social metacognitive interaction for each group

| | Pre-task | | | | | | | | Post-task | | | | | | | |
| | Ignored | | Accepted | | Shared | | Co-constructed | | Ignored | | Accepted | | Shared | | Co-constructed | |
	N	%	N	%	N	%	N	%	N	%	N	%	N	%	N	%
Group A	3	20.00	10	66.67	2	13.33	0	0.00	16	44.44	19	52.78	1	2.78	0	0.00
Group B	10	34.48	17	58.62	2	6.90	0	0.00	12	46.15	12	46.15	2	7.69	0	0.00
Group C	6	13.95	34	79.07	2	4.65	1	2.33	9	29.03	20	64.52	2	6.45	0	0.00
Group D	3	7.69	32	82.05	3	7.69	1	2.56	6	15.38	32	82.05	1	2.56	0	0.00
Group E	0	0.00	33	64.71	2	3.92	16	31.37	1	1.96	31	60.78	12	23.53	7	13.73
Group F	8	17.02	33	70.21	1	2.13	5	10.64	11	26.83	22	53.66	8	19.51	0	0.00
Group G	1	1.54	57	87.69	2	3.08	5	7.69	5	9.80	42	82.35	3	5.88	1	1.96
Group H	4	7.55	31	58.49	9	16.98	9	16.98	6	12.00	25	50.00	13	26.00	6	12.00
Group I	2	4.00	39	78.00	3	6.00	6	12.00	3	7.89	29	76.32	5	13.16	1	2.63
Group J	3	8.11	32	86.49	2	5.41	0	0.00	4	8.70	37	80.43	5	10.87	0	0.00
Group K	5	9.09	45	81.82	1	1.82	4	7.27	8	17.78	33	73.33	3	6.67	1	2.22
Total	45	9.30	363	75.00	29	5.99	47	9.71	81	17.84	302	66.52	55	12.11	16	3.52

metacognitive episodes were established during pre-task regulation. The least metacognitive episodes occurred in Group A (N=15) and most episodes occurred in Group G (N=65).

Subsequently, we investigated what types of intra-group metacognitive interaction groups exhibited (i.e., ignored, accepted, shared or co-constructed) during pre-and post-task regulation (see Table 5). During pre-task regulation, 363 (75.00%) metacognitive episodes were coded as accepted intra-group interactions. On average per group, accepted interaction occurred during 33 (SD=12.44) episodes. Group E (N=0) showed the least- and Group B (N=10) showed the most accepted interactions. 47 episodes (9.71%) were coded as co-constructed intra-group metacognitive interactions during pre-task discussions. On average per group, co-constructed interactions occurred during 4.27 (SD=4.90) episodes. Showing a great diversity between groups, Groups A, B and J (N=0) showed the least- and Group E (N=16) showed the most co-constructed interactions. Ignored intra-group interactions occurred in 45 (9.30%) metacognitive episodes, almost just as often as co-constructed interactions. On average group members ignored each other during 4.09 (SD=2.98) episodes, with again a great diversity between groups. For instance, Group E never ignored group members, while Group B ignored group members during 10 episodes. During pre-task regulation, only 29 (5.99%) metacognitive episodes showed any shared intra-group interaction.

Of all metacognitive episodes during post-task regulation, 302 (66,52%) showed accepted intra-group interaction. On average per group, accepted interactions occurred during 27.45 (SD=8.78) episodes. Group B (N=12) showed the least- and Group G (N=42) showed the most accepted interactions. During 81 (17.84%) metacognitive episodes, ignored intra-group interactions occurred. On average per group, ignored interactions occurred during 7.36 (SD=4.39) episodes, showing a great diversity between groups. Group E (N=1) showed the least- and Group A (N=16) showed the most ignored interactions. Shared intra-group interactions occurred in 55 (12.11%) metacognitive episodes. On average per group, shared interactions occurred during 5.00 (SD=4.24) episodes. Groups A and D (N=1) showed the least- and Group H (N=13) showed the most shared interactions. During post-task regulation, only 16 (3.52%) metacognitive episodes showed any co-constructed intra-group interaction.

## Depth and sharedness of regulation across the lessons

In addition to the observed differences between the groups in Tables 4 and 5, we have plotted two interaction graphs for each group of which four exemplary graphs will be explored further in this paragraph. Group A and B were chosen because they displayed the most ignored intra-group interactions, while Group E and H were chosen because they displayed the most co-construction compared to the other groups (see Table 5).

When looking at the blue graphs on the left below in Fig. 3, Group E and Group H displayed a larger darker blue area, indicating more co-constructive interactions compared to other groups. In contrast, Group A and Group B displayed a larger lighter blue area, indicative of more ignored interactions over the lessons (see left-side y-axis for the percentiles sharedness of interactions on a scale of 0 to 1). The line graph plotted over these blue area graphs displays the individual scores on strategy reflection and shows an upward line between the second and third lessons for most students across the groups. However, after the third lesson, students' individual strategy reflection scores seem to decline for most students, except for one student in Group B who displays a stable line across the first three lessons, scoring 1 point each lesson and then scoring the maximum of 4 points on the last lesson (see right-side y-axis of the graph for the individual scores on a scale of 0 to 4).

Group A
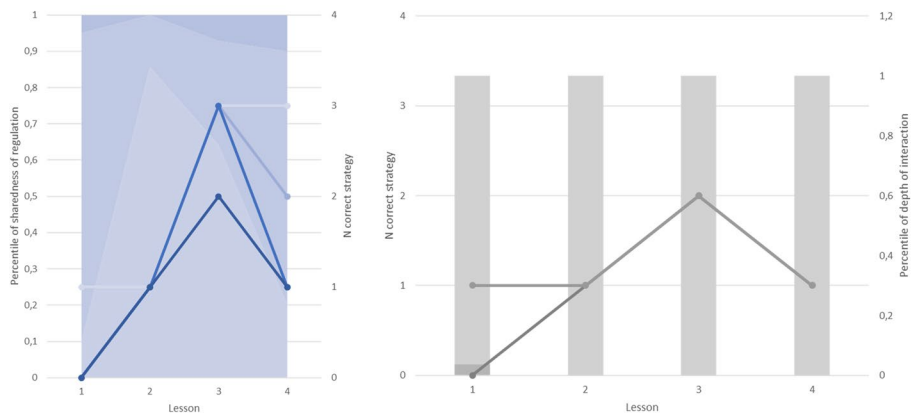
Group B
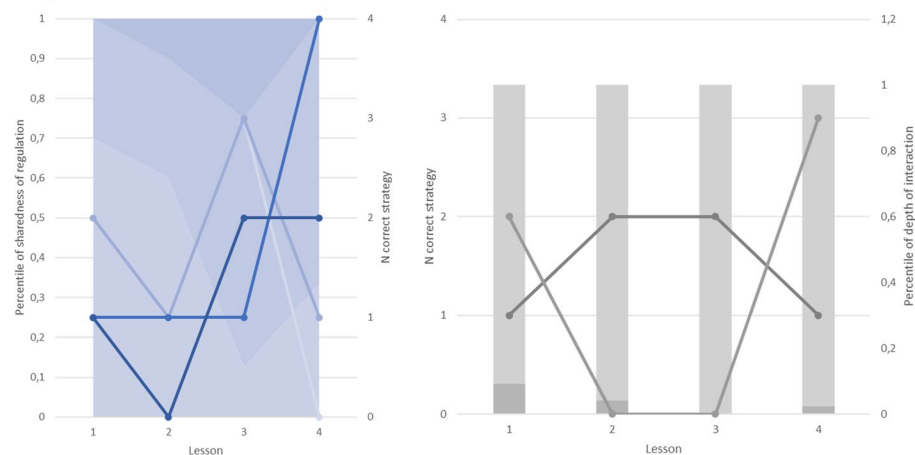
Fig. 3 *Note.* For both graphs, the x-axis displays the number of each lesson. For the blue graphs, the left side y-axis displays on a scale from 0 to 1 the percentiles for the sharedness of interaction, while the right side y-axis displays on a scale from 0 to 4, the number of correct answers given on the individual reflection question. For the grey graphs the left side y-axis displays on a scale from 0 to 4, the number of correct answers given on the group's strategy selection and reflection questions, while the right side y-axis displays on a scale from 0 to 1 the percentiles for the depth of groups' discussion.

When looking at the grey bar charts on the right in Fig. 3, Groups E and Group H show slightly more dark grey in the bars across the lessons, indicating that they engaged in more deep-level metacognitive activities compared to Group A and Group B. Group A and Group B also displayed some deep-level metacognitive activities, albeit it mainly being in the first lessons and then simmering down (see right-side y-axis of the graph for the percentiles of depth of groups' discussion on scale from 0 to 1). The line graph plotted over these grey bar charts displays the group scores on strategy selection (i.e., dark grey line) and strategy reflection (i.e., light grey line). Except for Group B (i.e., scores 3 points during the 4th lesson), none of the groups gained more than 2 points when reflecting on a strategy. Group A and Group E also got identical scores

Group E
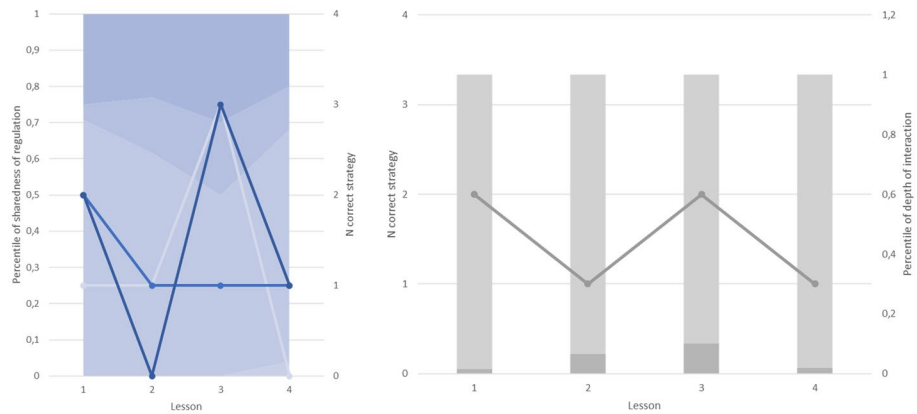
Group H
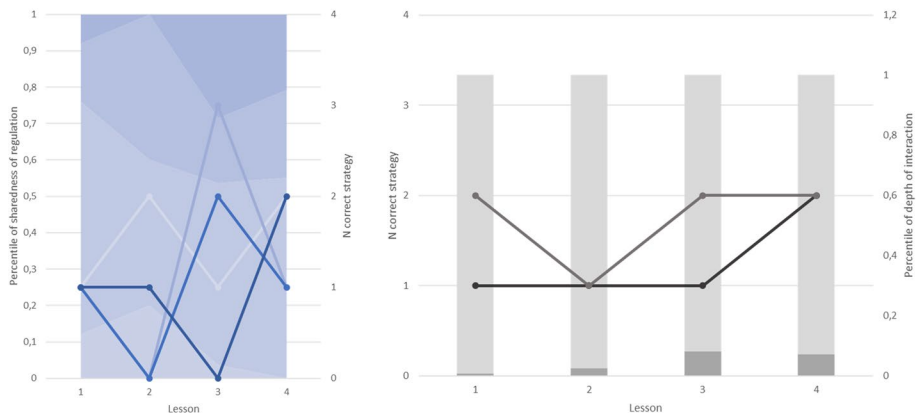
Fig. 3 (continued)

on strategy selection and reflection, indicating that they gave the same group answer before and after completing the task (see left-side y-axis for the group scores on strategy selection and strategy reflection on a scale of 0 to 4).

Overall, the blue line graphs from the individual scores appeared to be roughly similar in shape as the line graphs from the group scores. Also, all four groups seemed to peak in the third lesson, while the sharedness and depth of regulation differed between these groups during this lesson (for full results, see Appendix B).

Association of shared regulation with regulation of TOR

Finally, we examined whether different types of intra-group interaction were associated with low- and deep-level metacognitive activities. Table 6 shows how each type of intra-group metacognitive interaction influenced the depth of metacognitive activities. During pre-task

Table 6 Correlations between types of intra-group interactions and low- and deep-level metacognitive activities

	Pre-task				Post-task			
	Ignored	Accepted	Shared	Co-constructed	Ignored	Accepted	Shared	Co-constructed
Low-level	-0.374	0.739**	0.325	0.758**	-0.562	0.417	0.864**	0.680*
Deep-level	-0.364	0.142	0.696*	0.837**	-0.33	0.058	0.000	-0.059

* Correlation is significant at the 0.05 level (2-tailed)

** Correlation is significant at the 0.01 level (2-tailed)

regulation, the frequency of shared ($r=0.70$, $p<0.05$) and co-constructed ($r=0.84$, $p<0.01$) intra-group metacognitive interaction was significantly correlated with the frequency of deep-level metacognitive activities. In other words, shared and co-constructed interaction co-occur more frequently in deep-level metacognitive interactions.

An illustration of how co-constructed interaction can facilitate deep-level metacognitive activities is shown in Table 7. At the start of this example a student states what strategy the group should use ("I think it is targeted reading."), to which the writer only replies with a confirming response to write down this suggestion ("I'll fill that in."). However, another student questions what this strategy entails ("Targeted is reading a paragraph, right?"), opening the conversation for another student to respond and add why they should use this strategy instead of another ("Yes, but you have to explain something here, so therefore it is not searched."). This answer on why they should use this 'targeted' strategy is then questioned, and the student is asked about their expectations on where the answer can be found in the text ("So in how many places will the answer be then?"), to which the student replies with a number ("That could be three."), referring to the question the writer must answer during planning. This conversation prompts another student to respond by trying to clarify the previous comment ("Plan economy… wait…"), although the response remains relatively vague. Therefore, another student responds with a more explicit suggestion about where the information in the text might be ("Maybe the answer it is in the introduction."). This response is followed by a clarification about what the other student tried to say earlier, now linking task demands about what they should look for in the text to an exact location by looking at headers ("Wait, I think 'the influence of the governments', because a plan economy is, what is that called again, a dictatorship. So, it is about the government, so I think it's in there."). In this metacognitive episode, through discussing strategy selection, group members construct a better understanding of the task and text. Each metacognitive activity triggers a response from a group member to question or better explain an earlier comment. By building on each other's comments, these students were challenged to verbalize their assumptions about the task in relation to the text, setting an example of how co-constructed interaction can build up to deep-level metacognitive activities.

Table 6 also shows low-level metacognitive activities were significantly correlated with accepted ($r=0.74$, $p<0.01$) and co-constructed ($r=0.76$, $p<0.01$) intra-group interactions. An illustration of how accepted intra-group interactions facilitate low-level metacognitive activities is shown in Table 8. In this example a student opens with a question as suggested by the script ("Where is the relevant information in the text to answer the question?"), to which a group member replies with a vague answer ("Multiple places."). Prompted by this vague comment, a student asks for clarification ("How many places?"), which leads the group member to ask for more information about the task ("How many statements are there?"). The task guard then answers this previous question ("There are five statements."), which prompts the group member to respond with

Table 7 An example of co-constructed social metacognitive activities as initiated by the collaboration script during pre-task

Speech utterance	Main code	Sub-level code		
"I think it is targeted reading."	Metacognitive activitiy	Planning	Formulating a problem-solving plan	Low-level
"I'll fill that in."	Relational activitiy			
"Targeted is reading a paragraph, right?"	Metacognitive activitiy	Planning	Formulating a problem-solving plan	Low-level
"Yes, but you have to explain something here, so therefore it is not searched."	Metacognitive activitiy	Planning	Connecting task demands with planning	Deep-level
"So in how many places will the answer be then?"	Metacognitive activitiy	Orientation	Exploring content	Low-level
"That could be three."	Metacognitive activitiy	Orientation	Exploring content	Low-level
"Plan economy… wait…"	Metacognitive activitiy	Orientation	Exploring content	Low-level
"Maybe it is in the introduction."	Metacognitive activitiy	Orientation	Exploring content	Low-level
"Wait, I think 'the influence of the governments', because a plan economy is, what is that called again, a dictatorship. So it is definitely about the government, so I think it's in there."	Metacognitive activitiy	Planning	Connecting task demands with planning	Deep-level

Table 8 An example of accepted social metacognitive activities as initiated by the collaboration script during pre-task regulation

Speech utterance	Main code	Sub-level code		
"Where is the relevant information in the text to answer the question?"	Metacognitive activitiy	Scripted question		
"Multiple places."	Relational activitiy	Orientation	Exploring content	Low-level
"How many places?"	Metacognitive activitiy	Orientation	Exploring content	Low-level
"How many statements are there?"	Metacognitive activitiy	Orientation	Exploring task demands	Low-level
"There are five statements."	Metacognitive activitiy	Orientation	Exploring task demands	Low-level
"I think five places."	Metacognitive activitiy	Orientation	Exploring content	Low-level
"Five places, yes."	Metacognitive activitiy	Orientation	Exploring content	Low-level

an answer to the earlier question of how many places the relevant information would be ("I think five places."). Even though this answer remains very vague, it provides an answer to the question the writer has to answer during planning, namely *"On a scale from 1 to 10; 1 being in 1 place, and 10 being in many places—Where can you find the relevant information in the text to answer this assignment?"*, and in the end the other group members accept this answer with a simple reply ("Five places, yes."). This metacognitive episode shows the importance of the quality of group interaction. The group members do not build on previous comments, questions stated by group members are simply being answered in this example. None of the group members object to these answers or ask for more clarification, thereby metacognitive activities remain low-level.

For post-task regulation, Table 6 shows that low-level metacognitive activities were significantly correlated with shared ($r=0.86, p<0.01$) and co-constructed ($r=0.68, p<0.05$) intra-group interactions. An illustration of how shared intra-group interactions facilitate low-level metacognitive activities is shown in Table 9. In this example a student opens with a statement following another metacognitive episode where the group determined where the relevant information was located in the text. In the current episode, as is shown in Table Shared, the group starts discussing strategy reflection, a student opens this discussion by stating what reading strategy would have been best according to him/her, and by sharing what reading strategy they has used during task execution (*"Okay and then just focused reading I guess I don't know what most … I've read intensively myself."*). This statement is then followed by a question from a group member asking what the correct strategy would have been according to the script (*"What was the correct reading strategy?"*). The leader answers the question by giving the most efficient reading strategy according to the script (*"The first to … euuhm targeted reading… yes targeted."*). The text guard then follows this answer by stating what strategies group members have used, as is shown on their screen (*"Two have read intensive and two have targeted."*), to which another group member replies by sharing what reading strategy they have used (*"Yes, in the end I read intensively because… I read intensively anyway."*). This prompts another group member to also share their used strategy and thoughts on what strategy would have been best (*"Yes, I did too, but I do think that targeted fits better, I think I just did it wrong."*). After this, the group switches to another topic, and the metacognitive episode ends. This example shows how students agree on which strategy would have been best, but linger on sharing their own thoughts and actions, without constructing a shared idea

Table 9 An example of accepted social metacognitive activities as initiated by the collaboration script during post-task regulation

Speech utterance	Main code	Sub-level code	Strategy evaluation	Low-level
"Okay and then just targeted reading I guess I don't know what most ... I've read intensively myself."	Metacognitive activitiy	Reflection	Strategy evaluation	Low-level
"What was the correct reading strategy?"	Metacognitive activitiy	Scripted question		
"The first to ... euuhm targeted reading... yes targeted."	Metacognitive activitiy	Reflection	Strategy evaluation	Low-level
"Two have read intensive and two have targeted."	Metacognitive activitiy	Reflection	Strategy evaluation	Low-level
"Yes, in the end I read intensively because I read intensively anyway."	Metacognitive activitiy	Reflection	Strategy evaluation	Low-level
"Yes, I did too, but I do think that targeted fits better, I think I just did it wrong."	Metacognitive activitiy	Reflection	Strategy evaluation	Low-level

of what would have been the best strategy. Since there are no group members who question the metacognitive statements of another, this discussion remains at a low-level.

As expected, ignored metacognitive group interactions were not correlated with low-level nor with deep-level regulation (see Table 6). For example, a group member tried to open the conversation on which strategy they should use ("why are we going to use this strategy? We already see a header…"), which was ignored by group members discussing something off task. These group members completely ignored this attempt to select a strategy, which is a good example of how ignored intra-group interactions do not usually elicit any kind of metacognitive activity.

Individual task representation, strategy selection and reflection

To establish whether students improved in individual TOR regulation, we examined whether students improved on task representation, strategy selection, and strategy reflection. With respect to task representation, a paired-samples t-test showed that students significantly improved their task representation on the post-test, a statistically significant mean increase of 0.829 points, 95% CI [0.269, 1.389], t(37)=2.994, p 0.005, d=0.47. On average, students significantly improved on their task representation as indicated by a mean increase of 0.512 points 95% CI [0.112, 0.912], t(37)=2.588, p 0.013, d=0.40, which shows that during the post-test students more often indicated that a high-level task required more inferencing as compared to the pre-test. No significant differences were found for low- and intermediate tasks (see Table 10). Regarding strategy selection, a paired-samples t-test showed that students did not significantly improve on selecting a reading strategy. Finally, regarding strategy reflection, a paired-samples t-test showed that students significantly improved on reflecting which strategy could be used best in hindsight, a statistically significant mean increase of 0.737 points, 95% CI [0.156, 1.317], t(37)=2.572, p 0.014, d=0.42. On average students reflected significantly more often, with a mean increase of 0.816 points 95% CI [0.450, 1.181], t(37)=4.524, p<0.001, d=0.73, which shows that an intensive reading strategy would be best for a high-level task after the intervention on the post-test compared to the pre-test. No significant difference was found for reflection on low- and intermediate-level tasks (see Table 10).

Discussion

This design study examined shared regulation of TOR when working with a collaboration script among prevocational students, and investigated the ensuing improvement of individual pre- and post-task regulation during TOR. Results suggest that by creating roles with different responsibilities and thereby providing each member with different information to fulfill the task during pre- and post-task regulation, the collaboration script possibly elicited metacognitive regulation in the groups. A closer examination of the depth of metacognitive regulation, however, indicated that most contributions were low-level, while deep-level metacognitive contributions were scarce during pre-task regulation and virtually non-existent during post-task regulation. The sharedness of metacognitive episodes was mostly accepted intra-group interaction. Both during pre- and post-task regulation, group members merely accepted each other's suggestions. Only one-tenth of these episodes were co-constructed during pre-task regulation and substantially less during post-task regulation. Thus, even though the script seems to elicit metacognitive regulation, it did not successfully stimulate groups to engage in deep-level and shared metacognition. Surprisingly following the intervention, students did improve on individual metacognitive regulation, suggesting they did become more aware of different task complexities. In particular, they appear to have

Tabel 10 Pre- and post-test scores on task representation, strategy selection and strategy reflection, including effect sizes

		Pretest		Posttest		t(37)	p	Cohen's d
		M (SD)	95% CI	M (SD)	95% CI			
Task representation		3.98 (1.49)	[3.50, 4.45]	4.80 (1.27)	[4.40, 5.21]	2.994	0.005	0.468
	low-level task	2.24 (0.80)	[1.99, 2.50]	2.46 (0.79)	[2.22, 2.71]	1.388	0.173	0.217
	Intermediate-level task	0.78 (0.82)	[0.52, 1.04]	0.88 (0.78)	[0.63, 1.12]	0.561	0.578	0.088
	High-level task	0.95 (0.97)	[0.64, 1.26]	1.46 (0.93)	[1.17, 1.76]	2.588	0.013	0.404
Strategy selection		3.41 (1.45)	[2.96, 3.87]	3.71 (1.33)	[3.29, 4.13]	1.074	0.290	0.174
	low-level task	0.76 (0.86)	[0.48, 1.03]	0.71 (1.01)	[0.39, 1.02]	-0.380	0.706	-0.062
	Targeted	1.15 (1.12)	[1.16, 1.87]	1.29 (0.81)	[1.04, 1.55]	-1045	0.303	-0.169
	Intensive	1.15 (0.94)	[0.85, 1.44]	1.71 (0.96)	[1.41, 2.01]	3.240	0.003	0.526
Strategy reflection		3.39 (1.30)	[2.98, 3.80]	4.15 (1.20)	[3.77, 4.52]	2.572	0.014	0.417
	Searching	0.80 (0.87)	[0.53, 1.08]	0.83 (0.89)	[0.55, 1.11]	0.147	0.884	0.024
	Targeted	1.49 (0.93)	[1.20, 1.78]	1.46 (1.00)	[1.15, 1.78]	-0.487	0.629	-0.079
	Intensive	1.10 (0.86)	[0.83, 1.37]	1.85 (0.88)	[1.58, 2.13]	4.524	<0.001	0.734

become better at recognizing characteristics of high-level tasks and also better at reflecting on the best strategy for these tasks.

Depth and sharedness while working with a multi-level collaboration script

The main contribution of the study lies in the innovative design of the collaboration script, which provides different information to group members on tablets during face-to-face collaboration. The macro-level script aimed to support students active contributions to the discussion during pre-and post-task regulation of TOR. The script did elicit externalization of students' metacognitive regulation by making students dependent on each other for information (Kirschner et al., 2018). This suggests that the script did support information exchange by implementing roles with different resources. However, the transactivity within the group interaction was low. Students mainly just shared regulation and accepted each other's contributions without further justification. This implies that students did not elaborate on other group members' metacognitive contributions and the script did not successfully support the sharedness of metacognitive regulation (Järvelä & Hadwin, 2015; Molenaar et al., 2014). Previous research suggests that the depth of metacognitive regulation, also referred to as foci, mainly influence group performance (De Backer et al., 2016; Iiskala et al., 2021) regardless of the form of intra-group regulation. Nevertheless, it has previously been noted that deep-level metacognitive regulation is more common in forms of intra-group interaction where students build on each other's comments (Molenaar et al., 2014; Iiskala et al., 2021), which was also found in the present study. Where the macro script was developed to elicit sharedness, the micro script was developed to elicit depth in regulation. Given that most utterances were metacognitive in nature, the findings suggest that the micro-level script did encourage metacognitive regulation.

However, most metacognitive contributions were low-level and students did not combine task and text characteristics in their discussion. Students' contributions showed that they have knowledge about strategies, but they did not elaborate on why a particular strategy was relevant for a task. Nor did they explicate which information in the text prompted them to consider a particular strategy. A possible explanation is that the structured nature of the scripted questions was not challenging enough for students and therefore the script did not trigger many high-level metacognitive activities (Iiskala et al., 2015). However, within a few groups and during some occasions, high-level utterances did occur. For example, when looking at the blue graphs for each group in Appendix B, Group E and Group H displayed more co-constructive interaction compared to other groups. These two groups also displayed slightly more deep-level interaction in the grey graphs. However, these groups achieved average scores on group strategy selection and reflection in both graphs. A possible explanation might be that the students in these groups find the reading tasks they encounter more complex than the other groups, and therefore engage in more intra-group interaction in order to understand the assignment (Iiskala et al., 2021). In fact, this applies to other groups as well. Even though most groups displayed less overall co-constructive interaction during the lessons, on the occasions (e.g., group C during lesson 4 and group I during lesson 1) that they did display more co-construction, their group scores on the questions were comparably lower. This potentially suggests groups only engage in shared or co-constructed metacognitive interaction when they face challenges and feel group collaboration may support individual regulation (Iiskala et al., 2015). This may also help explain the results that students improved specifically in recognizing high-level tasks and reflecting on the best strategy to apply for these tasks, because even though overall group discussions involved mostly low-level metacognitive regulation, scores on the post-test did show that students improved on individual task representation and strategy reflection. In particular, with respect to high-level reading tasks,

students improved in recognizing task characteristics of high-level tasks and were better able to, in hindsight after completing the task, reflect on the appropriate strategy. Therefore, we cautiously suggest that even though the collaboration script did not trigger deep-level metacognition, it may have helped students to improve on some critical elements regulating TOR, such as task representation and strategy reflection.

Another explanation for the variability of sharedness and depth of intra-group interaction are individual differences. In the present study, we did not investigate the impact of learner characteristics on intra-group interaction. However, results from the blue interaction graphs do display differences in intra-group interaction. This could be explained by individual differences, as previous research has shown that student participation usually varies within groups (Iiskala et al., 2015). Individual learner characteristics, such as metacognitive knowledge or reading ability, could have influenced the depth of interactions and/or sharedness of the groups' interaction. Students had to create their own representation of a task and text to meaningfully contribute to the groups' discussion. In doing so, each student brought prior knowledge regarding reading strategies, the content, self-efficacy and (intrinsic) motivation, into a collaboration. Indeed, De Backer et al. (2022) were able to distinguish three regulation profiles, such as the 'all-round-oriented and affirming regulator' (AOAR), the 'social-oriented and elaborating regulator' (SOER) and the 'individual-oriented and passive regulators' (IOPR), through the clustering of data, gathered from different learner characteristics, such as metacognitive regulation, prior knowledge, task performance, motivation and self-efficacy. AOAR and SOER identified students were intrinsically motivated to learn, with the difference being that AOAR's had moderate self-efficacy beliefs and performance, while SOER's were high achievers with complementary self-efficacy beliefs. IOPR's were not intrinsically motivated to learn, they doubt their abilities and perform poorly on the content. In contrast to our design study, this study was conducted among university students. Hence it would be interesting to find out if the same profiles apply for our prevocational students. In line with De Backer and colleagues (2022), we would argue that gaining more insight into individual differences offers an opportunity to differentiate instruction given to students based on their personal strengths and weaknesses. The script in the present study, did not leave much space for spontaneous discussions, which may have hampered the groups' natural socially shared regulation (Wise & Schwarz, 2017). For example, the questions during pre- and post-task regulation were aimed at creating a task representation, strategy selection and strategy reflection and may therefore not have challenged students to construct their own metacognitive thinking.

Limitations

A major limitation of the present study is its small sample size. For this design study, we have prioritized executing the design study in an authentic classroom situation, which restricted the size. This situated design study does align with ideas that learners' interaction is inextricably linked with the social context in which learning takes place (Grau & Whitebread, 2012). Additionally, the extensiveness of the collected data and the comprehensive examination of the depth and sharedness of groups' discussions also led to size restrictions. Based on the results of this design study, we formulated a number of improvements for a follow up experimental study to ascertain possible assumptions about causal relationships. Since previous studies indicate that students benefit from structure during collaboration around TOR (Okkinga et al., 2021; Hämäläinen et al., 2008), we would recommend keeping the role and information division of the script in a new study to maintain interdependency among group members. However, the script and the questions asked of each role must be critically examined and the difficulty level of the TOR tasks should be higher. One way to improve the script is to formulate the questions in the scriptlets in a less

structured form. A study by Molenaar et al. (2014) showed promising results for eliciting socially shared regulation with problematizing scaffolds which are more open in nature. These problematizing scaffolds consisted of questions that prompted students to explain and articulate their metacognitive thinking, leading to more co-constructed metacognitive activities. Students in this study were also less inclined to ignore group members when provided with problematizing scaffolds. With respect to more difficult TOR tasks, we aim to include more challenging tasks in a future study since in the current study students did seem to value group collaboration. Therefore, in our future research we will adjust the design of a collaboration script with less structured questions and examine the influence of such a revised script on metacognitive regulation and social sharedness on regulation of TOR. In addition, we would recommend examining the influence of individual learner characteristics on intra-group interaction and gather information regarding the influence of prior knowledge concerning reading strategies, the content, self-efficacy and (intrinsic) motivation.

Conclusions

The paper concludes by arguing that despite the limitation of the small sample size, the comprehensiveness of the data suggests several directions for future research. For example, the study examined the depth and social sharedness of metacognitive regulation while working with a collaboration script that was designed to offer support on a micro- and macro-level. While the depth of the discussion was mainly low-level, the nature of the utterances shows that it is not unlikely that the micro-level script supported students in verbalizing metacognitive activities. In addition, the macro-script seemed to be able to support intra-group interaction by creating an information dependency. However, given the limitations of the present study, we present these results with some caution and suggest probing these results further in an experimental study with more power to establish a causal relation between the script and depth and sharedness of regulation. Nonetheless, all groups engaged in accepted interaction most of the time, even though co-construction did occur a few times in some groups. Follow up studies should therefore explore the influence of more complex tasks on the depth and sharedness of metacognitive regulation (Iiskala et al., 2015). Another direction is to examine the influence of less structured scripts on group discussions in comparison to more structured scripts, to determine whether less structured scripts leave more opportunity for natural socially shared regulation (Wise & Schwarz, 2017). These future studies should also incorporate more analyses regarding individual characteristics and the influence these characteristics have on participation within the group (De Backer et al., 2022). In addition, calls should be made to execute these studies in more authentic (classroom) situations, including multiple measurements over time. To conclude, with the innovative design of the collaboration script, combined with the comprehensive decomposition of groups' conversations collected in an authentic classroom setting, the present study sheds more light on the socially shared regulation of prevocational students, an underexposed group of students, in a computer-supported collaborative learning setting.

Appendix A

Macro-level script

The macro-level script in this study fosters shared pre- and post-task regulation and individual reading following Järvelä et al. (2016) and combines the components 'play', 'scene' and 'roles' (following Fisher et al., 2013):

- 'Play' details the main goal for collaboration, specifically the shared regulation of TOR.
- 'Roles' sets the condition for sharedness of regulation by creating information dependency among group members. In our study four roles are defined, see Fig. 4, that of (A) the leader, (B) the writer, (C) the task guard and (D) a text guard. The leader (A) guided discussions and had information about the task and reading strategies. The writer (B) was in charge of writing down group decisions on shared task representation, strategy selection and strategy reflection. The task guard (C) was given the task and was instructed to discuss task characteristics with group members. Finally, the text guard (D) was given the text and was instructed to discuss content characteristics with group members. These roles transgress 'scenes' as set by the macro script.
- 'Scene' creates a sequence of activities that supports shared regulation of TOR (i.e. 'play') during pre- and post-task regulation. Figure 4 shows how TOR is divided in four phases, 1. Planning, 2. Task execution, 3. Reflection and 4. Reflection. To facilitate shared regulation, while also leaving room for individual differences in reading, only phase 1. Planning and 4. Reflection are synchronized so that the group can discuss pre-and post-task regulation of TOR.

Micro-level script

The micro-level script in this study consists of 'scriptlets' and is built to trigger students to externalize their perceptions of the task and consequently build a shared understanding of task complexity and strategy selection (following Järvelä & Hadwin, 2015). These 'scriples' are integrated in the script to support pre- and post-task regulation (Fischer et al., 2013):
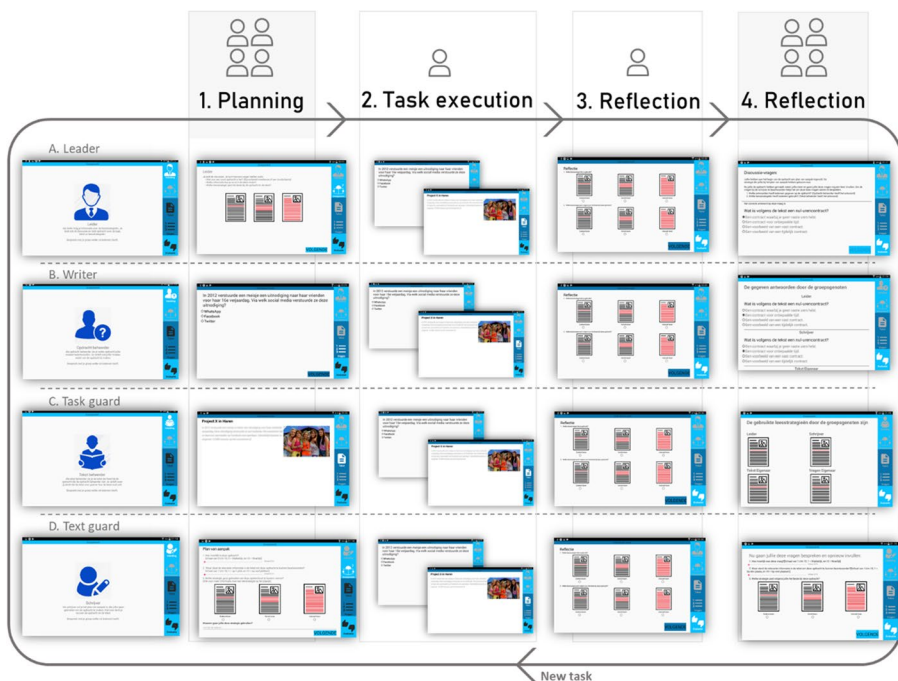
Fig. 4 Task flow of the 'Y-read?' collaboration script in icons and in screenshots

- 'Scriptlets' arrange activities within a 'scene' in order for students to construe knowledge about the sequencing of metacognitive regulation during TOR. For example in the pre-task regulation phase students create a shared task representation in order to select a reading strategy. The content of the 'scriptlets' are dependent on the 'role' (i.e. leader, writer, task guard or text guard) and the 'scene' (i.e. planning or reflection).

     During the planning phase students create a representation of the task based on these characteristics (Rouet et al., 2017; Winne & Hadwin, 1998). This ideally involves making predictions about what information is required to fulfil the task and if this information could be found on a local or global level. In addition students would have to assess the level of inferencing necessary to process this information (Van Steensel et al., 2013). Accordingly, they then use this information to select an appropriate reading strategy. The 'scriptlets' in the planning phase arranges the activities in a way that students jointly orientate on characteristics of the task and text before selecting a reading strategy to accomplish the task. See Fig. 5 for an example of the 'Y-read?'-environment and the information each role was provided with.

     Following planning the 'scene' guides the students to task execution (see Fig. 4), where students individually read the text and complete the task. In addition, students individually reflect on their used strategy and what they thought would have been the best strategy in hindsight. Students had to reflect on the task and strategy use individually, to stimulate students to create their own perceptions of a task before entering the group discussion again.

     After task completion students are guided to the group reflection (see Fig. 4, phase '4. Reflection'). During post-task group reflection, students jointly evaluate and explicitly discuss why their perception of the task was correct or incorrect in hindsight. In addition they have to align their arguments and evaluate how they reflect on the task and strategy use in hindsight. See Fig. 6 for an example of the 'Y-read?'environment and the information each role was provided with.

Appendix B

For each group we have plotted two interaction graphs. The graphs display the shareness and depth of group discussions. In addition the lines display individual and group performance on selecting and/or reflecting on a strategy.

   The blue graphs display shareness of each groups interaction for each lesson.

- The x-axis displays the number of each lesson.
- The left side y-axis displays on a scale from 0 to 1 the percentiles for the shareness of interaction.
- The right side y-axis displays on a scale from 0 to 4, the number of correct answers given on the individual reflection question.

   The area chart displays the shareness of groups' intra-group interaction. From dark blue to light blue, the darkest shade of blue represents co-constructed interaction, followed by shared interaction, thereafter accepted interaction, while the lightest shade of blue represents ignored interaction. In addition the lines show each group members individual reflection on what they thought would have been the best strategy (see Fig. 5 in the method section, column '3. Reflection'). For each correct given answer on the question "which reading strategy would have been best in hindsight for this task?" a student was granted a point, with a maximum of four points for each lesson.

   Legend:

Fig. 5 An example of the 'Y-read?'-environment and the information each role was provided with during ▶ planning

▮	Co-constructed interaction
▮	Shared interaction
▮	Accepted interaction
▮	Ignored interaction

Each line represents the number of correct answers for each group member on the question: *"which reading strategy would have been best in hindsight for this task?"*.
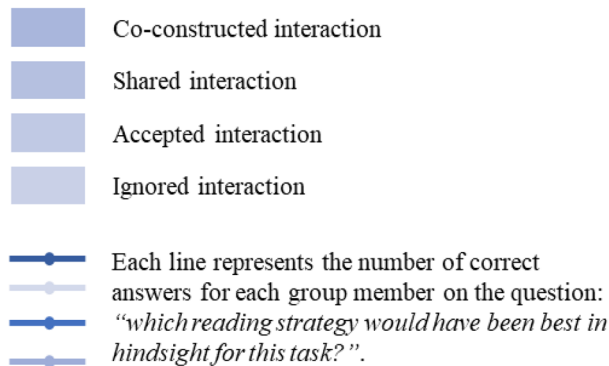
The grey graphs display the depth of group's discussion for each lesson.

• The x-axis displays the number of each lesson.
• The left side y-axis displays on a scale from 0 to 4, the number of correct answers given on the group's strategy selection and reflection questions.
• The right side y-axis displays on a scale from 0 to 1 the percentiles for the depth of groups' discussion.

The bars in the grey graphs show the depth of groups' discussion for each lesson in percentiles, while the lines show the groups correct answer on the selected and reflected strategy (see Fig. 5 in the method section, column '1. Planning' and '4. Reflection'). For each correct given answer on the questions: "Which reading strategy is best for the assignment and the text?" and "which reading strategy would have been best in hindsight for this task?" the group was granted a point, with a maximum of four points for each lesson.

Legend:

▮	Low-level metacognitive activities
▮	Deep-level metacognitive activities
—●—	Correct selected strategy
—●—	Correct reflected strategy
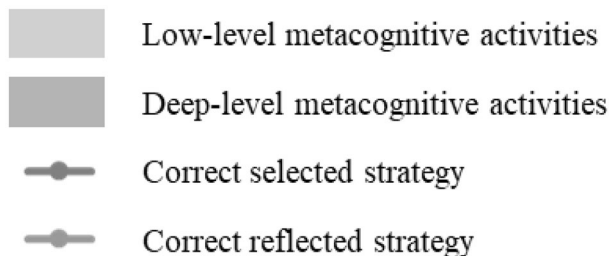
Leader

Leider

Je leidt de discussie. Je kunt hiervoor vragen stellen zoals:
- Wat voor een soort opdracht is het? (Bijvoorbeeld meerkeuze of een invulschema)
- Welke informatie kun je nu al in de tekst vinden?
- Welke leesstrategie past het beste bij de opdracht en de tekst?

VOLGENDE

Groepswerk

Leider · Inleiding · Groepswerk · Tekst · Vragen · Evaluatie

Writer

Planning

1. How difficult will this assignment be?
Scale from 1 to 10; 1 = in 1 place, and 10 = in many places
Schaal 0/10

2. Where can you find the relevant information in the text to answer this assignment?
Scale from 1 to 10; 1 = in 1 place, and 10 = in many places
Schaal 0/10

3. Which strategy are you going to use to execute the task?
(Click on the picture for more information about a reading strategy)

Search strategy   Targeted strategy   Intensive strategy
○   ○   ○

4. While reading the text…
☐ I can probably literally find the information I am looking for.
☐ I probably have to think about the information to make the assignment.

Next

Introduction · Planning · Text · Task · Reflection

Task guard

What is a zero-hours contract according to the text?
○ A contract where you have no fixed hours.
○ An open-ended contract.
○ An example of a permanent contract.
○ An example of a temporary contract.

Next

Introduction · Planning · Text · Task · Reflection

Text guard

Working as an employee

Different types of contracts

Disadvantages of a flexible contract

Introduction · Planning · Text · Task · Reflection

Leader

Discussie-vragen:

Jullie hebben aan het begin van de opdracht een plan van aanpak ingevuld. De strategie die jullie bij het plan van aanpak hebben gekozen was:

Nu jullie de opdracht hebben gemaakt weten jullie meer en gaan jullie deze vragen nog een keer invullen. Om de vragen bij de schrijver te beantwoorden helpt het om deze twee vragen samen te bespreken:

1. Welke antwoorden heeft iedereen gegeven op de opdracht? (Opdracht beheerder heeft het antwoord)
2. Welke leesstrategieën heeft iedereen gebruikt? (Tekst beheerder heeft het antwoord)

Het correcte antwoord op deze vraag is

Wat is volgens de tekst een nul-urencontract?

⦿ Een contract waarbij je geen vaste uren hebt.
◯ Een contract voor onbepaalde tijd.
◯ Een voorbeeld van een vast contract.
◯ Een voorbeeld van een tijdelijk contract.

VOLGENDE

Writer

Planning

1. How difficult will this assignment be?
Scale from 1 to 10; 1 = in 1 place, and 10 = in many places
Schaal 0/10

2. Where can you find the relevant information in the text to answer this assignment?
Scale from 1 to 10; 1 = in 1 place, and 10 = in many places
Schaal 0/10

3. Which strategy are you going to use to execute the task?
(Click on the picture for more information about a reading strategy)

Search strategy ◯   Targeted strategy ◯   Intensive strategy ◯

Next

Task guard

De gegeven antwoorden door de groepsgenoten

Leider

Wat is volgens de tekst een nul-urencontract?

◯ Een contract waarbij je geen vaste uren hebt.
⦿ Een contract voor onbepaalde tijd.
◯ Een voorbeeld van een vast contract.
◯ Een voorbeeld van een tijdelijk contract.

Schrijver

Wat is volgens de tekst een nul-urencontract?

◯ Een contract waarbij je geen vaste uren hebt.
⦿ Een contract voor onbepaalde tijd.
◯ Een voorbeeld van een vast contract.
◯ Een voorbeeld van een tijdelijk contract.

Tekst Eigenaar

Text guard

De gebruikte leesstrategieën door de groepsgenoten zijn

Leider   Schrijver

Tekst Eigenaar   Vragen Eigenaar

Fig. 6 An example of the 'Y-read?'-environment and the information each role was provided with during reflection

Group A
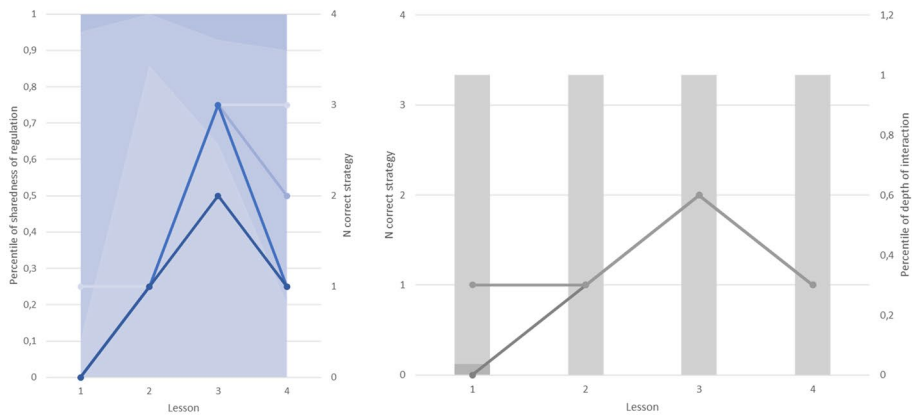
Group B
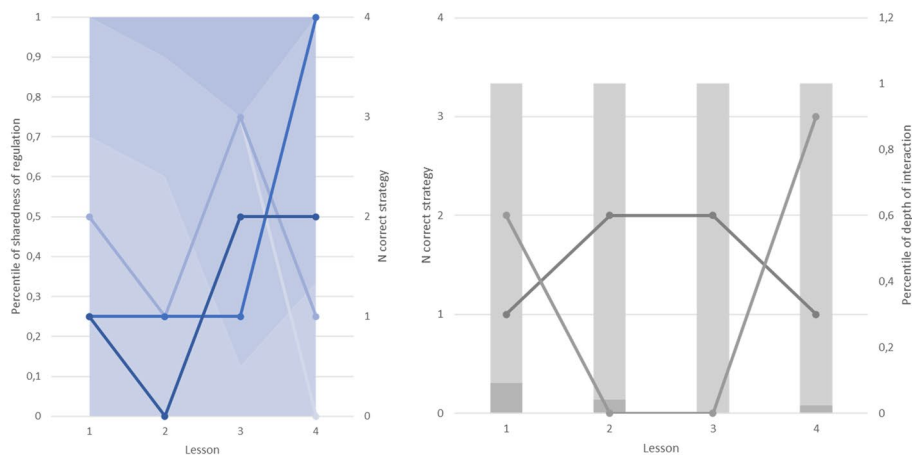
Group C
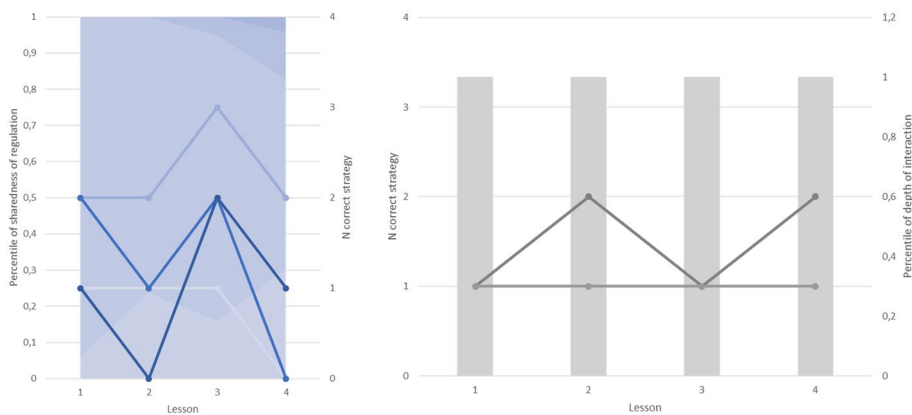
Group D
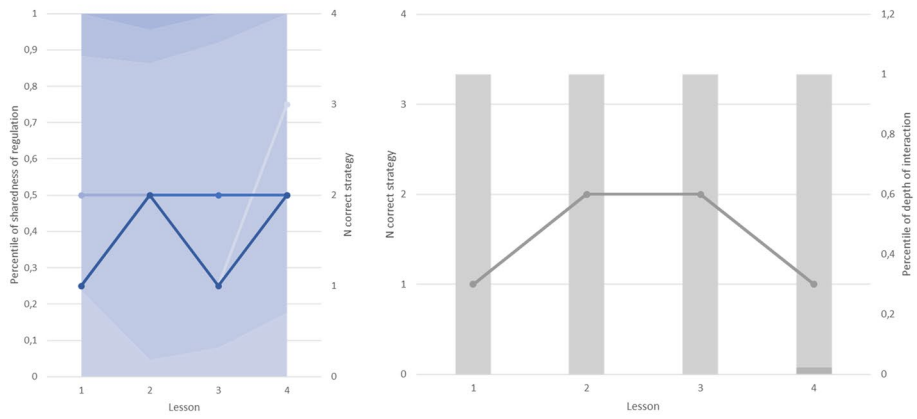
Group E
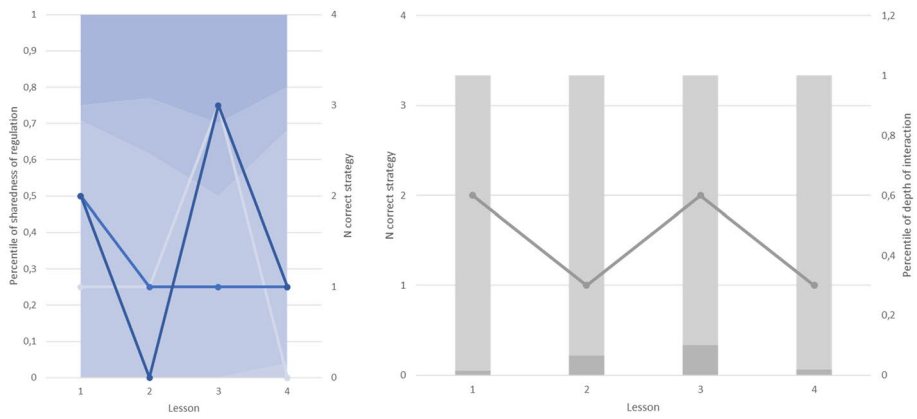
Group F
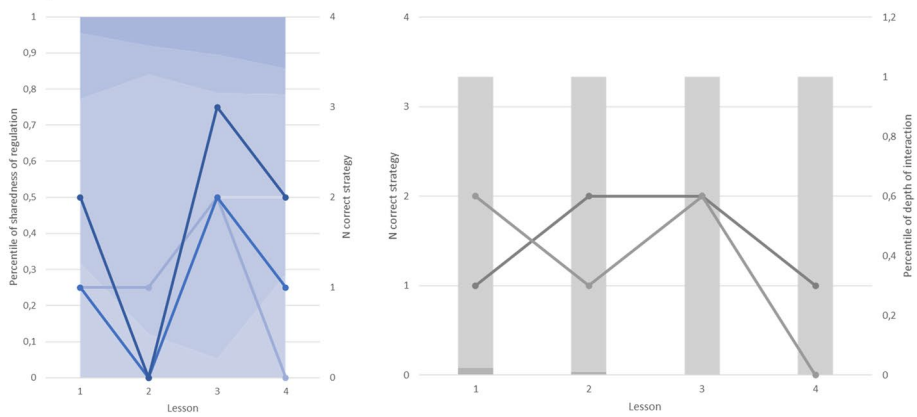
Group G
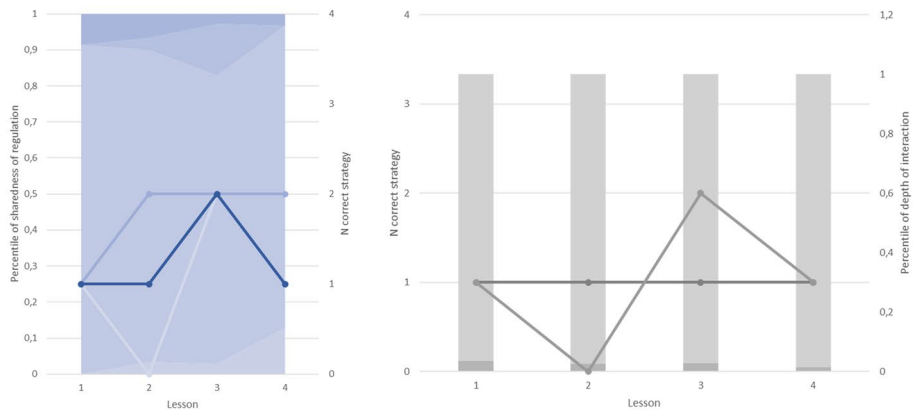
Group H

Group I

Group J

Group K
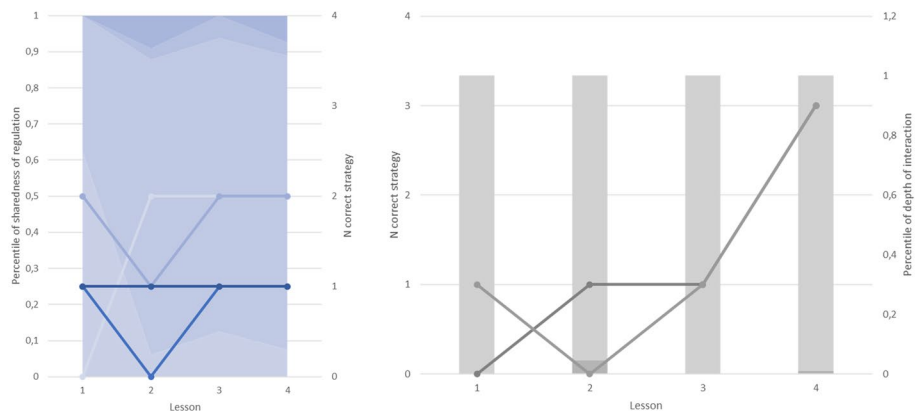
Authors' contributions All authors contributed to the study conception and design. Material preparation, data collection and analysis were performed by Jolique Kielstra, Inge Molenaar and Roel van Steensel. The first draft of the manuscript was written by Jolique Kielstra and all authors commented on previous versions of the manuscript. All authors read and approved the final manuscript.

Availability of data and material Not applicable'.

Code availability 'Not applicable'.

Declarations

Conflicts of interest/Competing interests The authors declare that they have no conflict of interest.

References

Anmarkrud, Ø., McCrudden, M. T., Bråten, I., & Strømsø, H. I. (2013). TOR of multiple documents: Online comprehension processes and offline products. *Instructional Science, 41*, 873–894. https://doi.org/10.1007/s11251-013-9263-8

Butler, D. L. (2002). Qualitative approaches to investigating self-regulated learning: Contributions and challenges. *Educational Psychologist, 37*(1), 59–63. https://doi.org/10.1207/S15326985EP3701_7

De Backer, L., Van Keer, H., Moerkerke, B., & Valcke, M. (2016). Examining evolutions in the adoption of metacognitive regulation in reciprocal peer tutoring groups. *Metacognition Learning, 11*, 187–213. https://doi.org/10.1007/s11409-015-9141-7

De Backer, L., Van Keer, H., De Smedt, F., Merchie, E., & Valcke, M. (2022). Identifying regulation profiles during computer-supported collaborative learning and examining their relation with students' performance, motivation, and self-efficacy for learning. *Computers & Education, 179*, 104421. https://doi.org/10.1016/j.compedu.2021.104421

De Milliano, I., van Gelderen, A., & Sleegers, P. (2016). Types and sequences of self-regulated reading of low-achieving adolescents in relation to reading task achievement. *Journal of Research in Reading, 39*(2), 229–252. https://doi.org/10.1111/1467-9817.12037

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

Gil, L., Martinez, T., & Vidal-Abarca, E. (2015). Online assessment of strategic reading literacy skills. *Computers & Education, 82*, 50–59. https://doi.org/10.1016/j.compedu.2014.10.026

Grau, V., & Whitebread, D. (2012). Self and social regulation of learning during collaborative activities in the classroom: The interplay of individual and group cognition. *Learning and Instruction, 22*(6), 401–412. https://doi.org/10.1016/j.learninstruc.2012.03.003

Hämäläinen, R., & Oksanen, K.,&Häkkinen, P. (2008). Designing and analyzing collaboration in a scripted game for vocational education. *Computers in Human Behavior, 24*(6), 2496–2506. https://doi.org/10.1016/j.chb.2008.03.010

Iiskala, T., Volet, S., Lehtinen, E., & Vauras, M. (2015). Socially shared metacognitive regulation in asynchronous CSCL in science: Functions, evolution and participation. *Frontline Learning Research, 3*(1), 78–111. https://doi.org/10.14786/flr.v3i1.159

Iiskala, T., Volet, S., Jones, C., Koretsky, M., & Vauras, M. (2021). Significance of forms and foci of metacognitive regulation in collaborative science learning of less and more successful outcome groups in diverse contexts. *Instructional Science, 49*(5), 687–718. https://doi.org/10.1007/s11251-021-09558-1

Järvelä, S., & Hadwin, A. (2015). Promoting and researching adaptive regulation: New Frontiers for CSCL research. *Computers in Human Behavior, 52*, 559–561. https://doi.org/10.1016/j.chb.2015.05.006

Järvelä, S., Kirschner, P. A., Hadwin, A., Järvenoja, H., Malmberg, J., Miller, M., & Laru, J. (2016). Socially shared regulation of learning in CSCL: Understanding and prompting individual-and group-level shared regulatory activities. *International Journal of Computer-Supported Collaborative Learning, 11*(3), 263–280. https://doi.org/10.1007/s11412-016-9238-2

Kirschner, P. A., Sweller, J., Kirschner, F., & Zambrano, J. (2018). From cognitive load theory to collaborative cognitive load theory. *International Journal of Computer-Supported Collaborative Learning, 13*(2), 213–233. https://doi.org/10.1007/s11412-018-9277-y

Lombard, M., Snyder-Duch, J., & Bracken, C. C. (2002). Content analysis in mass communication: Assessment and reporting of intercoder reliability. *Human Communication Research, 28*(4), 587–604. https://doi.org/10.1111/j.1468-2958.2002.tb00826.x

Malmberg, J., Järvenoja, H., & Järvelä, S. (2013). Patterns in elementary school students′ strategic actions in varying learning situations. *Instructional Science, 41*(5), 933–954. https://doi.org/10.1007/s11251-012-9262-1

Miller, M., & Hadwin, A. (2015). Scripting and awareness tools for regulating collaborative learning: Changing the landscape of support in CSCL. *Computers in Human Behavior, 52*, 573–588. https://doi.org/10.1016/j.chb.2015.01.050

Molenaar, I., van Boxtel, C., & Sleegers, P. (2011). Metacognitive scaffolding in an innovative learning arrangement. *Instructional Science, 39*, 785–803. https://doi.org/10.1007/s11251-010-9154-1

Molenaar, I., Sleegers, P., & van Boxtel, C. (2014). Metacognitive scaffolding during collaborative learning: A promising combination. *Metacognition and Learning, 9*, 309–332. https://doi.org/10.1007/s11409-014-9118-y

OECD. (2003). *The PISA 2003 assessment framework: Mathematics, reading, science and problem solving knowledge and skills*. OECD. https://doi.org/10.1787/9789264101739-en

Okkinga, M., van Gelderen, A. J. S., van Schooten, E., van Steensel, R., & Sleegers, P. J. C. (2021). Implementation quality of principles of reciprocal teaching in whole-classroom settings: A two-year study with low-achieving adolescents. *Reading Psychology*. https://doi.org/10.1080/02702711.2021.1887019

Paans, C., Molenaar, I., Segers, E., & Verhoeven, L. 2020. Children's macro-level navigation patterns in hypermedia and their relation with task structure and learning outcomes. *Frontline Learning Research*, *8*(1), 76–95. https://doi.org/10.14786/flr.v8i1.473

Panadero, E. (2017). A review of self-regulated learning: Six models and four directions for research. *Frontiers in Psychology, 8*, 1–28. https://doi.org/10.3389/fpsyg.2017.00422

Radkowitsch, A., Vogel, F., & Fischer, F. (2020). Good for learning, bad for motivation? A meta-analysis on the effects of computer-supported collaboration scripts. *International Journal of Computer-Supported Collaborative Learning, 15*, 5–47. https://doi.org/10.1007/s11412-020-09316-4

Rouet, J., Vidal-abarca, E., Erboul, B., & Millogo, V. (2001). Effects of information search tasks on the comprehension of instructional text. *Discourse Processes, 31*(2), 163–186. https://doi.org/10.1207/S15326950DP3102_03

Rouet, J. F., Britt, M. A., & Durik, A. M. (2017). RESOLV: Readers' representation of reading contexts and tasks. *Educational Psychologist, 52*(3), 200–215. https://doi.org/10.1080/00461520.2017.1329015

Salmerón, L., Naumann, J., García, V., & Fajardo, I. (2016). Scanning and deep processing of information in hypertext: An eye-tracking and cued retrospective think-aloud study. *Journal of Computer Assisted Learning, 33*(3), 222–233. https://doi.org/10.1111/jcal.12152

Serrano, M. Á., Vidal-Abarca, E., & Ferrer, A. (2018). Teaching self-regulation strategies via an intelligent tutoring system (TuinLECweb): Effects for low-skilled comprehenders. *Journal of Computer Assisted Learning, 34*(5), 515–525. https://doi.org/10.1111/jcal.12256

Strijbos, J. W., & De Laat, M. F. (2010). Developing the role concept for computer-supported collaborative learning: An explorative synthesis. *Computers in Human Behavior, 26*(4), 495–505. https://doi.org/10.1016/j.chb.2009.08.014

Taylor, K., & Baek, Y. (2019). Grouping matters in computational robotic activities. *Computers in Human Behavior, 93*, 99–105. https://doi.org/10.1016/j.chb.2018.12.010

Teasley, S. (1997). Talking about reasoning: How important is the peer in peer collaboration?. In L. B. Resnick, R. Saljo, C. Pontecorvo, & B. Burge (Eds.), *Discourse, tools and reasoning: Essays on situated cognition* (vol. 160, pp. 361–384). Springer. https://doi.org/10.1007/978-3-662-03362-3_16

Van Steensel, R., Oostdam, R., & Van Gelderen, A. (2013). Assessing reading comprehension in adolescent low achievers: Subskills identification and task specificity. *Language Testing, 30*(1), 3–21. https://doi.org/10.1177/0265532212440950

Vidal-Abarca, E., Gilabert, R., & Rouet, J. F. (1998, July). *The role of question type on learning from scientific text* [Paper presentation]. Meeting on Comprehension and Production of Scientific Texts, Aveiro, Portugal.

Vidal-Abarca, E., Mañá, A., & Gil, L. (2010). Individual differences for self-regulating task-oriented reading activities. *Journal of Educational Psychology, 102*(4), 817–826. https://doi.org/10.1037/a0020062

Volet, S., Summers, M., & Thurman, J. (2009). High-level co-regulation in collaborative learning: How does it emerge and how is it sustained?. *Learning and Instruction, 19*(2), 128-143. https://doi.org/10.1016/j.learninstruc.2008.03.001

Weinberger, A., Reiserer, M., Ertl, B., Fischer, F., & Mandl, H. (2005). Facilitating collaborative knowledge construction in computer-mediated learning environments with cooperation scripts. In R. Bromme, F.W.

Hesse, & H. Spada (Eds.), *Barriers and biases in computer-mediated knowledge communication* (vol. 5, pp. 15–37). Springer. https://doi.org/10.1007/0-387-24319-4_2

Winne, P., & Hadwin, A. F. (1998). Studying as self-regulated learning. In D. Hacker, J. Dunlosky, & A. Graesser (Eds.), *Metacognition in educational theory and practice* (pp. 277–304). Lawrence Erlbaum.

Wise, A. F., & Schwarz, B. B. (2017). Visions of CSCL: Eight provocations for the future of the field. *International Journal of Computer-Supported Collaborative Learning, 12*(4), 423–467. https://doi.org/10.1007/s11412-017-9267-5

Zimmerman, B. J. (2008). Investigating self-regulation and motivation: Historical background, methodological developments, and future prospects. *American Educational Research Journal, 45*(1), 166–183. https://doi.org/10.3102/0002831207312909