

# Design and effects of representational scripting on group performance

## Citation for published version (APA):

Slof, B., Erkens, G., Kirschner, P. A., & Jaspers, J. (2010). Design and effects of representational scripting on group performance. *Educational Technology, Research and Development*, 58(5), 589-608.  
<https://doi.org/10.1007/s11423-010-9148-3>

## DOI:

[10.1007/s11423-010-9148-3](https://doi.org/10.1007/s11423-010-9148-3)

## Document status and date:

Published: 01/10/2010

## Document Version:

Peer reviewed version

## Document license:

CC BY-SA

## Please check the document version of this publication:

- A submitted manuscript is the version of the article upon submission and before peer-review. There can be important differences between the submitted version and the official published version of record. People interested in the research are advised to contact the author for the final version of the publication, or visit the DOI to the publisher's website.
- The final author version and the galley proof are versions of the publication after peer review.
- The final published version features the final layout of the paper including the volume, issue and page numbers.

[Link to publication](#)

## General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal.

If the publication is distributed under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license above, please follow below link for the End User Agreement:

<https://www.ou.nl/taverne-agreement>

## Take down policy

If you believe that this document breaches copyright please contact us at:

[pure-support@ou.nl](mailto:pure-support@ou.nl)

providing details and we will investigate your claim.

Downloaded from <https://research.ou.nl/> on date: 30 Nov. 2022

Open Universiteit  
[www.ou.nl](http://www.ou.nl)



# Design and Effects of Representational Scripting on Group Performance

*B. Slof<sup>1</sup>, G. Erkens<sup>1</sup>, P. A. Kirschner<sup>2</sup>, & J. Jaspers<sup>1</sup>*

*<sup>1</sup>Utrecht University, P.O. Box 80.140, 3508 TC Utrecht, the Netherlands*

*<sup>2</sup>Open University of the Netherlands, P.O. Box 2960, 6401 DL Heerlen, the Netherlands*

*Correspondence address first author:*

*B. Slof*

*Utrecht University*

*Faculty of Social Sciences*

*Department of Educational Sciences*

*Research Centre Learning in Interaction*

*PO Box 80.140, 3508 TC Utrecht, the Netherlands*

*Tel: +31 30 2533445 | Fax: +31 30 2532352 | E-mail: B.Slof@uu.nl*

**Abstract:** This study investigated the effects of representational scripting on non-expert student learning while collaboratively carrying out complex learning-tasks. The premise underlying this research is that effective cognitive activities would be evoked when complex learning-tasks are structured into phase-related part-tasks and are supported by providing students with part-task-congruent external representations for each phase; representational scripting. It was hypothesized that this approach would lead to increased individual learning and better complex learning-task performance. In groups, 96 secondary education students worked on a complex business-economics problem in four experimental conditions, namely one condition in which the groups received representations that were part-task-congruent for all three phases and three conditions in which the groups received one of these representations for all three phases (i.e., part-task-incongruent for two of the three phases). The results indicate that groups receiving part-task-congruent representations in a phased order performed better on the complex learning-task, though this did not result in increased individual learning.

**Keywords:** *External representations, Complex learning-tasks, Computer Supported Collaborative Learning, Representational scripting*

## Introduction

Research on computer supported collaborative learning (CSCL) has shown that computer technology can provide support for students collaboratively carrying out complex learning-tasks. For such learning-tasks, groups of students often carry out different kinds of activities in two different dialogue spaces. In the *content space*, students carry out cognitive activities to deal with the phase-related part-task such as orienting themselves to the problem, finding solutions to the problem and evaluating the solutions found. In the *relational space*, students carry out communicative activities such as making their own knowledge and ideas explicit to others, creating shared understanding with the other group members, and negotiating multiple perspectives with the others (Barron, 2003; Janssen, 2008). To this end, students can be:

- stimulated to externalize their knowledge and ideas through chat and representational tools (e.g., Fisher, Bruhn, Gräsel, & Mandl, 2002),
- provided with scaffolding for their learning through scripting and representational guidance tools that structure the learning process (Reiser, 2004, Suthers, 2006), and/or
- offered offloading possibilities through the availability of storage spaces for contributions, external representations, and/or external information sources, all of which leave more working memory capacity for (part or whole) task completion (e.g., Hollan, Hutchins, & Kirsh, 2000).

These studies, though very valuable and informative, neglect the fact that complex learning-tasks are usually composed of fundamentally different phase-related part-tasks, each of which needs to be supported by different tools for them to be properly carried out (e.g., Van Bruggen, Boshuizen, & Kirschner, 2003). This is not only the case for CSCL, but is true of all complex learning-tasks. For carrying out such learning-tasks, students need to be supported in (1) dealing with the phase-related part-task demands and carrying out activities endemic to the different part-tasks in the proper sequence, (2) acquiring and applying well-suited problem representations for each part-task, and (3) combining these different problem representations into a whole (Ploetzner, Fehse, Kneser, & Spada, 1999; Spector, 2008; Van Merriënboer, Kester, & Paas, 2006; Van Merriënboer & Kirschner, 2007).

The study reported on here is aimed at designing a CSCL-environment to support students in successfully carrying out complex learning-tasks. By scripting the completion-process with representational tools (i.e., *representational scripting*), collaborative cognitive activities beneficial for carrying out the different part-task are evoked. The goal of this study is to determine whether integrating *scripting* and *representational tools* leads to a better use of available computer technology, and specifically whether this representational scripting affects both complex learning-task performance and individual learning gains in CSCL. In this article we speak of students, since the CSCL-environment is intended to be used in an educational setting where students learn collaboratively by carrying out complex learning-tasks, but the design of the representational scripting and its use might also be beneficial for all those involved in carrying out complex learning-tasks that have a part-task structure.

## Representational Scripting

Representational scripting entails the integration of scripting and representational tools whereby the different phase-related part-tasks of a problem-solving process are made explicit and are sequenced for students which in turn leads to the evocation and application of specific problem representations. This part-task-related support is intended to guide students when carrying out complex learning-tasks, leading to more successful complex learning-task performance and better solutions.

*Scripting* is intended to structure the completion-process to make it more efficient and effective. According to Dillenbourg (2002) a script is “a set of instructions regarding to how the group members should interact, how they should collaborate and how they should solve the problem” (p. 64). Such scripting entails the segmentation of a complex problem in distinct phases for discussion, solution and evaluation, with distinct purposes of each phase for the problem solving process (Beers, Boshuizen, Kirschner, & Gijsselaers, 2005; Dillenbourg; O’Donnell & Dansereau, 1992). The script structures the complex learning-task by dividing it into a sequence of ontologically distinct problem-phases (i.e., problem orientation, problem solution, solution evaluation) so that they can be provided with representations congruent with the part-task demands and activities required for each phase (Duffy, Dueber, & Hawley, 1998; Van Bruggen, et al., 2003).

The *representational tools* are meant to provide different views (i.e., problem representations) of the knowledge domain in which the complex learning-task is situated. Visualizing the domain by providing external representations (ERs) influences students’ cognitive behavior through their representational guidance (Ertl, Kopp, & Mandl, 2008; Suthers, 2006). Due to its ontology (i.e., objects, relations, and rules for combining them) every ER offers a restricted view of the domain making it easier to express certain aspects of that domain (Brna, Cox, & Good, 2001; Van Merriënboer & Kirschner, 2007). By matching the representational guidance of the ERs with the phase-related part-tasks, student understanding and part-task-specific activity should increase. However, an ER is seldom effective for all part-task demands and activities (Schnotz & Kirschner, 2008; Van Bruggen, et al., 2003). Carrying out complex learning-tasks requires students to create different perspectives of the whole learning-task (i.e., different problem representations) which necessitates providing multiple ERs to support them in creating these representations. To effectively do this, one must avoid or neutralize the difficulties students encounter when combining multiple ERs, namely problems translating from and coordinating between different kinds of representations (Ainsworth, 2006), and incongruence between representation and phase-related part-task (Vekiri, 2002). This necessitates that the representational guidance of a specific ER must be congruent (i.e., matched) with the part-task demands and activities of a specific problem phase.

## Matching ERs and Part-task Demands

Non-expert students carrying out complex learning-tasks without guidance rely primarily on surface features (i.e., using objects referred to in the problem) instead of the underlying principles of the domain, and tend to employ weak problem-solving strategies such as working via a means-ends strategy towards a solution instead of strong ones that are carefully tailored to the specific structure of the domain (Simon, Langley, & Bradshaw, 1981). An important reason for this is that these students have problems with creating and combining suitable

knowledge representations, lacking a well developed understanding required for carrying out the complex learning-task (Ploetzner, et al., 1999; Seufert, 2003). Without such an understanding, students are often not able to create a meaningful problem representation. This is problematic because the ease with which a problem can be solved often depends on the quality of the problem representation. Different problem representations initiate different kinds of operators which can act to produce new information that supports problem solvers in coming to a solution to the problem (Chi, Glaser, & Rees, 1982; Jonassen, 2003). To overcome these difficulties, students need to be made aware of the different problem phases and their required problem representations and be supported in creating and combining these representations (Ainsworth, 2006; Bredeweg & Forbus, 2003; Frederiksen & White, 2002; Ploetzner, et al., 1999). As described in Table 1 this might be accomplished through scripting the problem-solving process by sequencing and making the part-tasks explicit so that they could be foreseen with ontologically congruent ERs in the representational tools (i.e., representational scripting). In the following two paragraphs the different problem phases and their part-task demands and the different ERs are described in more detail.

Table 1: Representational Scripting; Matching ERs and Phase-related Part-task Demands

Problem phase	Part-task demand	ER
Problem orientation	Determining core concepts and relating them to the problem	Showing concepts and their interrelationship
Problem solution	Proposing multiple solutions to the problem	Showing causal relation between the concepts and possible solutions
Solution evaluation	Determining suitability of the solutions and coming to a final solution to the problem	Showing mathematical relation between the concepts and enabling manipulation of their value

### ***Problem phases and their part-task demands***

Solving a complex problem is frequently regarded as a three-phase process, namely (1) orienting to the problem, (2) finding one or more possible problem solutions, and (3) evaluating the solutions so as to choose the best one (Duffy, et al., 1998; Van Bruggen, et al., 2003). Each of these phases requires the creation of a specific qualitative or a quantitative problem representation. *Qualitative representations* provide an overview of the relevant concepts and their interrelationships in the knowledge domain and/or their underlying causal principles. When the interrelationships are *quantitatively* specified, as is often the case in business-economics for example, students are more restricted in creating a suitable problem representation because their attention is more focused on the mathematical relationships between specific concepts. This may be detrimental for the first two phases of problem solving (i.e., problem orientation, problem solution) because it hinders them in thinking about multiple solutions. Furthermore, quantitative representations can only be understood and applied if the students have a well developed qualitative understanding of the knowledge domain. When understood, quantitative representations enable students to evaluate their proposed solutions, something qualitative representations do not allow.

In the *problem orientation phase*, students need to construct a cognitive bridge between their initial mental model and the mental model to be created (Chi, et al., 1982; Jonassen, 2003). This phase involves a part-task which focuses on constructing a global problem representation, becoming aware of the problem itself and of the important concepts of the knowledge domain, and becoming aware of the constraints and criteria for solution

and evaluation (e.g., this concept should affect this concept and that is something that will help to achieve the goal). For creating such a problem overview, a qualitative problem representation containing the relevant concepts is more appropriate than a quantitative one for supporting students in broadening the problem space. The *problem solution phase*, which follows the orientation phase, is where the students apply the underlying causal principles of the knowledge domain to produce concrete solutions. The part-task in this phase is more structured than in the previous phase and focuses on combining the concepts of the domain into principles and making causal relationships between the problem and the proposed solutions explicit (e.g., if this concept is increased, then this concept decreases). Here students might create a number of possible solutions and then reason about the advantages and disadvantages of each. The main advantage of these activities is that the solutions come in a rather straightforward, often causal, way from this which makes the completion-process more efficient and effective (e.g., Jonassen & Ionas, 2008). The problem representation remains qualitative, but contains - along with the central concepts of the problem - causal information (i.e., if this, then that) which supports students in finding multiple solutions to the problem. During the third and final phase, the *solution evaluation phase*, it is more appropriate that students relate the solutions they arrived at to their consequences so as to determine their suitability. This should enable students to reach a final and suitable problem solution. This part-task focuses on calculating the proposed solutions and gaining insight into their quantitative effects (e.g., increasing this concept doubles that concept, but also increases it to a level that is unrealistic).

### ***Part-task congruent external representations***

External representations (ERs) support students in creating different problem representations (i.e., qualitative and quantitative ones) through their differences in *representational guidance*. In order to be beneficial for problem solving it is important that the representational guidance of a specific ER is congruent with part-task demands and activities of a specific problem phase (Schnotz & Kürschner, 2008; Van Bruggen, et al., 2003). The representational guidance of an ER is provided by its ontology, which is specified through its expressiveness and processability (see Table 2). *Expressiveness* refers to what the ER can represent, namely concepts and their interrelationships (i.e., specificity), and how accurately they are represented (i.e., precision). *Processability* refers to the differences in processing the information from the ER caused by the difference in expressiveness, and which determines the number and quality of inferences that can be made. Less expressive (i.e., less specific and less precise) ERs have the advantage of being highly processable (Larkin & Simon, 1987) making it easy to make many inferences from them (i.e., elaboration). Such ERs guide students in elaborating on the concepts of the knowledge domain and in relating them to the problem (e.g., Jonassen, 2003). These ERs, however, do not have much expressive power (Cox, 1999); the inferences made from them cannot be very specific and precise. For this, the order of the ER is important.

The *order* of an ER (Frederiksen, White, & Gutwill, 1999) determines the quality of the inferences (i.e., kind of reasoning). A *zero order* ER supports reasoning about concepts and in relating this reasoning to the problem in qualitative way. It is highly processable, but not very expressive. A *first order* ER is more expressive - and thus specific and precise - which supports reasoning about causal relationships and guides discussion and/or thought about possible solutions. A *second order* ER is the most expressive guide and supports quantitative inference-making enabling negotiation and/or determination of suitability of the proposed solutions. When the representational guidance of the ER is congruent with (i.e., matched to) the ontological demands of the

part-task of a problem phase, students are supported in carrying out the required part-task demands and activities of that phase. A mismatch, on the other hand, means that the ER is incongruent with the part-task and, therefore, may hinder students carrying out complex learning-tasks. Reasons for this could be that the available ER is not expressive enough because it contains only global information, or that it is too hard to process because students do not have enough prior domain knowledge to properly grasp it and make use of the ERs' expressiveness.

Table 2: Congruence between External Representations and Phase-related Part-task Demands

Phase-related part-task	ER	Representational guidance			
		Expressiveness		Processability	
		Specificity	Precision	Elaboration	Order
Problem orientation	Conceptual	Low	Undirected relations	Unstructured	Zero
Problem solution	Causal	Middle	Causal directed relations	Quasi-structured	First
Solution evaluation	Simulation	High	Model directed relations	Structured	Second

## Design and Expectations

This study focuses on how the design of a CSCL-environment that scripts problem-solving behavior by providing ontologically distinct ERs affects both complex learning task-performance and individual learning gains. To this end, four experimental conditions were defined. In triads, students in all conditions had to collaboratively solve a case-based problem in business-economics which was divided into three problem phases, each coupled with different ERs. To study the effects of the representational scripting, the ERs were either matched or mismatched to the different problem phases (see Table 3).

In Table 3: Overview of the Experimental Conditions three

Condition	Phase-related part-task/ER			Match/mismatch
	Problem orientation	Problem solution	Problem evaluation	
Conceptual	Conceptual ER	Conceptual ER	Conceptual ER	Match for the orientation phase only
Causal	Causal ER	Causal ER	Causal ER	Match for the solution phase only
Simulation	Simulation ER	Simulation ER	Simulation ER	Match for the evaluation phase only
Match	Conceptual ER	Causal ER	Simulation ER	Complete match

mismatch conditions, student groups received either a static ER (i.e., conceptual or causal ER) or a dynamic ER (i.e., simulation) which matched only one of the part-tasks, namely problem orientation, problem solution, and solution evaluation, respectively. The scripting structured the problem-solving process in three phases, but only one of the three ERs was available to the students for solving the problem. In other words, there was a phase-mismatch where the ER which was provided ontologically matched only one of the three phases and there was a mismatch for the other two. In the fourth condition, student groups received all three ERs in a phased order receiving the ER most suited to each problem phase. Here, thus, there was a match between all three ERs and all three part-tasks. Student groups in this condition received the complete array of representations. Due to the presumed match between ERs and phase-related part-tasks, student understanding and part-task-related activity should increase, allowing them to come up with better solutions for the problem. It was, therefore, hypothesized students in the match condition (*H1*) create a better developed understanding (i.e., learning gains) and (*H2*) will

arrive at a better solution to the problem (i.e., complex learning-task performance), because their knowledge has progressively evolved from qualitative to quantitative.

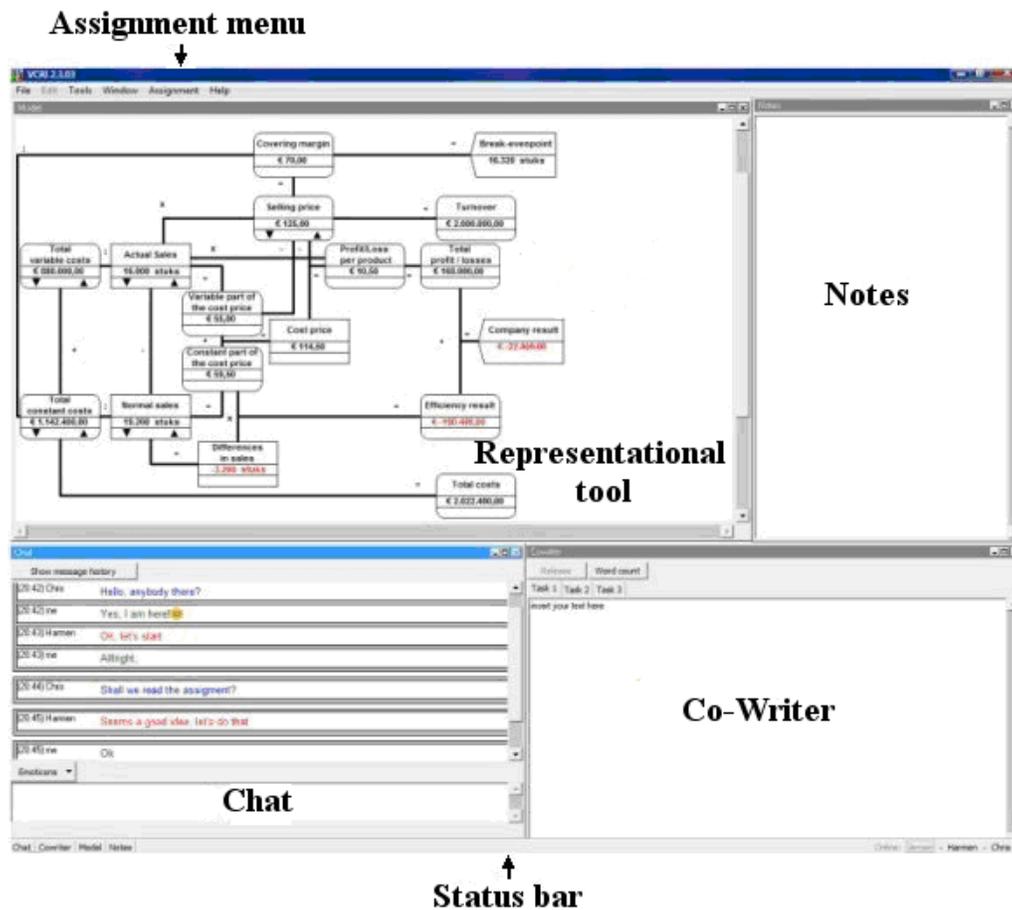
## Method

### Participants

Participants were students from six business-economics classes in three secondary education schools in the Netherlands. The total sample consisted of 96 students (59 male, 37 female). The mean age of the students was 16.67 years ( $SD = 0.77$ ,  $Min = 15$ ,  $Max = 18$ ). Students were, within classes, randomly assigned to a total of 32 triads, which were equally divided between the four experimental conditions.

### Design of the CSCL-environment and the complex learning-task

Students collaborated in a CSCL-environment called *Virtual Collaborative Research Institute* (VCRI, see Figure 1). VCRI is a groupware application for supporting the collaborative performance of complex learning-tasks, inquiry-tasks and research projects (Jaspers, Broeken, & Erkens, 2005). For this study, five tools that are part of the VCRI were augmented with representational scripting. All tools, except the Notes tool, were shared among group members.



**Fig. 1** Screenshot of the VCRI-environment

## VCRI

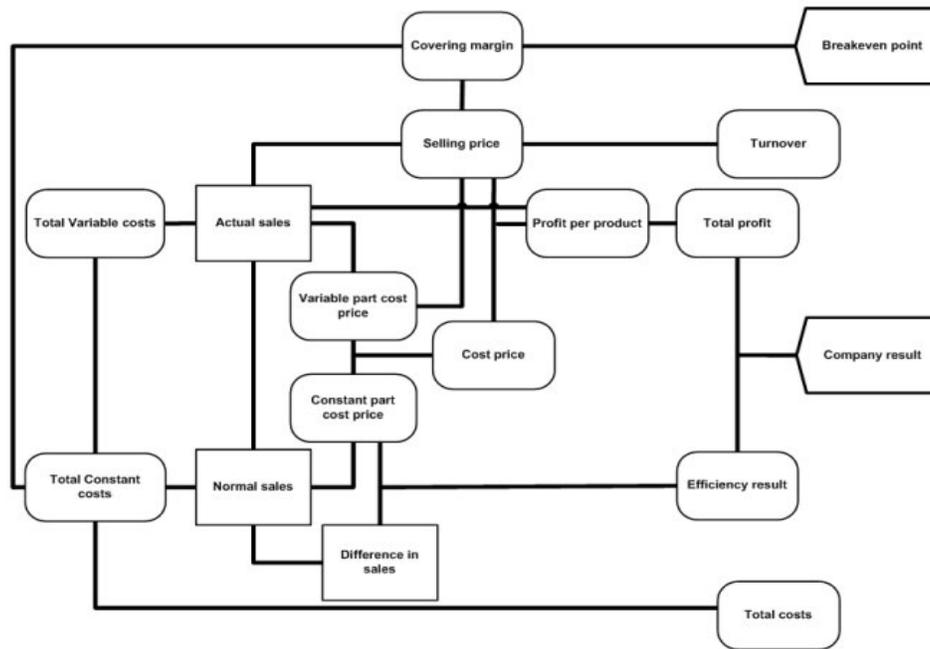
The *chat tool* enables synchronous communication and supports students in externalizing and discussing their ideas and knowledge. The chat history is stored automatically and can be re-read. Students can find the description of the complex learning-task and its phase-related part-tasks in the *Assignment menu*. Besides this, additional information sources such as a definition list, formula list, and clues for solving the problem were also available in the assignment menu. The *Co-writer* is a shared text-processor where students can formulate and revise their answers to the part-tasks. The *Notes tool* is an individual notepad that allows students to store information and to structure their own knowledge and ideas before making them explicit. The *Status bar* is an awareness tool that displays which group members are logged into the system and which tool a group member is currently using. All students in all conditions had access to these tools and information sources, and were, thus, information equivalent. The conditions only differed in the way that the ERs guided the students in creating and combining different problem representations.

### *Representational scripting and phase-related part-tasks*

All groups worked on a complex business-economics problem in which they had to advise an entrepreneur about changing her/his business strategy in order to make the business more profitable (i.e., achieve a better company result). To provide a suitable advice, students had to carry out three different phase-related part-tasks, namely (1) determine the main concepts responsible for the company's results and relate them to the problem, (2) determine how certain interventions (i.e., changes of the business strategy) affect company results, and (3) compare these consequences and formulate a final advice based on this comparison. Through the use of *scripting*, the complex learning-task was divided into three phases (i.e., problem orientation, problem solution, solution evaluation) each focusing on one of the part-tasks. All groups were 'forced' to carry out the part-tasks in a predefined order; they could only start with a new part-task after finishing the earlier part-task. When group members agreed that a part-task was finished, they had to 'close' that phase in the *assignment menu*. This 'opened' a new phase, which had three consequences for the groups, namely they (1) received a new part-task (2) had to enter their new answers in a different window of the *Co-writer* and could not alter, but could still see, their prior answers, and (3) received an ER. For the three mismatched conditions, the ER did not change. Only in the fourth, matched, experimental condition, did the students get a new ER which was ontologically matched to the demands of that phase-related part-task. A description of the different phases and the matching ERs for the fourth experimental conditions follows. All other experimental conditions received the part-tasks in the same order (i.e., used the same script), but did not receive different ERs.

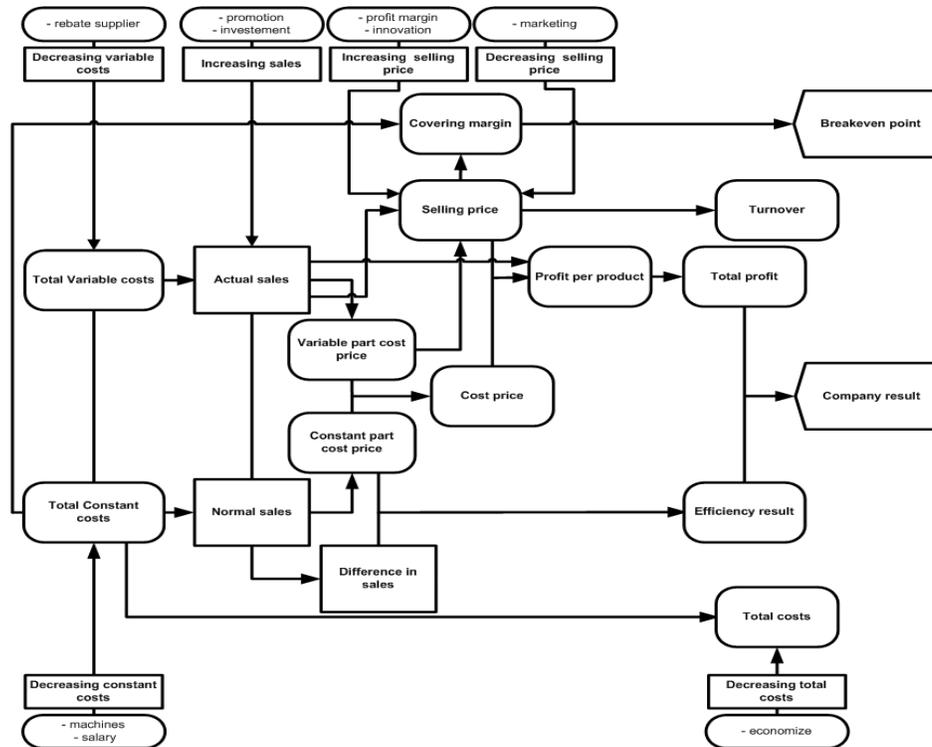
The *problem orientation phase* focused on creating a global problem representation by asking students to explain what they thought the problem was, and describing what the most important concepts were for coming to an advice. During this phase, students received the *conceptual ER* (i.e., a static representation of the knowledge domain; see Figure 2), which made two aspects salient, namely the core concepts needed to carry out this part-task and which core concepts were related to which other core concepts. Students could, for example, see that the 'company result' is determined by the 'total profit' and the 'efficiency result'. This should make it easier for them to create an overview of all relevant concepts (i.e., to broaden the problem space), which should

support them in finding multiple solutions to the problem in the following phase. The low expressiveness of the conceptual ER supports the creation of a global problem representation which can be elaborated on in the following problem phases that contain part-tasks that require the support of more expressive ERs, that is: a qualitative casual and a quantitative problem representation.



**Fig. 2** Conceptual ER

The *problem solution phase* aimed at creating a scientific problem representation (i.e., explicating the underlying business-economics principles) by asking students to formulate several solutions to the problem. During this phase, students received the *causal ER* (i.e., a static representation of the knowledge domain; see Figure 3), in which the causal relationships - visible through the arrows showing direction of the relationship between the concepts - were specified. The causal ER also contributed to increasing student understanding by providing them with possible interventions (i.e., changes of the business strategy), each of which had a different effect on the company results. This should make it easier to explore the solution space and therefore should support students in finding multiple solutions to the problem. Students could, for example, see that receiving a rebate from a supplier affects the ‘variable part cost price’, which in turn affects the ‘cost price’. The conceptual ER is not expressive enough for this part-task because the relations in that ER were not specified and the students did not receive any information about possible solutions. This means that they had to produce the advice themselves, without having sufficient understanding of the underlying principles of the knowledge domain. The simulation ER used in the following phase has a quantitative character which supports testing the proposed advices, but is difficult to process without a properly developed qualitative understanding.



**Fig. 3** Causal ER

The *solution evaluation phase* aimed at increasing understanding of the knowledge domain with the aid of a quantitative problem representation. Students were asked to determine the financial consequences of their proposed solutions, and to formulate a final advice for the entrepreneur by negotiating the suitability of the solutions with each other. During this phase, students received a *simulation ER* (i.e., a dynamic representation of the knowledge domain; see Figure 4) which enabled them to manipulate the value of the concepts by clicking on the arrows in the boxes. When the value of a certain concept was increased or decreased, the simulation model automatically computed the value of all other concepts. The results obtained here should facilitate determining and negotiating the suitability of the proposed solutions and coming to a final advice. Students could, for example, test how a supplier rebate (i.e., decrease of the total variable costs) affects the ‘cost price’ and how this in turn affects the ‘company result’. Only the simulation ER is capable of providing this kind of support, because the relationships between the concepts in this ER were specified as equations (i.e., weight of the relationship).

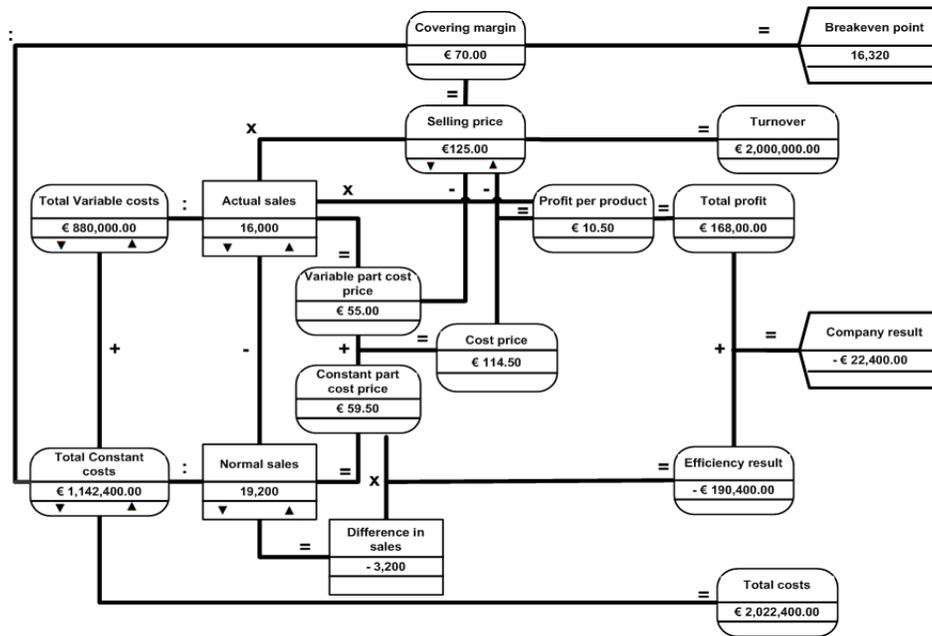


Fig. 4 Simulation ER

## Procedure

In total, students devoted three, 70-minute, lessons to the completion of the complex learning-task during which each student worked on a separate computer in a computer room. Before the first lesson, students received an instruction about the (1) CSCL-environment, (2) group composition, and (3) complex learning-task and its phase-related part-tasks. The instruction made it clear to the students that their score on the post-test as well as their group answer to the problem (i.e., complex learning-task performance) would serve as grades and affect their GPA. Furthermore, a 45-minute pre-test was administered to determine prior domain knowledge and relevant personal information (e.g., age, sex). Thereafter, students worked on the complex learning-task in the computer room, where all actions and answers to the part-tasks were logged. During the lessons, the teacher was on stand-by for task-related questions and a researcher was present for technical support. After the final computer lesson, a 45-minute post-test was administered to determine the acquired domain knowledge of the students after the intervention.

## Measures

### *Learning gains*

Student recall and understanding of the knowledge domain was measured with a pre-test (20 items,  $\alpha = .60$ ) and a post-test (20 items,  $\alpha = .79$ ). Based on work of Gagné, Wagner, and Briggs (1992) a learning-task analysis was conducted which resulted in 17 business-economics concepts. According to Anderson and Krathwohl (2001), a knowledge domain consist of different knowledge dimensions which refer to the different ways (i.e., factual, conceptual, procedural) in which the concepts can be understood. *Factual knowledge* entails students being familiar with the concepts of the knowledge domain. *Conceptual knowledge* entails students understanding

the interrelationships between the different concepts of the domain. *Procedural knowledge* entails students knowing how to apply a certain technique or procedure and are capable of determining when applying that technique or procedure is appropriate. The multiple-choice items in both tests were drawn from the total pool of items and equally divided across the three knowledge dimensions and were, thus, unique. Because of the low reliability of the scores on the subscales of both tests (e.g.,  $\alpha \leq .50$ ) we did not test for student recall and understanding of the different knowledge dimensions. In the analyses, thus, we only made use of the overall scores on the pre-test and the post-test. Below an example for each type of question is provided (questions were translated from Dutch):

#### *Factual knowledge*

The cost price is the price that:

- a) a customer has to pay for a product.
- b) an entrepreneur has to pay to produce a product.
- c) an entrepreneur has to pay to store a product.
- d) an entrepreneur has to pay to produce and to sell a product.

#### *Conceptual knowledge*

Does an increase in selling price automatically lead to an increase in turnover?

- a) No, when the selling price increases this may lead to a decrease in actual sales and, thus, not automatically to an increased turnover.
- b) Yes, when the selling price increases this does not affect the actual sales and, thus, the turnover automatically increases.
- c) No, the turnover is mainly affected by the number of customers willing to buy the product, an increase in selling price, therefore, does not automatically lead to an increase in turnover.
- d) Yes, when the selling price increases the turnover automatically increases whether or not the actual increase or decrease.

#### *Procedural knowledge*

Entrepreneur Y has an electronics store and sells a wide variety of products such as TVs, stereos and computers. At the end of the week the entrepreneur has sold five computers with a selling price of € 1,550.00 each and six TVs with a selling price of € 1,350.00 each. What was the turnover for the entrepreneur for the selling of the computers?

- a) € 6,750.00
- b) € 7,750.00
- c) € 8,100.00
- d) € 9,300.00

#### *Complex learning-task performance*

To measure the effect of condition on group performance, an assessment form for each topic of the learning-task was developed (see Table 4). The whole learning-task was divided in three phases, and for each phase three

questions were asked (i.e., nine questions in total). The answers to each of the nine questions were evaluated based upon ‘suitability’, ‘elaboration’, ‘justification’, and ‘correctness’, which resulted in 36 metrics for these four topics. We also evaluated whether students used their answers from a prior phase and whether they altered their way of reasoning when they had to answer the questions asked in a following phase (i.e., ‘continuity’). This consisted of two items because there were two phase transitions (i.e., problem orientation to problem solution and problem solutions to solution evaluation). Furthermore, the ‘quality of the final advice’ that the students gave was evaluated on three aspects, namely how many (1) concepts and (2) financial consequence were incorporated, and (3) whether the answer was in line with the guidelines provided in the description of the complex learning-task. This resulted in a total of 41 items which all could be coded as ‘0’ (wrong), ‘1’ (passing) or ‘2’ (good); the higher the code, the higher the quality of the answer. Groups could maximally score 82 points ( $41 \times 2$  points) for their complex learning-task performance.

Table 4: Items and Reliability for the Complex Learning-task Performance.

Topic	Description	Items	$\alpha$
Suitability	Whether the groups’ answers were suited to the different part-tasks.	9	.61
Elaboration	Number of different business-economics concepts or financial consequences incorporated in the answers to the different part-tasks.	9	.53
Justification	Whether the groups justified their answers to the different part-tasks.	9	.73
Correctness	Whether the groups used the business-economics concepts and their interrelationships correctly in their answers to the different part-tasks.	9	.68
Continuity	Whether the groups made proper use of the answers from a prior problem phase.	2	.67
Quality advice	Whether the groups gave a proper final advice. - Number of business-economics concepts incorporated in the advice. - Number of financial consequences incorporated in the advice. - Whether the final answer conformed to the guidelines provided.	3	.71
Total	Overall score on the complex learning task-performance.	41	.89

## Data Analysis

When conducting studies in the field of CSCL, students are often working in groups. In such settings, researchers have to cope with several statistical concerns, namely (1) hierarchically nested datasets, (2) non-independence of dependent variables, and (3) differing units of analysis (e.g., Janssen, 2008). The latter concern is also relevant for this study because the dependent (e.g., post-test score) and independent (e.g., experimental condition) variables were measured at different levels, namely the individual and the group level respectively. *Multilevel analysis* (MLA) is a statistical technique suited to “appropriately grasp and disentangle the effects and dependencies on the individual level, the group level, and sometimes the classroom level” (Strijbos & Fischer, 2007, p. 391). To determine whether MLA was a suited technique for answering our research question we computed the amount of variance on the post-test score that could be accounted for by the group (e.g., intraclass correlation coefficient, Kenny, Kashy, & Cook, 2006). Of the total variance on post-test score 59% could be explained by the variance at the group level. This means that working in groups accounts for more variance on individual post-test scores than individual characteristics of the group members (e.g., age, sex). For this reason, MLA was used to determine the effect of experimental condition on post-test score. One-way MANOVA was

used for answering the second research question. Since there were specific directions of the results expected (see hypotheses) all analyses are one-sided.

## Results

### Learning gains

The overall mean score on the pre-test was 14.87 ( $SD = 2.33$ ; max = 20). The overall mean on the post-test score was 14.69 ( $SD = 2.40$ ; max = 20). The  $t$ -test showed that the overall post-test score of 90 students (not all 96 students were present when the pre-test and/or post-test were administered) was not significantly higher than the overall pre-test score ( $t(90) = 0.72, p > .05$ ). There were, thus, no individual learning gains. One-way ANOVA showed a significant main effect between condition on the pre-test score ( $F(3, 86) = 3.34, p < .05$ ). This means that students differed in the amount of prior knowledge and it was, therefore, necessary to correct for this. Table 5 shows the overall and condition means and standard deviations on students' pre-test and post-test scores.

Table 5: Means and Standard Deviations of Students' Pre-test and Post-test Scores for Conditions

Test	Conceptual condition ( $n_{student} = 22$ ) $M(SD)$	Causal condition ( $n_{student} = 24$ ) $M(SD)$	Simulation condition ( $n_{student} = 21$ ) $M(SD)$	Match condition ( $n_{student} = 23$ ) $M(SD)$	Overall conditions ( $N_{student} = 90$ ) $M(SD)$	$p$ -value
Pre-test	15.20 (1.85)	14.95 (2.76)	13.69 (2.20)	15.72 (2.05)	14.66 (2.37)	$p < .05$
Post-test	13.70 (2.96)	15.00 (1.90)	14.47 (2.18)	15.50 (2.48)	14.47 (2.45)	$p < .05$

MLA revealed that students in the match condition scored significantly higher than those in the other conditions ( $\beta = 1.93, p = .04$ ). When comparing the conditions separately, a trend was found; students in the match condition scored higher than students in the conceptual condition ( $\beta = 1.89, p = .07$ ). Differences between the other conditions were not significant. The model fit the data ( $\chi^2(3) = 28.65, p = .00$ ) and could, therefore, be used to account for the differences in variance on the post-test score.

These results are not completely in line with our first hypothesis. Students in the match condition only scored higher on the post-test in comparison to students in the conceptual condition. Furthermore, there were no learning gains.

### Complex learning-task performance

One way MANOVA on the total score on the complex learning-task performance showed a significant difference for condition ( $F(3, 28) = 1.72, p = .03$ ; Wilks' Lambda = 0.33; partial  $\eta$  squared = .31). Bonferroni post hoc analyses showed that groups in the match condition scored significantly higher than groups in both the conceptual ( $p = .00$ ;  $d = 2.19$ ) and the simulation condition ( $p = .04$ ;  $d = 1.26$ ). Differences between other conditions were not significant. Table 6 shows the overall and condition means and standard deviations of the scores on the complex learning-task performance.

Table 6: Means and Standard Deviations of the Complex Learning-task Performance for Conditions

Topic	Conceptual condition	Causal condition	Simulation condition	Match condition	Overall conditions
-------	----------------------	------------------	----------------------	-----------------	--------------------

	( $n_{group} = 8$ )	( $N_{group} = 32$ )				
	<i>M (SD)</i>	<i>p -value</i>				
Suitability (max 18)	12.25 (2.49)	15.12 (1.64)	13.88 (3.36)	15.75 (2.42)	14.25 (2.76)	$p < .05$
Elaboration (max 18)	6.38 (3.74)	8.89 (2.70)	6.37 (2.83)	8.38 (2.33)	7.50 (3.03)	<i>n.s.</i>
Justification (max 18)	3.50 (1.69)	6.88 (3.56)	4.12 (2.70)	7.50 (2.62)	5.50 (3.12)	$p < .05$
Correctness (max 18)	5.50 (2.45)	8.25 (3.69)	7.12 (1.96)	9.25 (2.05)	7.53 (2.87)	$p < .05$
Continuity (max 4)	2.50 (1.41)	3.12 (1.13)	3.00 (1.31)	3.62 (0.52)	3.06 (1.16)	<i>n.s.</i>
Final answer (max 6)	2.75 (1.04)	4.88 (1.64)	5.12 (2.48)	4.25 (1.28)	4.25 (1.87)	<i>n.s.</i>
Total score (max 82)	32.88 (10.40)	47.13 (12.30)	39.62 (0.39)	48.75 (7.27)	42.09 (11.68)	$p < .05$

When the results for the dependent variables were considered separately, using one-way ANOVAs with Bonferroni post hoc analyses, condition effects were found for suitability ( $F(3, 28) = 2.99, p = .03$ ), justification ( $F(3, 28) = 4.23, p = .01$ ) and correctness ( $F(3, 28) = 2.99, p = .03$ ). The mean scores indicated that there were several significant differences between conditions. First, groups in the match condition scored significantly higher on suitability than groups in the conceptual condition ( $p = .01; d = 1.45$ ) and a trend was found in comparison to the groups in the simulation condition ( $p = .07; d = 0.77$ ). Second, groups in the match condition scored significantly higher on justification than groups in both the conceptual ( $p = .01; d = 1.53$ ) and the simulation condition ( $p = .02; d = 1.29$ ). Finally, groups in the match condition scored significantly higher on correctness than groups in the conceptual condition ( $p = .03$ ; one sided;  $d = 1.83$ ) and a trend was found in comparison to the groups in the simulation condition ( $p = .06; d = 1.04$ ).

These results confirmed our second hypothesis, namely that groups that received an ontologically congruent ER for each phase-related part-task scored higher on the complex learning-task performance.

## Conclusion and Discussion

This study shows that combining the advantages of scripting with representation tools - *representational scripting* - supports students in collaboratively carrying out complex learning-tasks, leading to better complex learning-task performance. Structuring the complex problem-solving task into ontologically distinct problem-phases and providing the phase-related part-tasks with part-task-congruent representations was expected to increase group performance in terms of group complex learning-task performance and individual learning gains. The design of representational scripting did indeed result in better scores on the complex learning-task performance. The groups in the match (i.e., part-task-congruent) condition outperformed the groups in both the conceptual and simulation conditions, their answers were more suited for a specific part-task, contained more justifications, and were more often correct. No differences were found between the match and the causal condition. Apparently the causal representation provided more support than both the conceptual and the simulation representation did, but in combination these three representations resulted in a higher score on group complex learning-task performance. The results concerning complex learning-task performance confirmed our expectation and are in line with those of others (Jonassen, 2003; Ploetzner, et al., 1999), who also stress the importance of sequencing and interrelating qualitative and quantitative aspects of the knowledge domain during problem solving. As is the case with many other researchers (Ertl, et al., 2008; Fischer, et al., 2002; Schnotz & Kürschner, 2008; Suthers, 2006), our results stress the importance of providing ERs during collaborative problem solving. The representational guidance that they provide is able to guide student discourse and/or activities towards particular task content. However, in contrast to these studies, we provided multiple ERs where representational guidance was matched to different kinds of part-tasks that had to be carried out when dealing

with complex learning-tasks. To our knowledge such an approach has not been used in other studies. Ertl, et al. (2008), for example, did use a condition in which scripting was applied to structure the problem-solving process and an ER was provided to further support the students. Their design, however, did not enable them to compare the effects with those of another condition in which scripting and another or multiple ERs were used. It was also expected that gradually shifting from a conceptual to a simulation representation would result in higher individual scores on the post-test (i.e., learning gains). Students in the match condition indeed outperformed students in the non-matched conditions. However, this difference was only significant when comparing students in the match condition to students in the conceptual condition.

Although most of the results are in line with our expectations, there were, however, two contrasting findings that require further discussion. First, students' pre-test and post-test score did not differ significantly from each other resulting in no learning gains. This result might be explained by the (1) design of the representational scripting and/or (2) measurement of the learning gains. The design of the representational scripting was primarily aimed at supporting students in applying domain knowledge in order to come to better and richer solutions and might, therefore, be less suited for knowledge acquisition. According to Kirschner, Sweller, and Clark (2006), carrying out complex learning-tasks is an instructional method based on the epistemological content (i.e., methods and processes) instead of the pedagogical content (i.e., acquiring knowledge) of a knowledge domain. Although both the epistemological and the pedagogical content include factual, conceptual and procedural knowledge, students do not necessarily use the same cognitive processes (Anderson & Krathwohl, 2001). That is, recalling and grasping the meaning of concepts, principles and procedures is often regarded as prerequisite for the higher-order cognitive processes required for carrying out complex learning-tasks. Such learning-tasks consist of part-tasks demanding students to apply their understanding of the domain in order to analyze the problem, come up with proper solutions and evaluate their suitability and might be less supportive for acquiring more domain knowledge. Furthermore, the pre-test and the post-test measured recall and understanding of the knowledge domain. Both tests were, therefore, only useful for determining learning gains in terms of acquired domain knowledge. The tests did not enable students to demonstrate whether they were better able to apply their understanding of the domain, an ability which also can be regarded as a form of learning gains. This also could be an explanation for the lack of differences in learning gains. Second, individual learning gains and learning-task performance of students in the causal condition was very similar to what was found in the match condition. Students in both conditions received the causal ER, which showed all relevant concepts, solutions and their causal interrelationships, providing students with multiple qualitative perspectives on the knowledge domain. It seems therefore important to recognize that causal reasoning is beneficial for complex learning-task performance (Jonassen & Ionas, 2008). However, it does not completely explain the lack of differences. Perhaps combining the causal ER with both the conceptual and the simulation ER hinders complex learning-task performance when students experience difficulties integrating the different ERs. When students do not know how to use an ER and/or combine multiple ERs, they might choose to stick with the familiar one and make no attempt to integrate the different ERs (Ainsworth, 2006). Furthermore, students in the causal condition did not receive additional ERs and their cognitive processes were, therefore, less focused on applying their understanding of the domain in comparison to students in the match condition. In this respect students in the causal condition could devote more attention to understanding the concepts and their causal relationship which might have made the post-test more suited for them in comparison to students in the

match condition. An alternative explanation might be that collaboration requires interaction in both the content space and the relational space from all group members. If the whole group is not able to carry out these activities, the collaboration process may hinder students in successfully carrying out their complex learning-task (e.g., Barron, 2003).

## Implications and Future Research

This study has several implications for learning-environment design (e.g., CSCL-environment) for supporting students in carrying out complex learning-tasks. The results indicate that complex learning-task completion is facilitated and complex learning-task performance is better when the different phase-related part-tasks are made explicit and properly sequenced, and part-task congruent ERs (i.e., domain specific content schemes) are provided. Using multiple ERs can provide different perspectives of the knowledge domain and, when properly matched to the task demands and activities, the complementary function (see Ainsworth, 2006) of those ERs can gradually increase student understanding and phase-related part-task activity. However, several limitations have to be taken into account when interpreting the results and the implications of this study. First, conducting studies within schools for a longer period of time has the advantage of taking place in a more ecologically valid research setting than the laboratory. The current study was integrated into the curriculum of the participating schools and student' scores on the post-test as well as the complex learning-task performance affected their GPA. Unfortunately this also had as a result that there were acceptable but still low reliability scores for the instruments used. When tailoring the measurement of the learning gains to the specifics of the curriculum there are often no suitable standardized measurement instruments available. These instruments, therefore, needed to be developed in cooperation with the teachers which made them more (ecologically) valid for measuring the individual and group learning gains. Although this is how teachers usually work with and assess their students, this approach could compromise the reliability of the instruments compromising the generalization of the results. Since our study was conducted in six classes divided over three different schools this concern appears to not be substantial in our study, but cannot be ruled out completely. Second, this study took place in the field of business-economics. Although there are many other domains (e.g., physics, urban planning, meteorology) in which qualitative and quantitative problem representations are required, the effects of a particular design depends on the characteristics of the problem and the involved knowledge domains. When designing tools, representational scripting, and/or learning environments one should carefully take this into account. The effect of the design of representational scripting does, therefore, not automatically apply to all complex learning-tasks. Third, condition effects were found for complex learning-task performance and learning gains, but when one inspects the standard deviations it appears that there are also differences between groups within the conditions. The present results of this study are solely focused on the question whether a difference in characteristics of representational scripting affects complex learning-task performance and individual learning.

Finally, at least one issue remains unclear, namely how the design of representational scripting lead to the results obtained. We, for example, cannot yet explain the lack of significant differences between the causal and the match condition. To resolve this issue, additional research into the effects of representational scripting should be carried out to investigate the results and the collaboration process in multiple settings (i.e., CSCL and face-to-face) for multiple problems and in a diversity of knowledge domains. We are, therefore, currently analyzing the log-files (i.e., dialogue-protocols) to determine what students talked about (i.e., content space) and

how students coordinated their collaboration process (i.e., relational space). These analyses should provide insight into the completion-process and how it was affected by the design of the representational scripting.

## References

- Ainsworth, S. (2006). DeFT: A conceptual framework for considering learning with multiple representations. *Learning and Instruction, 16*, 183–198.
- Anderson, L. W., & Krathwohl, D. R. (2001). *A taxonomy for learning, teaching, and assessing: A revision of Bloom's taxonomy of educational objectives*. New York: Longman.
- Barron, B. (2003). When smart groups fail. *Journal of the Learning Sciences, 12*, 307–359.
- Beers, P. J., Boshuizen, H. P. A., Kirschner, P. A., & Gijssels, W. H. (2005). Computer support for knowledge construction in collaborative learning environments. *Computers in Human Behavior, 21*, 623–643.
- Bredeweg, B., & Forbus, K. (2003). Qualitative Modeling in Education. *AI Magazine, 24*(4), 35–46.
- Brna, P., Cox, R., & Good, J. (2001). Learning to think and communicate with diagrams: 14 questions to consider. *Artificial Intelligence Review, 15*, 115–134.
- Chi, M. T. H., Glaser, R., & Rees, E. (1982). Expertise in problem solving. In R. Sternberg (Ed.), *Advances in the psychology of human intelligence* (pp. 7-75). Hillsdale, New Jersey: Lawrence Erlbaum Associates.
- Cox, R. (1999). Representation construction, externalised cognition and individual differences. *Learning and Instruction, 9*, 343–363.
- Dillenbourg, P. (2002). Over-scripting CSCL: The risks of blending collaborative learning with instructional design. In P. A. Kirschner (Ed.), *Three worlds of CSCL: Can we support CSCL?* (pp. 61-91). Heerlen, The Netherlands: Open Universiteit Nederland.
- Duffy, T., Dueber, B., & Hawley, C. (1998). *Critical thinking in a distributed environment: A pedagogical base for the design of conferencing systems* (CRLT Technical report No 5-98). Bloomington, IN: Indiana University, Center for Research on Learning and Technology.
- Ertl, B., Kopp, B., & Mandl, H. (2008). Supporting learning using external representations. *Computers & Education, 51*, 1599–1608.
- Fischer, F., Bruhn, J., Gräsel, C., & Mandl, H. (2002). Fostering collaborative knowledge construction with visualization tools. *Learning and Instruction, 12*, 213–232.
- Frederiksen, J. R., & White, B. Y. (2002). Conceptualizing and constructing linked models: Creating coherence in complex knowledge systems. In P. Brna, M. Baker, K. Stenning, and A. Tiberghien (Eds.), *In the role of communication in learning to model*. (pp. 69-96). Mahwah, New Jersey: Lawrence Erlbaum Associates.
- Frederiksen, J. R., White, B. Y., & Gutwill, J. P. (1999). Dynamic mental models in leaning sciences: The importance of constructing derivational linkages among models. *Journal of Research in Science Teaching, 36*, 806–836.
- Gagné, R. M., Briggs, L. J., & Wagner, W. W. (1992). *Principles of instructional design*. (4<sup>th</sup> ed.). Forth Worth: Harcourt Brace Jovanovich.
- Hollan, J., Hutchins, E., & Kirsh, D. (2000). Distributed Cognition: Toward a new foundation for Human-Computer Interaction research. *ACM Transactions on Computer-Human Interaction, 7*(2), 174–196.

- Janssen, J. (2008). *Using visualizations to support collaboration and coordination during computer-supported collaborative learning*. Ph.D. thesis, Utrecht University, The Netherlands.
- Jaspers, J., Broeken, M., & Erkens, G. (2005). Virtual Collaborative Research Institute (VCRI). Version 2.2. Utrecht, The Netherlands: Utrecht University.
- Jonassen, D. H. (2003). Using cognitive tools to represent problems. *Journal of Research on Technology in Education*, 35, 362–381.
- Jonassen, D. H., & Ionas, I. G. (2008). Designing effective support for causal reasoning. *Educational Technology Research and Development*, 56, 287–308.
- Kenny, D. A., Kashy, D. A., & Cook, W. L. (2006). *Dyadic data analysis*. New York/London: The Guilford Press.
- Kirschner, P. A., Sweller, J., & Clark, R. E. (2006). Why minimal guidance during instruction does not work; An analysis of the failure of constructivist, discovery, problem-based, experiential, and inquiry-based teaching. *Educational Psychologist*, 4(2), 75–86.
- Larkin, J. H., & Simon, H. A. (1987). Why a diagram is (sometimes) worth ten thousand words. *Cognitive Science*, 11, 65–100.
- O'Donnell, A. M., & Dansereau, D. F. (1992). Scripted cooperation in student dyads: a method for analyzing and enhancing academic learning and performance. In R. Hertz-Lazarowitz & N. Miller (Eds.), *Interaction in cooperative groups: The theoretical anatomy of group learning* (pp. 120-141). New York: Cambridge University Press.
- Reiser, B. J. (2004). Scaffolding complex learning: The mechanisms of structuring and problematizing student work. *Journal of the Learning Sciences*, 13, 273–304.
- Ploetzner, R., Fehse, E., Kneser, C., & Spada, H. (1999). Learning to relate qualitative and quantitative problem representations in a model-based setting for collaborative problem solving. *Journal of the Learning Sciences*, 8, 177–214.
- Schnotz, W., & Kürschner, C. (2008). External and internal representations in the acquisition and use of knowledge: visualization effects on mental model construction. *Instructional Science*, 36, 175–190.
- Simon, H. L., Langley, P. W., & Bradshaw, G. (1981). Scientific discovery as problem solving. *Syntheses*, 47, 1-27.
- Spector, J. M. (2008). Cognition and learning in the digital age: Promising research and practice. *Computers in Human Behavior*, 24, 249–262.
- Strijbos, J.-W., & Fischer, F. (2007). Methodological challenges for collaborative learning research. *Learning and Instruction*, 17, 389–393.
- Seufert, T. (2003). Supporting coherence formation in learning from multiple representations. *Learning and Instruction*, 13, 191–203.
- Suthers, D. D. (2006). Technology affordances for intersubjective meaning making: A research agenda for CSCCL. *Computer-Supported Collaborative Learning*, 1, 315–337.
- Van Bruggen, J. M., Boshuizen, H. P. A., & Kirschner, P. A. (2003). A cognitive framework for cooperative problem solving with argument visualization. In P. A. Kirschner, S. J. Buckingham-Shum, & C. S. Carr (Eds.), *Visualizing Argumentation: Software tools for collaborative and educational sense-making*. (pp. 25-47). London: Springer.

- Van Merriënboer, J. J. G., Kester, L., & Paas, F. (2006). Teaching complex rather than simple tasks: Balancing intrinsic and germane load to enhance transfer of learning. *Applied Cognitive Psychology, 20*, 343–352.
- Van Merriënboer, J. J. G., & Kirschner, P. A. (2007). *Ten steps to complex learning. A systematic approach to four-component instructional design*. New Jersey: Lawrence Erlbaum Associates.
- Vekiri, I. (2002). What is the value of graphical displays in learning? *Educational Psychology Review, 14*, 261–312.

## Tables

Table 1: Representational Scripting; Matching ERs and Phase-related Part-task Demands

Problem phase	Part-task demand	ER
Problem orientation	Determining core concepts and relating them to the problem	Showing concepts and their interrelationship
Problem solution	Proposing multiple solutions to the problem	Showing causal relation between the concepts and possible solutions
Solution evaluation	Determining suitability of the solutions and coming to a final solution to the problem	Showing mathematical relation between the concepts and enabling manipulation of their value

Table 2: Congruence between External Representations and Phase-related Part-task Demands

Phase-related part-task	ER	Representational guidance			
		Expressiveness		Processability	
		Specificity	Precision	Elaboration	Order
Problem orientation	Conceptual	Low	Undirected relations	Unstructured	Zero
Problem solution	Causal	Middle	Causal directed relations	Quasi-structured	First
Solution evaluation	Simulation	High	Model directed relations	Structured	Second

Table 3: Overview of the Experimental Conditions

Condition	Phase-related part-task/ER			Match/mismatch
	Problem orientation	Problem solution	Problem evaluation	
Conceptual	Conceptual ER	Conceptual ER	Conceptual ER	Match for the orientation phase only
Causal	Causal ER	Causal ER	Causal ER	Match for the solution phase only
Simulation	Simulation ER	Simulation ER	Simulation ER	Match for the evaluation phase only
Match	Conceptual ER	Causal ER	Simulation ER	Complete match

Table 4: Items and Reliability for the Complex Learning-task Performance.

Topic	Description	Items	$\alpha$
Suitability	Whether the groups' answers were suited to the different part-tasks.	9	.61
Elaboration	Number of different business-economics concepts or financial consequences incorporated in the answers to the different part-tasks.	9	.53
Justification	Whether the groups justified their answers to the different part-tasks.	9	.73
Correctness	Whether the groups used the business-economics concepts and their interrelationships correctly in their answers to the different part-tasks.	9	.68
Continuity	Whether the groups made proper use of the answers from a prior problem phase.	2	.67
Quality advice	Whether the groups gave a proper final advice. - Number of business-economics concepts incorporated in the advice. - Number of financial consequences incorporated in the advice. - Whether the final answer conformed to the guidelines provided.	3	.71
Total	Overall score on the complex learning task-performance.	41	.89

Table 5: Means and Standard Deviations of Students' Pre-test and Post-test Scores for Conditions

Test	Conceptual condition ( $n_{students} = 22$ )	Causal condition ( $n_{students} = 24$ )	Simulation condition ( $n_{students} = 21$ )	Match condition ( $n_{students} = 23$ )	Overall conditions ( $N_{students} = 90$ )	$p$ -value
	$M (SD)$	$M (SD)$	$M (SD)$	$M (SD)$	$M (SD)$	
Pre-test	15.20 (1.85)	14.95 (2.76)	13.69 (2.20)	15.72 (2.05)	14.66 (2.37)	$p < .05$
Post-test	13.70 (2.96)	15.00 (1.90)	14.47 (2.18)	15.50 (2.48)	14.47 (2.45)	$p < .05$

Table 6: Means and Standard Deviations of the Complex Learning-task Performance for Conditions

Topic	Conceptual condition ( $n_{groups} = 8$ )	Causal condition ( $n_{groups} = 8$ )	Simulation condition ( $n_{groups} = 8$ )	Match condition ( $n_{groups} = 8$ )	Overall conditions ( $N_{groups} = 32$ )	$p$ -value
	$M (SD)$	$M (SD)$	$M (SD)$	$M (SD)$	$M (SD)$	
Suitability (max 18)	12.25 (2.49)	15.12 (1.64)	13.88 (3.36)	15.75 (2.42)	14.25 (2.76)	$p < .05$
Elaboration (max 18)	6.38 (3.74)	8.89 (2.70)	6.37 (2.83)	8.38 (2.33)	7.50 (3.03)	<i>n.s.</i>
Justification (max 18)	3.50 (1.69)	6.88 (3.56)	4.12 (2.70)	7.50 (2.62)	5.50 (3.12)	$p < .05$
Correctness (max 18)	5.50 (2.45)	8.25 (3.69)	7.12 (1.96)	9.25 (2.05)	7.53 (2.87)	$p < .05$
Continuity (max 4)	2.50 (1.41)	3.12 (1.13)	3.00 (1.31)	3.62 (0.52)	3.06 (1.16)	<i>n.s.</i>
Final answer (max 6)	2.75 (1.04)	4.88 (1.64)	5.12 (2.48)	4.25 (1.28)	4.25 (1.87)	<i>n.s.</i>
Total score (max 82)	32.88 (10.40)	47.13 (12.30)	39.62 (0.39)	48.75 (7.27)	42.09 (11.68)	$p < .05$

