Faculty of Behavioral, Management and Social Sciences

Combining direct instruction on the Control-of-Variables strategy with task segmentation: Is there a positive synergistic effect?

Erwin van Harmelen

MSc. Thesis
June 2017

Supervisors:
Dr. Hannie Gijlers
Prof. dr. Ard Lazonder

Faculty of Behavioral, Management and Social Sciences,
Department of Instructional Technology
University of Twente
P.O. Box 217
7500 AE Enschede
The Netherlands

Table of contents

Preface **3**

Abstract **4**

Introduction **5**

Inquiry-based learning *6*

Control-of-Variables Strategy (CVS) *7*

Research questions and hypothesis *8*

Method **9**

Research context *9*

Participants *10*

Materials *10*

Procedure *15*

Results **15**

Discussion **25**

Conclusion **27**

References **28**

Preface

Before you lies my master thesis, “Combining direct instruction on the Control-of-Variables strategy with task segmentation: Is there a positive synergistic effect?”. The research discussed in this thesis was conducted at three elementary schools situated in Enschede, the Netherlands. The research was conducted between September 2016 and June 2017 as part of the master Educational Science and Technology at the department of Instructional Technology of the University of Twente, located in the Netherlands.

A special word of thanks goes out to my supervisor Dr. Hannie Gijlers. During each phase of conducting my research and writing the thesis Hannie was always ready and available to provide me with valuable feedback, whilst still letting me discover enough on my own. In addition, I want to thank my second supervisor Prof. dr. Ard Lazonder for providing me with very thorough feedback during the final weeks of writing this thesis. This feedback allowed for me to identify several loopholes in my line of reasoning and increase the quality of this thesis. In addition, I want to thank Casper de Jong for completing this endeavour simultaneously. Having Casper around enduring the same challenges as myself and offering support and feedback was an important source of motivation for me to be able to complete this thesis. For crafting the slopes and the materials to create an electric circuit I want to thank my father, John. When I would have had to solely rely on my own crafting skills this research would certainly have stranded in the design phase.

Hereby I also thank Ms. Anouck Haven and Ms. Astrid Dijkstra for offering the opportunity to study next to working as an elementary school teacher at the Prinseschool. Without their flexibility and support I would not have been able to complete this master. Thanks goes out to my colleagues of OBS Prinseschool, OBS Het Palet and OBS Glanerbrug-Zuid for allowing me to conduct my research in their classes.

Finally, I want to thank my family and friends, especially my wife Amy and son Abel, for providing me with the necessary distraction, support and love during the last three years and for being understanding when I needed to study.

I hope you enjoy reading this thesis.

Erwin van Harmelen

Losser, June 2017

Abstract

Children's ability to design unconfounded experiments in inquiry-based learning has long been identified as both crucial and problematic. Research on two types of support for designing unconfounded experiments, direct instruction and task segmentation, have shown to independently increase children's ability to design unconfounded experiments. This study focused on whether combining direct instruction (DI) and task segmentation (TS) led to a higher percentage of unconfounded experiments, a higher gain in CVS-knowledge, a higher number of variables investigated and higher flow-scores when compared to these types of support used individually and a control condition. Quantitative results showed DI and TS combined did not score significantly higher on percentage of unconfounded experiments and CVS knowledge acquisition. In addition, no significant differences were found between the combined condition and the other conditions concerning flow and the number of variables investigated. The DI-condition did outperform the other conditions on the CVS knowledge acquisition. Additionally, the TS-condition outperformed the other conditions on percentage of CVS-experiments. Qualitative results showed children struggled with dealing with incorrect circuits and applying CVS, even when their scores on the CVS-test was high. No positive synergistic effect of combining DI and TS was found in this study. However, results support previous findings concerning the regulative effects of task segmentation. Furthermore, results of this study concerning the learning gains of the direct instruction condition on knowledge of the CVS support prior findings concerning the effectiveness of direct instruction on the CVS. However, in contrast to prior findings, the DI condition did not outperform other conditions on percentage of uncounfounded experiments.

Keywords: Control-of-Variables strategy, task segmentation, experimentation, inquiry-based learning.

Introduction

In 2013 the Techniekpact 2020 (Ministerie van Onderwijs, 2013) was signed between various stakeholders in the Netherlands. The Techniekpact consisted of agreements and key priorities focusing on increasing the number of students in the Netherlands that choose for Science, Technology, Engineering and Mathematics (STEM) education and professions. One of the means to support this development was ensuring the STEM curriculum in primary and secondary education provided enough basic knowledge for students to choose STEM-education. According to the Dutch Ministry of Education, making students in both elementary and secondary education skilful in inquiry-based learning is one of the means through which these goals should be attained. In addition, the Dutch Ministry of Education made it mandatory for all elementary schools to offer STEM-education to children by 2020 (Ministerie van Onderwijs, 2013).

Inquiry-based learning has long been the focus of research in science education (e.g. de Jong & Van Joolingen, 1998; Minner, Levy & Century, 2009). Children in primary education struggle with various processes in the inquiry cycle (Minner, et al., 2009), such as drawing invalid inferences (De Jong, 2006) and setting up unconfounded experiments (Chen & Klahr, 1999). These difficulties are one of the main reason why open-inquiry environments, in which children are not supported during inquiry, lead to a lower increase in conceptual understanding compared to direct instruction (e.g. Kirschner, Sweller, & Clark, 2006; Mayer, 2004). It is therefore important to sufficiently support children to ensure they are successful when working on inquiry-based learning tasks. (e.g. de Jong & Van Jooling, 1998; Hmelo-Silver, Duncan, & Chinn, 2007; Lazonder & Harmsen, 2016).

One of the main challenges for children in elementary education during inquiry-based learning is creating unconfounded experiments by correctly isolating variables (e.g. de Jong, 2006; Klahr & Nigam, 2004). Inhelder and Piaget (1958) were amongst the first to conduct research on children's experimentation skills through use of the Control-of-Variables strategy (CVS). This strategy entails that the experiments that children conduct should be focused on manipulating the variable of interest whilst keeping other variables constant. Support by means of a short instruction on CVS prior to the inquiry-task (Klahr & Nigam, 2004) can strengthen children's understanding of CVS leading to more unconfounded experiments designed by children. Scaffolding can also be used to support children in designing unconfounded, valid, experiments. Usage of task segmentation, which is also referred to as task structuring, has been found to increase children's comprehension of CVS thus allowing them to design more unconfounded experiments. (e.g. Kuhn & Dean, 2005; Lazonder & Egberink, 2014; Lazonder & Kamp, 2012; Lazonder & Wiskerke-Drost, 2015).

Previous research on CVS mainly focused on comparing a condition in which children are assisted by a specific type of support against children in an unsupported control condition. Little is known about whether combining direct instruction on CVS prior to the inquiry-task with an inquiry-task that is segmented leads to even higher gains in knowledge on CVS and improves the percentage of unconfounded experiments children design. Tabak (2004) pointed out that combining various types of support, such as instruction by a teacher and using scaffolding, can have a synergistic effect, strengthening the support offered to children during learning tasks. Recent research conducted on the effectiveness of direct instruction on CVS (e.g. Klahr & Nigam, 2004; Lazonder & Egberink, 2014) and task segmentation (e.g. Kuhn & Dean, 2005; Lazonder & Kamp, 2012) used computer-simulations as learning environments, limiting the number of design choices students have when designing experiments. Whether using physical materials yields similar effects when these means of support are used is still unknown. A recent meta-analysis conducted by Lazonder and Harmsen (2016) identified that the focus of the bulk of educational research on support of inquiry-based learning has been on whether a difference in learning gains concerning domain knowledge could be observed between conditions. Research focusing on performance during the learning activity has received significantly less attention. Although methodologically sound qualitative studies on inquiry-processes exist (e.g. Schauble, 1996) qualitative research focused specifically on the experiences of children when the focus of the inquiry-task is acquiring the CVS has received little attention in past research on this topic.

Therefore, this study focused on how a combination of direct instruction on CVS prior to the inquiry-task and a task-segmented worksheet where variables have already been isolated influenced elementary children's understanding of CVS and the percentage of unconfounded experiments designed, when working on a guided inquiry task on electric circuits. In addition, this study combined quantitative and qualitative measurements providing valuable new insights into whether a positive synergistic effect of combining direct instruction on CVS and task segmentation exists.

Inquiry-based learning

Inquiry-based learning finds its roots in the constructivist view on learning. The origins of this view lie in the works of Jean Piaget and Lev Vygotsky and is defined as learning by doing.

According to Minner et al., (2009), inquiry-based learning tasks are characterized as consisting of three elements. Firstly, students are responsible for their own learning and are allowed to make decisions on how they learn. Secondly, students engage with the content through use of logic and deduction. Finally, inquiry instruction fosters students' curiosity and enthusiasm, in turn increasing students' motivation. Minner et al. (2009) state, challenging students to actively participate in the investigation process by creating their own experiments leads to an increase in conceptual understanding for science domains. A broad base of research exists that supports the claim that inquiry-based learning can lead to a higher gain in conceptual knowledge and increase children's scientific reasoning skills when compared to traditional methods (e.g. Furtak, Seidel, Iverson & Briggs, 2012; Lazonder & Harmsen, 2016).

Although support exists for the use of inquiry-based learning there has also been criticism. Kirschner, Sweller and Clark (2006) argue that the minimally guided approach for inquiry-based learning, also known as open discovery learning, provides insufficient support for learners. This in turn leads to students not mastering important concepts and skills, that would have been mastered through use of direct instruction. Indeed, open-discovery learning has shown to be inferior to direct-instruction (Mayer, 2004). One of the possible explanations for this can be found in the Cognitive Load Theory (Sweller, 1988). Sweller pointed out that when learning, students use their working memory for information processing. The capacity of the working memory is limited. When learning tasks are too demanding this leads to cognitive overload, resulting in students being unable to complete the task, or comprehend all parts of the task. Open-discovery learning puts a high strain on the working memory of students, in turn leading to cognitive overload (Mayer, 2004). However not all inquiry-based methods leave students without instructional guidance and support. Guided inquiry-learning provides students with scaffolded learning environments, thus allowing children to successfully complete inquiry-based learning tasks (Hmelo-Silver, Duncan & Chinn, 2007; Lazonder & Harmsen, 2016).

In inquiry-based learning a distinction can be made between various processes which make up the inquiry-cycle. According to de Jong (2006) the following cognitive processes are involved in inquiry-based learning: orientation, hypothesis generation, experimentation, reaching conclusions, evaluation, planning, and monitoring. Each of these processes come with their own challenges, and in general, children and students struggle with these inquiry processes (De Jong & Van Joolingen, 1998). The focus of this study is specifically on the inquiry-process of experimentation. During the inquiry-process of experimentation, without the proper support students and children create confounded experiments, in turn leading to them make false inferences (e.g. de Jong, 2006; Klahr & Nigam, 2004).

Inquiry-based learning can be supported by scaffolds to increase performance, support the acquisition of inquiry-skills and improve domain-knowledge acquired (e.g. Pea, 2004; Lazonder & Harmsen, 2016). Scaffolding is a means to support children, allowing them to complete a task they would otherwise be unable to complete (e.g. Minner, Levy & Century, 2009; Lin, Hsu, Lin, Changlai, Yang, & Lai, 2012; Reiser, 2004). Scaffolds explain or take over more complex parts of a task and can be used when learners are not yet skilful enough themselves to perform a certain learning task (Pea, 2004; Reiser, 2004; Lazonder & Harmsen, 2016). Quintana et al. (2004) state that scaffolds in inquiry-based learning support students in setting appropriate goals, designing unconfounded experiments and developing necessary inquiry-skills. Scaffolds in inquiry-based learning focus on supporting both performance and learning (Tabak, 2004) and through working on a scaffolded task, learners improve their process skills and conceptual understanding (Pea, 2004). Reiser (2004) suggested scaffolding either structures or problematizes tasks for learners. When the focus of a scaffold is structuring, it simplifies the learning task to make it easier for learners to complete. When the focus of scaffolding is problematizing the scaffold directs the learner's attention to specific parts of the learning task that would otherwise be ignored. Tabak (2004) pointed out that synergy between various types of scaffolding and support (e.g. teacher coaching, software support) that address the same learning need can lead to a solid method of supporting students and children during learning tasks. For a positive synergistic effect between types of support it is important that materials share the same framework, structure and language. These features need to be consistent in the types of support that are being combined for a synergistic effect to be possible (Tabak, 2004).

To conclude, a distinction needs to be made between open discovery learning and guided inquiry, the latter being the focus of this study. Inquiry-based learning consists of various processes, one being experimentation. One of the problems students experience during experimentation is the creation of unconfounded experiments. Supporting students through scaffolding can compensate for these difficulties and reduce the cognitive load of the learning task.

Control-of-Variables Strategy (CVS)

In guided-inquiry it is important for children in elementary education to learn to isolate variables, allowing for unconfounded experiments and making causal inferences based on these experiments. Students across all ages struggle with designing unconfounded experiments for testing hypothesis (e.g. de Jong & Joolingen, 1998; Klahr & Nigam, 2004). An important strategy to master when designing unconfounded experiments is the Control-of-Variables Strategy (CVS). Chen and Klahr (1999) defined CVS as "a method for creating experiments in which a single contrast is made between experimental conditions and the ability to distinguish between confounded and unconfounded experiments" (p.1098). Children who are skilled at this strategy design valid, unconfounded experiments from which causal inferences can be made. Chen and Klahr stated that acquiring the skill of CVS is an important step in the development of scientific reasoning skills as it gives children insight in how to conduct research through use of experiments. Chen and Klahr were the first to use a specially designed test to analyse the skill-level of children concerning CVS. Children of various grades were given a problem statement and were then asked whether a certain experiment would be a valid means of investigating the problem. Direct instruction on how to apply CVS has since been found to be an effective means of support to increase children's ability in and comprehension of CVS (e.g. Chen & Klahr, 1999; Lazonder & Egberink, 2014). In direct instruction on CVS, children receive a training on how to design unconfounded experiments within a certain domain, prior to commencing with experimenting themselves. Chen and Klahr (1999) found explicit instruction concerning CVS to be more effective when acquiring knowledge on the CVS and designing unconfounded experiments compared to implicit instruction through use of probe questions, and a control condition which received no support. According to Chen and Klahr (1999) only explicit instruction which explained the rationale of the strategy was effective in the acquisition of CVS. In addition, using probe questions was found to be less effective in guiding children in discovering the CVS. Furthermore, Chen and Klahr state fourth grade children which received direct instruction on CVS within a certain domain (e.g. springs) could transfer use of the strategy to different domains (e.g. slopes or sinking).

Klahr and Nigam (2004) examined transfer of the CVS with third- and fourth-grade children. In an experiment with two conditions, discovery-learning CVS and direct instruction on CVS, they found a higher percentage of children in the direct instruction condition to have mastered the CVS. Furthermore, the children who were marked as having mastered CVS were more proficient at evaluating other children's science posters compared to children not mastering CVS. The initial direct-instruction and discovery learning phase took place in the domain of rolling objects. Afterwards students had to check for CVS validity of the results and procedures described by their peers on science posters concerning other domains.

Next to the explicit means of supporting the acquisition of CVS through direct instruction, task segmentation (TS) is an implicit means of scaffolding to help children use CVS when conducting experiments. Kuhn and Dean (2005) argued that although helping children acquiring the necessary CVS-skills through guided-discovery might be more time-consuming and labour-intensive in comparison to direct instruction, implicit instruction also provides children with the opportunity to gain a level of meta-strategic understanding which cannot be assumed to be present after one session of direct instruction. Sixth-graders were provided with hints to focus on a single variable to investigate, in essence segmenting their task. The condition that received such suggestions at the start of each lesson was able to produce more unconfounded experiments and valid inferences compared to an unsupported condition. Kuhn and Dean argued that the advantage of using implicit instruction over direct instruction lies in the long-term mastering of the CVS and the ability of children and students to recognize when the CVS should be used.

Lazonder and Kamp (2012) took using implicit instruction on the CVS one step further by splitting up a multi-variable task in four single variable research questions, allowing for focused design of experiments. Their so-called segmented inquiry task was examined in an experimental study with eighth grade elementary school children using a segmented inquiry-task or an unguided inquiry-task. Children working with the segmented inquiry-task designed more unconfounded experiments and gained more conceptual understanding than children who received an unsegmented version of the task (Lazonder & Kamp, 2012). In addition, children working with a segmented task were better at regulating their investigations and could draw more valid inferences from their experiments.

Lazonder and Egberink (2014) conducted an experiment where direct instruction on CVS was compared to task structuring. In the direct instruction condition a short lecture on how to isolate variables in a multi-variable task was given. For this lecture a computer simulation was used that was situated in the domain of rolling objects. In the task structuring condition children received a segmented version of the task that covered the variables in successive order. Lazonder and Egberink (2014) showed that both conditions were equally effective in teaching children how to design unconfounded experiments compared to unguided inquiry. No relationship was found between the use of either direct instruction or task structuring on the post-test measuring children's knowledge of the CVS. Lazonder and Egberink pointed out that task-structuring focuses on maximizing performance in order for children to be able to complete the inquiry, with the goal of attaining domain-knowledge. Task segmentation is effective in maximizing performance, but ineffective in the long term to learn about inquiry itself. (Lazonder & Egberink, 2014) These results were later replicated by Lazonder and Wiskerke-Drost (2015).

In short supporting children's understanding of CVS, both through task segmentation and direct instruction, have shown both independently to lead to better design of unconfounded experiments by elementary school children. Explicit direct instruction can provide children with the necessary comprehension of the workings of CVS, whereas implicit use of task segmentation helps students learn using the CVS on a more meta-strategical level through constraining the experimental space and having them focus their experimenting on a single variable.

Research questions and hypothesis

Working on a guided-inquiry task requires elementary school children to conduct experiments. To conduct unconfounded experiments from which valid inferences can be made, mastering the CVS is crucial. Previous research on supporting children's learning of CVS by means of direct instruction prior to experimenting or using a task-segmented worksheet have shown to increase understanding of CVS and the design of unconfounded experiments (e.g. Klahr & Nigam, 2004; Kuhn & Dean, 2005). Prior research on the acquisition of the CVS has focused solely on comparing the effectiveness of direct instruction with task segmentation (e.g. Lazonder & Egberink, 2014; Lazonder & Wiskerke-Drost, 2015) and has argued for the possible superiority of either direct instruction (Klahr & Nigam, 2004) or task segmentation (Kuhn & Dean, 2005). The present research focused on combining both types of support.

The focus of the quantitative part of this study was on determining whether a positive synergistic effect exists when combining the two types of support, direct instruction (DI) and task segmentation when compared to the individual types of support and an unsupported condition. Previous research found an increase in knowledge of the CVS and a higher percentage of unconfounded experiments when using either one of these types of support. It was therefore expected that receiving an instruction on the CVS in combination with using a task-segmented worksheet where the variables are covered in successive order will lead to a higher understanding of how to apply the CVS whilst working on a guided-inquiry task on electric circuits and will lead to a higher percentage of CVS experiments.

It was expected that a positive synergistic effect between direct instruction and task segmentation exists and will lead to children being able to research more distinct variables compared to the individual support conditions and the control condition. The rationale behind this expectation being that the highly regulative effect of task segmentation (Lazonder & Egberink, 2014), combined with an increased knowledge on how to apply the CVS, should allow students to be able to identify the

CVS in the segmented worksheet and subsequently be able to design unconfounded experiments in a more structured manner compared to the other conditions. This will allow for children to cover more variables in their investigations.

One characteristic of inquiry learning is that it fosters children's curiosity and enthusiasm, leading to increased motivation (Minner, Levy & Century, 2009). In addition, prior research on inquiry-learning, more specifically on experimenting, has shown it is challenging for children to create unconfounded experiments (e.g. de Jong & Joolingen, 1998; Klahr & Nigam, 2004). It is therefore important to determine whether differences exist between conditions on how children experience the support offered. An inquiry-based learning activity should be challenging, but should not lead to frustration for children. Therefore, workflow will be measured whilst working on the activity by using the flow-questionnaire (Rheinberg, Vollmeyer, & Engeser, 2003). Using the flow-questionnaire makes it possible to measure whether the students are sufficiently challenged during the activity and whether they find themselves capable of completing the activity.

Prior research focused on direct instruction in CVS through use of a simulation. In this study, physical materials will be used to see whether direct instruction on the CVS yields the same effect. The context of electric circuits was chosen because research has shown (e.g. Osborne, 1983; Jabot & Henry, 2007) that working with electric circuits is highly challenging to children.

Working with physical materials makes it possible for children to make measurement errors, for instance an incorrectly connected wire in a circuit, when conducting their experiments. The qualitative part of this study focused on identifying whether differences existed between conditions in children's perceptions concerning CVS and how children dealt with incorrect circuits whilst working on the guided-inquiry task. Combining these two methods will provide the ability to not only measure increase of CVS-knowledge and the quality of experiments between conditions but also how children perceive the increased support they receive thus establishing a comprehensive picture of the effectiveness of combining direct instruction on CVS and task segmentation. The first research question and subsequent hypotheses will be examined through quantitative methods whilst the second research question is the focus of the qualitative part of this study.

1. What is the effectiveness of combining direct instruction and task segmentation on knowledge of CVS, number of variables investigated, perceived flow and percentage of CVS experiments for elementary school children working on a guided inquiry-based learning task in the domain of electricity?

H1: The combination of direct instruction and task segmentation leads to higher gains in knowledge on how to apply CVS in comparison to the TS, DI and control condition.

H2: The combination of direct instruction and task segmentation leads to children having a higher percentage of CVS experiments in comparison to the TS, DI and control condition.

H3: The combination of direct instruction and task segmentation leads to a higher number of variables covered in the experiments in comparison to the TS, DI and control condition.

H4: The combination of direct instruction and task segmentation leads to a higher level of perceived flow in comparison to the TS, DI and control condition.

2. How do elementary school children reflect on CVS-performance and dealing with incorrect circuits when working on a guided inquiry-task with different types of support?

Method

Research context

The context in which this study takes place is the domain of electricity, and more specifically working with electric circuits. Teaching the domain of electricity to elementary children is challenging, as the concepts of electricity, such as voltage and resistance, are not observable when conducting experiments with light bulbs and batteries (Ibtisam, Bar & Galili, 2006).

Although several misconceptions concerning electric circuits, resistance and electric current exists, two of these misconceptions are important to take into account when designing the inquiry activity. Tiberghien and Delacote (1976) observed that most children are not aware of the requirement of the circuit to be closed for it to work. This type of misconception on electric circuits is called the one-polar model. Children that use the one-polar model assume that by connecting a cable to one pole of the battery a light-bulb will illuminate. A second misconception often made by children is that, when in a circuit with one battery and two lights, one of the lights should illuminate brighter than the second light (Jabot & Henry, 2007; Cepni & Keles, 2006). In this model, dubbed the attenuation model, children view electric current as flowing in one direction and lightbulbs each use a part of the current leaving less current for the other bulbs (Chiu & Lin, 2005).

Participants

The participants of this study were 102 sixth graders, 51 boys and 51 girls, with a mean age of 11.04 (*SD* = 0.465). Children came from four different classrooms. In each classroom children were randomly assigned to one out of four conditions, the control group, the TS-group, the DI-group and the DI+TS group. A total of nine children were excluded from the dataset because they did not pay attention during the instruction or refused to participate in the learning activities during experimentation. A total of 93 children were included in the final sample. Table 1 shows the number of children per condition and the number of boys/girls in each condition after exclusion. Parental consent was requested for participation in the experiment and for video recording during the experiment. For six children, no parental consent was obtained for video recordings.

Table 1

Number of Children in Each Condition.

Condition	*N*	Boys	Girls
Control	24	14	10
TS	24	11	13
DI	22	9	13
DI+TS	23	13	10
Total	93	47	46

Note: TS = Task Segmentation, DI = Direct Instruction, DI+TS = Direct Instruction + Task Segmentation.

Materials

Inquiry learning task and introduction video

During the experimental sessions children worked on an inquiry learning task on electrical circuits. The goal of this task was to investigate the effects of four distinct variables on the luminosity of a light in a basic circuit. The children could investigate the effect of the following variables: the number of batteries (one or two), the number of lightbulbs (one or two), the length of the circuit (using between two and six cables) and the position of the light within the circuit. Four electricity sets were constructed to use in this experiment. Each set included two lights, with metal pins on the side to attach the wire, two batteries, and six alligator clip cables. This set offered the opportunity to discover how the four variables relating to circuits interacted. Figure 1 shows the materials that could be used.

A 3-minute long video on circuits was developed to show children how to set up a basic circuit with one light and one battery. In addition, this video showed how to connect two batteries in series and showed children were not allowed to connect more than one cable at a time to one of the metal pins of either the batteries or the lights. All children were shown this video prior to commencing with the experiment individually.

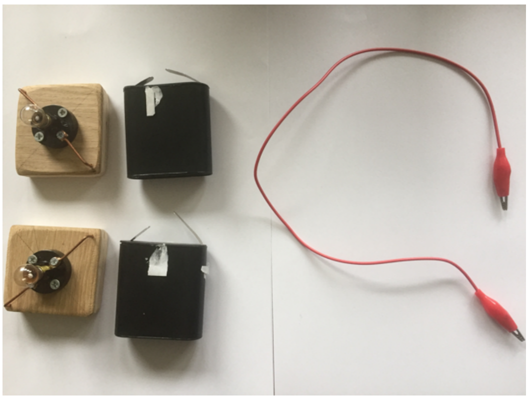

Figure 1: Children had two lightbulbs, two batteries and six alligtor clip cables available to them.

Task-segmented inquiry worksheet

The task-segmented inquiry worksheet consists of five research questions on electric circuits. The first question covered a basic circuit with one battery and one light and was used as a baseline for further experimenting. The following four questions covered one distinct variable at a time. Children were instructed to investigate the effect of four variables on the luminosity of a light. Each question offered room to write down conclusions. Children assigned to the TS and DI+TS condition used this worksheet during the experiment. An example of a research question is: "What happens to a light when you use one or two batteries?"

Unstructured inquiry worksheet

The unstructured inquiry worksheet consists of an open-inquiry question on electric circuits. The children received one research question which covered all four distinct variables previously mentioned at once. Similar to the task-segmented worksheet the first question covered a basic circuit with one battery, children were instructed to research the effect of the four variables on the luminosity of a light, and one light and was used as a baseline. Room to write down conclusions was provided for on the worksheet. Children assigned to the control and DI condition used this worksheet during the experiment. The research question children had to answer was: "What happens to a light when you use one or more than one battery, one or more than one light, change the position of the light in the circuit or use a short or long circuit."

Direct instruction on the CVS

Lazonder and Egberink (2014) developed a CVS-training for their research, which was used in this study. This training was situated in the domain of rolling objects and included a multi-variable experiment to teach children on how to control variables. Children had to roll a ball down a slope and discover how the variables weight of the ball, angle of the slope and position on the slope influenced how far the ball rolled. The slope had a high and low angled position. There were two balls, a heavy and a light ball. Finally, children had two positions on the slope from which to release the ball. The researcher of this study was also the teacher of this training. To keep the size of the training groups small, the group that received the CVS-training was split up between DI and DI+TS condition. This ensured a group size of between five and eight children per session. At the start of the training firstly the variables of the experiment and the overall research question "How far does the ball roll" were introduced. Subsequently, the researcher presented the children with an experiment which focused on the question "What is the influence of the type of ball (light/heavy) on how far the ball rolls". The two experiments conducted by the researcher to investigate this question were confounded, more than one variable was changed. The children were then asked if it could be stated with absolute certainty whether the results of the two experiments could be explained solely by the type of ball. Next, children gave suggestions for how modify the experiment to conduct a proper investigation of the research question and the correct experiment was conducted. After this introduction children were taught how to design unconfounded experiments. It was explained to them that firstly they need to have a subject or a question that can be answered through experimentation. Then it was explained that they need to choose a value for the subject of investigation, for instance a light or heavy ball. Finally, children were told that the other variables also needed a value, which has to be kept constant between experiments. After this introduction, children designed and conducted their own experiments to investigate the variables position on slope and angle of slope. These experiments were then carried out and feedback was given until the researcher was sure that all children comprehended the CVS. Figure 2 shows an example of an experiment children could design. This training was used for children assigned to the DI and DI+TS conditions.

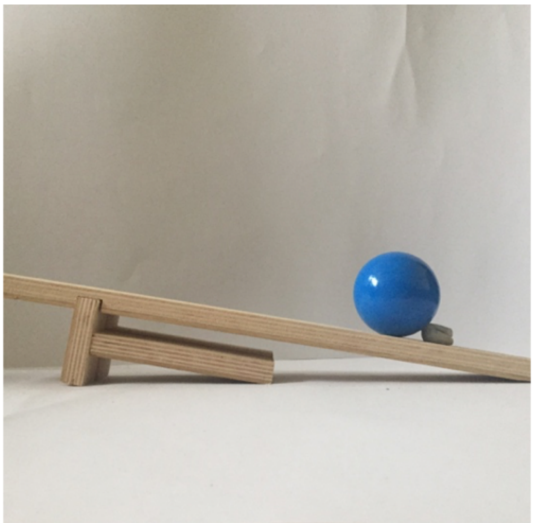

Figure 2: An example of a CVS-experiment with a light ball, low angle and low position.

Pre-test domain knowledge

This test consisted of five multiple-choice questions in the domain of electricity that children were expected to know after the inquiry-based task. The first question covered a basic circuit with one light and one battery. The other four questions each covered one of the four distinct variables could be investigated in the experiment. Each question was accompanied by two or three pictures. Figure 3 shows an example of one of the questions. One point was given for each correct answer, resulting in a total score ranging from zero through five. To ensure the target group could comprehend the items a primary school teacher checked the wording of the questions. In addition, a pilot was commenced with eight children of an eighth-grade class that was not included in the experiment to verify whether the target group understood and comprehended the questions. One question was changed based on observation of this pilot.

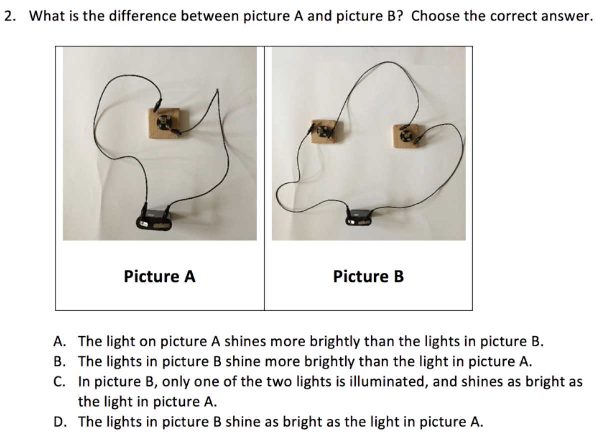

Figure 3: An example of a question of the knowledge test. This question concerned the influence of one or more batteries on the luminosity of a light.

CVS-test

A paper-and-pencil test was used to measure the ability of children to control variables in an experiment with multiple variables. This test was developed and validated by Lazonder and Egberink (2014) and is a Dutch version of the CVS-test based on the CVS-test developed by Chen and Klahr (1999). The CVS-test contained 9 questions concerning experiments with multiple variables. Cronbach's α was used to measure internal-consistency of the CVS-test. Testing showed an α of .84, which could be slightly improved to .85 by removing Question 5 (r = .293). Since an α of .84 indicates a high internal-consistency, the decision was made to include all nine questions for analysis. Scoring ranged from 0 through 15 points. One point was assigned when the question of whether the experiment was valid was answered correctly. When an experiment was invalid, one additional point was awarded when the experiment was improved correctly. Cohen's Kappa κ was used to calculate inter-rater reliability. A total of 52 tests, one-fourth of the total, were dual-coded. The coders had a high level of agreement, κ = .975. One student was absent during the pre- and post-CVS-test. Figure 4 shows an example of a question asked in the CVS-test.

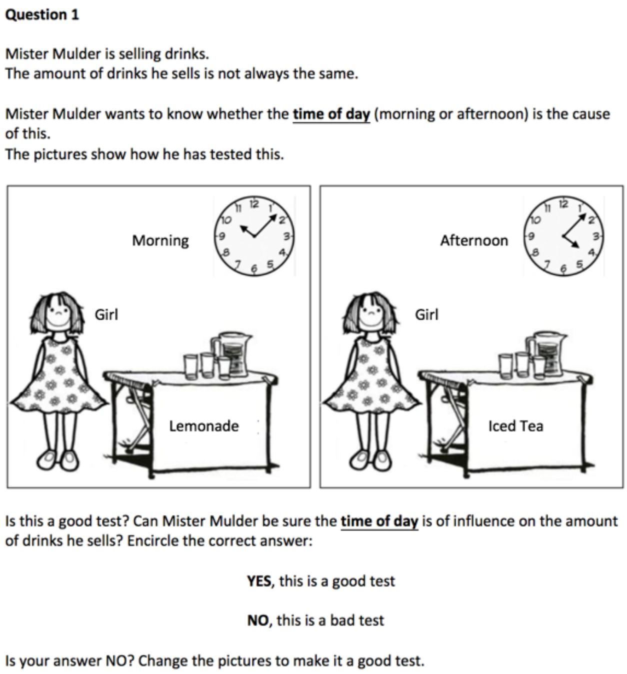

Figure 4: An example item from the CVS-test.

Flow-questionnaire

A Dutch translation of the Flow Short Scale (Rheinberg, Vollmeyer, & Engeser, 2003) was used to measure flow. This questionnaire consisted of nine items that were answered on a 7-point Likert-scale. An example item was: "I think this exercise is useful". Low scores on the flow-questionnaire indicated high agreement, high scores indicated low agreement. Children filled out this interview prior to commencing with the inquiry-activity, after eight minutes into the activity and at the end of the activity. Cronbach's α was used to measure internal-consistency of the flow-questionnaire. The Cronbach's α for the scale was .85, with no option for improvement by removing questions.

Video-coding

A coding scheme was developed to code the videos of the children working on the inquiry-task. ELAN (Lausberg & Sloetjes, 2009) was used to code the video's. ELAN is software developed at the Max Planck Institute for Psycholinguistics, The Language Archive, Nijmegen, The Netherlands and can be used to add segments of code to video-recordings. The coding process consisted of two steps, the first step focussed on identification of experiments and in the second round the experiments were classified. For each step of coding a total of 32 videos were dual coded to determine inter-rater reliability. Because the children were shown how to conduct a simple circuit of one light, one battery and two wires, and the first assignment of both worksheets was creating this simple circuit, the circuit with one battery and one light was used as a baseline from which children conducted other experiments by adding for instance one light or one battery. For each child, it was determined whether this baseline-experiment was present, however this experiment was not coded as an experiment and thus did not count towards the total number of experiments. The first step focused on determining whether the behaviour of the children could be marked as an experiment or showed other behaviour not related to the experiment. When a change was made in the experimental setup this qualified as a unique experiment, for instance when a light was added or removed in the circuit. Cohen's Kappa κ was used to calculate inter-rater reliability and reached κ = .860. During the second step a code was assigned to the behaviours categorized as experiments coded during the first step. Table 2 shows the codes that could be assigned to the experiments. A distinction was made between errors that could be made within the limits of the assignment concerning incorrectly connected circuits (IC) and errors that were made that fell outside of the assignment as well as experiments that were partly visible on the video-recording (OTHER). An experiment was coded as CVS when either only one variable was changed between the experiment and the baseline experiment, or when two experiments were conducted as a pre-post-test. When an experiment was coded as CVS, the coders had to make a distinction between CVS codes 1 through 4, relating to the four distinct variables that children were asked to investigate. In addition, CVS 5 was added for experiments which followed the rationale of CVS, but were not covered in the assignment. An example of experiments categorized as CVS5 is comparing a circuit with two batteries followed by two lights, with a circuit that has a battery, a light, a battery and a light. Cohen's Kappa κ for the second step reached κ = .937.

Table 2

Explanation of Codes Assigned to Experiments.

Code	Explanation
CVS1	An unconfounded experiment which measures the effect of one or more batteries.
CVS2	An unconfounded experiment which measures the effect of one or more lights.
CVS3	An unconfounded experiment which measures the effect of a circuit with more/less wires.
CVS4	An unconfounded experiment which measures the effect of the position of the light within the circuit.
CVS5	An unconfounded experiment which logically falls under CVS, but does not cover one of the above variables.
NCVS	A confounded experiment where more than one variable is changed.
IC	An experiment where either through an error in the experimental setup or through incorrectly connecting wires the circuit does not work correctly.
OTHER	Covers parallel circuits (not allowed) and experiments where it is not possible to determine what code to assign since part of the experiment is not visible.

Interview-script

The interview was set up as a cued-retrospective, semi-structured interview. After children completed the experimentation, the researcher watched the video-recordings and selected two children per condition to be interviewed. The interview took place one day after the actual experiment. Interview questions were inspired by children's behaviour during the experiment and focussed on the following issues: Frequency of CVS / Non-CVS, dealing with incorrect circuits, frequency of experiments. However, since interview-questions were inspired by the behaviour of the children during the experiment, other interesting behaviours could be included as well. In the interview scripts the questions were accompanied by time-stamps. These time-stamps coincided with when the behaviour was shown on the video. Children watched the selected video-segment, and subsequently the researcher asked the question. An example question is: "Your light was not working in this experiment. Did you discover why?"

Procedure

The experimental sessions and interviews took place during two regular schooldays. At the first day of the experiment all children completed the CVS-test and the domain knowledge test. Next, all children watched the introduction video on the domain of electricity, followed by a short explanation of the flow-questionnaire. In a separate classroom three separate tables were set up to accommodate one child per table. A video camera was present at each of these tables to film the children. On each table one electricity set was available for experimentation. Children worked individually on the experimental task for a maximum of twenty minutes. To prevent validity-threats the sequence of children working on the experimental task was such that children that did not get CVS-instruction made the inquiry-tasks on electricity first. To keep the groups that were instructed low in size, the CVS-instruction was given twice, first the DI-group was taught and afterwards the DI+TS group was taught, after which these children also worked individually on the inquiry-task. After all the children finished working on the inquiry-task the CVS-test was made again. During day two, which took place the day after the children worked on the experiment, two children per condition were interviewed individually on their actions during the inquiry by watching segments of the video-recordings and answering questions based on the interview-scripts.

Results

Prior to answering the research questions, it was tested whether there were initial differences between conditions. Subsequently, further analysis of each dependent variable in the study was conducted to test the hypotheses and describe the children's statements during the interviews.

A total of 91 children participated in the knowledge-test prior to the experiment. Two children were absent during the knowledge pre-test. The children were randomly assigned to conditions so no a priori differences were expected. A Kolmogorov-Smirnov test was used to test for normality on the main dependent variable knowledge test. The result of this test was significant, D(91) = .171, $p < .001$, indicating that the data was not normally distributed. Therefore, a Kruskal-Wallis test was performed to determine whether there were a-priori differences in prior knowledge between conditions, $\chi^2(3, N = 91) = 3.290$, $p = .349$. In short, the children had similar scores on the prior knowledge test concerning circuits.

The results of a Kolmogorov-Smirnov test on the scores of the CVS-pre-test D(92) = .155, $p < .001$, indicated that that the data was not normally distributed. A Kruskal-Wallis test showed that there was no statistically significant difference in pre-test scores on the CVS-test between conditions, $\chi^2(3, N = 91) = .768$, $p = .857$.

Learning gain in CVS

Children's knowledge of the CVS was tested prior to and after the experiment. Learning gains of were computed by subtracting post-CVS from pre-CVS scores. Table 3 shows the mean scores by condition Kolmogorov-Smirnov test was used to test for normality on the main dependent variable CVS-gain. The Kolmogorov-Smirnov test was significant, D(91) = 0.193, $p < .001$, indicating that the data was not normally distributed. A Kruskal-Wallis H test showed that there was a statistically significant difference in children's gain score on the CVS test between conditions, $\chi^2(3, N = 91) = 10.183$, $p = .017$. To test our expectation that children in the three treatment conditions would outperform their peers from the control-condition, three Mann-Whitney U tests were conducted. Bonferroni-adjusted post-hoc α for the Mann-Whitney U tests was α =.017. The Mann-Whitney U test showed the DI-condition scored significantly higher on the CVS-test compared to the control condition, $U = 120$, $p = .002$, $r = .45$. There was no significant difference between the TS and control condition, $U = 245$, $p = .505$, and between the DI+TS and control condition, $U = 255$, $p = .650$. Further testing showed that TS ($U = 152$, $p = .033$, $r = .32$) and DI+TS ($U = 136.5$, $p = .012$, $r = .38$) scored significantly lower than the DI-condition.

Table 3

Results of the CVS-test.

	Pretest			Posttest			Gain		
	M	*SD*	*N*	*M*	*SD*	*N*	*M*	*SD*	*N*
Control	9.210	4.863	24	9.540	5.082	24	0.330	2.582	24
TS	9.670	4.219	24	10.780	4.078	23	0.830	2.823	23
DI	9.100	4.011	21	11.670	3.498	21	2.570	2.336	21
DI+TS	8.570	4.110	23	9.300	4.646	23	0.740	2.700	23
Total	9.140	4.273	92	10.290	4.423	91	1.080	2.713	91

Note: TS = Task Segmentation, DI = Direct Instruction, DI+TS = Direct Instruction + Task Segmentation.

Number and percentage of experiments.

From the coded video data, the number of experiments was extracted. Table 4 shows the mean scores for each condition on the number of experiments conducted. Please note that we only have video data for the 87 children (out of 93) whose parents provided permission for video recording.

Table 4

Mean Number of Experiments.

Condition	*M*	*SD*	*N*
Control	8.170	4.997	23
TS	7.500	3.189	22
DI	8.150	3.964	20
DI+TS	6.090	3.115	22
Total	7.470	3.929	87

Note: TS = Task Segmentation, DI = Direct Instruction, DI+TS = Direct Instruction + Task Segmentation.

The number of experiments classified as CVS, NON-CVS, IC and OTHER were computed as a percentage relative to the total number of experiments for each participant. Table 5 shows the mean percentages of each condition. A Kolmogorov-Smirnov test was used to test for normality on the main dependent variables CVS and NON-CVS. The result of the Kolmogorov-Smirnov CVS: D(87) = .136, $p < .001$ and NON_CVS: $D(87) = .213$, $p < .001$ were significant, indicating that the data was not normally distributed. The Kruskal-Wallis H tests showed that there was no statistically significant difference between conditions for the variable NCVS: $\chi^2(3, N = 87) = 6.681, p = .083$. A Kruskal-Wallis test for CVS showed a significant difference between conditions, $\chi^2(3, N = 87) = 8.185, p = .042$. Based on the expectation that the three treatment conditions would outperform the control condition, three Mann-Whitney tests were conducted with a Bonferroni-adjusted post-hoc α of α = .017. Results of the Mann-Whitney tests showed that the TS-condition had a significantly higher percentage of CVS-experiments than the DI condition, (U = 114.5, p = .008, r = .41). The DI+TS condition (U =158.5 , p = .046, r = .30) and control condition (U = 161.5, p = .036, r = .31) did not score significantly higher than the TS-condition on percentage of CVS-experiments. Since the expectation was that the DI condition would score a lower percentage of NON-CVS in comparison to the TS condition an additional Mann-Whitney U test was conducted. Bonferroni-adjusted post-hoc α for the Mann-Whitney U tests was α = 0.017. The Mann-Whitney showed the percentage of NON-CVS was significantly lower for the TS condition compared to the DI condition, U =122.5 , p = .011, r = .39.

Table 5

Mean Scores of Codes Assigned to Experiments as a Percentage of Total Experiments Conducted.

	CVS		NON-CVS		IC		Other	
	M	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
Control	0.551	0.339	0.113	0.114	0.112	0.241	0.228	0.337
TS	0.733	0.306	0.103	0.158	0.100	0.139	0.063	0.167
DI	0.449	0.337	0.229	0.203	0.199	0.209	0.115	0.192
DI+TS	0.519	0.387	0.178	0.257	0.207	0.252	0.087	0.210
Total	0.566	0.353	0.153	0.193	0.153	0.217	0.124	0.242

Note: TS = Task Segmentation, DI = Direct Instruction, DI+TS = Direct Instruction + Task Segmentation.

Covering all variables

Children's experiments were analysed to determine to what extent the aspects of the general research question (see Table 2) were covered. A variable was computed which ranged from zero to four, based on the number of unique variables (see Table 6). The result of a Kolmogorov-Smirnov on the variable covering all variables, D(87) = .204, $p < .001$, was significant, indicating that the data was not normally distributed. A Kruskal-Wallis H test showed that there was no statistically significant difference between the four conditions in the number of variables investigated, $\chi^2(3) = 4.683$, $p = .197$.

Table 6

Mean Number of Variables Investigated.

	M	*SD*	*N*
Control	2.090	1.411	23
TS	2.730	1.202	22
DI	1.900	1.334	20
DI+TS	1.900	1.659	22
Total	2.160	1.430	87

Note: TS = Task Segmentation, DI = Direct Instruction, DI+TS = Direct Instruction + Task Segmentation.

Flow-questionnaire.

Table 7 shows the mean total scores for each condition on the flow-questionnaire, with low scores (e.g. 1, 2, 3) indicating a high agreement. For each of the measurement-points a total score was computed by adding the scores on items one through nine and dividing this total by nine. As one child did not complete the flow-questionnaire, there were 92 children in the analysis. A Kolmogorov-Smirnov test was used to test for normality on the main dependent variable flow-questionnaire. The result for initial flow: $D(92) = 0.096$, $p < .001$, intermediate flow: $D(92) = 0.149$, $p < .001$, final flow: $D(87) = 0.205$, $p < .001$ were significant, indicating that the data was not normally distributed. Three Kruskal-Wallis tests were then conducted to see whether the conditions differed significantly on each measurement. The Kruskal-Wallis H test showed that there was no statistically significant difference between conditions on initial flow, $\chi^2(3, N = 92) = 1.449$, $p = .694$, intermediate flow, $\chi^2(3, N = 92) = 1.428$, $p = .699$. final flow, $\chi^2(3, N = 92) = 2.873$, $p = .412$. Subsequently three new variables were computed to analyse changes in flow across measurement points. Kolmogorov-Smirnov tests showed that the difference between initial flow and intermediate flow, $D(92) = 0.105$, $p = .014$, difference intermediate and final flow, $D(91) = 0.179$, $p < .001$, and initial and final flow, $D(87) = 0.105$, $p = .015$ were not normally distributed. The Kruskal-Wallis H test showed that there was no statistically significant difference between conditions on the difference between measurements, initial to intermediate, $\chi^2(3, N = 92) = 0.658$, $p = .883$, intermediate to final, $\chi^2(3, N = 92) = 7.043$, $p = .071$, and initial to final, $\chi^2(3, N = 92) = 0.857$, $p = .836$.

Table 7

Mean Scores on the Flow-questionnaire.

	Initial flow		Intermediate flow		Final flow	
	M	*SD*	*M*	*SD*	*M*	*SD*
Control	2.240	0.898	1.884	0.948	1.606	0.771
TS	2.251	1.191	2.245	1.255	1.990	1.424
DI	2.161	0.519	1.975	0.750	1.825	0.844
DI+TS	2.363	0.668	1.990	0.931	2,029	1.017
Total	2.329	0.865	2.025	0.987	1.862	1.048

Note: N=92, TS = Task Segmentation, DI = Direct Instruction, DI+TS = Direct Instruction + Task Segmentation.

Children's perceptions

In the next sections, excerpts from the children's interviews are presented to illustrate how children perceived working on the guided inquiry-task within the different conditions. A total of eight children from each condition were interviewed. These excerpts will be used to highlight the differences between conditions concerning use of the CVS and dealing with incorrect circuits. The final section provides excerpts that were not the focus of the initial research question but do provide valuable insights into the process of inquiry. In the excerpts the time-stamp is also provided, where applicable.

Use of CVS

Quantitative analysis showed that children in the DI condition performed better than children in the other conditions concerning their understanding of CVS. However, this superior understanding of CVS did not necessarily translate into more unconfounded experimenting. This is illustrated by the excerpts from post-experimentation interviews with the children. In excerpt 1, a child with a relatively high score experiences problems with the design of unconfounded experiments. Children also seemed to struggle with the transfer from the CVS-lesson to the context of working with electric circuits, as excerpt 2 shows. Finally, excerpt 3 shows that, even with low scores on the CVS-test, some children would still be able to perform well in designing unconfounded experiments.

Excerpt 1, Elise, DI condition

Elise scored eight points on the pre-test and twelve points on the post CVS-test. Elise conducted a total of 9 experiments, 2 of which were CVS, one was a IC and six were Non-CVS. Elise covered one of the four distinct variables in her experiments. Excerpt 1, shown in Table 8, describes one of the experiments Elise designed. In this experiment, which was categorized as non-CVS, Elise went from one to two batteries, and added cables. Because of this, Elise struggled to determine the correct cause of the light shining brighter.

Table 8

Excerpt 1: Episode from the interview with Elise

Person	Utterance
Interviewer	15:00 – 15:20 In this experiment you have two batteries, one light and six cables. What did you discover here?
Elise	Well, euhm, these two still needed to be connected here. Those already illuminated, but the other two did not illuminate yet.
Interviewer	And concerning the brightness, did you discover anything?
Elise	Yes, they burned brighter, it was really bright.
Interviewer	And what was the cause of that?
Elise	Maybe because I used more cables.
Interviewer	The fact that you used more cables caused the light to shine brighter?
Elise	Yes

Excerpt 2, Shreya, DI condition

Shreya was absent during the pre-and post CVS-test. Shreya attended the CVS-lesson and watched the instruction video. Shreya conducted a total of 4 experiments, 3 of which were CVS and one being Non-CVS. Shreya covered three of the four distinct variables in her experiment. Excerpt 2, shown in Table 9, describes the experiment coded as non-CVS. She started out with an experimental setup with two lights and two batteries. Because of the fact she manipulated two variables Shreya struggled in finding the correct conclusion. In addition, even when specifically referencing to the prior CVS-lesson Shreya does not seem to notice the flaw in her experimental setup.

Table 9

Excerpt 2: Episode from the interview with Shreya

Person	Utterance
Interviewer	3:10 – 3:20 What was the goal of this experiment, since you added a second battery?
Shreya	I wanted to check whether it would be brighter with two batteries.
Interviewer	And what did you conclude?
Shreya	I think the two lights burned brighter with two batteries.
Interviewer	And can you know for sure this is caused by the two batteries. Or could it also be

Interviewer	caused by the two lights?
Shreya	I think it is caused by the extra battery.
Interviewer	Did you remember the lesson you received prior to the task? Where you could only a
Interviewer	adjust one thing at a time?
Shreya	Yes
Interviewer	And did you only adjust one thing or multiple things.
Shreya	Multiple things.

Excerpt 3, Floris, TS condition.

Floris scored 7 points on the pre-CVS-test and six points on the post CVS-test. Floris conducted a total of 9 experiments, 7 of which were CVS and two were non-CVS. Floris covered three of the four distinct variables in his experiment. Even though Floris scored below average on the CVS-test and did not improve during the experiment, Excerpt 3, shown in Table 10, shows Floris being able to apply CVS correctly in his experiments. This excerpt described an experiment with two batteries and one light. Floris not only designs his experiment correctly, he also adds a post-test to experiment to check whether his findings were correct.

Table 10

Excerpt 3: Episode from the interview with Floris

Person	Utterance
Interviewer	What was your conclusion for this experiment?
Floris	Brighter light
Interviewer	And what was the cause?
Floris	Two batteries.
Interviewer	You discovered that two batteries is brighter or less bright than one?
Floris	Brighter.
Interviewer	3:50 – 4:00 Why did you perform this experiment right after the first experiment?
Floris	What is brighter, one battery or two batteries.
Interviewer	And why did you do this second experiment?
Floris	To test, because the first was already with two batteries and afterwards one.
Interviewer	The goal was to compare them?

Floris	Yes.

Dealing with incorrect circuits

In absolute numbers, out of a total of 650 experiments conducted by children, 97 experiments were coded as incorrect circuits. The following excerpts show that incorrect circuits were made independent of condition and level of comprehension of CVS.

Excerpt 4 & 5, Heleen, TS condition.

Heleen earned the maximum score on both the pre- and post CVS-test. Heleen conducted a total of 12 experiments, 10 of which were CVS, one was a IC and one was Non-CVS. Heleen covered the four distinct variables in her experiment. The following excerpts, presented in Table 11 and 12, concerns two experiments conducted, one with time-stamp 5:00 – 5:10 and one with time-stamp 17:20 – 17:30. At time-stamp 5:00 – 5:10 Heleen tried to conduct an experiment with two batteries and one light. The experiment did not work because she connected the two plus poles of the two batteries. During the second time-stamp Heleen stated that because she learned much during the experiments she tried to give it another go, and connected the experiment correctly, resulting in a functioning circuit. The TS condition did not give any specific support for dealing with incorrect circuits. Although Heleen had a maximum score on the CVS-test, she still struggled with a flaw in her experimental setup. However, she could correct this error, whether this was by luck or insight acquired during the guided-inquiry task cannot be determined.

Table 11

Excerpt 4: Episode from the interview with Heleen

Person	Utterance
Interviewer	5:00 – 5:10 Another experiment, you try to connect 2 batteries and one light bulb.
Interviewer	What was the purpose of this experiment?
Heleen	To see whether the light would shine brighter.
Interviewer	Ok, so to see whether the light would shine brighter with two batteries compared to one?
Heleen	Yes
Interviewer	Did you discover why it did not work?
Heleen	I did not, I thought I made an error.
Interviewer	And did you check whether you made an error?
Heleen	Yes, I did at the end. It worked at the end.
Heleen	After it did not work I also checked whether everything was connected.
Interviewer	And did you discover at first why it did not work?

Heleen	No

Table 12

Excerpt 5: Episode from the interview with Heleen

Person	Utterance
Interviewer	17:20 – 17:30 The cool thing is that you also say wow during this experiment.
Heleen	Yes, because the light was so much brighter.
Interviewer	Ok, and this is the experiment that did not work before.
Interviewer	Why did you try it again at the end?
Heleen	I thought it might work since I knew more than at the beginning.
Interviewer	You thought maybe I made an error at the beginning, and because of the many experiments conducted you gave it a go again at the end?
Heleen	Yes!
Interviewer	And what was your conclusion concerning the difference between one or two batteries?
Heleen	When there are more batteries and one light the light shines brighter.

Excerpt 6 & 7, Pien, DI+TS condition.

Pien scored five on the pre-CVS-test and three on the post-CVS test. Pien conducted a total of six experiments, two of which were CVS, two were IC and two were Non-CVS. Pien only covered one out of a possible four distinct variables in her experiment. Table 13 and Table 14 describe two episodes of the interview with Pien. Pien was assigned to the DI+TS condition, the condition with the highest level of support. However, the first experiment Pien attempted to do was a Non-CVS experiment, which did not work resulting from incorrectly connecting the batteries. From 10:00 onward, Pien kept attempting to get the circuit working, but was unable to do so.

Table 13

Excerpt 6: Episode from the interview with Pien

Person	Utterance
Interviewer	3:00 – 5:30 It took a while for you to set up your experiment. Were you in doubt about something?
Pien	I was struggling with making a circuit with two batteries and two lights.
Interviewer	Ok, and did you discover why the lights did not shine?

Pien	No
Interviewer	You did not discover the error that you made here?
Pien	No

Table 14

Excerpt 7: Episode from the interview with Pien

Person	Utterance
Interviewer	10:30 – 10:40 Why did you repeat this experiment?
Pien	Because I was still trying to figure out how to get two batteries working.
Interviewer	Ok, so you were still attempting to include the two batteries in the circuit?
Pien	Yes

Alternative experiments

During video-coding it became apparent that children employed various techniques for checking their answers. In addition, the CVS 5 code was assigned to experiments that could be categorized as CVS but were not required for completing the experiment. Excerpt 8 gives an example of experimental design by a child to verify their answers. Excerpt 9 shows an example of an alternative experiment, various children conducted experiments beyond the scope of the exercise.

Excerpt 8, Zoe, TS-condition

Zoe scored seven on the pre-CVS-test and five on the post-CVS test. Zoe conducted a total of four experiments, all four of them were categorized as CVS. Zoe covered two of the four distinct variables. During the experiment, Zoe had two occasions where she made two experimental setups. Excerpt 5, shown in Table 15, shows the reasoning behind one of these experiments. Children sometimes struggled to determine whether the luminosity differed between experiments. To check whether differences existed several children conducted a pre-post-test, whilst others like Zoe created two experimental setups.

Table 15

Excerpt 8: Episode from the interview with Zoe

Person	Utterance
Interviewer	14:10-14:20 Why did you do these two experiments simultaneously? You have one experiment with a battery, two cables and a light, and a second experiment with a battery, four cables, and a light.
Zoe	To see whether the light was brighter. As you can see I am holding my hand behind

	both lights to see whether there is a difference.
Interviewer	And was there a difference when more cables were added?
Zoe	A little bit of a difference.
Interviewer	Ok, and here you were checking?
Zoe	With more cables, what would happen.

Excerpt 9, Antonia, control condition

Antonia scored nine on the pre-CVS-test and five on the post-CVS test. Antonia conducted a total of twelve experiments, nine of which were CVS, one was IC and two were Non-CVS. Antonia covered all the distinct variables in her experiments. What is remarkable is that, although she was clearly able to design unconfounded experiments, her CVS-test score decreased. Excerpt 9, shown in Table 16, describes an alternative experiment. In this experiment, Antonia makes a comparison between a circuit with two batteries and then two lights or a circuit with one battery, one light, one battery, one light. Working with physical materials offered the opportunity to design experiments beyond the scope of the research question. This behaviour was seen in all the experimental conditions. However, children in the control condition clearly struggled with determining when the guided inquiry-task was finished.

Table 16

Excerpt 9: Episode from the interview with Antonia

Person	Utterance
Interviewer	7:20-7:30 / 10:10-10:20 This seems to be the same experiment, with 2 lights and 2 batteries, twice. What was the reason for conducting this experiment twice?
Antonia	Well, euhm, as you can see, first the light is connected to the light, and second the light is connected to the battery. That is what I wanted.
Interviewer	Ok, so what you were investigating, if I understand correctly. In the first experiment, you had a battery connected to a battery and a light connected to a light. Then in the second experiment you tried to discover what would happen when your circuit would be battery, light, battery light?
Antonia	Yes
Interviewer	And was there any difference?
Antonia	No

Discussion

The aim of this study was to investigate whether combining a task-segmented worksheet with a CVS-instruction would lead to children having a better understanding of the CVS, a higher percentage of unconfounded experiments, a higher flow whilst working and a higher number of variables covered in experiments. In addition, the qualitative part of this study focused on how the children reflected on the guided inquiry-task regarding to CVS and dealing with incorrect circuits. Four conditions were compared to each other: (a) unstructured inquiry (e.g. control-condition), (b) task segmentation, (c) direct instruction and (d) direct instruction + task segmentation.

Results showed that the DI+TS did not perform significantly better concerning understanding of CVS and percentage of CVS experiments. In addition, concerning the number of variables covered and perceived flow the DI+TS condition did not outperform the other three conditions. No positive synergistic effect of combining direct instruction with task segmentation was found. A possible explanation could be that both types of support were left intact in their original form instead of for instance providing ques in the segmented-worksheet that reminded the children of the lesson they received prior to commencing with the inquiry-task. Although empirical research on synergistic effects of a single type of support compared with the combination of two types of support is scarce, a study conducted by Zydney (2010) reported similar effects when combining two support tools. The combined condition was outperformed by the condition that used only a single scaffold. Furthermore, Zydney stated these two tools interfered with each other's performance. Although the language and structure of both the direct instruction and the task segmentation was consistent, no further adaptations to these support types were made. Further research where for instance by using heuristics the worksheet is linked to the instruction on CVS might allow for a better functioning combination of these support techniques. Bjork and Bjork (2011) offer a second possible explanation. According to Bjork and Bjork (2011), encountering desirable difficulties when working on a learning task likely induces more transfer-appropriate processing leading more understanding of the to be learned strategy or content. The combination of direct instruction and task segmentation might have led to children encountering less desirable difficulties, it offered children less opportunity to practice the CVS, in turn causing them to gain little knowledge on the CVS. An alternative explanation might be that combining both types of support induces such high cognitive load (Sweller, 1994) that children to performed worse compared to the other conditions. Several scholars advocated against the use of inquiry-based methods for this reason (e.g. Mayer, 2004; Kirschner, Sweller & Clark, 2006) due to the high strain it allegedly puts on cognitive load when no support is offered. The same might be the case when too much support is offered. The CVS-training trains children in dealing with multiple variables but the task segmentation already offers structured guidance in how to deal with these. This can possibly be conflicting to children concerning whether to decide on which variable to manipulate themselves or follow the order already stated on the worksheet.

Although the DI+TS condition did not perform as predicted there are two other interesting findings worth mentioning. The DI-condition performed significantly better on the CVS post-test, on average scoring a modest 1.5 points higher than other conditions. Furthermore, the TS condition completed a significantly higher percentage of unconfounded experiments when compared to the DI-condition. This is a remarkable find, since children scored significantly higher in the DI-condition on knowledge of CVS. Prior research conducted by Lazonder and Kamp (2014) did not find any increase in knowledge of the CVS for either the direct-instruction or the task segmentation condition. An explanation for this finding is this study used smaller groups for CVS-instruction in comparison to prior research where the CVS was taught whole class. Qualitative analysis of the interviews gives us a possible explanation for the TS condition scoring higher on percentage of unconfounded experiments. Even when explicitly asked, children struggled transferring the content of the CVS-lesson to the domain of electric circuits. Since the task-segmented worksheet offers a high level of guidance and restricted further experimenting, children could design a higher percentage of CVS experiments because it was clearer where to start and which variables to manipulate. This regulative effect of using task-segmentation was also found in prior research (Lazonder & Kamp, 2012; Lazonder & Wiskerke-Drost, 2015). Additionally, since the unguided worksheet only offered one research question, covering all four variables, it was difficult for children in the DI condition to determine when they were finished with the guided inquiry-task. A second explanation is that this experiment was situated in the domain of electricity, more specifically electric circuits. Although the CVS-training has proven to increase the level of unconfounded experiments in the past working in other domains (e.g. Chen & Klahr, 1999; Klahr & Nigam, 2004; Lazonder & Egbertink, 2014), physically working with electric circuits might be more demanding than other domains, thus leading to the CVS-training being less effective when looking at the quality of experiments conducted. Previous research has shown that children struggle with comprehending electricity and working with electric circuits (e.g. Osborne, 1983; Jabot & Henry, 2007). Transferring the knowledge of CVS from the domain of rolling objects to the more challenging domain of electricity might have been difficult. Other forms of guidance through scaffolding might be more effective when working with electric circuits. A third possible explanation relates to the use of physical materials in this study. Research in combining a virtual laboratory with physical experiments when teaching students on electric circuits has shown to increase domain knowledge when compared to solely using virtual laboratory and lab work (e.g. Zacharia, 2007; Zacheria & de Jong, 2014). However, these studies did not take into account the level of CVS used by students. When working in a simulation-based environment with the goal of training children in CVS, children are restricted to investigating the variables presented to them in an off/on manner. This confining of choices on the one hand allows children to gain more domain knowledge (Zacheria, 2007), but on the other hand might lead to children falling behind in experimental skills when compared to children working solely with physical experiments. When the goal of the science lesson is acquiring inquiry-skills, more specific designing unconfounded experiments, physicality might be more of influence compared to when the goal is acquiring domain knowledge.

No significant differences were found between conditions concerning unique variables researched, and higher levels of perceived flow. A possible explanation for finding no significant difference on unique variables researched is that the number of possible experiments children could set up, was limited. Even through trial and error most children could create unconfounded experiments for at least two of the four possible distinct variables. In sum, although task segmentation was found to be highly regulative concerning the percentage of experiments that were CVS, this did not automatically lead to them completing a higher percentage of their assignment. This finding is in line with previous research conducted by Lazonder and Kamp (2012). Concerning the levels of flow most children reported enjoying working with physical materials during the interviews. Additionally, several children reported not having worked with electric circuits before. This novelty effect could have led to children being overtly positive when filling out the flow-questionnaire.

Qualitative analysis offered several additional explanations for the results. Some children struggled with the basics of constructing a circuit, even though this was covered in the instructional video shown prior to the guided inquiry-activity. Some children were unable to create a closed circuit. Several children could overcome incorrect circuits, whereas others got stuck by a faulty experimental setup. Furthermore, the fact that children were working with physical materials offered the opportunity to conduct experiments beyond the scope of the research question. Prior research conducted by Lazonder and Kamp (2012) also reported this effect. As mentioned previously, working in a simulation in some occasions restricts decision making compared to working with physical materials. Zacheria and de Jong (2014) found that when comparing virtual manipulatives with physical manipulatives undergraduate students often ran into process-related problems that prevented them from properly forming conceptual models. Whether having the opportunity to make incorrect circuits, thus leading to measurement errors, affects children’s acquisition of experimentation skills could be the focus of future research.

The main limitation of this study was that the quality of inferences derived by children from their experiments were not included in this study. Including the quality of inferences made could have offered a more comprehensive picture on the effectiveness of direct instruction on CVS and task-segmentation. Although results of this study can be informative concerning the process of experimenting, making claims on whether the children gained any domain-knowledge is beyond the scope of this study.

Future research could explore which types of support is most appropriate for inquiry-learning tasks in the domain of electricity. Additionally, when differences between domains exist, more research is needed into whether the CVS-instruction leads to a higher quality of experiments when working on domains that have not been previously studied in this regard. Furthermore, future research should focus on whether the differences that exists between simulations and working with physical materials lead to differences in mastering CVS and conducting unconfounded experiments. This research should especially focus on the role of measurement errors as it increases the desirable difficulty of the learning material (Bjork & Bjork, 2011). Does including the possibility for making measurement errors, when working in guided inquiry, make children better at performing experiments? Moreover, combining direct instruction on CVS with task segmentation should be used in a different, less demanding domain to whether the absence of a positive synergistic effect found in this study persist in a different context. Finally, combining virtual manipulatives and physical manipulatives have shown to increase domain knowledge, whether this is also positively influences the level of comprehension of the CVS when compared to only virtual or physical experimentation is still unclear.

This study contributes to theory as well as practice. This study supports prior findings concerning the regulative effect of task segmentation. In addition, results concerning the gain in domain knowledge on the CVS in the direct instruction condition are in line with previous findings on the effect of direct instruction on the CVS. On a theoretical level a remarkable outcome of this study was that even though children in the direct instruction condition performed significantly better on knowledge of CVS, this did not show in the percentage of unconfounded experiments they designed. Working in the challenging domain of electricity might require more or different types of support for the CVS to be correctly applied. However, this study also showed that direct-instruction on CVS can be more effective when children are taught in smaller groups. Also, this study showed that combining types of support is a delicate process and will not always lead to increased performance and learning of the CVS. This study provides a starting point on the notion that not all of science domains are equally fitting for teaching the CVS and that children might struggle in transferring use of the CVS more when domains are more challenging in terms of designing experiments. Furthermore, results from this study suggest that physicality might be of influence when acquiring the CVS.

On a practical level, several implications can be drawn from this study. When teaching children how to work using the CVS, teachers should consider which domain to use for this purpose as some domains are more appropriate than others. Furthermore, prior to commencing with the activity teachers should be aware of possible measurement errors children might come across and provide support accordingly. In addition, teachers should be aware that combining support types does not always lead to a positive synergistic effect. Even when a support is highly regulative, this not always leads to the desired outcome, for instance learning the CVS.

Conclusion

The results of this study provide new insights in combining support techniques with the aim of increasing the knowledge eight grade elementary school children have of the CVS. Guided inquiry, through use of scaffolds and support, has shown to match direct instruction in effectiveness (e.g. Hmelo-Silver, Duncan & Chinn, 2007; Lazonder & Harmsen, 2016). However, this research shows combining support techniques does not automatically lead to an increase in performance during experimentation and acquisition and use of the CVS. No positive synergistic effect was found when combining direct instruction and task segmentation. Future studies focused on combining direct instruction with task segmentation could focus on having the instruction and inquiry task situated in domains which are more closely related in terms of the complexity of conducting the experiments.

Concerning task segmentation this study supports previous findings concerning the positive regulative effect of segmenting a task. The results of this study provide new insights on the effect of CVS-training when compared to previous studies using direct instruction on CVS. This study supports previous findings concerning the effectiveness of direct instruction for increasing the knowledge of the CVS. However, this study shows that training children in mastering CVS not always leads to more unconfounded experimenting. The effectiveness of a CVS-training might be domain dependant and dependant on whether the experiment is conducted in a simulation or a physical environment. Some domains might be more fitting for teaching CVS-skills compared to others. Replicating this study in other domains could provide more insight as to whether this difference truly exists. Furthermore, whether combining a simulation and a physical environment also leads to higher gains in knowledge on CVS could be the focus of future research.

References

Azaiza, I., Bar, V., & Galili, I. (2006). Learning electricity in elementary school. *International Journal of Science and Mathematics Education*, *4*(1), 45-71.

Bjork, E. L., & Bjork, R. A. (2011). Making things hard on yourself, but in a good way: Creating desirable difficulties to enhance learning. *Psychology and the real world: Essays illustrating fundamental contributions to society*, 56-64.

Burton, H., & Sinclair, R. J. (2000). Attending to and remembering tactile stimuli: a review of brain imaging data and single-neuron responses. *Journal of Clinical Neurophysiology*, *17*(6), 575 -591.

Cepnu, S., & Keles, E. (2006) Turkish students' conceptions about the simple electric circuits. *International Journal of Science and Mathematics Education*, *4*(2), 269-291.

Chen, Z., & Klahr, D. (1999). All other things being equal: Acquisition and transfer of the control of variables strategy. *Child Development*, *70*, 1098-1120.

ELAN, [Computer software]. (2017). Max Planck Institute for Psycholinguistics, The Language Archive, Nijmegen, The Netherlands, retrieved from: ttp://tla.mpi.nl/tools/tla-tools/elan/

Furtak, E., Seidel, T., Iverson, H., & Briggs, D. (2012). Experimental and Quasi-Experimental Studies of Inquiry-Based Science Teaching: A Meta-Analysis. *Review of Educational Research*, *82*(3), 300-329.

Hmelo-Silver, C., Duncan, R., & Chinn, C. (2007). *Scaffolding and Achievement in Problem Based and Inquiry Learning: A Response to Kirschner, Sweller and Clark (2006). 42*(2), 99-107: Educational Psychologist.

Inhelder, B., & Piaget, J. (1958). *The growth of logical thinking from childhood to adolescence.* New York: Basic Books.

Jabot, M., & Henry, D. (2007). Mental models of elementary and middle school students in analyzing simple battery and bulb circuts. *School Science and Mathematics*, *107*(1), 371-381.

De Jong, T. (2006). Technological advances in inquiry learning. *Science*.

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

Kirschner, P., Sweller, J., & Clark, R. (2006). *Why Minimal Guidance During Instruction Does Not Work: An Analysis of the Failure of Constructivist, Discovery, Problem-Based, Experiential, And Inquiry-Based Teaching. 41*(2), 75-86: Educational Psychologist.

Klahr, D., & Nigam, M. (2004). The Equivalence of Learning Paths in Early Science Instruction. *Psychological Science*, *15*(10), 661-667.

Kuhn, D., & Dean, D. (2005). Is Developing Scientific Thinking All About Learning to Control Variables? *Psychological Science*, *16*(11), 866-870.

Lausberg, H., & Sloetjes, H. (2009). Coding gestural behavior with the NEUROGES-ELAN system. *Behavior research methods*, *41*(3), 841-849.

Lazonder, A., & Egberink, A. (2014). Children's acquisition and use of the control-of-variables strategy: effects of explicit and implicit instructional guidance. *Instructional Science*, *42*(2), 291-304.

Lazonder, A., & Harmsen, R. (2016). Meta-Analysis of Inquiry-Based learning: Effects of Guidance. *Review of Educational Research*, 1-38.

Lazonder, A., & Kamp, E. (2012). Bit by bit or all at once? Splitting up the inquiry task to promote children's scientific reasoning. *Learning and Instruction*, *22*, 458-464.

Lazonder, A., & Wiskerke-Drost, S. (2015). Advanced Scientific Reasoning in Upper Elementary Classrooms: Direct Instruction Versus Task Structuring. *Journal of Science Education and Technology*, *24*(1), 69-77.

Lin, T. C., Hsu, Y. S., Lin, S. S., Changlai, M. L., Yang, K. Y., & Lai, T. L. (2012). A Review of Empirical Evidence on Scaffolding for Science Education. *International Journal of Science and Mathematics Education*, *10*, 437-455.

Mayer, R. (2004). Should there be a three-strike rule against pure discovery learning? The case for guided methods of instruction. . *American Psychologist*, *59*, 14-19.

Ministerie van Onderwijs. (2013). *Nationaal Techniekpact 2020.* retrieved from: https://www.rijksoverheid.nl/binaries/rijksoverheid/documenten/convenanten/2013/05/13/nationaal-techniekpact-2020/techniekpact-2020.pdf

Minner, D., Levy, A., & Century, J. (2009). Inquiry-Based Science Instruction - What Is It and Does It Matter? Results from a Research Synthesis Years 1984 to 2002. *Journal of Research in Science Teaching*, *47*(4), 474-496.

Osborne, R. (1983). Towards modifying children's ideas about electric current. *Research in Science & Technological Education*, *1*(1), 73-82.

Pea, R. (2004). The Soclai and Technological Dimensions of Scaffolding and Related Theoretical Concepts for Learning, Education, and Human Activity. *Journal of the Learning Sciences*, *13*(3), 423-451.

Quintana, C., Reiser, B., Davis, J., Krajcik, E., Fretz, S., & Duncan, R. (2004). A scaffolding design framework for software to support science inquiry. *The Journal of the Learning Sciences*, *13*(3), 235-244.

Reiser, B. (2004). Scaffolding Complex Learning: The Mechanism of Structuring and Problematizing Student Work. *The Journal of the Learning Sciences*, *13*(3), 273-304.

Rheinberg, F., Vollmeyer, R., & Engeser, S. (2003). Die Erfassung des Flow-Erlebens [The assessment of flow experience]. In J. Stiensmeier-Pelster & F. Rheinberg (Eds.), Diagnostik von Selbstkonzept, Lernmotivation und Selbstregulation [Diagnosis of motivation and self- concept] (pp. 261–279). Gottingen: Hogrefe.

Schauble, L. (1996). The development of scientific reasoning in knowledge-rich contexts. *Developmental Psychology*, *32*(1), 102.

Sweller, J. (1988). Cognitive load during problem solving: Effects on learning. *Cognitive science*, *12*(2), 257-285.

Sweller, J. (1994). Cognitive load theory, learning difficulty, and instructional design. *Learning and instruction*, *4*(4), 295-312.

Tabak, I. (2004). Synergy: A Complement to Emerging Patterns of Distributed Scaffolding. *The Journal of the Learning Sciences*, *13*(3), 305-335.

Tiberghien, A., & Delacote, G. (1976). Manipulation of the presentation of electric circuits among young children, aged 7-12 years. *Revue Françoise de Pedagogy*, *34*, 32-44.

Zacharia, Z. C. (2007). Comparing and combining real and virtual experimentation: an effort to enhance students' conceptual understanding of electric circuits. *Journal of Computer Assisted Learning*, *23*(2), 120-132.

Zacharia, Z. C., & de Jong, T. (2014). The effects on students' conceptual understanding of electric circuits of introducing virtual manipulatives within a physical manipulatives-oriented curriculum. *Cognition and instruction*, *32*(2), 101-158.

Zydney, J. M. (2010). The effect of multiple scaffolding tools on students' understanding, consideration of different perspectives, and misconceptions of a complex problem. *Computers & Education*, *54*(2), 360-370.