Supporting Gifted Students in Inquiry-Based Learning Processes

Geertje M. Verduijn-Meijer

University of Twente

Author Note

Geertje M. Verduijn-Meijer, Student number 1127268, University of Twente; Dr. T. H. S. Eysink, 1st Supervisor, Department of Instructional Technology, University of Twente; J. Ter Vrugte, MSc, 2nd Supervisor, Department of Instructional Technology, University of Twente

A master thesis submitted to the Department of Instructional Technology of the Faculty of Behavioural Sciences, in partial fulfilment of the requirements for the degree of Master of Educational Science and Technology, Educational Design & Effectiveness

Correspondence concerning this master thesis should be addressed to Geertje M. Verduijn-Meijer, Student number 1127268, University of Twente, Postbus 217, 7500 AE, Enschede. E-mail: g.m.meijer@student.utwente.nl

15th of April 2016

Abstract

This two-part study aimed at researching the inquiry-based learning processes of gifted elementary students, and discovering which learning processes need support. Study 1 examined which learning processes gifted students spontaneously show when they are in an inquiry-based learning setting, and which learning processes might be improved by instructive support. Fourteen gifted elementary students worked on a guided learning task in an inquiry-based learning setting while thinking aloud. The think-aloud protocols were coded and analysed, based on a learning processes scheme and a domain knowledge scheme. Based on the analysed think-aloud protocols, the Inquiry Twister, an overview with inquiry steps, was designed. The Inquiry Twister is used in Study 2 to support seventeen gifted elementary students in the same inquiry-based learning task used in Study 1. Like in Study 1, think-aloud protocols were collected in Study 2, while gifted students were working on the supported inquiry-based learning task. The coded and analysed think-aloud protocols were used to evaluate whether the Inquiry Twister supported the students. The results of Study 1 indicate that the gifted students spontaneously exhibited mainly transformative learning processes in an inquiry-based learning setting, and barely showed regulative learning processes. In Study 2, in which their learning processes were externally regulated, the gifted students showed significantly more retrieving of prior knowledge, long-term planning, and reflection on knowledge. However, the scaffold did not increase the students' domain knowledge. Future research should reveal whether an improved Inquiry Twister, combined with training in when and why to use this scaffold, increases students' domain knowledge.

Keywords: inquiry-based learning; giftedness; learning processes; instructional support; scaffolds

Table of Contents

Supporting Gifted Students in Inquiry-Based Learning Processes 5
1.1 Learning Characteristics of Gifted Students 5
1.2 Inquiry-based Learning as an Instructional Strategy for Gifted Students 5
1.3 Learning Processes in Inquiry-based Learning 6
1.4 Transformative Learning Processes 6
1.4.1 Formation of a Research Question and a Hypothesis 6
1.4.2 Experienced Difficulties in Formation of a Research Question and a Hypothesis 6
1.4.3 Expected Formation of a Research Question and a Hypothesis by Gifted Students 7
1.4.4 Design and Performance of an Experiment 7
1.4.5 Experienced Difficulties in Designing and Performing an Experiment 7
1.4.6 Expected Design and Performance of an Experiment by Gifted Students 8
1.4.7 Data Analysis and Inferences 8
1.4.8 Experienced Difficulties in Data Analysis and Inferences 9
1.4.9 Expected Data Analysis and Inferences by Gifted Students 9
1.5 Regulative Learning Processes 10
1.5.1 Experienced Difficulties in Regulative Processes 11
1.5.2 Expected Regulative Processes of Gifted Students 11
1.6 Scaffolds to Support the Learning Processes 11
Study 1: Method 12
2.1 Participants 12
2.2 Domain 12
2.3 Materials 13
2.3.1 Think-aloud Method Instruction 13
2.3.2 Inquiry-based Learning Task 13
2.3.3 Inquiry-based Learning Task Instruction 13
2.3.4 Post-test 13
2.3.5 Interview 13
2.4 Procedure 14
2.5 Coding and Data Analysis 14
2.5.1 Time spent on the Inquiry-based Learning Task 14
2.5.2 Think-aloud Protocols 14
2.5.3 Behavioural Strategies 14
2.5.4 Design and Performance of Experiments 14
2.5.5 Data Analysis and Interpretation by Students 15
2.5.6 Students’ Domain Knowledge 16
2.5.7 Students’ Views on Supporting Inquiry-based Learning 16
2.5.8 Inter-rater Reliability 16
Study 1: Results 16
3.1 Time spent on the Inquiry-based Learning Task 16

3.2 Learning Processes 17
3.3 Design and Performance of Experiments 18
3.4 Data Analysis and Interpretation by Students 19
3.5 Students' Domain Knowledge 19
3.6 Students' Views on Supporting Inquiry-based Learning 19
Study 1: Conclusion 19
4.1 Findings 19
4.2 Implications 20
Study 2: Method 20
5.1 Participants 20
5.2 Domain 21
5.3 Materials 21
5.3.1 Inquiry-based Learning Task 21
5.3.2 Inquiry Twister 21
5.3.3 Inquiry Twister Instruction 21
5.3.4 Post-test 21
5.3.5 Interview 21
5.4 Procedure 21
5.5 Coding and Data Analysis 22
5.5.1 Inter-rater Reliability 23
Study 2: Results 23
6.1 Time spent on the Inquiry-based Learning Task 23
6.2 Learning Processes 23
6.3 Design and Performance of Experiments 25
6.4 Data Analysis and Interpretation by Students 27
6.5 Students' Domain Knowledge 27
6.6 Students' Views on Supporting Inquiry-based Learning 28
6.7 Students' Use of and Views on the Inquiry Twister 28
Study 2: Conclusion 29
Discussion 29
8.1 Study 1 29
8.2 Study 2 30
8.3 Future Research 32
General Conclusion 33
References 34
Appendix A: Post-test 39
Appendix B: Learning Processes Scheme 40
Appendix C: Domain Knowledge Scheme 42

Supporting Gifted Students in Inquiry-Based Learning Processes

The current State Secretary of Education wants to give gifted students the chance to develop their abilities better by offering challenging education that meets their learning needs (Dekker, 2013). If education insufficiently meets the learning needs and characteristics of gifted students, underachievement and socio-emotional problems may occur (Doolaard & Harms, 2013; Rayneri, Gerber, & Wiley, 2003; Wellisch & Brown, 2012).

Inquiry-based learning is a promising instructional strategy, but in practice many non-gifted students experience difficulties in the inquiry-based learning processes (Eysink & de Jong, 2012; De Groof, Donche, & Van Petegem, 2013; Zion, Michalsky, & Mevarech, 2005). Little is known about the inquiry-based learning processes of gifted students. Therefore, this research explores the inquiry-based learning processes of gifted Dutch students. An exploration of the inquiry-based learning processes of gifted students is done in light of the learning processes described below, including the known difficulties commonly experienced by students.

1.1 Learning Characteristics of Gifted Students

Gifted students show great talent in one or more domains. They often have high IQ-scores (above 130), show great curiosity and a drive to analyse the world that results in explorative behaviour (Silverman, 2003; Webb, 1994; Wellisch & Brown, 2012), have a broad and intense range of interests, and show eagerness and motivation to learn (Eysink, Gersen, & Gijlers, 2015; Silverman, 2003; Wellisch & Brown, 2012). Research shows that gifted students have a preference for complexity (e.g., Shore, 2010). They love complex tasks that include some unknown aspects and have to be accomplished by sophisticated, creative problem solving strategies, preferably multiple ones. In problem solving, the creativity and inventiveness of gifted students becomes clear. They like to explore new ways of doing things and often come up with original ideas, creations, and explanations (Shavinina, 2009; Webb, 1994). Furthermore, gifted students prefer intuitive, imaginative, visual, tactile, and kinaesthetic ways of processing new information (Altun & Yazici, 2010; Oakland, Joyce, Horton, & Glutting, 2000; Pyryt, Sandals, & Begoray, 1998).

To satisfy the learning needs of gifted students, learning materials need to meet certain criteria. Bonset and Bergsma (2002) have described these criteria. Learning materials should have a high degree of difficulty and complexity and involve new, interesting, and challenging topics. These topics should involve authentic and realistic problems as well as abstract concepts and generalisations. The subjects should be offered by means of open learning tasks with a variety of information sources, which stimulate an inquisitive attitude in students. Finally, it is important that learning materials appeal to students' autonomy and promote metacognitive skills.

1.2 Inquiry-based Learning as an Instructional Strategy for Gifted Students

Based on the criteria mentioned above, inquiry-based learning could be an appropriate learning strategy for gifted students. Although little research has been done, several researchers have indicated that inquiry-based learning is likely to meet the learning needs of gifted students (e.g., Eysink et al., 2015; Shore, 2010).

In short, inquiry-based learning is a specific type of problem-based learning, namely an inductive and systematic research approach to learning. Domain knowledge and inquiry skills are acquired simultaneously in inquiry-based learning (van Graft & Kemmers, 2007). Students are challenged to acquire domain knowledge by actively investigating phenomena in a multivariable context. They have to identify the causes and effects of these phenomena by asking relevant, authentic, and researchable questions. Students perform experiments to collect data, search for data based explanations, and draw conclusions to answer their research questions. Finally, the students evaluate and reflect on their inquiry and results, and communicate the latter to others (van Graft & Kemmers, 2007). Inquiry-based learning has shown advantages over traditional education (Eysink & de Jong, 2012). Through inquiry-based learning students are supposed to acquire deep knowledge and understanding of the subject matter (Njoo & de Jong, 1993).

1.3 Learning Processes in Inquiry-based Learning

Inquiry-based learning comprises two types of learning processes: transformative and regulative (Njoo & de Jong, 1993). The transformative learning processes concern the performance of the inquiry. Based on the SDDS-model of Klahr and Dunbar (1988), three transformative learning processes can be distinguished: formation of a research question and a hypothesis, design and performance of an experiment, and data analysis and inferences. These learning processes do not have to take place linearly (De Groof et al., 2013). Regulative learning processes involve the executive control of the inquiry process and comprise processes like planning and monitoring (Njoo & de Jong, 1993). Regulative learning processes take place simultaneously with all transformative learning processes.

The two types of learning processes are further elaborated below. The known difficulties commonly experienced by students are also described. In addition, expectations of how the inquiry-based learning processes take place for gifted students are given, based on what already is known about these students.

1.4 Transformative Learning Processes

1.4.1 Formation of a research question and a hypothesis. Inquiry-based learning starts with the exposure of students to a problem, phenomenon, object, or organism (van de Keere & Vervaet, 2013; van Graft & Kemmers, 2007). If the research topic is in their zone of proximal development, the students' wonder, curiosity, and motivation is stimulated (van Graft & Kemmers, 2007). The students' exploration of the research subject will raise questions. In addition, their prior knowledge will be retrieved. In the formation of a research question and a hypothesis it is essential that students use their prior knowledge. According to De Groof et al. (2013), by using prior knowledge students link their existing knowledge to the research subject, and integrate new knowledge with existing knowledge. In this way, the inquiry-based learning process will not result in a jumble of isolated facts. Furthermore, De Groof et al. (2013) have mentioned that using prior knowledge will have a positive influence on the inquiry process. Moreover, substantive domain knowledge and inquiry skills, including the formulation of a hypothesis, are mutually influential (Klahr & Dunbar, 1988; Zimmerman, 2000).

The questions raised during the exploration of the research subject will lead to a central research question, preferably formed by the students themselves (De Groof et al., 2013; van de Keere & Vervaet, 2013; van Graft & Kemmers, 2007). After the formulation of a good research question, students are supposed to formulate a testable hypothesis (Klahr & Dunbar, 1988; van de Keere & Vervaet, 2013). In a hypothesis the emphasis is on the relation between at least two variables, and the testing of a theory by experiments. A hypothesis is not limited to one specific inquiry and could be confirmed by more inquiries. During the testing of the hypothesis by experimenting, students translate a hypothesis into observable predictions. A prediction is an expected outcome of a specific inquiry that will be verified or falsified by experiments (de Jong, 2006; Gauw, 2011; van de Keere & Vervaet, 2013).

It is crucial in this first phase of inquiry-based learning that students see the need to form a research question as the starting point of their learning process. According to De Groof et al. (2013), forming research questions is important to actively involve students in the inquiry-based learning processes. Furthermore, students need to know what constitutes an effective, and researchable question and hypothesis, and how to formulate these (De Groof et al., 2013).

1.4.2 Experienced difficulties in the formation of a research question and a hypothesis. It is common among students (at the age of ten) to not have a specific research goal in mind when conducting inquiries. They experiment and see what happens. Therefore, they rarely make informative comparisons in the analysis phase (Kuhn, 2010). It appears that students find it hard to formulate good research questions and hypotheses (De Groof et al., 2013; Zion et al., 2005). De Groof et al. (2013) indicate that students frequently formulate non-effective questions, aimed at testing all the variables at once.

Furthermore, students often do not know how to formulate a syntactically correct hypothesis (De Groof et al., 2013; de Jong & van Joolingen, 1998). For example, they do not know that a hypothesis comprises variables and the relations between these variables. So in formulating hypotheses, “students frequently fail to specify variables of interest and the relationship among these variables” (Zion et al., 2005, p. 958). This could also be due to students’ incomplete and incorrect prior knowledge because of the mutual relation between domain knowledge and inquiry skills (De Groof et al., 2013; Gauw, 2011; Land, 2000).

Besides the difficulty in formulating hypotheses, students tend only to formulate plausible hypotheses and to avoid precisely formulated hypotheses that are likely to be rejected (De Groof et al., 2013; Zimmerman, 2007; Zion et al., 2005). Furthermore, students tend to formulate hypotheses with variables assumed to be causal (De Groof et al., 2013; Zimmerman, 2007). Other problems are formulating only one hypothesis, rather than testing alternative hypotheses, or making predictions instead of hypotheses (Gauw, 2011).

1.4.3 Expected formation of a research question and a hypothesis by gifted students. Gifted students could have a great advantage compared to non-gifted students in this orientation phase because of their advanced insight and understanding of problems (Barfurth, Ritchie, Irving, & Shore, 2009; Shore, 2010; Shore & Kanevsky, 1993). In addition, they have a broader knowledge base and sophisticated and high speed information processing (Freeman, 2003; Shore & Kanevsky, 1993; Steiner & Carr, 2003). The extensive knowledge of gifted students is highly interconnected, and new information is immediately linked to their prior knowledge, due to their quick (re)organisation and categorisation of information (Shore & Kanevsky, 1993). This enables them to see relationships between objects or phenomena that they can use when specifying the variables of interest (Freeman, 2003).

Therefore, it is expected that gifted students will be able to formulate hypotheses, at least if they do not become sloppy and inaccurate in their problem orientation due to their high speed of information processing (Diezmann & Watters, 1997). According to Bishop (2000), gifted students could, like other students, have problems focusing. Another pitfall of gifted students may be that they set the bar too high for themselves, due to their preference for complexity and their perfectionism (Diezmann & Watters, 1997; Webb, 1994). They could formulate a research question that is too complex and therefore not testable.

1.4.4 Design and performance of an experiment. To test their hypotheses, students design and perform an experiment in which they search for evidence to verify or falsify a hypothesis against an alternative hypothesis. Students then translate the hypothesis into observable predictions (de Jong, 2006; Gauw, 2011; van de Keere & Vervaet, 2013). To test their hypotheses, De Groof et al. (2013) have indicated that students have to know what constitutes solid evidence, which evidence is necessary to give a substantiated answer to the research question, and which valid strategy is appropriate to collect this evidence. Students need to be aware that there are different strategies to collect evidence and that observations will serve as evidence.

Strategies to test a hypothesis comprise the manipulation and isolation of variables. There are different strategies for manipulating and isolating variables. However, the ‘Control-of-Variables’ strategy (CVS), also known as the ‘Vary-One-Thing-At-a-Time’ strategy (VOTAT) is considered to be the only valid method to draw valid, unconfounded inferences (Zimmerman, 2007). In the CVS strategy one variable is changed while the other variables are kept constant. An important factor for using valid strategies is appropriate prior knowledge (Zimmerman, 2000).

Finally, data gathering involves making observations (De Groof et al., 2013; Klahr & Dunbar, 1988; van de Keere & Vervaet, 2013). It is important that students take care to register, structure, and synthesize the data. They could do this by taking notes, drawing diagrams or schemes, composing tables, graphics, or models, et cetera. The processing of data is important for data analysis and making inferences (De Groof et al., 2013; van de Keere & Vervaet, 2013). According to Manlove, Lazonder, and de Jong (2006), data processing promotes the active integration of new knowledge with students’ prior knowledge. In addition, monitoring is stimulated.

1.4.5 Experienced difficulties in designing and performing an experiment. Students experience several difficulties in designing and performing experiments. First, it is possible that students do not know what to do when performing an experiment (De Groof et al., 2013; Quintana et al., 2004). They could lack the strategic knowledge to select inquiry activities and coordinate them, or they could be distracted by less important activities, like management activities that need to be done during the inquiry (Quintana et al., 2004). De Groof et al. (2013) also have mentioned a lack of searching strategies to gather information from the internet. Additionally, students often have difficulty evaluating the quality of gathered information.

It appears that the research goal influences the strategy selected. Students tend to focus on causal hypotheses, inferences, and factors, whether they are warranted or not (Gauw, 2011; Zimmerman, 2007). When a hypothesis comprises a positive outcome, people tend to use invalid strategies like the ‘Hold-One-Thing-At-a-Time’ strategy (van de Keere & Vervaet, 2013; Zimmerman, 2000; Zimmerman, 2007). They want to keep the presumed causal variable constant in order to maintain a positive outcome. In case of a hypothesis with a negative outcome, people are more likely to use valid strategies like the CVS strategy to detect the variable that causes the negative outcome (van de Keere & Vervaet, 2013; Zimmerman, 2000; Zimmerman, 2007). When it comes to students, this could be due to their developing epistemology and metacognitive understanding of the purposes of experimentation (Zimmerman, 2007). Various research (e.g., De Groof et al. (2013) has shown that students often see inquiry-based learning as seeking facts and aimed at expected results, so they plan their experiments to produce desired effects and reduce undesired effects. Because they often want to produce a desired outcome, they design confounded, uninformative experiments (De Groof et al., 2013; Zimmerman, 2007; Zion et al., 2005).

Besides difficulties concerning inquiry strategies, the data gathering could be unreliable because students may make imprecise and unreliable observations (Land, 2000). This may be due to students’ prior knowledge (Land, 2000; Zimmerman, 2007). Finally, it appears that students rarely register the data, and that they frequently lack the skills to process it (De Groof et al., 2013).

1.4.6 Expected design and performance of an experiment by gifted students. Research has shown that gifted students are better than non-gifted students at acquiring new problem-solving strategies (Steiner, 2006; Steiner & Carr, 2003). Furthermore, they have more declarative knowledge about strategies and therefore possess a broader repertoire of strategies (Steiner, 2006; Steiner & Carr, 2003). Strategies are more appropriately used, because gifted students “often have a better and quicker understanding of which strategies are appropriate for the situation” (Steiner, 2006, p. 64). If it turns out that the chosen strategy is not the right one, they “switch to another appropriate strategy” (Shore & Kanevsky, 1993, p. 138). Due to their flexibility and insight in the use of strategies, it is expected that gifted students will use more valid strategies (Barfurth et al., 2009; Shore & Kanevsky, 1993; Steiner & Carr, 2003). They probably will also be quicker in switching from invalid to valid strategies.

Furthermore, gifted students commonly use sophisticated higher-level strategies (Steiner, 2006; Steiner & Carr, 2003). Although this could be considered positive, one remark should be made. Steiner (2006) showed that “gifted children relied on higher level strategies even when the lower level strategies were just as effective” (p.72). This could be a pitfall for them if they do not want to use easier strategies and hold on to their preferred higher-level strategies (Diezmann & Watters, 1997).

It is expected that gifted students will not have problems with data gathering. Their observations are more reliable than those of non-gifted students, because they observe in a highly objective manner (Shavinina, 2009). Furthermore, gifted students are better at distinguishing relevant from irrelevant information (Gorodetsky & Klavir, 2003; Shore, 2010; Steiner & Carr, 2003). However, a problem may arise when many data have to be gathered. In such a case, gifted students may become sloppy and inaccurate, because they do not like routines (Webb, 1994).

1.4.7 Data analysis and inferences. After processing the gathered data, they are analysed and evaluated by the students. Data analysis consists of encoding and representing the data as independent and distinct from prior theories, to consider implications for the hypothesis and prior theory (Kuhn, 2010; Kuhn & Pearsall, 2000). Encoding allows patterns in the data to be sought to determine the extent to which they match the hypothesis (Kuhn, 2010).

Data could be congruent or discrepant with the hypothesis and prior theories. Discrepant data can result in new understanding and conceptual change in which prior theories are revised (Kuhn, 2010; van der Keere & Vervaet, 2013). For conceptual change to occur, students have to experience that existing theories are inadequate to explain a phenomenon. Furthermore, a new concept has to be clear, sensible, plausible, and immediately usable (van de Keere & Vervaet, 2013).

In adequate inferential processing, theory and data are mindfully coordinated, which means that the implications of the data for the hypothesis and theory are clear (Kuhn, 2010; Kuhn & Pearsall, 2000; Zhang, Chen, Sun, & Reid, 2004). This requires epistemological insight (Kuhn, 2010; Kuhn & Pearsall, 2000). Students have to recognise the hypothesis and data as distinct knowledge sources. They have to be able to consider the hypothesis as potentially false, and the data as the means of falsifying the hypothesis. Students have to make inferences that are justified by the data, by means of the CVS strategy and multiple observations (Kuhn, 2010; Zimmerman, 2007). This involves causal as well as non-causal inferences. Furthermore, Zimmerman (2000) has indicated that when making inferences students should consider alternative hypotheses, because "evidence may relate to competing hypotheses" (p. 118).

Finally, at the end of the inquiry-based learning process students communicate and discuss their findings and conclusions with others (De Groof et al., 2013; van de Keere & Vervaet, 2013; van Graft & Kemmers, 2007). To this end, students have to be able to argue scientifically, to report, and to present their inquiry (De Groof et al., 2013). Van de Keere & Vervaet (2013) have pointed out that communication takes place during the whole inquiry process when the students cooperate with peers. Students critically reflect on their inquiries by discussing strategies, data, and inferences. In this way students acquire an epistemological insight in science: Knowledge is built by people through peer review (De Groof et al., 2013; van de Keere & Vervaet, 2013). Another advantage is that cooperative inquiry-based learning leads to better results (Keselman, 2003; Manlove et al., 2006).

1.4.8 Experienced difficulties in data analysis and inferences. Although students understand that theories are formed by research, they may confuse theory and evidence, especially in the case of causalities (Kuhn & Pearsall, 2000; Reiser et al., 2001; van de Keere & Vervaet, 2013). This could result in false inclusion and exclusion inferences (Zimmerman, 2000). Students regularly mistakenly determine a variable as causal when it co-occurs with the desired outcome (Kuhn, Black, Keselman, & Kaplan, 2000). Furthermore, they ignore non-causal factors, which results in incorrect encoding, misinterpretation, or distortion of evidence to focus on causes (Gauw, 2011; Keselman, 2003; Zimmerman, 2007). Another example of confusing theory and evidence is students' tendency to unconsciously modify their prior theory to fit the data (van de Keere & Vervaet, 2013; Zimmerman, 2007).

A common problem mentioned by various researchers is that students show low-level strategies of data analysis and tend to ignore, reject, or misinterpret data that do not fit their prior beliefs (e.g., Zimmerman, 2007; Zion et al., 2005). Misinterpretation of data that are discrepant with prior beliefs may occur if students cannot think of alternative hypotheses (De Groof et al., 2013). This can also occur by students inadequately representing prior theories, the data, or both, which prevents the students from constructing relations between them (Kuhn, 2010). However, even if conceptual change occurs, students can simultaneously rely on prior, intuitive theories (van de Keere & Vervaet, 2013; Zimmerman, 2007).

In addition, students tend to make judgments based on inconclusive or insufficient data (Zimmerman, 2000). For example, they accept a hypothesis after one confirmative experiment (Gauw, 2011; Zimmerman, 2000). The reverse also happens: students reject hypotheses when it is not warranted to do so (De Groof et al., 2013).

Finally, it appears that students frequently vacillate in their inferences (Zimmerman, 2000). They also find it hard to distinguish between everyday argumentation (based on power and persuasion to win) and scientific argumentation (based on evidence and probability to gain insight) (De Groof et al., 2013).

1.4.9 Expected data analysis and inferences by gifted students. Gifted students could have an advantage in analysing data and making inferences because of their advanced insight and reasoning abilities (Shore, 2010; Silverman, 2003). They excel at seeing relationships between objects or phenomena (Eysink et al., 2015; Freeman, 2003). Gifted students' data interpretation might be more reliable, because they tend to interpret the world in a highly objective manner (Shavinina, 2009). Furthermore, gifted students can distinguish relevant and irrelevant information better than non-gifted students (Barfurth et al., 2009; Gorodetsky & Klavir, 2003; Steiner & Carr, 2003).

Another quality of gifted students is that they are perfectionists and thus have high expectations of themselves (Silverman, 2003; Webb, 1994). Their self-criticism may be helpful in reviewing their data analysis and findings. However, gifted students are sensitive to criticism or peer rejection (Webb, 1994). One could imagine that this could hinder them from discussing and reviewing their findings with peers.

In communication, gifted students will probably have fewer difficulties with scientific argumentation because of their advanced insight and reasoning abilities. According to Barfurth et al. (2009), they are better at explaining their strategies and evaluating their thinking processes.

1.5 Regulative Learning Processes

Regulative learning processes simultaneously take place with all transformative learning processes. Self-regulating of one's inquiry process belongs to metacognition (Shore & Kanevsky, 1993). Hattie and Timperley (2007) have defined it as an "interplay between commitment, control, and confidence. It addresses the way students monitor, direct, and regulate actions toward the learning goal. It implies autonomy, self-control, self-direction, and self-discipline" (Hattie & Timperley, 2007, p. 93). According to Zimmerman (1990), self-regulation is a cyclic process in which the students constantly monitor the effectiveness of their strategies and react to it. This ensures effective learning.

Metacognition involves metacognitive knowledge and skills (Leader, 2008; Snyder, Nietfeld, & Linnenbrink-Garcia, 2011; Veenman, 2013). Metacognitive knowledge consists of declarative knowledge (knowledge about what strategies there are), procedural knowledge (knowledge how to use a strategy), and conditional knowledge (knowledge when to use each strategy) (Leader, 2008; Snyder et al., 2011). According to Veenman (2015), the development of metacognitive skills starts at the age of approximately eight years. Metacognitive skills involve the systematic use of metacognitive, motivational, and behavioural strategies to accomplish the research goal (Zimmerman, 1990).

Three metacognitive strategies can be distinguished in inquiry-based learning: planning, monitoring, and evaluation (Manlove et al., 2006). In planning, students design an experiment to test their researchable assumptions, including selection of appropriate materials and measuring instruments (van de Keere & Vervaet, 2013). Planning comprises problem orientation (analysing the task and available resources), goal setting (goals and sub-goals), and strategic planning (Manlove et al., 2006). The retrieving of prior knowledge is also part of planning (Eysink & de Jong, 2010). By monitoring, students ensure that they are making progress towards the research goal. This involves the monitoring of comprehension and task performance, based on the goals and sub-goals (Manlove et al., 2006; Snyder et al., 2011). A useful strategy is note taking (Manlove et al., 2006). Students react to the feedback of their monitoring in several ways, ranging from changes in self-perception to changes in strategy (Zimmerman, 1990). During evaluation, students evaluate the inquiry process, outcomes, and products (De Groof et al., 2013; Hattie & Timperley, 2007; Manlove et al., 2006). They reflect on the quality of their planning, its execution, and their collaboration, and they assess the outcomes, the inferences, and their understanding (De Groof et al., 2013; Manlove et al., 2006). According to De Groof et al. (2013), reflection helps students maintain control and focus in their inquiry.

In addition to metacognitive strategies, motivational and behavioural strategies are important. Motivational strategies relate to self-efficacy, self-attribution of success and failure, and task interest (Zimmerman, 1990). Self-regulating students show extraordinary effort and persistence. De Groof et al. (2013) have indicated three conditions necessary to have good motivational learning processes. First, students must have a perception of efficacy and they must feel competent. Second, students must be process-oriented and they have to see the need for inquiry-based learning. Finally, there has to be a safe inquiry-based learning environment.

Behavioural strategies involve the selection, structuring, and creation of optimal learning environments by students (Zimmerman, 1990). Students self-instruct and self-reinforce during inquiry performance and they seek help, the right information, and the most optimal learning place.

The above elaboration of the regulative processes shows that metacognition is essential for effective inquiry-based learning (Manlove, Lazonder, & de Jong, 2007; Veenman, 2013). Moreover, stronger student metacognitive skills could compensate for weaker cognition in inquiry-based learning (van de Keere & Vervaet, 2013).

1.5.1 Experienced difficulties in regulative processes. It appears that students regularly experience difficulties in regulative processes such as planning, monitoring, and reflective thinking (De Groof et al., 2013; Greene, Moos, Azevedo, & Winters, 2008; Land, 2000). Students' incomplete domain knowledge could hinder deep evaluation and strategic use of information resources (Land, 2000). Furthermore, students often find it hard to remember their actions and they may fail to refine ineffective strategies, due to, for example, a lack of conditional knowledge (Greene et al., 2008; Land, 2000; Zion et al., 2005). This could be aggravated by the students often being unsystematic in documenting plans, designs, and data, and failing to consult such records (Gauw, 2011; Zimmerman, 2007; Zion et al., 2005). Often they are unaware of their memory limitations and the need to record results and outcomes (Zimmerman, 2007). Besides this, students do not always seek help because they do not know in which situations they should do so (Greene et al., 2008). Finally, with respect to motivational processes, students sometimes lack motivation, enthusiasm, and curiosity (De Groof et al., 2013; Greene et al., 2008).

1.5.2 Expected regulative processes of gifted students. On the one hand, it is expected that gifted students will have no problems with the regulative processes. Research has shown that gifted students display higher levels of metacognition than non-gifted students, and they show more and better insight, reflection, monitoring, and evaluation of their problem solving, metacognitive, and self-regulatory processes (Barfurth et al., 2009; Shore & Kanevsky, 1993).

On the other hand, gifted students could have difficulties in self-regulated learning processes (Freeman, 2003; van Haaren & Veenman, n.d.; Veenman, 2013), where they may suffer from either a production deficiency or an availability deficiency (Veenman, 2013). In case of a production deficiency, gifted students possess metacognitive skills but do not spontaneously use them. When gifted students suffer from an availability deficiency, they insufficiently possess metacognitive skills. Gifted students often do not have to use metacognition in regular education, because their intelligence is sufficient to accomplish tasks. Therefore, they do not develop these skills. So, metacognitive development could be impeded by an inadequate, unchallenging learning environment (Freeman, 2003; Sontag & Stoeger, 2015).

1.6 Scaffolds to Support the Learning Processes

The previously described problems experienced by students indicate that inquiry-based learning needs adequate support to be successful. According to many researchers teachers, as coaches and facilitators, play a key role by giving scaffolds to support the students in their learning processes (e.g., Velthorst, Oosterheert, & Brouwer, 2011). Scaffolds include "all devices or strategies that support students' learning" (van Merriënboer, Kirschner, & Kester, 2003, p. 5). By scaffolding inquiry-based learning, tasks become more manageable and within students' zone of proximal development (Hmelo-Silver, Duncan, & Chinn, 2007; Quintana et al., 2004). A scaffold supports students' learning of the way a task should be done and why the task should be done that way (Hmelo-Silver et al., 2007).

Depending on the student´s level, several degrees of scaffolding the inquiry-based learning processes can be distinguished. These degrees range from fully structured inquiry to unstructured inquiry (Colburn, 2000; Estes & Dettloff, 2008; Hackling, 2007). According to Eysink et al. (2015), inquiry-based learning is most effective for gifted students "when they are allowed to experiment themselves, but only when their inquiry-based learning process is structured by prompts to generate hypotheses, perform experiments, and draw conclusions from observations" (p. 10).

Much research has been done on scaffolding inquiry-based learning (Hmelo-Silver et al., 2007). Many researchers have focused on scaffolds that support a specific aspect of the inquiry-based learning processes (De Groof et al., 2013; Reid, Zhang, & Chen, 2003), although Zhang et al. (2004) have advocated an integrated approach. However, until now the designed scaffolds are mainly based on difficulties experienced by non-gifted students, and it is still unknown which specific scaffolds could support gifted students in their inquiry-based learning processes. To see which scaffolds gifted students might need, this two-part research explores how they perform an inquiry-based learning task in order to discover which inquiry-based learning processes are shown spontaneously by gifted Dutch learners, and which learning processes of gifted learners might be improved by the use of a scaffold.

In Study 1, gifted elementary students worked on a guided learning task in an inquiry-based learning setting while thinking aloud. Think-aloud protocols were coded and analysed, based on a learning processes scheme, to see which inquiry-based learning processes the students exhibited. As part of the regulative learning processes it is observed whether the students showed behavioural strategies in the form of categorisation of weights. Furthermore, it is explored whether the students experimented in a systematic way. The following aspects are taken into account: the amount of time they spent experimenting, the number of experiments they performed, the number of repeated experiments, the use of the CVS strategy, and the performance of multiple experiments to test their hypotheses. In addition, it is checked whether the students made correct observations during experimenting, whether they took notes and the type of notes they took, whether they drew correct conclusions, and whether they unjustly held on to hypotheses. Furthermore, for each student the level of domain knowledge is determined. Exploring the exhibited learning processes of gifted students and their experimental behaviour gives an impression of these students' inquiry-based learning behaviour. This impression then gives an indication of which inquiry-based learning processes might be improved by the use of a scaffold. Therefore, a teacher-independent scaffold is designed to guide other gifted students in Study 2 in their inquiry-based learning processes during the performance of the learning task. The results of Study 2 are compared to the results of Study 1 to see whether the designed scaffold indeed supported the gifted students and improved students' domain knowledge.

Study 1: Method

2.1 Participants

The participants in Study 1 were 14 gifted Dutch elementary students (3 girls, 11 boys), 10–11 years old. One of the students was almost ten years old. Of these fourteen students, eight students came from different schools. These eight students received part time education for gifted students at two different schools. The other six students received full time education for gifted students at one school. The students were selected on the basis of the following criteria:

- they were in fifth grade, at the age of 10-11;
- they scored in the highest ten percent on the CITO tests of mathematics, spelling, and comprehensive reading (A- or above); and
- they had a minimum IQ of 130.

For each student written parental consent was requested and given. The students completed the inquiry-based learning task during school time. They received a puzzle eraser for their participation. The students were tested during April and May 2014 (the second half of the second semester of grade five).

2.2 Domain

The learning task was about the Law of Moments, which is also known as the lever effect. A lever is a beam that is connected to the ground by a fulcrum (hinge or pivot). The distance between the force (mass) and the fulcrum is called the lever arm. The input force is called the effort and the output force is called the load. So, each system in which the Law of Moments plays a role comprises two lever arms, two forces (an effort and a load), and a fulcrum. There are three classes of levers, classified by the relative positions of the fulcrum, the input force, and the output force. In this learning task a balance beam was used. This is a primary lever with the fulcrum in the middle of the effort and the load. The lever will be balanced when the anticlockwise moment is equal to the clockwise moment, formulated as: (left lever arm multiplied by force) = (right lever arm multiplied by force). This is known as the Law of Moments. As a result of this law, to lift a great load a small effort on the lever requires a larger lever arm (and vice versa). This learning task triggered the students to discover the Law of Moments by balance scale problems. Balance scale problems can be about weight, balance, distance, conflict-weight, conflict-distance, and conflict-balance (Siegler, 1976).

In the Netherlands, children learn about the Law of Moments during the first three years of secondary education, i.e., grade six, seven and eight (http://ko.slo.nl/vakgebieden/00004/00001/00006/00002/). In primary school children learn the basics about leverage in terms of the transfer of energy, without learning the explicit formula (College voor Toetsen en Examens, 2014; http://tule.slo.nl/OrientatieOpJezelfEnWereld/bestand/P-L45.pdf).

2.3 Materials

2.3.1 Think-aloud method instruction. The instruction of the think-aloud method consisted of three parts: direct instruction, a video example, and an exercise using the Japanese tangram game.

2.3.2 Inquiry-based learning task. The inquiry-based learning task of van Klink, Wilhelm, and Lazonder (n.d.) was adapted and used for this research. The students were provided with three sets of weights that differed in volume (small, medium, large) and mass (50 g, 100 g, 150 g). Furthermore, the students received a balance beam, scrap paper with a fine liner for note taking, and two instruction sheets with the explanation of the learning task.

In the learning task the students had to find out how the volume, mass and, position of the weights affect balance. The nine weights could be placed at eight different positions on the balance beam: four to the left and four to the right of the fulcrum. The students had to select weights and place them on either side of the fulcrum to see how this affected the balance (tip to the right, tip to the left, or balance). Thus, the students could manipulate three independent variables: volume, mass and position. The mass and position of the weights affect the balance. Volume does not affect balance. After experimenting with the three variables, the students would know more about the relationship between volume, mass, position, and balance.

To ensure that the level was in the participants' zone of proximal development, the learning task was pre-tested with five gifted students, aged 10-11 years. As a result of this pre-test, some small improvements were made to the set up and the instruction sheets, and some parts of the research question and the rules for the learning task were explained in clearer terms.

2.3.3 Inquiry-based learning task instruction. The instruction of the inquiry-based learning task started by the researcher reading aloud the instruction sheets, while the students read along. The instruction sheets began with an example of inquiry-based learning. This example was about two boys who are curious to know why a small marble needs less time to come to a stop than a large marble when both are thrown onto a flat surface. Thereafter, the researcher showed the students the balance beam and the weights, highlighting the weights' different volumes and masses. Subsequently, the research question was read and the students were pointed to the possibility of taking notes during the inquiry. When the students were able to grasp the research question, they were pointed to the inquiry rules: a) always keep thinking aloud, and b) you may hang one weight on each side of the fulcrum. Finally, the students were instructed about the end of the inquiry-based learning task, namely the writing and explaining of their conclusions. The researcher also told the students about the post-test and the short interview after the learning task. The researcher told the students that they were not allowed to experiment or touch the balance beam and the weights anymore after finishing experimenting.

2.3.4 Post-test. To determine the students' knowledge level after experimentation, a post-test was done. This post-test consisted of nine questions which asked the students about the effects of the variables on the balance of the balance beam. The students had to answer whether each variable was important for the balance. Furthermore, the students were asked whether they had to use two masses of equal weights and positions to get the beam balanced, whether a weight weighed the same at each position, and finally where they should hang two different weights to balance the beam. For this final question, the students were asked whether they could calculate in advance the precise positions. See Appendix A for the post-test.

2.3.5 Interview. To find out students' views about supporting inquiry-based learning, a short interview was carried out in which the students were asked how they felt about the inquiry-based learning task. Besides the aspects that went well and less well, the students were asked about the kind of support they would like to get during inquiry-based learning. To determine whether the students thought they had learned something, the students were also asked to describe their prior knowledge regarding the inquiry-based learning task. The description of students' prior knowledge was done after the learning task to prevent the spontaneously shown inquiry-based learning processes from influencing by this prompt to retrieve prior knowledge.

2.4 Procedure

The students worked in one session on a learning task in an inquiry-based learning setting while thinking aloud. The students performed the learning task individually, without a time limit.

The meeting began with a short introduction in which the researcher informed the students of the aim of the research and the procedure. Thereafter, the students received instruction about the think-aloud method (see Section 4.3.1) and the inquiry-based learning task (see Section 4.3.3). After the students grasped the inquiry-based learning task and the procedure, they began to work on the learning task.

The students were videotaped while performing the learning task, so all their utterances were recorded. When they fell silent for approximately ten seconds, the researcher reminded the students to keep thinking aloud. To avoid influencing the learning process and interrupting the flow, for the most part the only interaction during the learning task were these reminders to keep thinking aloud (Boren & Ramey, 2000). If the students got stuck regarding understanding the research question, the researcher instructed them to read the research question again.

When the students thought they had experimented enough to answer the research question, they stopped the learning task. The students wrote down their findings on an empty sheet and explained them to the researcher. Hereafter, the researcher asked the students nine questions about the learning task. Finally, the researcher conducted a short interview with the students.

2.5 Coding and Data Analysis

2.5.1 Time spent on the inquiry-based learning task. For each student the amount of time spent on the learning task was measured. The time was measured from the moment a student started the task to the moment the student indicated readiness to answer the research question by asking for the conclusion sheet.

2.5.2 Think-aloud protocols. The think-aloud protocols were transcribed and coded by one coder. The protocols were segmented into utterances. To avoid interpretation and bias of the data, punctuation was not used (van Someren, Barnard, & Sandberg, 1994). Subsequently, all utterances were coded into learning processes and corresponding proceedings, according to the adapted learning processes scheme of Eysink and de Jong (2012). The learning processes coding scheme is given in Appendix B. In this scheme the two learning processes of inquiry-based learning (transformative and regulative) are distinguished. The levels of the learning processes are further elaborated and specified in proceedings, the actions belonging to a learning process. Examples of verbal utterances are given to clarify the meaning of the proceedings. One proceeding was defined as all utterances belonging to the same proceeding, until a student showed utterances that corresponded to another proceeding, or until a student performed another experiment. By dividing the learning processes into proceedings, specific parts of the learning processes could be analysed to gain more insight into the students' learning processes, and to see whether students had difficulties with one of the proceedings.

2.5.3 Behavioural strategies. To determine whether the students showed behavioural strategies, the video recordings were watched to see the frequency and percentage of students who categorised the weights by volume or mass. Moreover, the moment of categorisation was observed (during the instruction of the learning task, at the start of the learning task, halfway through the learning task, or at the end of the learning task). And it was examined whether the students kept the weights categorised during the learning task.

2.5.4 Design and performance of experiments. Several aspects regarding the design and performance of experiments by the students were analysed to see whether students were carrying out the experiments in a systematic way.

To see whether students made correct observations, all observations – all utterances belonging to the proceeding coded as 1.2.1 – were examined. If an observation was not a correct description of what was happening, the observation was labelled as incorrect. The mean frequency, standard deviation, and percentage of correct observations were determined.

Subsequently, the number of experiments students performed was measured. An experiment was defined as the action in which a student hung one or two weights on the lever and then observed the outcome of the experiment. In addition, the number of repeated experiments and the frequency of note-taking during the learning task were analysed.

Furthermore, the experiments performed by the students were analysed to see whether the students used the CVS strategy. Per student it was examined in each experiment how many and which variables were changed. Subsequently, it was determined whether an experiment belonged to a CVS series. At least two experiments were needed to belong to a CVS series: an initial experiment, with subsequent experiment(s):

- in which on one side of the lever one variable altered; or
- in which one variable is changed on both sides of the balance beam in the same manner, provided that in the previous experiment, the two weights were identical in volume (both weights the same volume), mass (both weights the same mass), or position (both weights the same position); or
- in which on one side of the lever one weight has been added or reduced; or
- in which two weights were reversed, provided that the weights had the same positions.

Because many students frequently did not say per experiment which variable they investigated, it was impossible to assess whether each student explicitly investigated each variable. Furthermore, it frequently happened that students did not hypothesise or did not draw conclusions. Regrettably, that made it impossible to take into account objectively the students' reasoning during the examination of the experiments. Therefore, only the students' experiments were examined, with the assumption that most of the students' experiments were performed purposefully.

Furthermore, it was checked whether the students performed multiple experiments to test their hypotheses. The number and percentage of the students who did this were calculated. It was also examined how many hypotheses were tested by multiple experiments. For each hypothesis it was checked how many experiments a student performed. The number of experiments performed to test a hypothesis was defined as the number of experiments a student performed before starting to investigate a new variable, hypothesis, or research question. In addition, the frequency of performing multiple experiments within the group of students who formed hypotheses was determined. The number and percentages of this group were considered the most interesting, because these data better indicated whether students test their hypotheses by multiple experiments.

The type of notes students took was also analysed. In the think-aloud protocols it was indicated when each note was made by the students. The coding of the notes was done similarly to the coding of the utterances: each note was coded according to the proceedings of the learning processes scheme (see Appendix B).

2.5.5 Data analysis and interpretation by students. It was investigated whether students interpreted data correctly. Interpretations of data included all utterances belonging to the proceeding coded as 1.3.1 (drawing conclusions). It was crucial that the conclusions were drawn after at least one experiment was performed, so that there were data. For each included utterance it was examined whether it was consistent with the Law of Moments. If a conclusion fitted the theory, the utterance was labelled as correct. Conclusions based on feeling of weights (e.g., this weight feels heavier than that one) were also considered as correct, because feelings are subjective. Furthermore, students' incorrect conclusions that were immediately self-corrected were also considered as correct. When a part of the conclusion was considered incorrect, the whole conclusion was labelled as incorrect data interpretation. So the observed conclusions in the think-aloud protocols had to be entirely correct. The mean frequency, standard deviation, and percentage of correctly interpreted data were also measured.

Furthermore, it was examined whether students unjustly held on to hypotheses (in case of discrepant data). All student hypotheses were examined, and observations were made to see whether students showed conceptual change by revising incorrect assumptions in case of discrepant data. Like the data interpretation, a hypothesis was considered correct when it was consistent with the Law of Moments. The mean frequency, standard deviation, and percentage of unjustly held assumptions were determined.

2.5.6 Students' domain knowledge. In addition to the learning processes, the average domain knowledge of the students was examined, based on a domain knowledge scheme (shown in Appendix C). Two types of domain knowledge were measured: spontaneous knowledge and knowledge shown during the post-test. This was done to see what knowledge students show spontaneously, and what students really knew when they were questioned. In this way it became clear whether there was a discrepancy between these two knowledge levels. The spontaneous knowledge was determined by the written conclusions of the learning task and the corresponding explanations of the students. The domain knowledge during the post-test was based on the answers given by the students to the researcher's nine questions.

The domain knowledge scheme consists of the three variables: volume, mass, and position. Each variable is elaborated in the conclusions. For each variable it was examined, per line, which conclusions were drawn by the students. For each conclusion the students drew, they got the corresponding points. The more a conclusion fitted the Law of Moments, the more points a student received. As can be seen in the domain knowledge scheme, some conclusions correspond to one another. The students could earn 13 points in total, so the knowledge levels could vary from zero to 13 points.

2.5.7 Students' views on supporting inquiry-based learning. In the short interview after the post-test, the researcher asked the students about their views on supporting inquiry-based learning and summarized their answers.

2.5.8 Inter-rater reliability. The inter-rater reliability was ensured by a second coder, who coded independently from the researcher at least ten percent of all data concerning the following aspects: the think-aloud protocols, the hypotheses tested by multiple experiments, the type of notes, correctness of the observations made by students, correctness of data interpretation by students, the hypotheses that were unjustly held by students, and the levels of domain knowledge. The inter-rater agreement was calculated with Cohen's Kappa. Table 1 shows the aspects that were coded by the second coder, including the number of coded entities and Cohen's Kappa.

Table 1

Inter-rater reliability

Aspect	Number of entities coded by 2nd coder	Cohen's Kappa
Think-aloud protocols	63 proceedings	.87
Multiple experiments to test a hypothesis	8 hypotheses	.79
Type of notes	31 notes	.73
Correctness of observations	23 observations	1.00
Correctness of data interpretation	18 conclusions	1.00
Hypotheses that were unjustly held by students	8 hypotheses	1.00
Domain knowledge (spontaneous knowledge level and knowledge level during post-test)	2 students	.85

Study 1: Results

For each analysis, a Shapiro-Wilk *W* test was performed to determine whether a data set was normally distributed (i.e., a *p*-value higher than 0.05). In case a data set was not normally distributed, a non-parametrical equivalent of a statistical test was used to analyse the data.

3.1 Time spent on the Inquiry-based Learning Task

There was no time limit, so all students could inquire as long as they wanted. Therefore, the inquiry time varied from 03:19 minutes to 17:31 minutes. The average time students spent on their inquiries during the learning task was 07:39 minutes (*SD* = 03:41).

3.2 Learning Processes

Table 2 shows the average number of observed learning processes of the students. As can be seen, not all students showed all learning processes. Three of the 14 students did not form any research question and three other students did not make any assumption. Six students made one or two hypotheses. Furthermore, eight students, or 57 percent, took notes. With respect to planning, the students showed only short term-planning. Although all students showed monitoring, this monitoring mainly took place at the end of the learning task when the students indicated they were ready to write down their conclusion. Regarding evaluation and reflection, one student evaluated his knowledge once. With respect to motivational processes, two students expressed task interest.

Furthermore, it was observed that six students, or 43 percent, categorised the weights. All of these six students categorised the weights during the instruction of the learning task. Three of the six students kept the weights categorised during the learning task.

Table 2

Average number of observed inquiry-based learning processes (M), and the number of students who showed the learning process (N)

	Learning process	*M*	*SD*	*N*
Transformative processes	**Formation of a research question and a hypothesis (1.1)**			
	Asking oneself content-related questions	2.50	2.44	11
	Formation hypotheses	.50	.65	6
	Making predictions	1.21	1.37	8
	Total	4.21	3.36	13
	Design and performance of an experiment (1.2)			
	Observing what is happening	8.93	3.99	14
	Making notes	2.21	2.91	8
	Total	11.14	4.14	14
	Data analysis and inferences (1.3)			
	Drawing conclusions and theorising	11.57	4.93	14
	Total	11.57	4.93	14
Regulative processes	**Planning (problem orientation, goal setting, and strategic planning of actions in the short or long term prior to the task) (2.1)**			
	Analysing task and research question(s)	.00	.00	0
	Orientating to learning environment	.00	.00	0
	Retrieving of prior knowledge	.00	.00	0
	Determining a strategy (long term)	.00	.00	0
	Directing (short term planning)	12.21	5.69	14
	Total	12.21	5.69	14
	Monitoring (monitoring ongoing learning to its alignment with an earlier plan) (2.2)			
	Comparing the extent of knowledge and comprehension to the research question(s)	2.21	1.37	14
	Comparing the performance of one's inquiry to a plan made earlier	.93	1.64	6
	Total	3.14	2.41	14
	Evaluation and reflection (after the inquiry) (2.3)			
	Reflecting on the inquiry process	.00	.00	0
	Reflecting on the learning environment, how one learned, and one's motivation	.00	.00	0
	Reflecting on one's knowledge	.07	.27	1
	Total	0.07	.27	1
	Motivational processes (2.4)			
	Showing self-efficacy and self-attribution of success and failure	.00	.00	0
	Showing task interest	.64	1.65	2
	Total	0.64	1.65	2

3.3 Design and Performance of Experiments

All students made correct observations during their inquiries. To provide insight into the design and performance of experiments by the students, Table 3 describes the average number of performed experiments, which of them were repeated experiments, and which experiments belonged to a CVS series.

Table 3

Performed experiments

	M	*SD*	Perc.
Experiments	19.86	8.08	100.00
Repeated experiments	1.43	1.60	7.19
Experiments that belonged to a CVS series	16.21	6.12	81.65

Note. Perc. = percentage of experiments.

As Table 3 shows, over 80 percent of the performed experiments belonged to a CVS series. Table 4 sorts the students by the percentages of their experiments that belonged to a CVS series. All students performed experiments that belonged to a CVS series. Except for one, at least 70 percent of each student's performed experiments belonged to a CVS series.

Table 4

Percentage of students and the percentage of their experiments that belonged to a CVS series

Percentage of experiments that belonged to a CVS series	Freq.	Perc.
< 60	0	.00
60-70	1	7.14
70-80	4	28.57
80-90	5	35.71
90-100	4	28.57

Note. Freq.= frequency of students; Perc. = percentage of students.

Furthermore, it was checked whether the students performed multiple experiments to test their hypotheses. The six students who formed hypotheses tested a hypothesis by, on average, 3.29 (*SD* = 2.43) experiments. Six out of the seven hypotheses, or 86 percent, were tested by multiple experiments. All students except one tested their hypotheses by multiple experiments.

The notes students took during the learning task were also examined. Based on the notes and the think-aloud protocols, five types of notes could be distinguished – corresponding to the learning processes – as can be seen in Table 5: hypothesis, observation, conclusion, and two forms of monitoring. Most notes students took were conclusions.

Table 5

Type of notes students made

Type of notes	*f*	Perc.
Hypothesis	2	6.45
Observation	6	19.35
Conclusion	19	61.29
Monitoring of knowledge	3	9.68
Monitoring of inquiry process	1	3.23

Note. Perc. = percentage of notes.

3.4 Data Analysis and Interpretation by Students

More than 90 percent of the conclusions they drew – on average 10.36 conclusions (*SD* = 4.33) – were correct. None of the students unjustly held on to hypotheses (in case of discrepant data).

3.5 Students' Domain Knowledge

Table 6 presents the students' average knowledge levels, specified in the variables investigated. To see whether there were differences between the students' knowledge levels– spontaneous and during the post-test – a Wilcoxon Signed Ranks Test was performed. In this test, the total spontaneous knowledge level and the knowledge level during post-test were compared. The students' knowledge level measured by the post-test was significantly higher than their spontaneous knowledge level (Z = -2.68, p = .004).

Table 6

Average domain knowledge of students

	Max.	*M*	*SD*
Spontaneous knowledge level			
Volume	1	.93	.27
Mass	1	1.00	.39
Position	11	3.64	2.53
Total	13	5.57	2.44
Knowledge level during post-test			
Volume	1	1.00	.00
Mass	1	1.00	.00
Position	11	5.64	2.02
Total	13	7.64	2.02

Note. Max.= maximum number of possible points.

3.6 Students' Views on Supporting Inquiry-based Learning

After completing the inquiry-based learning task, the students were asked to describe the kind of support they would like to have concerning inquiry-based learning. A teacher-dependent scaffold, such as giving hints about inquiry steps and providing research strategies, was mentioned by nine students (64 percent). Answering questions and giving explanation was mentioned by four students (29 percent). Teacher-independent scaffolds, such as scrap paper for note taking and an instruction sheet were mentioned by two students. One student did not know what to answer.

Study 1: Conclusion

4.1 Findings

The results indicate that the students mainly showed transformative learning processes during the inquiry-based learning task. Not all students showed these processes in the formation of research questions, hypotheses, and predictions. Eleven of the 14 students formed research questions, six students formed one or two hypotheses, and three students did not form any assumption. Furthermore, none of the students explicitly used prior knowledge towards this end. The students did show a systematic way of experimenting: All students used the CVS strategy in most of their experiments. Furthermore, all students made correct observations during their inquiries, and the students performed on average relatively few repeating experiments. However, only eight of the 14 students took notes while experimenting, and those were mainly conclusions. As expected the data analysis and interpretation by the gifted students could be considered reliable: Over 90 percent of the students interpreted all the data correctly, and none of the students unjustly held on to assumptions.

In contrast to the transformative processes, the regulative processes were minimally exhibited by the students. In terms of planning processes, the students showed only short term-planning. Almost all monitoring the students showed was at the end of their inquiry-based learning task, when they decided they were ready doing the learning task. Only one student showed any evaluation and reflection during the learning task. Motivational processes were also barely shown by the students;

only two students expressed task interest. Behavioural strategies, on the other hand, were shown by six students.

The students' level of domain knowledge during the post-test was significantly higher than their spontaneous domain knowledge. However, that both knowledge levels of the students were far from the highest knowledge level might indicate that the students could have learned more.

Finally, most students would like support in inquiry-based learning by means of hints about inquiry steps and the provision of research strategies.

4.2 Implications

To reach higher levels of domain knowledge in inquiry-based learning, gifted students might benefit from support in forming a research question and hypothesis, taking notes, and regulating their inquiry-based learning processes. This latter corresponds to the research of Eysink et al. (2015), who showed that gifted students benefit from support in regulative learning processes. Stoeger, Fleischmann, and Obergriesser (2015) also indicate that gifted students perform better after regulative support. Moreover, the students themselves indicated during the interview that they would appreciate support in the inquiry steps.

It might be that the students suffered from a production deficiency and did not spontaneously exhibit metacognitive skills (Veenman, 2013). It could be assumed that the students received ample opportunities to develop metacognitive skills, given that they received either full-time or part-time education for gifted students.

Therefore, Study 2 explores whether a teacher-independent scaffold aimed at supporting the students' regulative processes would result in higher knowledge levels. Regulative support might also trigger the students to form research questions and hypotheses.

Study 2: Method

Based on the results and conclusions of Study 1, the Inquiry Twister, a cyclic step-by-step plan, was designed to guide the students in Study 2 in regulating their inquiry-based learning processes. Think-aloud protocols were again collected while the students worked on the inquiry-based learning task, supported by the Inquiry Twister. Results were analysed to evaluate whether the scaffold supported the gifted students.

It has been hypothesised that the Inquiry Twister could be supportive for gifted students, because research shows that a model like the scientific thinking circle is an effective support for metacognition (Manlove et al., 2006; van de Keere & Vervaet, 2013). Van de Keere and Vervaet (2013) have mentioned studies in which students show significantly better inquiry skills regarding formulating hypotheses, designing experiments, processing data, and comparing data with hypotheses. Overall, students show more systematically and thoughtfully performed inquiries, and therefore longer inquiry times (van de Keere & Vervaet, 2013). Moreover, Eysink et al. (2015) have shown that gifted students also benefit from external regulation of their inquiry-based learning processes. They need guidance through the inquiry cycle, although it is crucial that the students are still actively engaged in all inquiry-based learning processes (Eysink et al., 2015).

Based on the studies mentioned by van de Keere and Vervaet (2013) it is expected that the students in Study 2 will exhibit more long-term planning, retrieving of prior knowledge, monitoring, and reflection and evaluation. Furthermore, it is expected that the students will form more research questions and hypotheses. Because it is expected that the students will perform their inquiries more thoughtfully, their inquiry times will be probably longer.

Furthermore, it is hypothesized that the students will reach higher levels of domain knowledge, since the gifted students in the study of Eysink et al. (2015) exhibit better results than their peers without metacognitive support.

5.1 Participants

The participants in this study were 17 gifted Dutch elementary students (9 girls, 8 boys), 10–11 years old. Of these seventeen students, eight students came from different schools. These eight students received part time education for gifted students at two different schools. The remaining nine students received full time education for gifted students at the same school. The students were selected based on the same criteria used in Study 1. All students were tested in June 2014, at the end of grade five.

5.2 Domain

Both the learning task and the domain (the Law of Moments) were the same as used in Study 1.

5.3 Materials

5.3.1 Inquiry-based learning task. The learning task, in which the students performed a guided inquiry, was the same as used in Study 1. The students received the same materials as the students in Study 1, as well as the Inquiry Twister.

5.3.2 Inquiry Twister. The Inquiry Twister was designed to help gifted students regulate their inquiry-based learning processes. The Inquiry Twister was based on the scientific thinking circle of Dejonckheere, van de Keere, and Tallir (van de Keere & Vervaet, 2013). Like the scientific thinking circle, the Inquiry Twister is a heuristic inquiry model. The model divided the inquiry-based learning task into manageable pieces by a cyclic step-by-step plan. The Inquiry Twister consisted of eight steps, which are displayed in Figure 1: a) in step one, the students should decide what variable they were going to investigate; b) in step two, the students should retrieve prior knowledge; c) step three should stimulate monitoring of knowledge, performance of multiple experiments to test hypotheses, and task interest; d) in step four, the students should form research questions and researchable assumptions; e) step five comprises the long term planning of experiments; f) step six involves the performance of experiments; g) in step seven, the students should evaluate their acquired knowledge; h) during step eight, the students should decide whether they could answer the research question. To ensure that the Inquiry Twister was understandable for the students, the scaffold was pre-tested with 5 gifted students (10–11 years old) who performed the inquiry-based learning task. After this pre-test no changes were made. The sessions of two of the pilot students were included in the data of Study 2, since these went well.

5.3.3 Inquiry Twister instruction. On the instruction sheet a third inquiry rule was added: “Follow the Inquiry Twister every time you inquire into the size, mass, or position.” Each step of the Inquiry Twister was briefly gone through by the researcher and the students.

5.3.4 Post-test. After the students gave their conclusions, the researcher asked them nine questions about the learning task. This questioning was meant to determine the students’ level of domain knowledge. The questions were the same as in Study 1.

5.3.5 Interview. After the post-test, the researcher briefly interviewed the students, asking the same questions as in Study 1. In addition, the researcher asked the students whether they experienced the Inquiry Twister as supportive.

5.4 Procedure

The procedure of this study was similar to the procedure of Study 1 (see section 2.4). As in Study 1, the session began by stating the aim of the research and the procedure. Thereafter, the students received instruction about the think-aloud method (see section 2.3.1) and the inquiry-based learning task (see section 2.3.3). The students were instructed in the same way as the students in Study 1. In addition, they received instruction for the Inquiry Twister (see section 5.3.3).

Next, the students began to work on the learning task, supported by the Inquiry Twister. Once the researcher briefly explained to a student that the variable mass meant the heaviness of the weight. After the learning task, the researcher gave the post-test and interviewed the students.

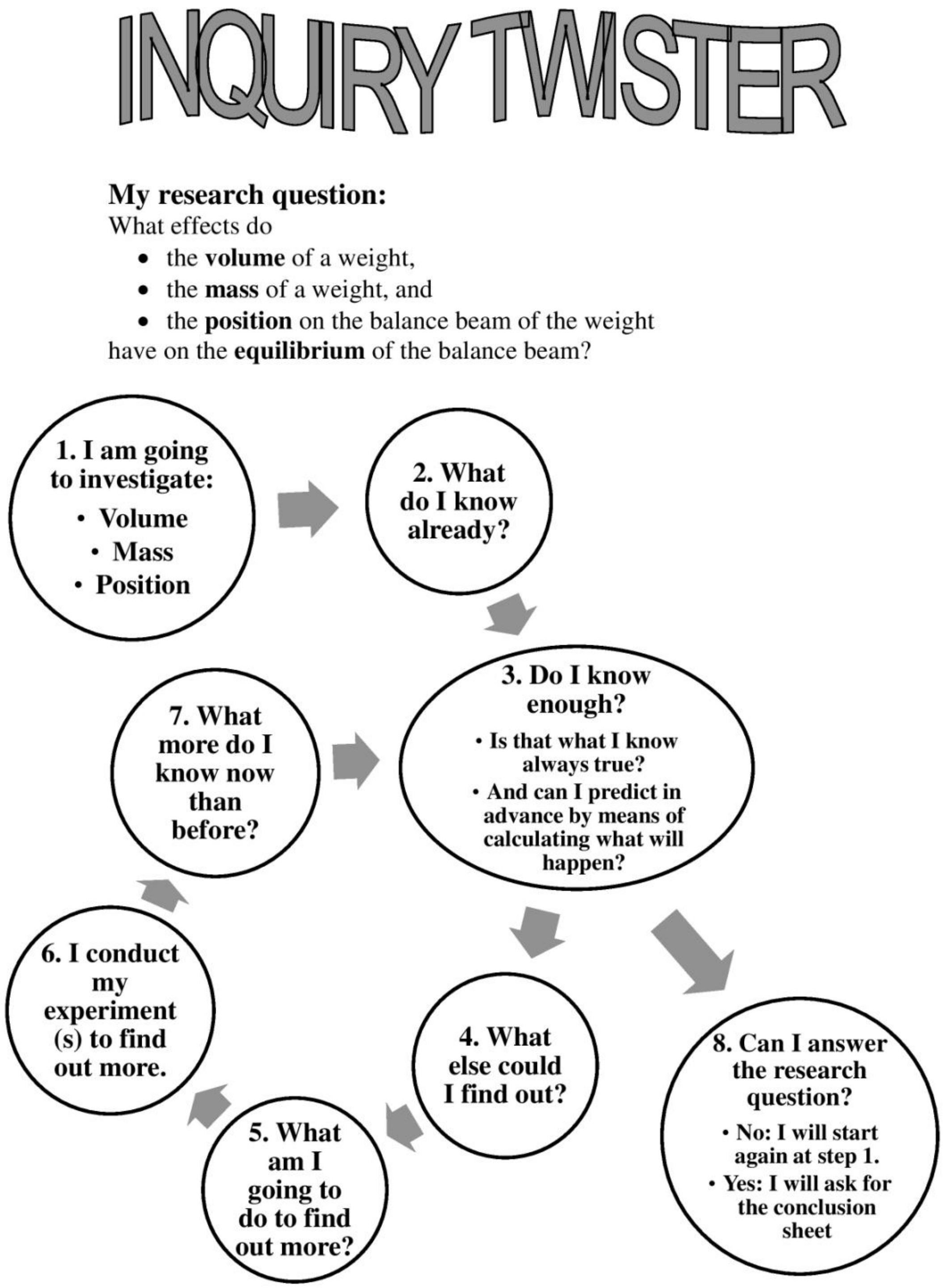

Figure 1. The Inquiry Twister.

5.5 Coding and Data Analysis

The coding and data analysis was done in the same way as in Study 1. As in Study 1, the following analyses were made: time spent on the learning task, think-aloud protocols, behavioural strategies, design and performance of experiments, data analysis and interpretation by students, average knowledge levels, and students' views. For a thorough description of the data analyses, see the method of Study 1. Furthermore, the results of this study were compared to the results of Study 1 to determine whether the Inquiry Twister supported the students.

5.5.1 Inter-rater reliability. The inter-rater reliability was ensured by a second coder, who coded independently from the researcher at least ten percent of all data concerning the following aspects: the think-aloud protocols, the type of research questions, the type of researchable assumptions, the hypotheses tested by multiple experiments, the type of notes, correctness of the observations made by students, correctness of data interpretation by students, the assumptions that were unjustly held by students, and the students' knowledge levels. The inter-rater agreement was calculated with Cohen's Kappa. Table 7 shows the aspects that were coded by the second coder, including the number of coded entities and Cohen's Kappa.

Table 7

Inter-rater reliability

Aspect	Number of entities coded by 2nd coder	Cohen's Kappa
Think-aloud protocols	86 proceedings	.85
Multiple experiments to test a hypothesis	12 hypotheses	.81
Type of notes	16 notes	.89
Correctness of observations	37 observations	1.00
Correctness of data interpretation	23 conclusions	1.00
Hypotheses that were unjustly held by students	12 hypotheses	1.00
Domain knowledge (spontaneous knowledge level and knowledge level during post-test)	2 students	.80

Study 2: Results

For each analysis, a Shapiro-Wilk *W* test was performed to determine whether a data set was normally distributed (i.e., a *p*-value higher than 0.05). In case a data set was not normally distributed, a non-parametrical equivalent of a statistical test was used to analyse the data.

6.1 Time spent on the Inquiry-based Learning Task

The average time students in Study 2 spent on their inquiries during the learning task was 07:39 minutes (*SD* = 05:04), the same average time spent as the students in Study 1 (*SD* = 03:41). The students' inquiry time in Study 2 varied from 02:18 minutes to 22:47 minutes.

6.2 Learning Processes

Table 8 presents, from Study 2 as well as Study 1, the average number of identified instances of engagement in the two processes of inquiry-based learning (transformative and regulative), which are further divided by learning processes and proceedings.

To see whether there were differences in the total number of observed learning processes between the two studies, Mann-Whitney Tests were done. In these tests, the type of study was the independent variable and the total number of utterances coded as falling within the learning processes the dependent variable. The data suggested that the students in Study 2 engaged significantly more in learning processes 1.2: design and performance of experiments (U = 63.00, Z = - 2.23, p = .026) and 2.3: evaluation and reflection (U = 48.00, Z = -3.24, p = .004).

To take a deeper look into the differences between the two studies concerning the learning processes, the number of observed proceedings was compared. Because not all data sets of the proceedings were normally distributed, Mann-Whitney Tests were done. The data indicated significant differences – in favour of the students in Study 2 – with respect to retrieving prior knowledge (U = 35.00, Z = -3,82, p = .000), determining a long-term strategy (U = 49.00, Z = -3.37, p = .005), and reflecting on knowledge (U = 48.00, Z = - 3.24, p = .004). This means, that the students in Study 2 more frequently retrieved prior knowledge, more often determined a long-term strategy, and showed more reflection on knowledge.

Table 8

Average number of observed inquiry-based learning processes (M), and the number of students who showed the learning process (N)

	Learning process	Study 1 *M*	Study 1 *SD*	Study 1 *N*	Study 2 *M*	Study 2 *SD*	Study 2 *N*
Transformative processes	**Formation of a research question and a hypothesis (1.1)**						
	Asking oneself content-related questions	2.50	2.44	11	4.06	4.13	15
	Forming hypotheses	.50	.65	6	.71	1.05	7
	Making predictions	1.21	1.37	8	1.71	1.83	11
	Total	4.21	3.36	13	6.47	5.32	15
	Design and performance of an experiment (1.2)						
	Observing what is happening	8.93	3.99	14	8.35	8.85	17
	Taking notes	2.21	2.91	8	.94	1.56	5
	Total	11.14	4.14	14	9.29	9.56	17
	Data analysis and inferences (1.3)						
	Drawing conclusions and theorising	11.57	4.93	14	11.18	7.89	17
	Total	11.57	4.93	14	11.18	7.89	17
Regulative processes	**Planning (problem orientation, goal setting, and strategic planning of actions in the short or long term prior to the task) (2.1)**						
	Analysing task and research question(s)	.00	.00	0	.24	.56	3
	Orientating to learning environment	.00	.00	0	.06	.24	1
	Retrieving of prior knowledge	.00	.00	0	1.35	1.37	12
	Determining a strategy (long term)	.00	.00	0	.82	.81	10
	Directing (short term-planning)	12.21	5.69	14	13.71	11.23	17
	Total	12.21	5.69	14	16.18	12.62	17
	Monitoring (monitoring ongoing learning compared to a plan made earlier) (2.2)						
	Comparing the extent of knowledge and comprehension to the research question(s)	2.21	1.37	14	3.12	1.83	17
	Comparing the performance of one's inquiry to a plan made earlier	.93	1.64	6	.94	1.35	8
	Total	3.14	2.41	14	4.06	2.79	17
	Evaluation and reflection (after the inquiry) (2.3)						
	Reflecting on the inquiry process	.00	.00	0	.00	.00	0
	Reflecting on the learning environment, how one learned, and one's motivation	.00	.00	0	.00	.00	0
	Reflecting on one's knowledge	.07	.27	1	.94	.83	11
	Total	0.07	.27	1	.94	.83	11
	Motivational processes (2.4)						
	Showing self-efficacy and self-attribution of success and failure	.00	.00	0	.06	.24	1
	Showing task interest	.64	1.65	2	.35	1.46	1
	Total	0.64	1.65	2	.41	1.46	2

As part of the regulative processes, it was observed that eight of the 17 students, or 47 percent, in Study 2 categorised the weights. As mentioned before, in Study 1 six of the 14 students, or 43 percent, categorised the weights. To see whether there were differences between Study 1 and Study 2 concerning the number of students who categorised the weights, a Pearson Chi Square Test was performed. No significant differences were found between the students in Study 1 and Study 2 in to the number of students who categorised the weights (X^2 = .06, p = 1.000). Approximately 45 percent of the students in both studies categorised the weights, mostly during the instruction of the inquiry-based learning task. Other moments were half way through the learning task and at the end of the learning task. To see whether there were differences between Study 1 and Study 2 concerning the moment of categorisation, a Fisher-Freeman-Halton Test was performed. No significant differences were found between the students in Study 1 and Study 2 concerning the moment of categorising the weights (Fisher's Exact Test 4.67, p = 1.000).

Four students in Study 2 and three students in Study 1 kept the weights categorised. Again a Fisher-Freeman-Halton Test was performed to see whether there were differences between the two studies concerning the number of students who kept the weights categorised during the learning task. No significant differences were found between the students in Study 1 and Study 2 with regard to the number of students who kept the weights categorised during the learning task (Fisher's Exact .36, p = 1.000).

6.3 Design and Performance of Experiments

Like the students in Study 1, the students in Study 2 made correct observations during their inquiries. Table 9 shows an overview of the design and performance of the experiments. The Table first shows the number of experiments students performed in Study 1 and Study 2. To see whether there were differences in the number of experiments performed, a Mann-Whitney Test was performed, with the type of study as the independent variable, and the number of experiments as the dependent variable. No significant differences were found in the number of performed experiments between Study 1 and Study 2 (U = 95.50, Z = -.93, p = .356).

Table 9

Performed experiments

	Study 1			Study 2		
	M	*SD*	Perc.	*M*	*SD*	Perc.
Number of experiments	19.86	8.08	100.00	18.18	14.87	100.00
Experiments that were repeated experiments	1.43	1.60	7.19	1.94	3.19	10.68
Experiments that belonged to a CVS series	16.21	6.12	81.65	14.35	13.03	78.96

Note. Perc. = percentage of performed experiments.

To determine whether there were differences concerning the frequency and exact percentage of repeated experiments between Study 1 and Study 2, Mann-Whitney Tests were performed, with the type of study as the independent variable and the frequency or percentage of repeated experiments as the dependent variable. No significant differences were found in the frequency or the percentage of repeating experiments (Frequency: U = 106.50, Z = -.53, p = .625; Percentage: U = 103.50, Z = -.65, p = .544).

Table 9 also shows the number of performed experiments that belonged to a CVS series. To examine whether there were differences concerning the frequency of experiments belonging to a CVS series, a Mann-Whitney Test was performed, with the type of study as the independent variable and the frequency of experiments belonging to a CVS series as the dependent variable. No significant differences were found. This implies that the number of performed experiments belonging to a CVS series were more or less equal in Study 1 and Study 2 (U = 84.50, Z = -1.37, p = .173).

To gain further insight regarding the percentages of experiments that belonged to a CVS series, Table 10 sorts the students by the percentages of their experiments that belonged to a CVS series. To see whether the exact percentages of the experiments that belonged to a CVS series differed between Study 1 and Study 2, a Mann-Whitney Test was performed. In this test the type of study was the independent variable, and the percentage of experiments belonging to a CVS series the dependent variable. No significant differences were found in the percentages of performed experiments that belonged to a CVS series (*U* = 83.50, *Z* = -1.41, *p* = .161).

Table 10

Number of students sorted by the percentage of their experiments that belonged to a CVS series

Percentage of experiments that belonged to a CVS series	Study 1		Study 2	
	Stud.	Perc.	Stud.	Perc.
< 60	0	.00	5	29.41
60-70	1	7.14	1	5.88
70-80	4	28.57	3	17.65
80-90	5	35.71	4	23.53
90-100	4	28.57	4	23.53

Note. Stud. = number of students; Perc. = percentage of participating students.

Furthermore, it was examined whether the students in Study 2 performed multiple experiments to test their hypotheses. The seven students who formed hypotheses tested a hypothesis by, on average, 3.75 (*SD* = 3.31) experiments. All but one of the seven students performed multiple experiments to test their hypotheses. Nine of the 12 hypotheses, or 75 percent, were tested by multiple experiments. This percentage is less than the 86 percent of the seven hypotheses tested by multiple experiments in Study 1 by five of the six students who formed hypotheses.

To determine whether there were differences between Study 1 and Study 2 in the amount of hypotheses tested by multiple experiments, a Mann-Whitney Test was performed, with the type of study as the independent variable and the amount of hypotheses as the dependent variable. No significant differences were found (*U* = 117.00, *Z* = -.09, *p* = .953). The number of students who performed multiple experiments to test a hypothesis in Study 1 and Study 2 were compared by a Fisher's Exact Test. No significant differences were found (Fisher's Exact *p* = 1.000). Finally, the number of experiments students performed to test their hypotheses were compared. This was examined by a Mann Whitney Test, with the type of study as the independent variable and the number of multiple experiments as the dependent variable. No differences were found between the two studies (*U* = 113.00, *Z* = -.27, *p* = .830). Because no differences were found, one could conclude that the students in Study 1 and Study 2 showed the same pattern of testing hypotheses by multiple experiments.

The type of notes the students in Study 2 took were also analysed. Based on the think-aloud protocols of Study 2, two types of notes were added to the four types of notes distinguished in Study 1: prior knowledge and directing. Table 11 shows an overview of the type of notes taken by students in Study 1 and Study 2.

To see whether there were differences concerning the percentages of the type of notes between the two studies, Mann-Whitney Tests were performed, with the type of study as the independent variable and the type of notes as the dependent variable. No significant differences were found (Hypothesis: *U* = 110.50, *Z* = -1.10, *p* = .739; Observation: *U* = 102.00, *Z* = -1.16, *p* = .518; Conclusion: *U* = 80.00, *Z* = -1.77, *p* = .128; Prior knowledge: *U* = 112.00, *Z* = -.91, *p* = .799; Directing: *U* = 105.00, *Z* = -1.31, *p* = .597; Monitoring (2.2.1): *U* = 117.00, *Z* = -.19, *p* = .953; Monitoring (2.2.2): *U* = 110.50, *Z* = -1.10, *p* = .739). One could conclude that the students in Study 1 and Study 2 more or less took the same type of notes. As can be seen in Table 11, the students in both studies mainly wrote down conclusions.

Table 11

Type of notes students took

Type of notes	Study 1		Study 2	
	f	Perc.	*f*	Perc.
Hypothesis	2	6.45	0	.00
Observation	6	19.35	2	12.50

Table 11 Continued

Type of notes students took

Type of notes	Study 1		Study 2	
	f	Perc.	*f*	Perc.
Conclusion	19	61.29	10	62.50
Prior knowledge	0	.00	1	6.25
Directing	0	.00	2	12.50
Monitoring (2.2.1)	3	9.68	1	6.25
Monitoring (2.2.2)	1	3.23	0	.00

Note. Perc. = percentage of notes.

6.4 Data Analysis and Interpretation by Students

Nearly 89 percent of the conclusions in Study 2 were correct, or, on average, 9.82 (*SD* = 6.89) correct conclusions per student. This was somewhat less than the 91 percent of the conclusions that were correct in Study 1 (*M* = 10.36, *SD* = 4.33). To determine whether there were differences between Study 1 and Study 2 in the frequency and percentage of correctly interpreted data, an Independent-Samples T Test and a Mann-Whitney Test were performed. In the latter test, the type of study was the independent variable, and the frequency or percentage of correctly interpreted data the dependent variable. The data suggest no significant differences between the students in Study 1 and Study 2 in the amount of correctly interpreted data (t = .25, p = .80). The same applies to the percentage of correctly interpreted data in Study 1 and Study 2 (U = 96.50, Z = -.93, p = .38). So, these findings indicated that the students in both studies showed the same frequency and percentage of correctly interpreted data.

Furthermore, like the students in Study 1, none of the students in Study 2 held on to hypotheses despite discrepant data. One student did try to find an effect of the variable volume, but after further experimenting he concluded that there was no effect. Another student had an assumption before she began with the inquiry-based learning task. This student thought that it was impossible for her to discover the formula of the Law of Moments, because she did not know from which material the lever was made. Therefore, during her inquiries she did not try to discover the formula, and she gave the same reasoning during questioning after the learning task.

6.5 Students' Domain Knowledge

Table 12 presents the average level of domain knowledge of the students, specified in the variables investigated.

Table 12

Average domain knowledge of students

		Study 1		Study 2	
	Max.	*M*	*SD*	*M*	*SD*
Spontaneous knowledge level					
Volume	1	.93	.27	1.00	.00
Mass	1	1.00	.39	1.00	.35
Position	11	3.64	2.53	3.41	2.58
Total	13	5.57	2.44	5.41	2.58
Knowledge level during post-test					
Volume	1	1.00	.00	.94	.24
Mass	1	1.00	.00	1.00	.00
Position	11	5.64	2.02	5.59	1.97
Total	13	7.64	2.02	7.53	1.91

Note. Max.= maximum number of possible points.

To see whether there were differences between the average levels of domain knowledge among the students in Study 2 (spontaneous and during questioning), a Wilcoxon Signed Ranks Test was performed. In this test, the spontaneous knowledge level and the knowledge level during questioning were compared. Like the students in Study 1, the students in Study 2 showed a significantly higher knowledge level during questioning (Z = -3.16, p = .001).

To determine whether there were differences between the two studies regarding the average knowledge levels – spontaneous and during post-test – Mann-Whitney Tests were done. In these tests the study was the independent variable and the knowledge level the dependent variable. The data indicated no difference in the total spontaneous knowledge level (U = 106.50, Z = -.52, p = .625) or the total knowledge level during questioning (U = 113.00, Z = -.25, p = .830). So, the data suggested that the knowledge levels of both Study 1 and Study 2 were equal.

6.6 Students' Views on Supporting Inquiry-based Learning

After the learning task was done, like in Study 1 the students in Study 2 were asked to describe the kind of support they would like to have concerning inquiry-based learning. A teacher dependent scaffold like giving explanations was mentioned by seven students (41 percent). Giving hints was mentioned by five students (29 percent). Answering questions was mentioned by four students (24 percent). Providing research strategies was mentioned by two students (12 percent). Monitoring of students' inquiry by a teacher was mentioned by one student (7 percent). Teacher independent scaffolds like a step-by-step plan and cooperation with peers were each mentioned by one student. Finally, one of the students did not know what to answer. Unfortunately, the videotaped account of another student was unusable.

6.7 Students' Use of and Views on the Inquiry Twister

Ten students (59 percent) explicitly used the Inquiry Twister during the inquiry-based learning task by following all its steps. The remaining seven students barely seemed to use the scaffold, apart from glancing at it a few times.

After the learning task was complete, all the students in Study 2 were asked how they experienced performing the learning task while supported by the Inquiry Twister. Eight students (47 percent) found the inquiry steps of the Inquiry Twister useful, of which six students explicitly used the Inquiry Twister. According to one of the eight students, the scaffold could be useful to maintain an overview during the inquiry. Another student said the Inquiry Twister could be useful to have a more critical stance towards oneself. In this way, one could learn and discover more. And another student even thought he could not perform this learning task without the Twister.

However, several students were less positive about the Inquiry Twister. Five students (29 percent) would prefer not to use a scaffold like the Inquiry Twister, because they found it somewhat restrictive, as using it would not allow them to follow their own steps. Another reason was that they found the use of a scaffold like the Inquiry Twister unnecessary, as they already knew which steps to take. One student said she would read the Inquiry Twister in advance to get the inquiry steps into her mind, and subsequently not make any use of the scaffold again. Three of these five students explicitly used the Inquiry Twister.

Of all students, eleven students (65 percent) indicated that this scaffold did not have an added value to them in this inquiry-based learning task, although three of these students used the Inquiry Twister. But according to three others of these eleven students, the Inquiry Twister could be useful when performing another inquiry-based learning task. The view of one student was unknown.

So, eight students (47 percent) were positive about the Inquiry Twister, three students (18 percent) were moderately positive, and five students (29 percent) had a negative stance towards the use of the Inquiry Twister. Eleven students (65 percent) did not see the added value of using the Inquiry Twister during the inquiry-based learning task used in this research.

Study 2: Conclusion

The results of this study indicate that like the students in Study 1, the students in Study 2 showed mainly transformative learning processes. But compared to the students in Study 1, the students in Study 2 exhibited more regulative learning processes, especially with respect to planning and evaluation and reflection. With respect to planning, the students in Study 2 showed significantly more retrieving of prior knowledge and determining a long-term strategy. Furthermore, the students in Study 2 showed more evaluation and reflection during the inquiry-based learning task. They exhibited significantly more reflection on knowledge. However, despite the fact that the students showed more regulative processes, their average knowledge levels were not higher. Furthermore, most students had a positive or moderate positive stance view about a scaffold like the Inquiry Twister to support inquiry-based learning. However, at the same time most of the students did not see the added value of using the Inquiry Twister during this inquiry-based learning task.

Discussion

This research addressed the inquiry-based learning processes of gifted elementary students. Study 1 explored which learning processes gifted students spontaneously show when they are in an inquiry-based learning setting, and which learning processes might be improved by instructional support. Gifted Dutch elementary students worked on a guided learning task in an inquiry-based learning setting while thinking aloud. Based on the analysed think-aloud protocols, a teacher-independent instructional support was designed to guide the students in Study 2 in their inquiry-based learning processes. The think-aloud protocols of Study 2 were analysed to see the extent to which the designed scaffold was effective.

8.1 Study 1

The results of Study 1 indicate that the students mainly showed transformative learning processes. Despite the fact that not all students formed research questions and hypotheses, they showed a systematic way of experimenting. They made correct observations during their inquiries, frequently used the CVS strategy, and performed relatively few repeating experiments. However, only eight of the 14 students took notes while experimenting, which were mainly conclusions. This corresponds to Zimmerman (2000), who indicates that children’s awareness to take notes emerges between the ages of 10 and 13. The data analysis and interpretation by the gifted students could be considered reliable: over 90 percent of the students interpreted all the data correctly, and none of the students unjustly held on to assumptions. Unlike the transformative learning processes, the regulative learning processes were shown to a small extent by the students.

This could raise the question of whether the learning task was too easy, so that the gifted students did not need to use regulative processes. Both knowledge levels of the students were in accordance with the literature, in the sense that most 13-year olds and adults did not discover the formula (Siegler & Chen, 2002). Although the frequently used balance scale task of Siegler (1976) is not the same as the learning task used in this study and a comparison of the data was therefore inhibited, Siegler’s results give an indication about the level of knowledge of students. Siegler’s Rule III of his decision tree model (1976) includes a broad range of reasonings, although substantial differences can be distinguished regarding the levels of knowledge. For example, guessing could be considered a lower level of knowledge than having a faulty theory like the addition rule. Therefore, the domain knowledge scheme used in this study to determine the students’ level of domain knowledge (Appendix C) is more specific and detailed in order to distinguish knowledge levels more clearly. The fact that the students’ levels of domain knowledge were far from the highest knowledge level might indicate that the students could have learned more. Moreover, one student proved that it was possible to discover the formula (to reach the highest knowledge level). Furthermore, the learning task, adapted from the balance beam task of van Klink et al. (n.d.), was performed in that study by older students (grade eight). Therefore, the learning task cannot be considered too easy.

Another reason for the minimally exhibited regulative processes could be that the students had an availability deficiency, meaning they did not possess regulative skills and could therefore not exhibit them (van Haaren & Veenman, n.d., Veenman, personal contact, 2015; Veenman, Kok, & Blöte, 2005). However, in the study of van Haaren and Veenman (n.d.), in which they found that most gifted students suffer from an availability deficiency, the participating students were much older (grade 11) than the participating students in this study (grade five).

A more likely explanation for the minimally exhibited regulative processes is that the students suffered from a production deficiency and did not spontaneously exhibit developed metacognitive skills. It could be assumed that the students received ample opportunities to develop metacognitive skills, given that they received either full time or part time education for gifted students.

This finding of little spontaneous self-regulation by gifted students seems to contradict findings that their metacognition is outstanding (Greene et al. 2008; Snyder et al., 2011; Zimmerman & Martinez-Pons, 1990). However, the studies of Snyder et al. (2011) and Zimmerman and Martinez-Pons (1990) were on students' self-perceptions regarding self-regulated learning. According to Veenman, Van Hout-Wolters, and Afflerbach (2006), online methods to measure metacognition, such as think-aloud protocols, are more accurate and valid than off-line methods, such as post-tests and interviews. Apart from this, the students in the study of Snyder et al. (2011) were much older, as were most participating students in the research of Zimmerman and Martinez-Pons (1990). These two factors (method and age) might have influenced the findings. In the study of Greene et al. (2008), the context of searching for information in a hypermedia environment was likely to be more familiar to the participating students than the inquiry-based learning environment of this study. This also might have influenced the results and could explain the difference in findings.

The exhibited learning processes of the gifted students in Study 1 indicate that gifted students might benefit from support in forming a research question and hypothesis, taking notes, and regulating their inquiry-based learning processes. Therefore, instructional support – the so called Inquiry Twister – was designed to guide the students in Study 2 in regulating their inquiry-based learning processes. Positive findings were expected, as indicated by previous research (Eysink et al., 2015; Manlove et al., 2006; van de Keere & Vervaet, 2013; Stoeger et al., 2015). It was hypothesized that students in Study 2 would exhibit more long-term planning, retrieving of prior knowledge, monitoring, and reflection and evaluation. Furthermore, it was expected that the students would form more research questions and hypotheses. Because it was expected that the students would perform their inquiries more thoughtfully, their inquiry times were expected to be longer. Furthermore, it was hypothesized that the students would reach higher knowledge levels.

8.2 Study 2

Indeed, compared to the students in Study 1, the students in Study 2 exhibited more regulative learning processes, especially with respect to planning and evaluation and reflection. With respect to planning, the students in Study 2 showed significantly more retrieving of prior knowledge and long-term planning. With respect to evaluation and reflection, they showed more reflection on knowledge.

However, it should be mentioned that like the students in Study 1, the students in Study 2 barely showed task analysis and orientation to the learning environment. This could be explained by sufficient instruction by the researcher. To better see whether gifted students spontaneously exhibit task analysis and orientation to the learning environment, in future research the instruction of the inquiry-based learning task should be minimized or even omitted. It is expected that the less instruction students receive, the more they will show orientation processes.

Another observation was that like the students in Study 1, the students in Study 2 barely showed evaluation and reflection on the inquiry process. This might be because this was a one-time inquiry-based learning task. Perhaps the students therefore did not see the added value of reflecting on and evaluating their inquiry process. A special evaluation step added to the Inquiry Twister is likely to stimulate the students to reflect on their inquiry process. An explicit evaluation step might also stimulate monitoring by the students. Although the students in Study 2 displayed more monitoring, the difference was not significant. It may be that the increased evaluation and reflection on knowledge by the students in Study 2 served as monitoring.

It was expected that the students in Study 2 would form more research questions and hypotheses, but unfortunately they did not. This might be because like the students in Study 1, they were given the research question beforehand, which may have discouraged the students from forming other research questions. Furthermore, the formulation of step four in the Inquiry Twister might be improved to stimulate the students to form research questions. The fact that students did not form more hypotheses may be due to the lack of an explicit hypothesis step in the Inquiry Twister, in contrast to the research of Eysink et al. (2015) and van de Keere and Vervaet (2013). It seems that students need an explicit prompt to form hypotheses.

This also seems to be the case for taking notes as the students in Study 2 did not take more notes. Another interesting observation was that like the students in Study 1, the students in Study 2 mainly noted conclusions. A specific prompt in the Inquiry Twister is likely to stimulate the students to write down specific observations, which eventually may lead to higher knowledge levels. Ultimately, it is expected that the difference between students' spontaneous knowledge level and their knowledge level during the post-test will be smaller. This should be investigated by future research.

Strikingly, the average inquiry time of the students in both studies were equal (07:39 minutes), although there was a greater range of the inquiry times in Study 2, as shown by the higher *SD*. It could be that for the students in Study 2 the learning task was more concrete due to the Inquiry Twister. This might have saved them the time needed to think about the steps to be taken, which the students in Study 1 had to do for themselves. This saved time could have compensated for the extra time the students in Study 2 needed for long-term planning, retrieving of prior knowledge, and evaluation of knowledge. Besides, the students in Study 2 did not show more forming of research questions and hypotheses or note taking, in contrast to the research mentioned by van de Keere and Vervaet (2013), in which the students exhibited longer inquiry times. One could say that the students in Study 2 worked more efficiently and purposefully. This is in line with the literature, which indicates that gifted students show relatively more planning during problem solving tasks (e.g., Steiner, 2006). However, an important difference with the literature is that the gifted students in this research showed more thoughtful planning after they were supported by a scaffold, instead of doing this spontaneously. Future research should shed more light on this discrepancy.

Another remarkable aspect regarding the inquiry time is the short duration of the inquiry times exhibited by the students in both studies. This seems to conflict with gifted students' curiosity and explorative behaviour, as commonly described in the literature. One would expect that curious and explorative behaviour would result in longer inquiry times. The third and fourth step of the Inquiry Twister apparently did not trigger the students in Study 2 to exhibit more explorative behaviour or to attain the highest knowledge level. As stated before, the learning task was not too easy. Perhaps the students felt inhibited by only being allowed to hang one weight on each side of the balance beam. Perhaps without this inhibition the students would have shown more explorative behaviour and longer inquiry times. Another measure that might help is to set a minimum research time for the students. Future research should investigate this.

The hypotheses were partly fulfilled. The students in Study 2 showed more regulative processes, but they did not exhibit more forming of research questions and hypotheses, or longer inquiry times.

It was expected that the students' levels of domain knowledge in Study 2 would be higher. However, they were not, although the students showed more regulative processes. According to Veenman (2013; personal contact, 2015), a reason for the students' lack of increased domain knowledge could be that the students were unfamiliar with the use of the Inquiry Twister. The Inquiry Twister required extra effort of exhibiting metacognitive skills which may have caused lower levels of domain knowledge (Veenman, personal contact, 2015). The students did not receive any training beforehand in the use of the Inquiry Twister, in contrast to the study mentioned by van de Keere and Vervaet (2013), in which the participants were trained in the use of the scientific thinking circle. Van Haaren and Veenman (n.d.) found that one-off instruction combined with metacognitive hints did not result in higher learning gains or in more metacognitive skills. In accordance with this, it is not surprising that the single instruction and the use of the Inquiry Twister did not result in higher levels of domain knowledge, at least if one assumes that the students in Study 2 had an availability deficiency (lack of developed metacognitive skills).

But the students in Study 2 did show significantly more regulative processes, which conflicts with the finding of van Haaren and Veenman (n.d.). An explanation for these different findings could be that in the research of van Haaren and Veenman (n.d.), the metacognitive instruction was about the usefulness of metacognitive skills, while the instruction of the Inquiry Twister was about how to use the Inquiry Twister. This difference in the provided knowledge regarding how to use metacognitive skills might have influenced the results and could account for the differences that were found. Furthermore, the difference in age of the participating students in the two studies (grade 11 in the study of van Haaren and Veenman [n.d.], and grade five in this study) might explain the differences that were found.

Another, more likely explanation of the different findings regarding the research of van Haaren and Veenman (n.d.) is that the students in Study 2 suffered from a production deficiency, and did not spontaneously exhibit developed metacognitive skills, rather than having an availability deficiency. The instruction and the use of the Inquiry Twister then logically resulted in an increase of regulative processes. Hence, the lack of increased knowledge levels could be explained by the extra effort the Inquiry Twister required of exhibiting metacognitive skills (Veenman, 2013).

Students' views in Study 2 are in line with the possibility of a production deficiency, since most students – although they were positive about the scaffold – did not see the added value of the Inquiry Twister during this learning task. It seemed that most of these students found the learning task too easy to use the Inquiry Twister. However, Eysink et al. (2015) found that gifted students experience more flow when they are supported by a structured inquiry worksheet, corresponding to the scientific thinking circle. They concluded that certain support "is not at the expense of positive feelings toward the task" (p. 9). So, if gifted students see and experience the advantage of using a scaffold like the Inquiry Twister, they are likely to prefer this over unsupported inquiry-based learning tasks. Therefore, a training in when and why a scaffold like the Inquiry Twister should be used is recommended.

8.3 Future research

Future research should reveal what is the most optimal scaffold gifted elementary students need. Although this research indicates that students show more regulative processes when supported by a scaffold like the Inquiry Twister, it was not effective enough to increase students' domain knowledge. To do that, the Inquiry Twister could be improved by explicitly presenting the formation of a research question and hypotheses, note taking, and evaluation and reflection on the inquiry process in separate steps.

Future research should investigate whether an improved Inquiry Twister, combined with training in when and why to use this scaffold, will lead to higher levels of domain knowledge in gifted elementary students. If this is the case, gifted students indeed suffer from a production deficiency, since cues or prompts during a learning task could be sufficient for these students, and they only need training in when and why metacognitive skills should be applied (Veenman, 2013). However, in case an improved Inquiry Twister and the additional training does not lead to higher knowledge levels, it is likely that gifted elementary students suffer from an availability deficiency (Veenman, 2013). In that case, Veenman (2013) recommends that teachers and practitioners offer students an extensive, informed, and prolonged training in self-regulation by embedded instruction in the inquiry-based learning context. This training should be carried out according to the WWW&H rule (what to do, when, why and how) of Veenman (2013). Future research will shed more light on this. It would also be desirable to use a larger sample to increase the reliability and generalisability. The sample size of this research was small ($N = 31$), so the generalisability of the findings is limited. Another limitation is that in this research students' experience with inquiry-based learning is not taken into account. Future research should investigate the influence of this factor.

When gifted students receive the appropriate training and support in the regulative learning processes, it is expected that they will achieve higher learning gains, as the research of Eysink et al. (2015) indicates. Furthermore, the students will probably experience more flow, and are "focused on the task, highly concentrated, and feel in control" (Eysink et al., 2015, p. 2). In this way, inquiry-based learning could be an appropriate learning strategy for gifted students.

General Conclusion

When learning in an inquiry-based setting without scaffolds, gifted students mainly show transformative processes. They performed the experiments, data analysis, and interpretation well. However, gifted students might benefit from support in forming a research question and hypothesis, in taking notes and processing data, and in regulating their inquiry-based learning processes. This research indicates that a scaffold like the Inquiry Twister –an overview of inquiry steps– with a one-off instruction, was effective to the extent that the gifted students show more retrieving of prior knowledge, long term-planning, and evaluation and reflection on knowledge. However, the scaffold did not increase the students' knowledge levels. Future research should reveal whether an improved Inquiry Twister combined with training in when and why to use this scaffold will increase the domain knowledge of students. This will shed more light on what constitutes adequate instructive support for gifted students during inquiry-based learning.

Acknowledgements

I am very grateful to everyone who has contributed to the realisation of this thesis. I especially want to thank my husband and my first supervisor, who have supported me by giving me eye openers and confidence.

References

Altun, F., & Yazici, H. (2010). Learning styles of the gifted students in Turkey. *Procedia Social and Behavioral Sciences*, *9*, 198–202.

Barfurth, M. A., Ritchie, K. C., Irving, J. A., & Shore, B. M. (2009). A metacognitive portrait of gifted learners. In L.V. Shavinina (Ed.), *International Handbook on Giftedness* (pp. 397-417). Netherlands, Amsterdam: Springer.

Bishop, K. (2000). The research processes of gifted students: A case study. *Gifted Child Quarterly, 44*(1), 54-64. doi: 10.1177/001698620004400106

Bonset, H. & Bergsma, S. (2002). *Hoogbegaafde leerlingen en het vak Nederlands.* Studie en onderzoek binnen het project Nederlands VO 33. Enschede: Stichting Leerplanontwikkeling.

Boren, T., & Ramey, J. (2000). Thinking aloud: Reconciling theory and practice. *IEEE Transactions on Professional Communication*, *43*(3), 261-278. doi: 10.1109/47.867942

Colburn, A. (2000). An inquiry primer. *Science Scope*, *23*(6), 42-44.

De Groof, J., Donche, V., & Van Petegem., P. (2013). *Onderzoekend leren stimuleren: effecten, maatregelen en principes*. Leuven, Belgium: Acco.

de Jong, T. (2006). Scaffolds for scientific discovery learning. In J. Elen & R.E. Clark (Eds.), *Handling complexity in learning environments* (pp. 107-128). London: Elsevier Science Publishers.

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

Dekker, S. (2013). *Kamerbrief over toptalent in het funderend onderwijs*. Retrieved from http://www.rijksoverheid.nl/ministeries/ocw/documenten-en-publicaties/kamerstukken/2013/09/02/kamerbrief-over-toptalent-in-het-funderend-onderwijs.html

Diezmann, C. M., & Watters, J. J. (1997). Bright but bored: Optimising the environment for gifted children. *Australian Journal of Early Childhood*, *22*(2), 17-21.

Doolaard, S., & Harms, T. (2013). *Omgaan met excellente leerlingen in de dagelijkse onderwijspraktijk* (Rapportage BOPO-programmalijn Onderwijskwaliteit PO, Deelproject 5 ‘Differentiatie en excellentie’, 413-09-065). Retrieved from http://irs.ub.rug.nl/dbi/5147336f30e5c

Estes, F., & Dettloff, L. (2008). Inquiring Minds: Reaching Gifted Students with Challenging Science. *Understanding Our Gifted, 21*(1), 19-23.

Eysink, T. H., & de Jong, T. (2012). Does instructional approach matter? How elaboration plays a crucial role in multimedia learning. *Journal of the Learning Sciences*, *21*(4), 583-625.

Eysink, T. H., Gersen, L., & Gijlers, H. (2015). Inquiry learning for gifted children. *High Ability Studies, 26*(1), 63-74.doi: 10.1080/13598139.2015.1038379

Freeman, J. (2003). Scientific thinking in gifted children. In P. Csermely & L. Lederman (Eds.), *Science Education: Talent Recruitment and Public Understanding*’ (pp. 17-30). Netherlands, Amsterdam: IOS Press with NATO Scientific Affairs Division.

Gauw, M. (2011). *Ontwerp, ontwikkeling en evaluatie van lesmateriaal voor onderzoekend leren in het basisonderwijs* (Master's thesis, Twente University, Netherlands).

Gorodetsky, M., & Klavir, R. (2003). What can we learn from how gifted/average pupils describe their processes of problem solving? *Learning and Instruction 13,* 305–325.

Greene, J.A., Moos, D.C., Azevedo, R., & Winters, F.I. (2008). Exploring differences between gifted and grade-level students' use of self-regulatory learning processes with hypermedia. *Computers & Education, 50,* 1069–1083.

Hackling, M.W. (2007). Inquiry and investigation in primary science. In V. Dawson & G. Venville (Eds.), *The art of teaching primary science* (pp. 127-148). Crows Nest, NSW: Allen & Unwin.

Hattie, J., & Timperley, H. (2007). The power of feedback. *Review of educational research*, *77*(1), 81-112.

Hmelo-Silver, C. E., Duncan, R. G., & Chinn, C. A. (2007). Scaffolding and achievement in problem-based and inquiry learning: A response to Kirschner, Sweller, and Clark (2006). *Educational Psychologist, 42*(2), 99-107. doi: 10.1080/00461520701263368

Keselman, A. (2003). Supporting inquiry learning by promoting normative understanding of multivariable causality. *Journal of Research in Science Teaching, 40*(9), 898-921.

Klahr, D., & Dunbar, K. (1988). Dual space search during scientific reasoning. *Cognitive Science, 12,* 1-48.

Kuhn, D. (2010). What is scientific thinking and how does it develop? In U. Goswami (Ed.), *Handbook of Childhood Cognitive Development* (pp. 371-393). Oxford, England: Blackwell.

Kuhn, D., Black, J., Keselman, A., & Kaplan, D. (2000). The development of cognitive skills thatsupport inquiry learning. *Cognition and Instruction, 18,* 495-523.

Kuhn, D., & Pearsall, S. (2000). Developmental origins of scientific thinking. *Journal of cognition and Development*, *1*(1), 113-129.

Land, S. M. (2000). Cognitive requirements for learning with open-ended learning environments. *Educational Technology Research and Development, 48*(3), 61-78.

Leader, W. S. (2008). *Metacognition among students identified as gifted or nongifted using the discover assessment* (Doctoral dissertation). Retrieved from http://arizona.openrepository.com/arizona/bitstream/10150/193777/1/azu_etd_2738_sip1_m.pdf

Manlove, S., Lazonder, A.W., & de Jong, T. (2006). Regulative support for collaborative scientific inquiry learning, *Journal of Computer Assisted Learning, 22*, 87-98.

Manlove, S., Lazonder, A. W., & de Jong, T. (2007). Software scaffolds to promote regulation during scientific inquiry learning. *Metacognition and Learning, 2*, 141-155.

Njoo, M., & de Jong, T. (1993). Exploratory learning with a computer simulation for control theory: Learning processes and instructional support. *Journal of Research in Science Teaching, 30*(8), 821-844.

Oakland, T., Joyce, D., Horton, C., & Glutting, J. (2000). Temperament-Based Learning Styles of Identified Gifted and Nongifted Students. *Gifted Child Quarterly, 44*(3), 183-189. doi: 10.1177/001698620004400304

Pyryt, M. C., Sandals, L. H., & Begoray, J. (1998). Learning style preferences of gifted, average-ability, and special needs students: A multivariate perspective. *Journal of Research in Childhood Education, 13*(1), 71-76.

Quintana, C., Reiser, B. J., Davis, E. A., Krajcik, J., Fretz, E., Duncan, R. G., . . . Soloway, E. (2004). A scaffolding design framework for software to support science inquiry. *Journal of the Learning Sciences, 13,* 337–386.

Rayneri, L. J., Gerber, B. L., & Wiley, L. P. (2003). Gifted achievers and gifted underachievers: The impact of learning style preferences in the classroom. *The Journal of Secondary Gifted Education,14*(4), 197-204.

Reid, D. J., Zhang, J., & Chen, Q. (2003). Supporting scientific discovery learning in a simulation environment. *Journal of Computer Assisted Learning*, 19, 9-20.

Reiser, B. J., Tabak, I., Sandoval, W. A., Smith, B. K., Steinmuller, F., & Leone, A. J. (2001). BGuILE: Strategic and conceptual scaffolds for scientific inquiry in biology classrooms. In S. M. Carver & D. Klahr (Eds.), *Cognition and instruction: Twenty-five years of progress* (pp. 263–305). Mahwah, NJ: Erlbaum.

Shavinina, L.V. (Ed.). (2009). A Unique Type of Representation Is the Essence of Giftedness:Towards a Cognitive-Developmental Theory. In L.V. Shavinina (Ed.) *International handbook on giftedness* (pp. 231 – 260). New York, NY: Springer.

Shore, B. M. (2010). Giftedness is not what it used to be, school is not what it used to be, their future, and why psychologists in education should care. *Canadian Journal of School Psychology, 25*(2), 151169.

Shore, B. M., & Kanevsky, L. S. (1993). Thinking processes: being and becoming gifted. In K.A. Heller, F.J. Mönks, & A.H. Passow (Eds.), *International Handbook of research and development of giftedness and talent* (pp. 133-147). Oxford: Pergamon Press.

Siegler, R. S. (1976). Three aspects of cognitive development. *Cognitive Psychology, 8*(4), 481-520.

Siegler, R. S., & Chen, Z. (2002). Development of rules and strategies: Balancing the old and the new. *Journal of Experimental Child Psychology, 81*(4), 446-457.

Silverman, L. K. (2003). Characteristics of Giftedness Scale: Research and review of the literature. Retrieved from http://www.gifteddevelopment.com.

Snyder, K. E., Nietfeld, J. L., & Linnenbrink-Garcia, L. (2011). Giftedness and Metacognition: A Short-Term Longitudinal Investigation of Metacognitive Monitoring in the Classroom. *Gifted Child Quarterly, 55*(3), 181-193.

Sontag, C., & Stoeger, H. (2015). Can highly intelligent and high-achieving students benefit from training in self-regulated learning in a regular classroom context? *Learning and Individual Differences 41*, 43–53. doi:10.1016/j.lindif.2015.07.008

Steiner, H. H. (2006). A microgenetic analysis of strategic variability in gifted and average-ability children. *Gifted Child Quarterly, 50*(1), 62-74.

Steiner, H.H., & Carr, M. (2003). Cognitive Development in Gifted Children: Toward a More Precise Understanding of Emerging Differences in Intelligence. *Educational Psychology Review, 15*(3), 215-246.

Stoeger, H., Fleischmann, S., & Obergriesser, S. (2015). Self-regulated learning (SRL) and the gifted learner in primary school: the theoretical basis and empirical findings on a research program dedicated to ensuring that all students learn to regulate their own learning. *Asia Pacific Education Review, 16,* 257-267. doi: 10.1007/s12564-015-9376-7

van de Keere, K., & Vervaet, S. (2013). *Leren is onderzoeken.* Tielt, Belgium: LannooCampus.

van Graft, M., & Kemmers, P. (2007). *Onderzoekend en ontwerpend leren bij natuur en techniek: basisdocument over de didactiek voor onderzoekend en ontwerpend leren in het primair onderwijs*. Den Haag: Platform Bèta Techniek.

van Haaren, M. G. P., & Veenman, M.V.J. (n.d.). *Format eindrapportage Onderwijs Bewijs* (Project number 172). Retrieved from http://www.nro.nl/wp-content/uploads/2014/11/Eindrapportage-Hoogbegaafdheid-en-metacognitie-def_0.pdf

van Klink, M., Wilhelm, P., & Lazonder, A.W. (n.d.). Hands-on vs. Hands-off Manipulation in Physical, Video, and Simulation-based Inquiry Learning Environments. Unpublished manuscript, University of Twente, Netherlands, Enschede. Retrieved from http://www.ilo.gw.utwente.nl/ilo/attachments/025_ERMSIL%20definitief.pdf

van Merriënboer, J. J.G., Kirschner, P. A., & Kester, L. (2003). Taking the load off a learner's mind: Instructional design for complex learning. *Educational psychologist*, *38*(1), 5-13.

van Someren, M. W., Barnard, Y.F., & Sandberg, J.A.C. (1994). *The Think Aloud Method: A Practical Guide to Modeling Cognitive Processes*. London: Academic Press.

Veenman, M.V. J. (2013). Training metacognitive skills in students with availability and production deficiencies. In H. Bembenutty, T. Cleary & A. Kitsantas (Eds.), *Applications of self-regulated learning across diverse disciplines: A tribute to Barry J. Zimmerman* (pp. 299-324). Charlotte, NC: Information Age Publishing.

Veenman, M.V.J. (2015). Metacognitie: 'Ken uzelve'. Gebruik die kennis vooral om het eigen gedrag te sturen. *De Psycholoog*, *50*(4), 10-21.

Veenman, M. V. J., Kok, R., & Blöte, A. W. (2005). The relation between intellectual and metacognitive skills in early adolescence. *Instructional Science, 33*(3), 193-211.

Veenman, M. V. J, Van Hout-Wolters, B. H. A. M., & Afflerbach, P. (2006). Metacognition and learning: Conceptual and methodological considerations. *Metacognition and learning*, *1*(1), 3-14.

Velthorst, G., Oosterheert, I., & Brouwer, N. (2011). Onderzoekend leren: de nieuwsgierigheid voorbij. *Tijdschrift voor lerarenopleiders (VELON / VELOV), 32*(3), 32-38.

Webb, J. T. (1994). *Nurturing social emotional development of gifted children*. ERIC Clearinghouse. Retrieved from http://www.casenex.com/casenex/ericReadings/NurturingSocialEmotionalDev.pdf

Wellisch, M., & Brown, J. (2012). An integrated identification and intervention model for intellectually gifted children. *Journal of Advanced Academics, 23*(2), 145-167.

Zhang, J., Chen, Q., Sun, Y., & Reid, D. J. (2004). A triple scheme of learning support design for scientific discovery learning based on computer simulation: Experimental research. *Journal of Computer Assisted Learning, 20*(4), 269–282.

Zimmerman, B. J. (1990). Self-regulated learning and academic achievement: An overview. *Educational psychologist, 25*(1), 3-17.

Zimmerman, B. J., & Martinez-Pons, M. (1990). Student differences in self-regulated learning: Relating grade, sex, and giftedness to self-efficacy and strategy use. *Journal of educational Psychology, 82*(1), 51-59.

Zimmerman, C. (2000). The development of scientific reasoning skills. *Developmental Review, 20,* 99-149.

Zimmerman, C. (2007). The development of scientific thinking skills in elementary and middle school. *Developmental Review, 27,* 172-223.

Zion, M., Michalsky, T., & Mevarech, Z. R. (2005). The effects of metacognitive instruction embedded within an asynchronous learning network on scientific inquiry skills. *International Journal of Science Education*, *27*(8), 957-983.

Appendix A

Post-test

1. Is the volume of the weights important to the balance of the balance beam?
2. Is the mass (the heaviness) important to the balance of the balance beam?
3. Do you always have to use two weights with the same amount of mass when you want to balance the balance beam? For example: do you have to use two weights that both have a mass of 50 grams?
4. Could you also use two weights with different masses to balance the balance beam?
5. Is the position of the weights on the balance beam important to the balance of the balance beam?
6. Do weights always have to be hung at exactly the same positions on both sides of the fulcrum to balance the balance beam?
7. Does a weight weigh the same at each position on the balance beam?
8. Suppose you want to hang two weights with different masses (heaviness) on the balance beam. Where should you hang each weight on the balance beam to balance the beam?
9. Could you calculate at what precise positions two weights have to hang to balance the balance beam, if you know the volume and the masses of the two weights?

Appendix B

Learning Processes Scheme

Table B1

Learning processes scheme

Processes	Learning process	Code	Proceeding
Transformative processes	1.1. Formation of a research question and a hypothesis (HYP)	1.1.1	**The students ask themselves (open) content-related questions (e.g., based on an observation).** *"How could I make a balance with two different weights on each side of the fulcrum?"* *"What would happen if I move the left weight one position nearer to the fulcrum?"* *" What more can I find out ?"*
		1.1.2	**The students form hypotheses.** *"I think the volume of the weight has no effect, because the masses of different weights are the same."* *"I guess volume doesn't matter."* *"The mass probably is important for the balance."*
		1.1.3	**The students make predictions.** *"I think the right side will go down."* *"The blue weight is probably heavier."* *"I guess this will have the same effect."*
	1.2. Design and performance of an experiment (EXP)	1.2.1	**The students observe what is happening.** *"The lever is unbalanced."* *"The lever is balanced when I use two weights of the same mass, with different volume."* *"I don't see any difference."*
		1.2.2	**The students register the data (e.g., taking notes, diagrams, etc.)** *This will be tracked separately.*
	1.3. Data analysis and inferences (ANA)	1.3.1	**The students draw conclusions and theorise (based on observation(s)).** *"This weight feels heavier than that one."* *"The volume of the weights has no effect."* *"The weight at the end is heavier because of gravity."*
Regulative processes	2.1. Planning (problem orientation, goal setting, and strategic planning of actions in the short or long term prior to the task) (PLA)	2.1.1	**The students analyse the task and research question(s).** *"I have only received a question. So I have to find it out by myself."* *"Let's first read the research question."* *"Let's first read all steps of the Inquiry Twister."*
		2.1.2	**The students orientate themselves to the resources and the environment.** *"Let's see… There are four positions on each side of the lever."* *"There are three sizes of each weight."* *"Hey, the beam moves if I touch a side."*
		2.1.3	**The students retrieve prior knowledge.** *"What do I know already?"* *"A few years ago we had to make sums with a scale."* *"I know that two people with different masses could sit on a seesaw and balance it."*
		2.1.4	**The students determine a strategy (long-term planning)** *"First, I will make a balance with the same weights. Then I will use another weight on one side."* *"I'll investigate this by performing experiments."* *"What am I going to do to find out more?"*

Table B1 Continued

Learning processes scheme

Processes	Learning process	Code	Proceeding
Regulative processes	2.1. Planning (problem orientation, goal setting, and strategic planning of actions in the short or long term prior to the task) (PLA)	2.1.5	**The students direct themselves (short-term planning: planned actions that are performed straightaway)** *"Let's try again, this time I will use a greater amount of mass."* *"I am going to inquire..."* *"Now, I'll perform my experiments."*
	2.1. Monitoring (monitoring ongoing learning as compared to its alignment with an earlier plan) (MON)	2.2.1	**The students compare the extent of their knowledge and comprehension to the research question(s).** *"I already know that volume has no effect."* *"Now I only need to find out what volume has to do with it."* *"I think I'm ready to answer the research question by writing on the conclusion sheet."* *"Do I know enough about this? Is my knowledge always true? Can I predict in advance what will happen by means of calculating? Can I answer the research question?"*
		2.2.2	**The students compare the performance of their inquiry to an earlier plan, in view of the research question(s).** *"This did not quite go as I had planned."* *"I used these weights earlier in my second experiment."* *"I don't know how to find out what it is."* *"As we just saw in a previous experiment…"* *"Just checking everything to be sure."*
	2.3. Evaluation and reflection (after the inquiry) (EVA)	2.3.1	**The students reflect on the inquiry process.** *"The inquiry went well."* *"Next time I will orientate better on the research question before I start."* *"It was difficult to investigate the volume."*
		2.3.2	**The students reflect on the learning environment, how they learned, and on their motivation.** *"By having a real lever and real weights, I probably understand this much better than if I had to learn it from a book."* *"By finding out myself I was more focused."* *"This was fun and exciting!"*
		2.3.3	**The students reflect on their knowledge.** *"I knew that volume does not affect balance. This test confirmed my prior belief."* *"I did not learn anything."* *"What more do I know now than before?"*
	2.4. Motivational strategies (MOT)	2.4.1	**The students show self-efficacy and self-attribution of success and failure.** *"I think I can do this."* *"I have done a good job."* *"Stupid of me to try this again."*
		2.4.2	**The students show task interest.** *"I like doing this."* *"I'll try again, because I want to figure it out."* *"That's funny."*

Appendix C

Domain Knowledge Scheme

Table C1

Domain knowledge scheme

Variable	Knowledge			Points
Volume	1a. The student gives no conclusion about the variable. **(0 pts)**			
	2a. Volume affects balance. **(0 pts)**	2b. Volume does not affect balance. **(1 pt)**		
Mass	3. The student gives no conclusion about the variable. **(0 pts)**			
	4a. Mass does not affect balance. **(0 pts)**	4b. Mass affects balance. **(1 pt)**		
	5a. Masses must be equal on both sides of the fulcrum. **(0 pts)**	5b. Equivalent to 11a, 11b, 12a, 12b, and 12c. (Otherwise **1 pt**) Different masses could be used.		
Position	7. The student gives no conclusion about the variable. **(0 pts)**			
	8a. Position does not affect balance. **(0 pts)**	8b. Position affects balance. **(1 pt)**		
	9a. Weights must always be placed at equal positions. **(0 pts)**	9b. Equivalent to 5b, 11a, 11b, 12a, 12b, and 12c. Weights do not always have to be placed on equal positions.		
	10a. When the mass is further from the fulcrum, it is heavier. **(1 pt)**			
	11a. The greater amount of mass should be nearer to the fulcrum, referred to by Siegler (1976) as the conflict-distance balance scale problem. The student does not formulate this formula explicitly, but shows it by (an) example(s). **(2 pts)**	11b. The greater amount of mass should be nearer to the fulcrum, referred to by Siegler (1976) as the conflict-distance balance scale problem. The student explicitly formulates the formula. **(3 pts)**		
	12a. The student has an imperfect integrative theory such as the addition rule (Siegler & Chen, 2002). **(4 pts)**	12b. The student could solve conflict-balance problems by intuitively applying the Law of Moments (mass*position = mass*position), but does not formulate this law explicitly. **(5 pts)**	12c. The student could solve conflict-balance problems by applying the Law of Moments (mass*position = mass*position), and could formulate the correct rule. **(6 pts)**	
			Total of points:	