Effective Instruction for Gifted Children

A Comparison of Direct Instruction and Inquiry Learning

Marlies ten Kleij

University of Twente

Supervised by

Dr. T. H. S. Eysink

Dr. A. H. Gijlers

Zwolle, 3 juli 2012
Abstract

In the debate about whether direct instruction or inquiry learning is the better instruction method, a lot of research has been done. Up till now, this research focused on average students, whereas this study focuses on gifted children. When looking at the learning characteristics and preferences of gifted children, it is expected that inquiry learning is more suitable as instruction method for gifted children, than direct instruction. In the present study, 47 participants received either a direct instruction or an inquiry lesson about the domain ‘floating and sinking’. The learning outcomes of both instructions were compared to each other to see which instruction yielded better learning outcomes and it was tested if the learning outcomes remained the same over time. Also, participants were asked about their preference for direct instruction or inquiry learning. The results showed no differences in learning outcomes between conditions, but they did show that gifted children preferred inquiry learning over direct instruction. The lack of difference in learning outcomes might be explained by the pertinacity of the misconceptions about ‘floating and sinking’.
Effective Instruction for Gifted Children

A Comparison of Direct Instruction and Inquiry Learning

It is widely accepted that learning is an active process in which learners are trying to make sense of new information and are trying to gain coherent and organized knowledge (Mayer, 2004), however, there is much debate going on about which instructional method achieves active learning better. On one hand there are advocates of direct instruction (e.g. Sweller, Merrienboer & Paas, 1998; Mayer, 2004; Kirschner, Sweller & Clark, 2006; Sweller, Kirschner & Clark, 2007), on the other hand there are the advocates of inquiry learning (e.g. Klahr & Dunbar, 1988; Kuhn, Black, Keselman & Kaplan, 2000; Zimmerman, 2000, 2004; Hmelo-Silver, Duncan & Chinn, 2007; Kuhn, 2007).

Direct Instruction and Inquiry Learning

Direct instruction was the prevailing instructional method in education for a long time (Alfieri, Brooks, Aldrich & Tenenbaum, 2011). Direct instruction provides learners with information that fully explains the concepts and procedures that they are required to learn (Kirschner, Sweller & Clark, 2006). It is a highly structured and guided form of instruction (Mayer, 2004). Kirschner, Sweller and Clark (2006) state that we are skilful in an area, because our long-term memory contains a huge amount of information about that specific area. Before this information is stored in long-term memory, it has to be processed by our working memory. Working memory has two well-known limitations, namely it is limited in capacity and duration when dealing with novel information (Kirschner, Sweller & Clark, 2006). When processing information rather than merely storing it, the capacity is even smaller. These limitations disappear when dealing with familiar information, which is already stored in long-term memory. Thus, when learners are provided with all information about the concepts and procedures that are to be learned, the limitations of working memory do not apply anymore, because the given information can be stored in long-term memory and their
working memory is not overloaded (Kirschner, Sweller & Clark, 2006). If this information is not given, learners have to search the problem space for problem-relevant information. All problem-based searching places huge demands on working memory (Kirschner, Sweller & Clark, 2006).

A typical direct instruction lesson, according to Kirschner, Sweller and Clark (2006), starts with an explanation of the subject to be learned, giving all the information that is needed. This is followed by one or more examples or (partially) worked examples. After these, learners get the chance to practice what they learned by making assignments resembling the examples (Klahr & Nigam, 2004). The goals, materials, explanations, examples, and the pace of instruction are all teacher controlled (Dean & Kuhn, 2006).

In inquiry learning, students investigate a set of phenomena and draw conclusions about it (Kuhn, Black, Keselman & Kaplan, 2000). Described from the learner’s point of view: ‘let me investigate how it works’ (Eysink, de Jong, Berthold, Kollofel, Opferman & Wouters, 2009). This means that students have active experiences with the subject matter. They have to come up with several hypotheses and possible explanations for the results (Kuhn, Black, Keselman & Kaplan, 2000). Learners are required to use metacognition about their thinking strategies to manage their inquiry process (De Jong & Van Joolingen, 1998).

In the inquiry learning process, students use scientific reasoning (Zimmerman, 2007). Scientific reasoning can be divided into three phases (Zimmerman, 2000). First, students search for a hypothesis. Second, they conduct an experiment, and finally they evaluate the evidence. These three phases repeat themselves within the experiment and they lead to deep, conceptual understanding about the domain at hand (Eysink, et al., 2009).

**Guided Instruction**

Proponents of inquiry learning and proponents of direct instruction both agree that minimally guided instruction does not work (Mayer, 2004; Hmelo-Silver, Duncan & Chinn,
Effective instruction for gifted children

2007; Kirschner, Sweller & Clark, 2007). In inquiry learning, students should be guided through the inquiry process, to make sure they experiment in a structural manner. Previous research has shown that learners can have several difficulties in the inquiry learning process (de Jong & van Joolingen, 1998, Eysink, et al., 2009). For example, they may find it hard to formulate a hypothesis (de Jong, 2006), or they come to conclusions that cannot be made based on the evidence (Klahr & Dunbar, 1988). Therefore, students need support in inquiry learning to make it an efficient and effective learning process (Mayer, 2004). The support will help students to focus on the core of the learning material, instead of focusing merely on the inquiry learning skills (de Jong & van Joolingen, 1998). Support can be given, for instance, by asking students questions, by making students plan their actions upfront or by including assignments to break a complex task into easier, more ‘doable’ steps (de Jong, 2006).

Several studies investigated the difference between direct instruction and inquiry learning. For example, Klahr and Nigam (2004) investigated the difference between direct instruction and discovery learning in acquiring inquiry learning skills. They found that learners who received direct instruction about control-of-variables strategy outperformed learners who received discovery learning in direct assessment as well as transfer assessments (Klahr & Nigam, 2004). Dean and Kuhn (2006) also studied the acquisition of the control-of-variables strategy, but over a longer period of time. Their results showed that direct instruction is not necessary, nor sufficient to master the control-of-variables strategy, and this result remained the same over a longer period of time (Dean & Kuhn, 2006).

An example of a study comparing learning outcomes between different instructional methods is the study of Eysink et al. (2009). They compared four multimedia learning arrangements, namely hypermedia learning, observational learning, self-explanation based learning, and inquiry learning (Eysink et al., 2009). Their results showed that when students generate part of the subject matter, they perform better than when the subject matter is merely
presented to them. Also, students in this study who used inquiry learning, or explanation-based learning, ended up with good conceptual, situational, intuitive, and procedural knowledge, especially on far transfer. An explanation for this, according to Eysink, et al. (2009), is the level of cognitive activity that is asked of students when they generate hypotheses, conduct experiments, and draw conclusions based on their findings.

Alfieri, Brooks, Aldrich, and Tenenbaum (2011) performed a meta-analysis of studies about under which conditions unassisted discovery learning would yield better learning outcomes than explicit-instructional tasks, and under which conditions enhanced-discovery methods (such as inquiry learning) would work best. The results of the first part of their meta-analysis showed that unassisted discovery learning does not benefit learning. However, providing learners with worked examples or timely feedback works better than direct instruction (Alfieri, Brooks, Aldrich & Tenenbaum, 2011). The second part of the meta-analysis showed that enhanced-discovery conditions led to better learning outcomes for all age groups. The explanation given by Alfieri, Brooks, Aldrich, and Tenenbaum (2011) is that well-placed scaffolds free the learners from high demands on working memory so they can direct their efforts more to creative processes. Thus, enhanced-discovery tasks, which require learners to be actively engaged and constructive, to explain their own ideas, or which provide worked examples, seem optimal.

**Gifted children**

In most of the existing studies about direct instruction and inquiry learning there was no distinction between students based on competence level. However, when you look at literature about gifted children, it would make sense that inquiry learning would yield better learning outcomes for this particular group.

Although there is not a single definition of gifted children (Gallagher, 2008; Kaufman & Sternberg, 2008), most definitions of gifted children mention high intelligence. Often an IQ
of 130 or above (Newman, 2008). Also, there is a number of other characteristics of
giftedness which predict that inquiry learning would yield better learning outcomes with
gifted children. These will be described below. Metacognition is mentioned first, followed by
creative thinking. These are two characteristics that are very helpful in inquiry learning. Then,
other cognitive characteristics are mentioned, which help in inquiry learning and in learning
in general. Finally, preferences of gifted children are mentioned.

As mentioned before, metacognition is an important component of giftedness
(Robinson & Clinkenbeard, 2008), and it plays an important role in inquiry learning as well
(De Jong & Van Joolingen, 1998). Gifted children are not better than average students at all
aspects of metacognition, but they do have more factual knowledge about thinking strategies,
and they are better at transferring strategies to contexts that are far from the context in which
the strategies were learned (Robinson & Clinkenbeard, 2008; Sekowski, Siekanska &
Klinkosz, 2009). Although gifted children do not seem to use different cognitive strategies
than other children, they tend to acquire and process information and solve problems faster,
better, and at an earlier age than other children do (Robinson & Clinkenbeard, 2008). Gifted
children generally have a very good memory and they learn fast, this is connected to the
thinking strategies that are applied (Sekowski, Siekanska & Klinkosz, 2009), and it helps to
unload their working memory (Kirschner, Sweller & Clark, 2006).

Another characteristic of giftedness is creative thinking. Gifted children are good at
finding numerous solutions to the same problem, have original ideas, expressions, and
solutions, and their ability to take risks is higher than with average children (Sekowski,
Siekanska & Klinkosz, 2009) Also, according to Kanevsky (2011) gifted children like to find
creative solutions to difficult problems. Creative thinking helps gifted children in stating
hypotheses and explaining results, which is asked of learners in inquiry learning (De Jong &
Other cognitive characteristics of gifted children are that they have more schemas than other children of their age, especially in their fields of expertise (Porath, 2006). This means they can retrieve information from their long-term memory more easily and the demand on their working memory is smaller (Kirschner, Sweller & Clark, 2006). To keep developing these schemas, instruction for gifted children should be focused on conceptual understanding. Gifted children are good at making (causal) connections, they are good at analyzing problems, they have a good understanding of abstract concepts, they have a good time and work management, and they have the ability to focus on a task for a long period of time (Stichting Leerplan Ontwikkeling [SLO]; Sekowski, Siekanska & Klinkosz, 2009).

When looking at the preferences of gifted children, they rate experimenting more positive than other children (Kanevsky, 2011). Seeking help from their teacher is something gifted children do not like to do, they prefer working independently (Sekowski, Siekanska & Klinkosz, 2009; Kanevsky, 2011). Also, gifted children have higher intrinsic motivation to learn than their average peers. And they have positive attributions about success and failure (Robinson & Clinkenbeard, 2008). This helps gifted children in testing hypotheses and restating them if necessary (Yoon, 2009).

Research about best education for gifted children shows a number of criteria that the learning material should meet in order for gifted children to learn most of it. Research by VanTassel-Baska and Brown (2007) showed that the use of advanced curricula at an accelerated rate connects to the faster and better acquiring and processing of information and solving problems by gifted students (Robinson & Clinkenbeard, 2008). Also, multiple higher level thinking skills should be embedded (VanTassel-Baska & Brown, 2007) and learning material should call upon metacognitive skills of students (Bronckhorst, Drent, Hulsbeek, Steenbergen-Pernternman & van der Veer, 2001), like inquiry learning does (De Jong & Van Joolingen, 1998). This is connected to the fact that gifted children have more factual
knowledge about thinking strategies and that they can transfer these strategies easily to other contexts (Robinson & Clinkenbeard, 2008; Sekowski, Siekanska & Klinkosz, 2009).

VanTassel-Baska and Brown (2007) also found that the use of inquiry as a central strategy to promote gifted students learning was an essential feature of best practice for gifted students, along with the use of student-centered learning opportunities that are problem-based and relevant to the student. This is also what Bronckhorst et al. (2001) implied with their criteria that learning material for gifted children should have open assignments, have a high level of abstractness and complexity, stimulate a science attitude in students, appeal to creativity, call upon independence, and elicit a reflective attitude. This feature and these criteria connect to the cognitive characteristics mentioned before (Stichting Leerplan Ontwikkeling [SLO]; Porath, 2006; Sekowski, Siekanska & Klinkosz, 2009) and it also appeals to the preferences gifted children have (Robinson & Clinkenbeard, 2008; Sekowski, Siekanska & Klinkosz, 2009; Kanevsky, 2011).

Hypotheses

When looking at the characteristics of gifted children and the criteria for effective gifted education that will teach them most, it seems that inquiry learning would be a suiting way for gifted children to learn (Kanevsky, 2011). If you also look at the preferences of gifted children, inquiry learning looks like an even more suitable instruction method for gifted children. In this study, the learning outcomes of a direct instruction condition were compared to the learning outcomes of an inquiry learning condition. The considerations in the introduction about direct instruction, inquiry learning and gifted children lead to the following hypotheses.

First, gifted children who are in an inquiry learning setting will have higher learning outcomes than gifted children who are in a setting in which direct instruction is given. The inquiry learning setting meets the criteria for educating gifted children that Bronkhorst et al.
Effective instruction for gifted children

(2001) and VanTassel-Baska and Brown (2007) propose. The direct instruction setting does not have the affordance for students to be creative, think abstract, be independent, use metacognitive skills, or have a reflective attitude.

Second, the assumption is made that these higher learning outcomes remain over a period of time. This follows from previous studies about inquiry learning and direct instruction that showed that the knowledge that was learned still existed in time, because when learners understand something they can remember it better (Dean & Kuhn, 2006; Alfieri, Brooks, Aldrich & Tenenbaum, 2011).

Finally, it is assumed that gifted children prefer learning through inquiry learning over learning through direct instruction. Gifted children prefer experimenting (Kanevsky, 2011) and working independently (Sekowski, Siekanska & Klinkosz, 2009).

Method

Participants

Five elementary schools took part in the experiment. Participants were selected based on the following criteria: Participants had to be in the third or fourth grade of elementary school (i.e., in the age of seven through ten), and they had to be gifted (i.e., IQ above 130) or talented (i.e., IQ between approximately 120 and 130; or IQ unknown, but selected for a gifted children course at their school) (Newman, 2008). This distinction between gifted and talented was made for practical reasons, because not every school tests students on their IQ before letting them take part in a gifted children course. Before taking part in the experiment, the students’ parents were asked for their consent, which all parents gave.

In total, 49 students participated. Two participants were excluded, because one of them scored more than two standard deviations below the mean score on the post-test and retention test, and the other one on the pre-test and post-test. The remaining 47 participants had an average
Effective instruction for gifted children

age of 9.25 years: their age ranged from 7.25 through 10.60 years. There were 25 girls, and 22 boys.

In the direct instruction (DI) condition, there were 23 participants in total (ten boys and 13 girls) with an average age of 9.27 years. Ten of them were gifted, 13 were talented. The inquiry learning (IL) condition contained 24 participants with an average age of 9.31 years, 12 girls and 12 boys. Half of them were gifted, and the other half talented.

The experiment took place during school hours. The experiment was not part of the curriculum of the participants. Students did not receive a compensation for their participation.

**The domain**

The domain floating and sinking was chosen for the experiment based on the work of Yin, Tomita & Shavelson (2008), who composed a list of ten typical misconceptions about floating and sinking children have. These misconceptions are formed by limited observation or experience. To teach students about the correct conceptions about ‘floating and sinking’, these existing misconceptions should be acknowledged. In order to make the current study not too elaborate, three misconceptions were used, namely (a) big objects sink, small objects float; (b) heavy objects sink, light objects float; and (c) a large amount of water makes things float (Yin, Tomita & Shavelson, 2008). The misconceptions were used to assemble the misconception questionnaires (the pre-test, post-test, and retention test). By asking the participants questions related to these misconceptions, a clear view was obtained about whether participants had the same misconceptions as the participants in the study of Yin, Tomita, and Shavelson (2008). Also, it gave a good image of the influence of the instructions on the existing misconceptions.

**Materials**

**Thinking aloud instruction.** For the thinking aloud instruction, a tangram puzzle was used along with two examples, one example for the experimenter to finish, one for the
Effective instruction for gifted children

participant (see Figure 1). The experimenter made the first puzzle while thinking aloud. After the puzzle was completed, the experimenter discussed what she said to further explain thinking aloud. After this, it was the participant’s turn to complete example two with the tangram pieces. The experimenter prompted participants to keep thinking aloud, by asking questions such as “What are you thinking right now?”, and “Keep talking”. When the participants had completed the tangram puzzle, the experimenter discussed with them what went well and what could have be done.

![Figure 1. Tangram examples used in the thinking aloud instruction. The experimenter made the triangle, participants were asked to do the bunny figure.](image)

Misconception questionnaires. Participants were asked to fill out a questionnaire at three moments: before they received the instruction, right after they finished the instruction, and two weeks after they received the instruction. These questionnaires are the pre-test, post-test, and retention test respectively and will be referred to as the misconception questionnaires. The questionnaires were developed to see if participants had any misconceptions about floating and sinking, and whether the instructions influenced the existing misconceptions. The three misconceptions that were tested were (a) big objects sink, small objects float; (b) heavy objects sink, light objects float; and (c) a large amount of water makes things float (Yin, Tomita & Shavelson, 2008).

The first misconception was used in two forms. Participants were asked if a block ten times as big or ten times as small as the example block would still sink or float in questions
one and two. In questions five and six, participants were asked whether two identical blocks glued together or one of these blocks cut through half would sink or float (see Figure 2).

The second misconception was used in questions three and four. Here participants were asked if a block ten times as heavy or light as the example block would still sink or float.

The third misconception was used in question seven. Participants were asked whether a block would still sink or float when it was thrown in a bigger or smaller bowl.

Thus, the misconceptions were spread over seven questions. The seven questions each had two or three sub questions, which gave 17 items on each questionnaire in total.

5) These two blocks are exactly alike and they both float on the water.

a) If we glue the two blocks firmly together, will they sink or float?

   A. They will float
   B. They will sink
   C. You can’t tell, because

b) If we saw one of the blocks exactly in half, would it float or sink?

   A. It will float
   B. It will sink
   C. You can’t tell, because

Figure 2. The fifth question, with its two sub questions, of the pre-test, post-test and retention test.

**Own ideas test.** To get a broader picture of the ideas participants had about ‘floating and sinking’, they were asked to think about their own experiences and already existing ideas about why some objects float and other objects sink. Participants could write down anything they thought of, related to the domain. This open question was called ‘initial idea’. To see if the instruction changed the initial ideas of the participants, they were asked the same question again afterwards. This question was called ‘eventual idea’, in the IL condition this was the
sixth assignment, in the DI condition the question was asked after participants had solved all math problems.

**Additional open questions.** The post-test was complemented with four open questions about what participants thought about the instruction, what their favorite and least favorite subjects were in school, and whether they preferred to learn new subjects by doing it themselves or when their teacher told them.

**Experiment set.** In the IL condition, participants used an experiment set. This experiment set consisted of seven black blocks. Each block had its own volume, mass, and density (see Table 1). To discriminate between the blocks, they were color coded. Colors were chosen instead of numbers or letters to preclude the suggestion of a sequence. Together with the experiment set, participants in this condition used a regular kitchen scale to weigh the blocks, and a water tank to test the blocks.

<table>
<thead>
<tr>
<th>Block color</th>
<th>Volume</th>
<th>Mass</th>
<th>Density</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yellow</td>
<td>1157.63</td>
<td>1250.00</td>
<td>1.08</td>
</tr>
<tr>
<td>Green</td>
<td>1157.63</td>
<td>640.00</td>
<td>0.55</td>
</tr>
<tr>
<td>Purple</td>
<td>512.00</td>
<td>640.00</td>
<td>1.25</td>
</tr>
<tr>
<td>Red</td>
<td>512.00</td>
<td>375.00</td>
<td>0.73</td>
</tr>
<tr>
<td>Orange</td>
<td>274.63</td>
<td>200.00</td>
<td>0.73</td>
</tr>
<tr>
<td>Pink</td>
<td>91.25</td>
<td>200.00</td>
<td>2.19</td>
</tr>
<tr>
<td>Grey</td>
<td>91.25</td>
<td>70.00</td>
<td>0.77</td>
</tr>
</tbody>
</table>

**Instructions.** In the DI condition, participants received an explanation of density, and ‘floating and sinking’. The explanation consisted of a one page long text, in which a figure of a block was included to clarify what was written. The information was accompanied by two
examples, showing participants how to calculate the density of a block. They also received a sheet with seven math problems on it, to practice what they just read about. Participants were allowed to use a calculator when solving the math problems. The first math problem used the data from example one and was filled out already. The second math problem used the data from example two. To keep the two different instructions as equal as possible, the math problems that were used, including the two examples, contained the information of the blocks of the experiment set from the IL condition. The instruction was concluded with the question ‘eventual idea’.

In the IL condition, participants received six assignments, accompanied by a worksheet. The assignments guided participants through the experiments with the experiment set, so that the experiment was manageable for the participants. Assignment one asked the students to pick a small block, think about what would happen when they would threw it in the water (i.e. hypothesize), weigh the block, and drop the block in the water tank (i.e. conduct an experiment) to test if their initial idea was correct (i.e. evaluate the evidence). The steps were included in the assignment to show participants how they could experiment in a structured manner. Assignment two asked participants to do the same with one of the large blocks. In order not to take participants too much by the hand, the different steps were not mentioned here. Assignment three asked the participants to compare the blocks they tested in assignments one and two. They were asked to write down (a) which blocks they chose, (b) what the difference was between the two blocks, and (c) if something happened what they had not expected. This assignment helped participants in drawing conclusions about the available data. The fourth assignment said to write down any new ideas about why some things float and other things sink. This assignment was included to help participants in restating their hypotheses. Assignments one through four were included to introduce participants in the scientific thinking process and to help them conduct an experiment in a structural manner. In
the fifth assignment, participants attended the blocks again. They were asked to test the remaining blocks in the same way as they did in assignments one and two. During testing the blocks, participants had the opportunity to write down new hypotheses, conclusions or if something happened that stood out to them. Thus, the fifth assignment asked of the participants to finish the experiment on their own. The worksheet participants received consisted of a table they could fill out while they were doing the assignments. The table already included the measures (or volume) of the blocks. The sixth, and final, assignment was the same as the concluding question of the DI condition, namely the question ‘eventual idea’.

**Procedure**

After the experimenter had introduced herself to the class, she asked the participants one by one to come with her. The average duration of each experiment was approximately 45 minutes. The experimenter explained what the participant was about to do, namely participating in an instruction about floating and sinking.

Before starting with the instruction and questions about ‘floating and sinking’, participants received the thinking aloud instruction. First, the experimenter explained why participants were asked to think aloud. The given explanation was that the lesson would be filmed so that the experimenter could hear again at home what every participant exactly said. When the thinking aloud instruction was finished, the experimenter pinned a microphone to the clothes of the participants and started the filming. Participants began with putting their name, age, and the name of their school on the first page. Then they were asked to write down what they already knew about why some things float and other things sink (i.e. ‘initial idea’). This question was included to test if participants had any prior knowledge about the subject and to activate this prior knowledge. When this was done, participants were given the pre-test, which they filled out while thinking aloud.
After the pre-test, it was time for the instructions about ‘floating and sinking’, which took participants approximately 20 minutes in both conditions. In the DI instruction, participants were given the explanation about density and the first example. After they read this out loud, participants were asked if they wanted to start with the math problems or if they first wanted to read example two as well. Participants who chose to read example two, started with the math problems after that.

In the IL condition, participants were given the experiment set, the scale, the assignments, and the worksheet to fill out. The experimenter made sure that participants followed the order of the assignments, starting at assignment one.

After having answered all math problems (in the DI instruction), or after having experimented with all blocks and finishing the first five assignments (in the IL condition), participants were asked to write down what their thoughts were now about why some objects float and other objects sink (i.e. ‘eventual idea’). Subsequently, participants in both conditions were asked to fill out the post-test questions.

The retention test was filled out by the participants two weeks after they received the instruction about ‘floating and sinking’. The retention test was administered in a whole-class setting, participants were not asked to think aloud.

**Data-analysis**
To score the answers to the open questions before the instruction and after the instruction (i.e., ‘initial idea’ and ‘eventual idea’), a scoring scheme was developed (see Table 2). Participants who gave answers with multiple components, only received points for the component that was worth the most points. Scores for the pre-test, post-test and retention test were computed by awarding a correct answer with one point, incorrect answers yielded no points. Thus, participants could get a maximum score of 17 on each test.
Table 2

Scoring Scheme Used to Score Answers to the Questions ‘Initial Idea’ and ‘Eventual Idea’.

<table>
<thead>
<tr>
<th>Points</th>
<th>Answers</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>Answers referring to both dimensions (i.e., weight and size) or density (e.g., mass and volume, mass divided by volume).</td>
</tr>
<tr>
<td>2</td>
<td>Answers referring to material (e.g., wood floats, rock sinks) or substance (e.g., depends on what it’s made of), or in the DI condition when a term from the math problems is mentioned, but not in the right manner (e.g., mass times volume).</td>
</tr>
<tr>
<td>1</td>
<td>Answers referring to only one dimension (either size or weight or gravity)</td>
</tr>
<tr>
<td>0</td>
<td>Answers referring to misconceptions (e.g., it contains air, or depends on the amount of water), own experiences (e.g., I saw this on ‘Nieuws uit de Natuur’), or other incorrect answers.</td>
</tr>
</tbody>
</table>

The additional question from the pre-test ‘do you prefer learning new things by doing it yourself or when your teacher tells it to you’ could be answered with (a) do it myself, (b) when my teachers tells it, or (c) no preference. The answers were counted, to give an idea of the preferences of the participants for direct instruction (i.e., when my teacher tells it) or inquiry learning (i.e., do it myself).

Results

Reliability

The reliabilities of the pre, post, and retention test were \( \alpha = .49 \), \( \alpha = .61 \), and \( \alpha = .65 \) respectively, as measured with Cronbach’s alpha. A possible explanation for the low reliability on the pre-test might be that the existing knowledge about the domain differed per participant. Deleting items did not increase the reliability on the pre-test, post-test and retention test.
Gifted vs. talented

Because there were both gifted and talented participants, it was tested if the scores for these groups differed. In the DI condition, test scores of gifted participants on the pre-test and post-test were normally distributed ($W = .92, p = .41$, and $W = .95, p = .69$, respectively), their scores on the retention test were not normally distributed as measured with the Shapiro Wilk test of normality ($W = .80, p = .02$). Scores of talented participants in the DI condition were normally distributed, as measured with Shapiro Wilk test of normality, on the pre-test, post-test, and retention test ($W = .90, p = .16$; $W = .92, p = .27$; and $W = .96, p = .70$, respectively). A one-way ANOVA showed that there was no significant difference on the pre-test ($F(1,21) = .47, p = .49$) and post-test ($F(1,21) = .20, p = .66$) for gifted and talented children in the DI condition. A Mann-Whitney test showed that there was no significant difference between scores of gifted and talented children on the retention test in the DI condition ($U = 49.50, Z = -.32, p = .75$).

In the IL condition scores of the gifted participants were normally distributed as measured with Shapiro Wilk test of normality on the pre-test, post-test, and retention test ($W = .94, p = .53$; $W = .91, p = .23$; and $W = .90, p = .14$, respectively). Scores of the talented participants in the IL condition were normally distributed as measured with Shapiro Wilk test of normality on the post-test ($W = .92, p = .28$), but not on the pre-test and retention test ($W = .79, p \leq .05$, and $W=0.84, p \leq .05$, respectively). An independent samples t-test showed that there was no significant difference between scores of gifted and talented participants on the post-test ($t(22) = -.97, p = .34$). A Mann-Whitney test showed that there was no significant difference in the IL condition between scores of gifted and talented children on the pre-test ($U = 44.00, Z = -1.63, p = .10$) and the retention test ($U = 59.00, Z = -.76, p = .45$). There was no difference in scores between gifted and talented children in both conditions, therefore the participants are treated as one group in the following analyses.
**Main effect of condition on the misconception questionnaires.** The means and standard deviations of the misconception questionnaires are displayed in Table 3.

The scores on the pretest were normally distributed as measured with Shapiro Wilk test of normality: $W = .92$ ($p = .08$) for the direct instruction condition and $W = .92$ ($p = .06$) for the inquiry learning condition. Skewness was 1.00 for the DI condition and .63 for the IL condition.

*Table 3*

*Overview of the Means and Standard Deviations of the Scores of the Misconception Questionnaires and Open Questions.*

<table>
<thead>
<tr>
<th>Condition</th>
<th>DI condition</th>
<th>IL condition</th>
</tr>
</thead>
<tbody>
<tr>
<td>Misconception questionnaires</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pre-test</td>
<td>Mean 9.43</td>
<td>10.46</td>
</tr>
<tr>
<td></td>
<td>SD 2.00</td>
<td>2.77</td>
</tr>
<tr>
<td>Post-test</td>
<td>Mean 9.43</td>
<td>10.50</td>
</tr>
<tr>
<td></td>
<td>SD 2.17</td>
<td>2.52</td>
</tr>
<tr>
<td>Retention test</td>
<td>Mean 9.62</td>
<td>10.50</td>
</tr>
<tr>
<td></td>
<td>SD 2.69</td>
<td>2.75</td>
</tr>
<tr>
<td>Open questions</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Initial idea</td>
<td>Mean 0.91</td>
<td>0.87</td>
</tr>
<tr>
<td></td>
<td>SD 0.60</td>
<td>0.68</td>
</tr>
<tr>
<td>Eventual idea</td>
<td>Mean 1.96</td>
<td>1.58</td>
</tr>
<tr>
<td></td>
<td>SD 0.93</td>
<td>1.02</td>
</tr>
</tbody>
</table>
The scores on the post-test were also normally distributed as measured with Shapiro-Wilk test of normality for both the DI and the IL condition ($W = .95, p = .36$, and $W = .93, p = .120$, respectively). Skewness for the DI condition was ≤05 and for the IL condition .18. The retention test scores were normally distributed in the DI condition (skewness = .45; Shapiro-Wilk test of normality: $W = .95, p = .32$) and in the IL condition as well (skewness = .36; Shapiro-Wilk test of normality: $W = .92, p = .58$). A one-way ANOVA test showed that there was no significant difference between the scores of the participants in the two conditions on the pretest ($F(1,45) = 2.10, p = .15$), post-test ($F(1,45) = 2.40, p = .13$), and retention test ($F(1,43) = 1.20, p = .29$).

To see if the scores on the post-test were higher than scores on the pretest, a paired samples t-test was performed. The mean scores in the DI condition did not differ ($t(23) = .00, p = 1.00$) and there were no significant improvements in the IL condition ($t(23) = -.10, p = .46$). A paired samples t-test was used to compare scores on the posttest with scores on the retention test. There was no difference in the IL condition ($t(21) = .00, p = 1.00$) and no significant difference for the DI condition ($t(21) = -1.04, p = .16$).

**Individual misconceptions.** A distinction was made between the misconceptions on the misconception questionnaires to see if there was a difference between the scores (see Table 4). A one-way ANOVA shows that there was no difference between conditions on the pre-test for the first misconception ‘big objects sink, small objects float’ ($F(1,45) = .01, p = .926$), nor for the second misconception ‘heavy objects sink, light objects float’ ($F(1,45) = 1.13, p = .29$), nor for the third misconception ‘a large amount of water makes things float’ ($F(1,45) = 2.20, p = .15$). On the post-test there was no significant difference between conditions on the first misconception ($F(1,45) = .55, p = .46$), the second misconception ($F(1,45) = .01, p = .93$), and the third misconception ($F(1,45) = 3.90, p = .05$). In the DI condition there was no significant improvement from the pre-test to the post-test scores for
the first, second and third misconception as measured with a paired samples t-test ($t(22) = -.60, p = .56$, $t(22) = -.42, p = .68$, and $t(22) = -2.01, p = .06$, respectively).

Table 4

<table>
<thead>
<tr>
<th>Misconception</th>
<th>Condition</th>
<th>Direct instruction</th>
<th>Inquiry learning</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Pre-test</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Misconception 1</td>
<td>Mean</td>
<td>5.91</td>
<td>5.96</td>
</tr>
<tr>
<td></td>
<td>SD</td>
<td>.78</td>
<td>1.76</td>
</tr>
<tr>
<td>Misconception 2</td>
<td>Mean</td>
<td>2.52</td>
<td>2.29</td>
</tr>
<tr>
<td></td>
<td>SD</td>
<td>.85</td>
<td>.62</td>
</tr>
<tr>
<td>Misconception 3</td>
<td>Mean</td>
<td>1.30</td>
<td>1.58</td>
</tr>
<tr>
<td></td>
<td>SD</td>
<td>.70</td>
<td>.58</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Misconception 1</td>
<td>Mean</td>
<td>6.22</td>
<td>5.71</td>
</tr>
<tr>
<td></td>
<td>SD</td>
<td>2.09</td>
<td>2.60</td>
</tr>
<tr>
<td>Misconception 2</td>
<td>Mean</td>
<td>2.43</td>
<td>2.42</td>
</tr>
<tr>
<td></td>
<td>SD</td>
<td>.84</td>
<td>.58</td>
</tr>
<tr>
<td>Misconception 3</td>
<td>Mean</td>
<td>1.52</td>
<td>1.17</td>
</tr>
<tr>
<td></td>
<td>SD</td>
<td>.60</td>
<td>.64</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Misconception 1</td>
<td>Mean</td>
<td>6.18</td>
<td>6.30</td>
</tr>
<tr>
<td></td>
<td>SD</td>
<td>2.02</td>
<td>2.75</td>
</tr>
<tr>
<td>Misconception 2</td>
<td>Mean</td>
<td>2.14</td>
<td>2.35</td>
</tr>
<tr>
<td></td>
<td>SD</td>
<td>.71</td>
<td>.65</td>
</tr>
<tr>
<td>Misconception 3</td>
<td>Mean</td>
<td>1.33</td>
<td>1.46</td>
</tr>
<tr>
<td></td>
<td>SD</td>
<td>.58</td>
<td>.59</td>
</tr>
</tbody>
</table>

*Note. Misconception 1 is big objects sink, small objects float; misconception 2 is heavy objects sink, light objects float; misconception 3 is a large amount of water makes things float.*

For the first and second misconception there was no significant improvement from pre-test to post-test in the IL condition ($t(23) = .65, p = .52$, and $t(23) = -.72, p = .48$,
respectively). On the third misconception there was a significant difference between pre-test and post-test in the IL condition ($t(23) = 3.50, p \leq .05$). On the retention test there was no significant difference between conditions as measured with a one-way ANOVA for the first misconception ($F(1,43) = .03, p = .87$), for the second misconception ($F(1,43) = 1.09, p = .30$), and for the third misconception ($F(1,43) = .52, p = .48$). A paired samples t-test showed that there is no significant difference between the post-test and retention test in the DI condition for the first misconception ($t(21) = -.16, p = .88$), for the second misconception ($t(21) = 1.82, p = .08$), and for the third misconception ($t(20) = .72, p = .48$). In the IL condition, there was no significant difference between post-test and retention test as measured with a paired samples t-test on the first misconception ($t(22) = -1.72, p = .10$), on the second misconception ($t(22) = .37, p = .72$), and on the third misconception ($t(23) = -1.77, p = .09$).

**Open questions**

**Reliability.** For the open questions, the inter-rater reliability was measured using Cohen’s Kappa. A second rater scored 21% of the answers to ‘initial idea’ and ‘eventual idea’. The inter-rater reliability for the raters was found to be Kappa = .70 ($p \leq .05$).

**Main effect of condition on the open questions.** The scores (see Table 3 for means and standard deviations) on the open questions ‘initial idea’ and ‘eventual idea’ were not normally distributed in the DI condition (Shapiro-Wilk test of normality: $W = 0.76, p \leq .05$ and $W = .85, p \leq .05$, respectively), nor in the IL condition (Shapiro-Wilk test of normality: $W = .80, p \leq .05$ and $W = .86, p \leq .05$, respectively). Thus, the Wilcoxon signed ranks test was used to test the scores within the conditions. In the DI condition, there was a significant difference between the scores for ‘initial idea’ and ‘eventual idea’ ($Z = -3.50, p \leq .05$). In the IL condition there also was a significant difference between the scores on the open questions ($Z = -2.81, p \leq .05$).
Between both conditions, there was no significant difference, as measured with a Mann-Whitney U test, on the scores for ‘initial idea’ ($U = 266.00, Z = -.243, p = .81$) and on the scores for ‘eventual idea’ ($U = 218.00, Z = -1.29, p = .20$).

**Preference**

The answers to the question whether participants preferred to learn new things by themselves or when their teacher explained it to them (i.e. preference) were also analyzed. Table 5 gives an overview of how many times each answer was given. A Pearson chi-square test shows that there is no significant association between condition and given answer ($\chi^2 = .87, p = .647$).

<table>
<thead>
<tr>
<th>Given answers</th>
<th>Direct instruction</th>
<th>Inquiry learning</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Frequencies</td>
<td>Percentage</td>
<td>Frequencies</td>
</tr>
<tr>
<td>Doing it myself</td>
<td>14</td>
<td>60.9%</td>
<td>16</td>
</tr>
<tr>
<td>Both equally</td>
<td>5</td>
<td>21.7%</td>
<td>6</td>
</tr>
<tr>
<td>Instruction</td>
<td>4</td>
<td>17.4%</td>
<td>2</td>
</tr>
<tr>
<td>Total</td>
<td>23</td>
<td>100.0%</td>
<td>24</td>
</tr>
</tbody>
</table>

*Note. Preferences are the answers to the question 'do you prefer learning new things by doing it yourself or when the teacher tells it?*

**Time on task**

The time it took every participant to finish the instruction about ‘floating and sinking’ was measured, starting right after the pre-test was filled out till the last question ‘eventual idea’ was answered. The mean instruction times (see Table 6) were compared between the both conditions.
Table 6

*Overview of the Means and Standard Deviations of the Time on Task per Condition.*

<table>
<thead>
<tr>
<th></th>
<th>DI condition</th>
<th></th>
<th>IL condition</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>SD</td>
<td>Mean</td>
<td>SD</td>
</tr>
<tr>
<td>Instruction time</td>
<td>22.70</td>
<td>7.87</td>
<td>19.13</td>
<td>4.91</td>
</tr>
<tr>
<td>Total time</td>
<td>49.22</td>
<td>10.00</td>
<td>43.75</td>
<td>8.50</td>
</tr>
</tbody>
</table>

In the DI instruction, times were normally distributed (Shapiro Wilk test of normality: $W = .94, p = .19$). The times in the IL condition were also normally distributed as measured with Shapiro Wilk test of normality ($W = .94, p = .19$). An independent samples t-test showed that the mean times in both conditions did not significantly differ ($t (36.60) = 1.86, p = .07$).

**Conclusion and discussion**

The aim of the present study was to compare learning outcomes of gifted children in inquiry learning with learning outcomes of gifted children in direct instruction about the domain ‘floating and sinking’. It was hypothesized that inquiry learning would yield better learning outcomes and that the higher learning outcomes would still exist over a period of time. It was also hypothesized that gifted children preferred inquiry learning over direct instruction.

First of all, it seems that gifted children also have the misconceptions of which the misconception questionnaires consisted. These misconceptions did not disappear after receiving the instructions in either condition. A possible explanation for this is that the misconceptions students had before participating in this study, were so persistent that they were not changed by the given instruction. This is also suggested in the study of Yin, Tomita and Shavelson (2008). That the results did not decline over time confirms the pertinacity of the misconceptions.
In learning about ‘floating and sinking’, the existing concepts of students have to be changed. What is needed to accomplish conceptual change, is the reinterpretation of the misconceptions within a different explanatory framework (Vosniadou & Brewer, 1992). This is a process that occurs gradually and slow. In this study, participants received one lesson about the domain. This lesson probably initiated the process of conceptual change, but more instruction is necessary to complete this process (Vosniadou & Brewer, 1992).

That the process of conceptual change was initiated, is illustrated by the answers to the open questions. These answers imply that participants did have more knowledge about ‘floating and sinking’ after they received instruction. That there was a significant improvement on the open questions and not on the multiple choice questions, might be explained by the question style. The multiple choice questions asked participants about one situation and it forced them to choose if a particular object would either sink or float. The open questions on the other hand asked participants to think about a general rule why some objects float and other objects sink. This may have triggered participants to think about a more coherent concept. The open questions also gave participants the possibility to include both their existing ideas, as well as the knowledge they had just learned from the instruction.

However, this gain in knowledge did not differ between conditions. An explanation for this might be the type of inquiry learning instruction that was used in this study. Namely, it was rather structured by the assignments that were included. The assignments were included to help the participants to experiment in a structured manner. One might argue that due to the assignments, the inquiry learning task became too structured. This can also be a reason why participants did not perform better in the inquiry learning condition than in the direct instruction condition.

Another possible explanation of why there was no difference in learning outcomes between condition, is the way gifted children think. Where average children learn mostly
passive in direct instruction, without engaging in elaborative learning processes and higher-order thinking processes (Eysink & de Jong, 2011), direct instruction to gifted children, as well as inquiry learning, might lead to elaborative learning processes. The reason for this is that gifted children have more factual knowledge about thinking strategies, are better at transferring them to other contexts, and acquire and process information faster (Robinson & Clinkenbeard, 2008; Sekowski, Siekanska & Klinkosz, 2009).

That there was no difference in learning outcomes, might indicate that the difference between the conditions lies in the motivation of gifted children to learn and in the preference of gifted children for inquiry learning as instruction method. This is also suggested by the preference of the participants in both conditions to learn new things by themselves. This is consistent with what is already known about the way gifted children think (Porath, 2006; Robinson & Clinkenbeard, 2008; Sekowski, Siekanska & Klinkosz, 2009). It is also consistent with other research about the learning preferences of gifted children (Kanevsky, 2011).

**Practical implications**

Based on this study, it becomes evident that to establish conceptual change, more than one lesson is required. Another implication of this study is that in educating gifted children, their own way of thinking should be considered, along with the preferences of gifted children. Thus, it is interesting to examine the differences in learning outcomes between direct instruction and inquiry learning when an inquiry learning task is used that is more open, abstract, and creative to match the characteristics of gifted children more. To get a clear image of the results of such a study, it is recommended to use gifted participants only, because gifted children have their own way of thinking (Porath, 2006; Robinson & Clinkenbeard, 2008; Sekowski, Siekanska & Klinkosz, 2009). In this study, a heterogeneous group of both talented and gifted participants was used for practical reasons, but by using only gifted students as
participants, a clearer image can be formed about the learning styles, preferences, and motivation of gifted children. Where this study initiates examining the preferences of gifted children, more thorough research about preferences and motivation is needed to get a clear image of those preferences and motivation.
References


Effective instruction for gifted children

*Children: Psycho-Educational Theory, Research and Best Practices* (pp. 177-198).

New York, NY: Springer


Reis, S.M. & Renzulli, J.S. (2009). Myth 1: The gifted and talented constitute one single homogeneous group and giftedness is a way of being that stays in the person over time and experiences. *Gifted Child Quarterly, 53*, 233-235


In L.V. Shavinina (Ed), *International Handbook on Giftedness* (pp. 467-485). New York, NY: Springer


