UNIVERSITY OF TWENTE

Faculty of Behavioral, Management and Social Sciences

Department of Instructional Technology

# The Ability of the Inquiry Skills Test to Predict Students' Performance on Hypothesis Generation

## Bachelor Thesis

Rebecca Kahmann

1st Supervisor: Dr. Hannie Gijlers
2nd Supervisor: Siswa van Riesen

25th of July 2016

# Abstract

Prior domain knowledge and inquiry skills tend to be two important factors for successful inquiry learning (van Joolingen et al., 2007; Zimmerman, 2007). The current study investigated to what extent the outcomes on a pre-test on domain knowledge and on the sub-test 'Setting up Hypotheses' of the Inquiry Skills Test (IST) can predict students' performance on hypothesis generation within an online learning environment. The sample consisted of 53 students (13-16 years old, pre-university level) who worked with an online learning environment on buoyancy and Archimedes' principle. The students' interactions with the environment were logged and the logged data on the students' hypotheses were qualitatively analyzed. The students' average score of correctly stated hypotheses per total stated hypotheses was used as indicator for the students' performance on hypothesis generation. The results show that no significant relation could be found between neither the prior knowledge and performance on hypothesis generation nor on the outcome on the sub-test 'Setting up Hypotheses' of IST and the performance on hypothesis generation. Typical difficulties that students generally have with the inquiry skill of hypothesis generation, like choosing the appropriate variables and stating a testable hypothesis (de Jong, 2006; Njoo & de Jong, 1993), could also be indicated through the analyses on the data from the log files during the current study. It may be that the missing correlation between the sub-test 'Setting up Hypotheses' and the performance during the learning environment could not be found due to the fact that the named difficulties are not tested within the IST and that the requirements of the learning environment are higher than the requirements of the sub-test 'Setting up Hypotheses'. The study concludes with recommendations on the assessment of inquiry skills.

*Keywords*: inquiry learning, prior knowledge, hypothesis generation, inquiry skills test

# Abstract

Domein voorkennis en onderzoeksvaardigheden worden vaak als belangrijke factoren voor onderzoekend leren genoemd (Klahr & Dunbar, 1988; van Joolingen & de Jong, 1997; van Joolingen et al., 2007; Zimmerman, 2007). De huidige studie onderzoekt in welke mate de scores op een domeinvoorkennistoets en scores op de subschaal 'Hypothesen opstellen' van de Inquiry Skills Test (IST) de prestatie van leerlingen op hypothesen generatie binnen een online leeromgeving kunnen voorspellen. De steekproef bestond uit 53 leerlingen (13 t/m 16 jaar oud, 3-vwo) die met een online leeromgeving over het principe van zweven, drijven, zinken en de wet van Archimedes bezig waren. De interacties van de studenten binnen de leeromgeving werden door middel van logfiles opgeslagen. De data met informatie over de opgestelde hypothesen van de studenten werden kwalitatief geanalyseerd. De gemiddelde score van adequaat opgestelde hypothesen in vergelijking tot de volledige aantal opgestelde hypothesen diende als indicator voor de prestatie van de studenten op hypothesen generatie. De resultaten tonen aan dat noch een significante correlatie tussen de sub-test 'Hypothesen opstellen' van de IST en de prestatie op hypothesen generatie werd gevonden noch een correlatie tussen voorkennis en de prestatie op hypothesen generatie. Verder werden vaak gevonden moeilijkheden bij het opstellen van hypothesen ook door de analysen over de informaties van de logfiles geïdentificeerd, zoals het kiezen van passende variabelen of het opstellen van een toetsbare hypothese (de Jong, 2006; Njoo & de Jong, 1993). Het is mogelijk dat de missende correlatie tussen de sub-test 'Hypothesen opstellen' kan worden verklaard door het feit dat de boven genoemde moeilijkheden bij het opstellen van moeilijkheden niet door de IST worden getoetst en daardoor dat de eisen van de leeromgeving hoger zijn dan de eisen van de subtest 'Hypothesen opstellen'. Aanbevelingen voor het testen van onderzoeksvaardigheden worden afsluitend genoemd.

*Sleutelwoorden*: onderzoekend leren, voorkennis, hypothesen generatie, inquiry skills test

# Introduction

During the last decades, the focus on what should be learned has shifted from knowledge about scientific facts to methods and skills that are supposed to support students during the process of learning in their later lives (van Joolingen, de Jong & Dimitrakopoulout, 2007; Kirschner et al., 2006; de Jong, 2006). Especially within science education, movements toward constructivist learning methods have become popular (Pedaste et al., 2015). Inquiry-based learning is a constructivist approach which enables students to construct their own knowledge through questioning and exploring specific problem scenarios almost independently (de Jong, 2006; de Jong & van Joolingen, 1998). The approach provides the opportunity for students to learn in a practical manner but it proves to be only effective when it entails sufficient guidance (de Jong, 2006; Kirschner et al., 2006). Kirschner et al. (2006) describe a series of studies reviewed by Clark (1989) in which students with low experience and/or experimental ability, belonging to the less-guidance condition, scored lower on the posttests than on pretests. One could interpret this in a way that less able and experienced students can attain more misconceptions than before. Despite this fact, less able and/or experienced students in the less guidance condition still reported to like the experience of experimentation (Kirschner et al., 2006).  This suggests that students need prior knowledge when making use of inquiry learning because otherwise the method often proves to be ineffective (de Jong, 2006; de Jong & van Joolingen, 1998; Lazonder, Wilhelm & Hagemans, 2008).

Van Joolingen et al. (2007) differ between two forms of prior knowledge. The first is prior domain knowledge which influences the student's decisions on the content of hypotheses and first concept models. Prior domain knowledge enables students to make use of the theory-driven approach to come to a conclusion which is more used within the scientific method than the data-driven approach (Klahr & Dunbar, 1988; Lazonder et al., 2008). The second form of prior knowledge is prior process knowledge which entails knowledge about what the inquiry learning processes are and how to make use of these processes. Process knowledge entails several forms of skills such as meta-cognitive skills, scientific thinking skills and inquiry skills. Also process knowledge has a positive effect on the outcomes of inquiry learning. According to Zimmerman (2007), domain knowledge and inquiry skills influence each other. Students with an adequate domain knowledge for instance, prove to be more able to make use of inquiry skills. This in turn, enables students to learn more from the inquiry tasks and provides the students with more detailed domain knowledge (Zimmerman,

2007). It is the goal of the present study to investigate to what extent the prior domain knowledge but also the inquiry skills of students' correlate with students' actual inquiry performance within a learning environment.

Research has shown that prior domain knowledge has a crucial influence on the successful performance of students within inquiry learning (de Jong, 2006; Wilhelm & Beishuizen, 2003; Zimmerman, 2007). Wilhelm and Beishuizen (2003) discovered that not only the knowledge about variables but also a primary understanding of relations between these variables is crucial for student's learning processes. Within their study, students benefited more from inquiry learning in the concrete task condition than students in the abstract task condition. Within the concrete task condition, variables and their relations were mostly known already from everyday life. In the abstract task condition, students could not make implications about relations of variables on the basis of the variables. It became obvious within the study that students within the abstract condition (therefore little prior knowledge about the relations of variables) stated less hypotheses, experimentation plans and conclusions than students in the concrete condition. In accordance with these findings, Zimmerman (2007) stated in a review on scientific thinking skills that students with an adequate level of prior domain knowledge tend to be more able to make use of scientific strategies and inquiry skills which in turn results in higher learning processes and outcomes. Within the current study, it is therefore the question whether the prior domain knowledge of the students has an impact on students' performance within the learning environment next to inquiry skills.

Stokking and van der Schaaf (1999) present two different views on how to interpret inquiry skills. One option is to see inquiry skills as essential cognitive capabilities to the inquiry process. In this case, the individual phases of the inquiry circle (s. Pedaste et al., 2015) cannot be used as an equivalent for the inquiry skills. The reason therefore is that one form of cognitive capacity could be evident in several inquiry phases while one inquiry phase could also entail varying skills. The other option is to interpret skills as significant tasks belonging to an all-embracing topic which is inquiry in this context (Stokking & van der Schaaf, 1999). According to this view, the inquiry phases can be recognized as inquiry skills as the phases are also tasks that need to be achieved for successful inquiry learning. This view agrees with the common school curricula and learning goals which are set in schools and with the feedback teachers give about students' investigations. It is therefore more oriented towards the experience of students and teachers. Within the current study, inquiry skills are

regarded as the completion of the phases of inquiry learning which corresponds to the latter view from Stokking and van der Schaaf (1999).

Studies have found that adequately generated hypotheses were positive indicators for students' further performance within the learning environment and learning process, suggesting that this process is a crucial inquiry skill for successful inquiry learning (Gijlers & de Jong, 2009; Njoo & de Jong, 1993). Hypothesis generation is a process of formulating a statement which entails variables and relations concerning the stated problem (Pedaste et al., 2015; de Jong, 2006; Gijlers & de Jong 2009). The purpose of hypothesis generation is to set up a fully-specified hypothesis which enables testing the hypothesis through experimentation to either come to evidence or disproof of the hypothesis or to generate ideas for new formulations of hypotheses (Klahr & Dunbar, 1988). A model that provides hypothesis generation a prominent role within inquiry learning is the Scientific Discovery as Dual Search model (SDDS) developed by Klahr and Dunbar (1988). SDDS includes the basic assumption that successful scientific reasoning consists of search within two problem spaces which are interconnected, namely the hypothesis space and the experiment space. While the focus of the hypothesis space is on generating hypotheses that could be observed on the basis of prior knowledge or experimental results, the experiment space includes all experiments that could be realized. Three main processes that support search within and between the two spaces are 'search hypothesis space', 'test hypothesis' and 'evaluate evidence'. The current study will concentrate on the hypothesis space and process of 'search hypothesis space' to indicate the inquiry skill of hypothesis generation.

Van Joolingen and de Jong (1997) further elaborated SDDS with amongst others the aim to provide a more detailed structure of the hypothesis space. On the basis of the general agreement among researchers that hypotheses consist of statements about a relation between various variables van Joolingen and de Jong (1997) motivated the division of the hypothesis space into the variable space and the relation space. The variable space includes different possible levels of generality which can be depicted in a generality hierarchy. Within this hierarchy, hypotheses about more general variables also apply to less general variables which are connected to the former variables. How general variables need to be depends on the stated problem. The relation space bears three different levels of precision: qualitative relational, quantitative relational and quantitative numerical. While the qualitative relational level only describes that a relation between variables exists, the quantitative relational level also describes the direction of the relation and the quantitative numerical level depicts a precise

relation that could be translated into a mathematical formular. While qualitative relations may facilitate to understand the domain, quantitative relations are more precise and therefore may help to understand the precise patterns within the domain (van Joolingen & de Jong, 1997). Within the current study, the essential aspects of hypothesis generation as described by Klahr and Dunbar (1988) and van Joolingen and de Jong (1997) are used to analyze students' performance on hypothesis generation within an online learning environment.

Students seem to have significant problems to generate hypotheses on their own and work with these effectively. Gijlers and de Jong (2009) found that students provided with pre-defined hypotheses were more successful in inquiry learning than students who had to formulate their own hypotheses. The study investigated the effects of various instructional supports during the hypothesis generation process on the success of collaborative inquiry learning. Therefore, three conditions were set up. In the first condition, the support tool was a shared proposition table in which the students were provided with pre-defined hypotheses and had to answer questions about the appropriateness of use of the hypotheses. If the dyads answered differently to the answers, this was indicated by the program to facilitate discussions about the provided hypotheses. Finally, students had to choose a hypothesis to test through experimentation. Within the second condition, students worked with a shared proposition scratchpad which supported students in formulating hypotheses by providing building blocks for the hypothesis like variables, conditions and relations. The third condition was the control condition which provided no extra support for the students to formulate a hypothesis. The findings of the study showed that students from the shared proposition table condition outperformed students from the other conditions with regard to learning outcomes. This might suggest that students are able to choose appropriate hypotheses but struggle to formulate hypotheses on their own. It is one goal of the current study to investigate which sub processes of hypothesis generation exactly students find difficult to be hampered to generate hypotheses independently or with little support.

When students formulate their own hypotheses, they are confronted with several challenges. One fundamental challenge for students with regard to hypothesis generation is the structure of a hypothesis. Students tend to struggle with the question of what is essential to a testable hypothesis and can confuse hypotheses with other statements like predictions or general expectations (Njoo & de Jong, 1993; de Jong, 2006). When students do not know what testable hypotheses consist of, this can lead students to not finding the crucial variables for the stated problem and to having trouble with indicating a relation between the variables

(De Jong & van Joolingen, 1998). Research has also found that students can be guided by the consideration that they should not formulate a hypothesis which is likely to be rejected through the results, also called the fear of rejection (De Jong & van Joolingen, 1998). These considerations can impair students to generate testable hypotheses because a crucial element of testable hypotheses is that they are falsifiable.  The current study intends to investigate which of the named challenges can be indicated through the analysis of students' performance within the used learning environment.

Inquiry skills like hypothesis generation can be detected through various measurements like scoring schemes for students' activities within learning environments, scoring schemes for students' comments who think aloud, or multiple choice tests. The main advantages of a multiple choice test are that it is time efficient and objective compared to other measurements (Burns, Okey, & Wise, 1985). Horstink (2006) designed a test called the Inquiry Skills Test (IST) to provide a measurement for domain independent inquiry skills of students. It is one goal of the present study to investigate to what extent the IST can indicate the performance of students' hypothesis generation within a domain-specific learning environment.

The IST was designed to provide an indication for students' inquiry skills but also to enable the evaluation of curricula and learning material concerning inquiry-based learning (Burns et al., 1985; P. Wilhelm, personal communication, March 3, 2016). The test as it exists today, has been changed several times to improve its validity. The foundation of this process was set by Horstink (2006) who assembled the second version of "The Integrated Process Skills Test" (TIPS II) and parts of the "Watson Glaser Kritisch Denken Test" (WGKDT) to implement the IST. The TIPS II and WGKDT were selected because they fit best to the criteria which Horstink set up. These criteria were an appropriate target group which meant high school students in this case, the duration which should not exceed one school lesson, the reliability and validity and whether all inquiry skills and therefore phases were incorporated. Horstink (2006) chose the inquiry phases defining variables, setting hypotheses, designing experiments and drawing conclusions as important for determining inquiry skills.

The TIPS II fulfilled all the criteria above so that Horstink decided to use all items from the TIPSII for the IST. The TIPS II is developed by Burns et al. (1985) to provide a further valid instrument for process skills next to its ancestor, the TIPS. The TIPS II contains 36 items which are based on various subject fields so that the test is independent from school

curricula and students' knowledge of specific subject areas. Therefore, the TIPS II measures "students' ability to apply the logic required to conduct fair investigations" (Burns et al., 1985). As the IST was intended to be in Dutch, Horstink (2006) mentioned that the TIPS II was translated one-to-one without further testing as the items were not culturally loaded. Still, one should be careful when taking over a test to another country as in this case, even though the items themselves are culturally mostly independent, the school structure in the USA and the Netherlands differ from each other in at least two aspects. First, in the USA all students go to the same form of secondary school with separate courses providing different degrees of complexity in for example science. In the Netherlands, however, there are three forms of secondary schools which vary in the focus of the skills being learned in the subjects. While the VMBO focuses mostly on teaching practical skills, the VWO focuses mostly on teaching scientific skills. The HAVO can be seen as a school form which forms a compromise between the other two school forms. Second, science often exists as one school subject in the USA, while in the Netherlands a distinction is made between physics, biology and chemistry. Therefore, even though the items of the TIPS II itself might not be culturally loaded, researchers should be careful with the scoring system due to the fact that the school system in the USA differs to some degree from the school system in the Netherlands which can have an impact on the results which students achieve in the test.

Altogether, the IST – which is used during the current study to measure the inquiry skill of hypothesis generation - incorporates six sub tests (P. Wilhelm, personal communication, March 3, 2016). The focus of the current study is on the sub-test 'Setting up Hypotheses'. The sub-test 'Setting up Hypotheses' will be used to measure the inquiry skill hypothesis generation. The other five sub-tests are 'Operational Defining', 'Design Experiments', 'Identifying Variables', 'Conclusion' and 'Interpretation'. These will not be discussed further within the current study. The six sub-tests correlate to the four phases set as criteria for an appropriate test by Horstink (2006). The sub-tests 'Identifying Variables' and 'Operational Defining' are included together in the phase Definition of Variables. Furthermore, the sub-tests 'Conclusion' and 'Interpretation' are incorporated in the phase Conclusion.

The internal reliability of the IST implemented by Horstink (2006) was satisfying with an $\alpha=0.69$ but it also showed that there was no significant internal consistency of the sub tests. Follow-up studies developed 16 further items to improve the internal consistency of the test with the result that the internal reliability rose up to $\alpha=0.91$ (P. Wilhelm, personal

communication, March 3, 2016). Furthermore, follow-up studies showed that the WGKT sub tests remained to depict a low internal consistence and did not have a great impact on the reliability of the IST. Therefore, the decision was made to replace these items through newly developed items which fit better to the style of the TIPS II (P. Wilhelm, personal communication, March 3, 2016). The standardization of the IST is in progress, currently. The IST 2.1 is a test which measures domain independent inquiry skills on high school students and undergraduates and consists of 42 multiple-choice items. It indicates a first rating of students' inquiry skills and can also be used to evaluate educational methods (P. Wilhelm, personal communication, March 3, 2016).

Jeckmans (2014) could support Horstink's findings on the satisfying intern reliability of the IST and served further findings on the validity of the test and especially its sub-tests. With regard to the validity of the sub-test 'Setting up Hypotheses' it was found that the scores on the quality of stated hypotheses did not correlate significantly with the sub-test 'Setting up Hypotheses' but instead with the sub-tests 'Conclusion' and 'Design Experiments'. Jeckmans (2014) interpreted this result in the sense that the ability to formulate adequate hypothesis facilitates a student to design an appropriate experiment and conclusion. This does however not clarify the missing correlation between the sub-test 'Setting up Hypotheses' and the quality of the formulated hypotheses during the task. The current study aims to further investigate this issue by reinvestigating the correlation between the score on the sub-test and the students' performance on hypothesis generation during an inquiry learning task and by searching for factors that might have an influence on the correlation.

Until now, the IST has been widely used with learning environments which were independent from school curricula and therefore it can be questioned to what extent the IST can predict students' performance within a domain specific topic that is affiliated to school curricula. Within the present study, an online learning environment was developed which encourages students to do investigations about the buoyancy and Archimedes' principle. The buoyancy is a regular part of Dutch curricula and some schools use Archimedes' principle as additional learning content. It is one goal of this study to investigate whether the outcomes on the sub-test of the IST 'Setting up Hypotheses' correlates significantly with the students' performance in this domain-specific learning environment during the phase of hypothesis generation.  As already stated, research suggests that students need to have prior domain knowledge for inquiry learning to be effective and that prior domain knowledge can improve students' performance on hypothesis generation (de Jong, 2006; de Jong & van Joolingen,

1998; Klahr & Dunbar, 1988). But also process knowledge (of which inquiry skills are a component) supports students' performance during inquiry which in turn enables them to learn more from the tasks (van Joolingen & de Jong, 1997; van Joolingen et al., 2007). Most studies concentrate on one of the two concepts, either domain knowledge or inquiry skills. The present study, wants to put both together and see to what extent the inquiry skill of hypothesis generation and prior domain knowledge have an influence on students' performance on hypothesis generation in the learning environment. Furthermore, it will be investigated which difficulties students have when setting up hypotheses, even if they have sufficient inquiry skills and prior domain knowledge. Altogether, the current study will deal with the following research questions:

1. Is there a significant relationship between the prior process knowledge of the inquiry skill hypothesis generation measured with the IST and students' actual performance of hypothesis generation within a domain specific online learning environment?

2. Is there a significant relationship between prior domain knowledge and students' performance of hypothesis generation within a domain specific online learning environment?

3. Which difficulties with hypothesis generation can be recognized on the basis of analysis of students' performance on hypothesis generation?

# Methods

## Participants

The sample consisted of 53 ninth-graders of a Dutch secondary school who obtained a pre-university level of education. The participants were from two classes, 21 students were part of one class and 24 were part of the other class. The average age of the participants was $M=14.62$ (min=13, max=16). The parents of the participants received a letter with information about the study and consented to the participation of their children.

## Materials

### Inquiry Skills Test

In the current study, the IST 2.1 was used which was the most recent version of the IST at the moment the study was conducted. The IST 2.1 consisted of 42 items of which most were taken from the Dutch translation of the second version of the Test of Integrated Process Skills (TIPS II). The test-items were stated as multiple choice questions with four choice options of which one was correct. The stated problems within the items were relatively simple

and affiliated to everyday life topics like which factors could influence how fast salt dissolves in water. The difference and simplicity of the contents within the test were chosen to focus on the inquiry skills and to reduce the influence of prior domain knowledge (Zimmerman, 2007). The test was taken in a paper-pencil version. As it has been discussed in the introduction, the IST consists of six sub-test, namely 'Identifying Variables', 'Setting up Hypotheses', 'Operational Defining', 'Design Experiments', 'Conclusion' and 'Interpretation'. Each sub-test contained of seven items. The present study was focused on the sub-test Setting up Hypotheses. The IST consists of six sub-tests, namely 'Identifying Variables', 'Setting up Hypotheses', 'Operational Defining', 'Design Experiments', 'Conclusion' and 'Interpretation'. Each sub-test contained of seven items. The present study was focused on the sub-test 'Setting up Hypotheses'. 'Setting up Hypotheses' was chosen to measure the the inquiry skill of hypothesis generation. It could be argued that 'Identifying Variables' could also be taken into consideration for identifying students' skills in hypothesis generation. However, identfying variables is often also seen as a sub-process of an orientation phase in which the inquiry task is introduced and the prior knwoledge about the topic is activated (Pedaste et al., 2015). Horstink (2006) also regarded the sub-test 'Identifying Variables' as a separate inquiry skill occuring before 'Setting up Hypotheses'. Furthermore, hypothesis generation it is not only about choosing the right variables but more about integrating these variables into a fully specified hypothesis which facilitates testing the hypothesis (Klahr & Dunbar, 1988). Including the sub-test 'Identifying Variables' could provide the sub-process of choosing variables a more prominent role than other aspects of hypothesis generation. Because of these arguments it was chosen not to integrate the sub-test 'Identfying Variables' in the current study.

For being successful on the sub-test 'Setting up Hypotheses', a student needed to judge which of the four given hypotheses fits best to the given task. Skills that students needed to possess for this judgement were to detect adequate independent and dependent variables for the stated problem and in one case to detect how concrete the hypothesis was stated. Further insights about the testability of the hypothesis and the precision of the relation were helpful but not necessary to choose for the right answer, even though these concepts are are pointed out as necessary aspects for an adequate hypothesis by Klahr and Dunbar (1988) and van Joolingen and de Jong (1997). An example of an item of the IST is shown in figure 1.

Susan is studying food production in bean plants. She measures food production by the amount of starch produced. She notes that she can change the amount of light, the amount of carbon dioxide, and the amount of water that plants receive. What is a testable hypothesis that Susan could study in this investigation?

    A. The more carbon dioxide a bean plant gets the starch a bean plant produces.

    B. The more starch a bean plant produces the more light it needs.

    C. The more water a bean plant gets the more carbon dioxide it needs.

    D. The more light a bean plant receives the more carbon dioxide it will produce.

*Figure 1.* A sample item of the IST taken from the TIPSII as English translation (Okey et al., 1985).

The item from figure 1 belongs to the sub-test 'Setting up Hypotheses'. To choose the right answer A, the student needs to understand that the purpose of Susan's research is to get to know more about food production in bean plants. Therefore, the amount of starch needs to be the dependent variable within Susan's study. The amount of light, carbon dioxide and water on the other are all described as possible factors and therefore independent variables. The student therefore needs to search for a hypothesis with a relation between one of the three factors and the amount of produced starch.

**Pre-test on domain knowledge**

The pre-test measured the student's knowledge on buoyancy and Archimedes' principle and consisted of four research-questions which included several sub-questions. The sub-questions were short-answer items and therefore required students to fill in either a word or a number or short explanations for a phenomenon or a technical term. A student could achieve up to 34 points per test. To gain a positive result in the tests, the student needed to recall and apply basic knowledge of floatation and Archimedes' principle. One example of a question is:

"Indicate the volume, mass and density of a ball which sinks in a bucket of water, and explain why this ball sinks."

To answer this question correctly, the student needed to indicate appropriate values for mass, volume and density which could be chosen freely and to state that the ball sinks because the density of the ball is higher than the density of water. The student could gain three points for this task, one point for indicating the appropriate values and two points for the explanation.

**Learning environment**

The learning environment enabled the student to learn about floatation and Archimedes' principle by doing online experiments about these topics. The structure of the learning environment was similar for both topics. First, an introduction to the topic was given in terms of a video. For the topic of floatation, the video showed how different liquids and objects can float on each other. For Archimedes' principle, the story was shown of Archimedes' task to find out whether the king's crown would be of pure gold without destroying the crown at some point.

After the introduction, the student had to do three experiments on both topics. The structure for each experiment was identical. First, the student set up hypotheses with the support of a guidance tool. After that, the student could test the hypotheses through experimentation and evaluate the hypotheses in the discussion phase. These processes supported the student to decide to either reformulate the hypotheses or to use the gathered knowledge for the hypothesis generation of the following experiments.

At the beginning of each experiment, the student needed to set up one hypothesis or more with the help of a guidance tool named hypothesis scratchpad on the basis of a stated problem. Figure 2 shows the setup of the hypothesis scratchpad. The little letters in the figure are supposed to help referring to the elements of the hypothesis scratchpad. A hypothesis could be created by dragging the sentence elements (a) to the hypothesis box (b). A further hypothesis could be added by clicking on the plus sign (c). When clicking on the button with the shape of a question mark (d), the student could find instructions about how to use the hypothesis scratchpad and the meaning of the symbols depicted in the app.  A student could depict his or her level of confidence about a hypothesis by changing the portion of blue in the 'horse shoe'(e).
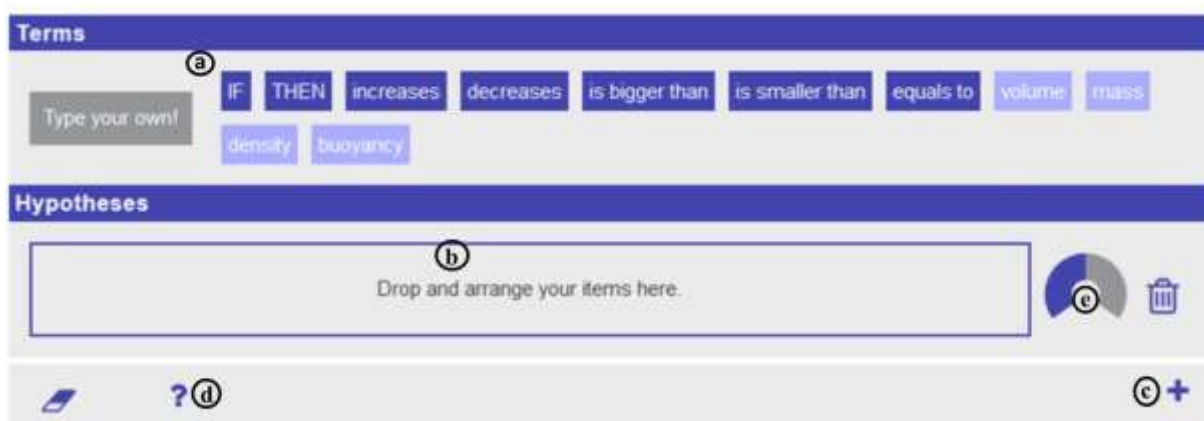


*Figure 2*. The layout of the hypothesis scratchpad within the experiments.

The different kinds of sentence elements supported students to show the necessary sub-processes of hypothesis generation. The sentence elements consisting of variables supported the student to choose for the appropriate independent and dependent variables. Furthermore, the words 'IF' and 'THEN' provided students with the hint to provide a hypothesis that consists of a relation between the variables. Sentence elements like 'increases' or 'is smaller than' facilitated to set up a relation with at least a quantitative relational level of precision. The testability of the hypothesis was a sub-process that could not be supported by the tool directly because it depended on how the student formulated the hypothesis. The student could decide independently to what extent the support of the hypothesis scratchpad was needed and even though support was given it was still up to the student to actually formulate a fully-specified hypothesis. If the student did not find the sentence elements as suitable as other formulations, it was possible to choose for different words by using the button 'type your own!' and then writing the favored formulation (a).

When being finished with the hypothesis, the student could test this hypothesis by constructing some experiments.  Therefore, a guidance tool called the Experiment Design Tool supported the student to indicate which variable should be varied, which variables should be kept constant and which variable should be measured. Furthermore, the values of the variables could be chosen in the Experiment Design Tool and also the expectation of the result could be stated. When all the necessary data had been filled in in the Experiment Design Tool, the student could take a look at the results. If the student felt confident enough that the stated hypothesis had been supported or disproved by the experiments, the student was encouraged to write his or her conclusion about the hypothesis. To further reflect on the stated hypotheses, a few questions were stated, like whether the student would change his or her level of confidence about the hypotheses after having done the experiments. For floatation as well as for Archimedes' principle three experiments were designed for the student which makes six experiments the student could work through within the learning environment.

## Procedure

The present study was conducted over four school lessons which were conducted a few days apart from each other. Each lesson had the extent of 45 minutes. The first lesson was filled with the introduction about the study and the study's goal. Furthermore, the pre-test was administered. The pre-test was used to define the base-line of the students' knowledge with regard to the topic. The students could choose their seats within the classroom freely. All seats were directed to the blackboard of the room and had an arm's length distance from each

other so that copying was prevented as far as possible. In case the students had questions about the tasks of the test, these could be formulated to the researcher who either gave some qualification about the formulation of a task or stated that the answer of the question would give too many clues about the task. Most of the students were finished with the pre-test after 20 to 30 minutes so that the lesson was ended when all students had handed in their pre-test by the researcher instead of after 45 minutes.

During the second lesson, the students filled in the IST. This questionnaire was used to measure the baseline of inquiry skills that the students had. The procedure with regard to questions about the test and seating arrangements was the same as for the pre-test. A few students were not able to finish the IST during the second lesson. Those students were able to finish it at the beginning of the following lesson.

In lesson three, first a short introduction about the learning environment was given, then the students worked with the learning environment individually. In this lesson, the students needed to work on the topic of buoyancy in three experiments. Each experiment included four phases, namely setting up hypotheses, designing experiments, making conclusions and reflecting on the experiment on the basis of discussion questions. In case of questions about the environment or struggles with the environment which had nothing to do with the content of the experiments, the researcher tried to support the students.

During lesson four, the students also worked with the learning environment individually like in lesson three, only that the topic was Archimedes' principle this time. When there were difficulties with the learning environment, the researcher also tried to support the students, just like in lesson three. For both lesson three and four, a computer with internet connection was needed for all students as the learning environment was online accessible.

## Coding and Scoring

The students' responses on the 42 items of the IST were scored as true or false. The score on the IST and the sub-tests told how much items a student answered correctly. The score on the pre-test was calculated per research question in the learning environment. It was investigated whether the answer was right or whether some elements were missing. If an element or a relation between elements had not been described sufficiently, half a point or more were detracted. For the first task of the pre-test, students could gain up to four points,

for the second and third task, up to nine points and for the fourth task they could gain up to twelve points. In total, this made a maximum score of 34 points on the pre-test.

All activities the students performed within the learning environment were stored through log files. From the log files the information about the hypotheses from all experiments within the learning environment was used to enable analyzing the students' performance on hypothesis generation. With the filtered information from the log files, the researcher could indicate which student set up how many hypotheses, to which experiment each hypothesis belonged and what the content of the hypothesis was.

To analyze the quality of the generated hypotheses, a coding scheme was created which included five criteria for an adequate hypothesis (s. appendix). The criteria of the coding scheme were based on the sub-skills needed for successful hypothesis generation within the learning environment as described in the section Materials and on the essential aspects of hypothesis generation presented by Klahr & Dunbar (1988) and van Joolingen and de Jong (1997). The first criterion for the hypotheses was that premises of a hypothesis were fulfilled. These premises were that the hypothesis was stated in the form of a sentence and included an independent and a dependent variable. The second criterion entailed that the hypothesis included an expectation (2) which was applicable to the domain of the experiment in quest (2.1), which was testable (2.2) and which entailed a relation between the variables (2.3) with a direction of effect (2.4). Criterion three required that the independent variable was adequately chosen on the basis of the stated problem and that if more than one variable was stated the interrelation between the independent variables was described. If the dependent variable was adjusted to the domain of the experiment in quest, criterion four was fulfilled. The last criterion was the uniqueness of the hypothesis in the sense that the hypothesis was not stated by the same student already. All criteria could be answered either with a 'yes' or a 'no'. If a criterion was fulfilled, it was answered with 'yes' which was coded with one point. For a 'no', no point was given. The researcher spoke of a hypothesis if one point was scored on the premises as well as on the uniqueness of the hypothesis. The hypothesis was seen as adequate when all criteria could be answered with a 'yes'. If one or more criteria had been answered with a 'no', the hypothesis was seen as not adequate. The resulting score of performance on hypothesis generation was received through calculating the percentage of adequate hypotheses per student.

# Results

During the present study, 15 of the 53 students either missed at least one of the sessions of the study or did not set up hypotheses within the learning environment. These 15 participants were therefore excluded from further analyses so that the remaining sample consisted of 38 participants.

The mean score on the pre-test of domain knowledge was considerably high with 24.5 points (s=5.34, min=9.0, max=31.0). The participants also scored considerably high on the IST (M=31.94, s=3.18, min=27.0, max=38.0) and on the sub-test 'Setting up Hypotheses' (M=5.16, s=1.18, min=2.0, max=7.0). However, the mean average of adequate hypotheses within the learning environment was lower than expected (M=27.84, s=33.24, min=0.0, max=100.0). The high standard deviation of this score may occur due to the fact that with 19 students the half of the participants did not set up any adequate hypothesis at all. The mean number of set up hypotheses was 3.53 (s=2.79, min=1, max=12) while with 26 students 68.4% of the students set up one to three hypotheses. Due to the low scores on the performance of hypothesis generation compared to the other scores, it is not a surprise that the correlation between the score on the sub-test 'Setting up Hypotheses' and the percentage of adequate hypotheses was neither strong nor significant ($r_s$= 0.19, p=0.25). This result disagrees with the research question which asks for a significant relation between the inquiry skill of hypothesis generation and the performance of hypothesis generation within the learning environment. The correlation between the score on the pre-test and the percentage of adequate hypotheses was also not significant ($r_s$= 0.28, p=0.09). Therefore, the result objects the research question that states the expectation of a significant correlation between prior domain knowledge and the performance of hypothesis generation in the learning environment. The results suggest that prior domain knowledge and the process knowledge of the inquiry skill hypothesis generation have no significant influence on the performance of hypothesis generation within a domain specific learning environment.

On basis of 150 statements that fulfilled criterion one and five from the scoring scheme and therefore were coded as hypotheses, it was investigated which criteria of the coding scheme were the most difficult to the students. Therefore, it was calculated per criterion how many of the 150 cases did not achieve the criterion in quest and the result was translated into a percentage score. The results suggested that testability (56.0%), setting up an adequate independent variable (31.33%) and the applicability to the domain (20.67%) were the three most common deficits that appeared when students set up hypotheses during the

current study. Table 1 shows per criterion the absolute amount of cases that did not fulfil the criterion and the percentages of not fulfilled cases in relation to the total cases that were scored on the basis of the criteria.

Table 1

*An overview of cases that did not fulfill the separate criteria and the percentages of cases of nonfulfillment*

| Criteria for adequate hypotheses | cases of nonfulfillment per criterion | Total scored cases per criterion | % |
|---|---|---|---|
| Expectation | 1 | 150 | 0.67 |
| Applicability to domain | 31 | 150 | 20.67 |
| Testability | 84 | 150 | 56.0 |
| Effect | 10 | 150 | 6.67 |
| Direction of effect | 15 | 150 | 10.0 |
| Adequate independent variable | 25 | 150 | 16.67 |
| Relation between independent variables | 8 | 150 | 5.33 |
| Adequate dependent variable | 47 | 150 | 31.33 |

All in all, the results suggest that against the expectations of the research questions there is no significant relation between the inquiry skill of hypothesis generation and the performance of hypothesis generation in a domain specific learning environment as well as between prior domain knowledge and the performance of hypothesis generation. Furthermore, it could be found that the most common difficulties that students have, have to do with the testability of a hypothesis, the applicability of the hypothesis to the domain in quest and stating the adequate dependent variable.

## Discussion

The current study investigated whether there is a significant relation between the inquiry skill of hypothesis generation measured by the IST and the performance of hypothesis

generation within a domain specific learning environment. As the results show, no significant correlation could be found. This outcome supports Jeckmans' study (2014) who also found that there was no significant correlation between the sub-test 'Setting up Hypotheses' and the performance on hypothesis generation within an inquiry task. This can be a sign for a low criterion validity of the sub-test 'Setting up Hypotheses' from the IST. The content of the sub-test's items suggests that the IST measures more the students' ability to judge which hypothesis fits best to the stated problem of the item than the inquiry skill of hypothesis generation. Students do not need to know what makes a hypothesis adequate explicitly but rather choose for the hypothesis they think is the most effective for the task. How students come to their judgement cannot be told with the information of the IST.

The investigation of the difficulties with hypothesis generation that occurred the most during the present study showed that students were especially challenged with generating a testable hypothesis, a hypothesis which was connected to the domain and/or a hypothesis with adequate independent and dependent variables. These findings support De Jong (2006) who reported that students have difficulties to determine appropriate variables and construct testable hypotheses. An essential problem is that some students cannot explicitly express what important elements of a hypothesis are which makes it hard to generate a testable hypothesis with the right variables (Njoo & de Jong, 1993). With the items of the IST it is not possible to detect to what extent students are able to handle these and other aspects of hypothesis generation. Most of the pre-defined hypotheses out of which the students need to choose within the sub-test are stated in a testable way and also depict a relation with a direction between the variables. Most of the time the students only need to look for the variables to be rightly chosen. In the learning environment of the current study on the other side, students needed to generate hypotheses on their own with the support of the hypothesis scratchpad. It can be suggested that the skills that students needed within the learning environment were more complex than the skills required in the sub-test 'Setting up Hypotheses' of the IST which might clarify the missing correlation between the two scores.

The findings of the current study suggest that another content of the sub-test 'Setting up Hypotheses' items or another format of the sub-test can increase the ability of the IST to predict the degree of students' inquiry skill of hypothesis generation. One possibility to test hypothesis generation in another format would be to change the multiple choice format into an open question format so that students would need to state their own hypothesis. An advantage of the open question format is that the requirement of the test would be close to the

requirements stated to a student within the phase hypothesis generation in an inquiry learning environment. This suggests that the test's possibility to indicate students' inquiry skill of hypothesis generation would increase. However, the new format of the items would make it hard to compare this sub-test with the other five sub-tests of the IST and also to integrate in the test. Another option is to state other multiple choice items which go deeper into the process of hypothesis generation than the current items. From the current study but also earlier research it is known that students have difficulties with setting up testable hypotheses and hypotheses with the accurate variables that fit to the stated problem. A possibility to test students' inquiry skills with a multiple choice item could be to look for their knowledge on the essential aspects of hypothesis generation. An item could for example provide a hypothesis with the question whether the hypothesis is appropriate to investigated on the stated problem or not and if not which point misses for the hypothesis to be appropriate. The advantage of this reconstruction would be that it could be integrated into the current IST. However, Gijlers & de Jong (2009) provide findings that the process on inquiry learning is higher when students are provided with pre-defined hypotheses than when being actively involved in the process of hypothesis generation. One possible clarification for this difference is that students struggle with the more complex requirements when generating hypotheses almost independently. This may suggest that the format of pre-defined answers in the IST would still hamper the IST from predicting performance on mostly independent hypothesis generation. Further research could investigate whether the change of the item's content actually increases the validity of the test or whether it only leads to measuring the students' ability to reflect on the appropriateness of a hypothesis.

The current study also investigated whether there is a significant relation between prior domain knowledge and the performance of hypothesis generation. Against the expectations of the researcher no significant correlation could be found. The findings of the present study contradict Wilhelm & Beishuizen (2003) who found in their study that students perform significantly higher during an inquiry task if they own prior knowledge about variables and relations between the variables. In the current pre-test on domain knowledge most students could indicate securely the important variables and relationships especially with regard to the topic of buoyancy. When relying on Wilhelm & Beishuizen (2003), the students should have had at least a few difficulties to set up adequate hypotheses for the three experiments on buoyancy. However, the participants of the current study were not able to make use of their domain knowledge to state relevant and testable hypotheses. It is possible

that little training with the learning environment and missing mastery of the inquiry skill hypothesis generation impaired students' ability to make use of their knowledge about the involved concepts and the relations between them and to state an appropriate hypothesis. It was the greatest challenge for students to set up a hypothesis which was testable. It is possible that students did use their prior knowledge to set up a hypothesis which is accurate in the relation between the required variables but that they failed to make the hypothesis sufficiently concrete and therefore testable.

One element of the study could have facilitated the missing correlation between the outcome on the sub-test 'Setting up Hypotheses' and the performance on hypothesis generation but also the correlation between the prior domain knowledge and the performance on hypothesis generation. The independent variables handled in the experiments were highly correlated which made it more complicated for the students to set up hypotheses and conduct experiments. It could be observed that most of the hypotheses were about the first three experiments. The domain of these experiments, buoyancy, was already familiar to the students through earlier lessons. Despite this fact, students had problems to set up adequate hypotheses for these experiments. The interrelation between factors like the density of an object and the density of the liquid which to together have an effect on the buoyancy could have confused students. This could have lead students to struggle with setting up adequate hypotheses. When adding a further learning content like water displacement, this could have resulted in students to give up setting up hypotheses and instead solely conducting experiments without further ideas of what to investigate exactly.

A further issue of the study was that one of the two classes had problems with the internet connection in the sense that the internet connection was not strong enough to let more than twenty students work with the internet, simultaneously. After two hardly productive lessons with the online learning environment, it was found that the internet connection worked better with other laptops so that the students got a third lesson to work with the learning environment. The problems with the internet connection for one class probably have had some impact on the performance scores and number of hypotheses that were set up. However, it is not clear how influential exactly this incident was and in which manner it had an influence on students' learning process and further work with the learning environment. Therefore, this incident was not taken into further consideration within the current study and the data from the class with the internet issues remained included for data analysis.

Research has provided evidence for the reliability and validity of the IST in its entity (Horstink, 2006; Jeckmans, 2014; P. Wilhelm, personal communication, March 3, 2016). However, the current study showed that despite the high reliability and construct validity of the complete test, there is a point of the IST which could be improved, namely the validity of the sub-test 'Setting up Hypotheses'. It has been found that the items of the sub-test do not measure what is required by students in the learning environment. Therefore, further studies could investigate whether the validity of the sub-test increases when other items are added that look more deeply at the essential aspects of hypothesis generation (such as testability or the appropriateness of variables). It could be an enrichment to the test if the items of the sub-test 'Setting up Hypothesis' gained more information value.

The current study has shown that even though evidence indicates the IST to be a valid instrument for measuring inquiry skills it has certain shortcomings with regard to the sub-test 'Setting up Hypothesis'. Despite the fact that students scored high on the sub-test, there was evidence that they had difficulties with essential aspects of hypothesis generation, such as choosing for the appropriate variables and developing a hypothesis which is testable. The coding scheme of the current study enabled the researcher to indicate these challenges. This knowledge can be used by developers and teachers to indicate at which point of the inquiry learning process students require support from the environment but also from the teacher. Furthermore, teachers and developers of inquiry learning environments can use this study as a guide for what to look for when trying to indicate students' performance on inquiry skills. When only the general process knowledge of inquiry skills is important to know, a multiple choice test like the IST is sufficient and appropriate to use. However, when a teacher wants to evaluate how students perform on the single phases of inquiry learning and which challenges they are confronted with, it is more advisable to work with scoring schemes which look at the essential aspects of single phases. Admittedly, using a scoring scheme could be more time consuming and also less objective than a standardized test but it does serve more precise information about the students' learning progress. Therefore, environment developers and teachers are advised to reflect which insights they want to gain with testing students' inquiry skills and on the basis of this reflection choose for the measurement which facilitates the achievement of the needed insights the most.

# References

Burns, J.C., Okey, J.R. & Wise, K.C. (1985). Development of an integrated process skill test: TIPSII. *Journal of Research in Science Teaching*, 22, 169-177. doi: 10.1002/tea.3660220208

Gijlers, H. & de Jong, T. (2009). Sharing and confronting propositions in collaborative inquiry learning. *Cognition and Instruction*, 27, 239-268. doi:10.1080/07370000903014352

Horstink, M. (2006). *Constructie en validatie van een test voor het meten van inquiry skills.* [Construction and validation of a test for the measurement of inquiry skills.] Bachelorthesis, University of Twente. Retrieved from http://essay.utwente.nl/55431/

Jeckmans, M. M. J. F. (2014). *Procesvalidatie van de Inquiry Skills Test.* Bachelorthesis, University of Twente. Retrieved from http://essay.utwente.nl/65949/

de Jong, T. (2006). Technological Advances in Inquiry Learning. *Science*, 312, 532-533. doi: 10.1126/science.1127750

de Jong, T. & van Joolingen, W. R. (1998). Scientific discovery learning with computer simulations of conceptual domains. *Review of Educational Research*, 68, 179-201. doi: 10.3102/00346543068002179

van Joolingen, W. R. & de Jong, T. (1997). An extended dual search space model of scientific discovery learning. *Instructional Science*, 25, 307-346. doi: 10.1023/A:1002993406499

van Joolingen, W. R., de Jong, T. & Dimitrakopoulou, A. (2007). Issues in computer supported inquiry learning in science. *Journal of Computer Assisted Learning*, 23, 111-119. doi: 10.1111/j.1365-2729.2006.00216.x

Kirschner, P. A., Sweller, J., & Clark, R. E. (2006). Why minimal guidance during instruction does not work: An analysis of the failure of constructivist, discovery, problem-based, experiential, and inquiry-based teaching. *Educational Psychologist*, 41, 75-86. doi: 10.1207/s15326985ep4102_1

Klahr, D., & Dunbar, K. (1988). Dual space search during scientific reasoning. *Cognitive science*, 12, 1-48. doi: 10.1207/s15516709cog1201_1

Lazonder, A. W., Wilhelm, P., & Hagemans, M. G. (2008). The influence of domain knowledge on strategy use during simulation-based inquiry learning. *Learning and Instruction*, 18, 580-592. doi: 10.1016/j.learninstruc.2007.12.001

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*, 821-844. doi: 10.1002/tea.3660300803

Pedaste, M., Mäeots, M., Siiman, L. A., de Jong, T., van Riesen, S. A. N., Kamp, E. T., Manoli, C.C., Zacharias C. Zacharia, Z.C. & Tsourlidaki, E. (2015). Phases of inquiry-based learning: Definitions and the inquiry cycle. *Educational research review*, 14, 47-61. doi: 10.1016/j.edurev.2015.02.003

Stokking, K. M., & van der Schaaf, M. F. (1999). *Beoordelen van onderzoeksvaardigheden van leerlingen: Richtlijnen, alternatieven en achtergronden*. Utrecht: Universiteit Utrecht, Onderwijskunde, ISOR.

Wilhelm, P., & Beishuizen, J. J. (2003). Content effects in self-directed inductive learning. *Learning and Instruction*, 13, 381-402. doi: 10.1016/S0959-4752(02)00013-0

Zimmerman, C. (2007). The development of scientific thinking skills in elementary and middle school. *Developmental Review,* 27, 172-223. doi:10.1016/j.dr.2006.12.001

# Appendix

## Coding Scheme

| Construct | Stelling | Codes |
|---|---|---|
| 1. Voorwaarden voor hypothese | Er wordt een hypothese gesteld die voldoet aan de voorwaarden van een hypothese (bevat ten minste 1 afhankelijke, en 1 onafhankelijke variabele) | • 0=nee<br>• 1=ja |
| 2. Verwachting | Er wordt een verwachting uitgesproken | • 0=nee<br>• 1=ja |
| 2.1 toepasbaarheid op onderzoeksvraag | De verwachting bevat elementen uit de onderzoeksvraag Onderzoeksvraag en mogelijke elementen:<br>**onderzoek 1**:<br>Welke eigenschappen van het object hebben invloed op het drijfvermogen op het object als je het in een bak water geeft?<br>Eigenschappen vd bal (m, V, p)<br>Dichtheid van water<br>Drijfvermogen van de bal<br>**Onderzoek 2:**<br>Hoe bepalen de massa en volume van een object de dichtheid van de bal?<br>Massa vd bal<br>volume vd bal<br>dichtheid vd bal<br>**Onderzoek 3:**<br>In hoeverre hebben de eigenschappen van een vloeistof invloed op het drijfvermogen van een object dat in een bak met de vloeistof wordt gegeven?<br>Eigenschappen vd bal<br>Eigenschappen van vloeistof<br>Drijfvermogen van de bal<br>**Onderzoek 4:**<br>Hoe wordt de hoeveelheid waterverplaatsing (zowel massa als volume) beïnvloed door de drie eigenschappen massa, volume en dichtheid van een bal die *drijft*?<br>Bal drijft<br>eigenschappen van de bal<br>Hoeveelheid waterverplaatsing (in massa en/of volume)<br>**Onderzoek 5:**<br>Hoe beïnvloeden de massa, het volume en de dichtheid van een bal die *zweeft* de hoeveelheid massa en volume van de waterverplaatsing?<br>Bal zweeft | • 0=nee<br>• 1=ja |

| | | |
|---|---|---|
| | Eigenschappen van de bal<br>Hoeveelheid waterverplaatsing (in massa en/of volume)<br>**Onderzoek 6:**<br>Hoe beïnvloeden de massa, het volume en de dichtheid van een bal die *zinkt* de hoeveelheid massa en volume van de waterverplaatsing?<br>Bal zinkt<br>Eigenschappen van de bal<br>Hoeveelheid waterverplaatsing (in massa en/of volume) | |
| 2.2 toetsbaarheid van verwachting | De verwachting is toetsbaar met de omgeving/het lab (de benodigde materialen zijn aanwezig in het lab, de meeteenheden worden weergegeven in het lab) | • 0=nee<br>• 1=ja |
| 2.3 relatie | De verwachting geeft een relatie tussen elementen aan (correlationeel of causaal verband) | • 0=nee<br>• 1=ja |
| 2.4 richting | De verwachting bevat een richting (positief of negatief verband) | • 0=nee<br>• 1=ja |
| 3. passende onafhankelijke variabele | De onafhankelijke variabele bestaat uit een of meerdere elementen van de volgende<br>**Onderzoek 1**<br>Massa, volume of dichtheid van de bal (evtl. In verband met eigenschappen van water)<br>**Onderzoek 2**<br>Massa of volume van de bal<br>**Onderzoek 3**<br>Massa, volume of dichtheid van de vloeistof (evtl. In verband met eigenschappen van de bal)<br>**Onderzoek 4**<br>Massa, volume of dichtheid van een drijvende bal<br>**Onderzoek 5**<br>Massa, volume of dichtheid van een zwevende bal<br>**Onderzoek 6**<br>Massa, volume of dichtheid van een zinkende bal | • 0=nee<br>• 1=ja |
| 3.1 verband tussen onafhankelijke variabelen | Als er meer dan 1 variabele is, wordt er een verband tussen de onafhankelijke variabelen weergegeven i.p.v. dat een opsomming van elementen wordt weergegeven | • 0=nee<br>• 1=ja/ niet van toepassing |
| 4. Passende afhankelijke variabele | De afhankelijke variabele sluit aan bij de onderzoeksvraag; per onderzoek is de bij de onderzoeksvraag passende afhankelijke variabele<br>**Onderzoek 1**<br>Drijfvermogen van de bal<br>**Onderzoek 2**<br>Dichtheid van de bal | • 0= nee<br>• 1= ja |

| | **Onderzoek 3**<br>Drijfvermogen van de bal<br>**Onderzoek 4**<br>Hoeveelheid waterverplaatsing (in massa of volume)<br>**Onderzoek 5**<br>Hoeveelheid waterverplaatsing (in massa of volume)<br>**Onderzoek 6**<br>Hoeveelheid waterverplaatsing (in massa of volume) | |
|---|---|---|
| 5.  Unieke<br>    hypothese | Eerste hypothese die zo of op een soortgelijke manier is gesteld | 4.  0=nee<br>5.  1= ja |

**Scoring**

*Geen hypothese*= stelling 1 en/of 5 is met nee beantwoord

*Hypothese*= tenminste 1 en 5 zijn met ja beantwoord

*Adequate hypothese*= alle stellingen zijn met ja beantwoord

Per student wordt het percentage van adequaat ontworpen hypothesen (aantal adequate hypothese/ aantal hypothesen*100)