Direct Instruction versus Guided Inquiry-Based Learning in a Physics Practical

Maarten-Harm Verburg
Supervisor: Ralph Meulenbroeks
Second examiner: Arthur Bakker

Student number 4006887
45 ECTS

Utrecht University, Freudenthal Institute

Keywords: Direct Instruction, Guided Inquiry-Based Learning, secondary school, physics practical, ionizing radiation

Abstract

In education, there is an ongoing discussion on the effects of teaching methods such as Direct Instruction (DI) and Guided Inquiry-Based Learning (GIBL). Until now, there is little research on the effects of DI and GIBL in science practicals. The differences in conceptual understanding between practicals based on DI and GIBL were investigated. This research focused on a physics practical on ionizing radiation for Dutch secondary school students varying between grade 9 and 12 (15 to 18 years old). The research question then became: In the case of a high school practical on ionizing radiation for grade 9 to 12 students, what are the differences in change of conceptual understanding between Guided Inquiry Based Learning and Direct Instruction? In total, 479 secondary school students participated in pre- and post-tests. In addition, four groups participated in a semi-structured group interview. In the pre- and post-tests, no significant difference in conceptual understanding was found between the DI and GIBL practicals. The group interviews need to be investigated more thoroughly before general conclusions can be drawn.

Introduction

Within education, there is an ongoing discussion about the effects of various educational teaching and learning methods. Two of the methods that are central to the debate are Direct Instruction (DI, also referred to as *closed* in this research) and Inquiry-Based Learning (IBL, also referred to as *open* in this research). On the one hand, DI leans much more towards teaching in fragments and teacher-directed education as argued by Stahl et al. (1998, as cited in Ryder, Burton & Silberg, 2006). On the other hand, IBL is a spectrum of more student-directed education (Caps & Crawford, 2013). Within the spectrum of IBL, Guided Inquiry-Based Learning (GIBL, also referred to as *open*) suits best for this research and will be explained in the theoretical background.

So far, most research has focused on the effects of DI and GIBL in regular classroom settings. In a regular classroom, setting students usually combine the use of school books, theoretical exercises and teachers' instruction to gain new knowledge. Science practicals however, offer students the chance to combine gaining new knowledge with existing knowledge, creating deeper conceptual understanding while actively experimenting with little face-to-face instruction. In the practicals, students often are also required to work together to solve more complex problems, than in for instance work books. Until now, there is limited research on the effects of DI and GIBL in science practicals (Hattie, 2008). This research will therefore focus on the effects of DI and GIBL on conceptual understanding in a physics practical.

Aim research

This study aims to investigate the effects of teacher-directed DI and student-directed GIBL on the conceptual understanding of 9th to 12th grade (15 to 18-year-old) students in a science practical about ionizing radiation. It is hoped that more knowledge about the effectiveness of these teaching methods on students' conceptual understanding, can be used to further improve science practicals and to improve the designing of educational materials in physics to better fit the wants and needs of teachers and students.

Theoretical background

Since the definition of both Direct Instruction (DI) and Guided Inquiry Based Learning (GIBL) vary with different sources, our definition will be given below. Furthermore, we will hypothesize on the expected outcomes in the practical investigated.

Direct instruction

Direct Instruction as a teaching method was introduced little over 50 years ago by Siegfried Engelmann and colleagues. Direct Instruction was first implemented as Direct Instruction System for Teaching and Remediation (DISTAR). The program addressed reading language and math (Magliaro, Lockee & Buton, 2005).

Stahl et al. (1998, as cited in Ryder, Burton & Silberg, 2006) stated that Direct Instruction is based on three core principles, which are listed below with their implication in the practical investigated.

- Language is broken down into components taught in isolation
- Learning is teacher-directed
- Students have little input

The three core principles by Stahl et al. (1998, as cited in Ryder, Burton & Silberg, 2006) approach the definition of Direct Instruction, introduced by Carnine (2000), that is: “Direct Instruction is an approach to teaching. It is skills-oriented, and the teaching practices it implies are teacher-directed. It emphasizes the use of small-group, face-to-face instruction by teachers and aides using carefully articulated lessons in which cognitive skills are broken down into small units, sequenced deliberately, and taught explicitly” (see Carnine, 2000, pp. 5-6; Traub, 1999).

The definition above refers to DI used in traditional/classroom settings. However, for the present research (a practical) face-to-face instruction is nearly absent. The instructions are presented in the form of a written step by step worksheet. Therefore, the definition of DI in this research is modified to be: *Direct Instruction is an approach to teaching that is skills-oriented and the teaching practices it implies are teacher-directed. It emphasizes the use of small-group, step-by-step and ready-made worksheets and aides using carefully articulated lessons in which cognitive skills are broken down into small units, sequenced deliberately, and taught explicitly.*

For the DI version of the practical, students work with experiment-based worksheets which are available on the website of the practical involved (Utrecht University, 2018). The assignments are split in separate parts: aim, setup, measurements, elaboration and notes. The parts of the assignment guide students through the experiment, therefore agreeing with the first core principle.

Since the worksheets are designed beforehand, students work towards a ready-made and unidirectional solution. This approach is also used in various physics school books used in upper level secondary school (Van Dalen et al., 2013; Flokstra & Groenewold, 2015). On the one hand, providing straightforward worksheets saves time for teachers, since there is a singular solution for every question (teacher-directed). On the other hand, providing students with a static assignment form, leaves little room for students’ own initiative and/or input.

Guided Inquiry-Based Learning

For ages, educators have tried to introduce learning by experimenting, because Direct Instruction precludes the experience of actual doing research in class (Gooding, 1985). One of the movements to improve the amount of research and the level of inquiry in education (Ciappetta, 2007; Gooding, 1985; Herman & Pinard, 2015) introduced Inquiry-Based Learning (IBL). (Morris, 1995).

Definition Inquiry-based Learning & Guided Inquiry-Based Learning

The American National Research Council (NRC) (2000) defined five essential features of IBL in all school levels, which are listed below.

- Learners are engaged by scientifically oriented questions.
- Learners give priority to evidence, which allows them to develop and evaluate explanations that address scientifically oriented questions.
- Learners formulate explanations from evidence to address scientifically oriented questions.
- Learners evaluate their explanations in light of alternative explanations, particularly those reflecting scientific understanding
- Learners communicate and justify their proposed explanations.
(National Research Council, 2010 p24).

The features defined by the NRC can be related to the review by Pedaste et al. (2015). In the review, 32 research articles regarding Inquiry-Based Learning were investigated to define general inquiry phases in IBL. Five distinct general inquiry phases were defined, which are listed in Table 1.

From the general definition of Inquiry Based Learning, IBL can be subdivided to different levels of Inquiry. Capps & Crawford (2013) designed a matrix to categorize the level of inquiry in science lessons can be determined, which is added in Appendix A. The phases of general inquiry by Pedaste et al. (2015) and the elements from the matrix designed by Capps & Crawford (2013) can be connected to the different phases of the GIBL worksheets, of which an example can be found online (Utrecht University, 2018). Table 1 summarizes the phases of the inquiry by Pedaste et al.

Furthermore, the matrix by Capps & Crawford (taking level 3 out of 4 in Appendix A) is related with the suitable phases used in the GIBL worksheets students use in the IRP. In the practical, the worksheets guide students through the experiments with tips and hints, but students themselves must design the research question, hypothesis and work plan. Guiding is the recurring word in the level three of inquiry by Caps & Crawford. Therefore, we will use the term Guided Inquiry Based-Learning (GIBL) in this research.

Phases of inquiry (NRC, 2000)	Elements Matrix (Capps & Crawford, 2013)	Phases GIBL worksheets IRP
Orientation	• Student guided in determining the tools and techniques needed.	Aim assignment
Conceptualization	• Student guided in designing and conducting an investigation. • Student guided in posing their own question.	Setup & research question

Investigation	• Student directed to collect certain data	Hypothesis, work plan & research
Conclusion	• Student guided in process of formulating explanations from evidence • Students guided in determining how evidence supports explanation or guided to other resources or alt explanations • Student guided in using math skills to answer a scientific question	Processing data & reporting
Discussion	• Student guided in the development of communication.	Processing data & reporting

Table 1: overview of the phases of inquiry (NRC, 2000) connected to the phases of the open assignment forms used in IRP.

Effectiveness DI versus GIBL in other research

The effectiveness of Inquiry-Based Learning over a Direct Instruction approach or vice versa has been often researched without a conclusive outcome (Furtak, Seidel, Iverson, & Briggs, 2012; Savelsbergh et al., 2016). If we look at research focusing on IBL only, it has become clear that IBL can improve students conceptual understanding (Edelson, Gordin, & Pea, 1999; Gormally, Brickman, Hallar, & Armstrong, 2009; Minner, Levy, & Century, 2010) and that it can improve students' attitude towards science (Furtak, Seidel, Iverson, & Briggs, 2012; Gibson & Chase, 2002; Nooijen, 2017; Savelsbergh et al., 2016; Sjøberg & Schreiner, 2010). Furthermore, IBL can also improve students 'minds-on' experience when performing experiments; students are not only executing the steps but also are forced to better overthink the steps in IBL (Hart et al., 2000; Millar & Abrahams, 2009)

At the same time, sceptical voices remain in educational research when it comes to the effectiveness of Inquiry-Based Learning (Kirschner, Sweller & Clark, 2006). The research by Kirschner, Sweller & Clark shows that minimal guidance (fully student-initiated IBL) during instruction is not worth studying, therefore this research focuses on a practical involving Guided Inquiry-Based Learning.

Conceptual understanding

For this research, we focus on the deeper conceptual understanding of students during the practical. New conceptual understanding on ionizing radiation will be combined with existing conceptual understanding to finish the practical successfully. The themes central to the practical are on ionizing radiation for grade 9 to 12 students in secondary school. In Dutch education, this correspondents to school levels 4 to 5 HAVO (preparing for university of applied sciences) and 4 to 6 VWO (preparing for university). In the syllabi of both school levels, the practical connects to the subdomains B2 of HAVO and subdomains B2 and E2 on ionizing radiation (College voor Toetsen en Examens, 2016a & 2016b).

Related research

The research that has been done so far provides limited research on the conceptual understanding of students in a practical setting outside of a regular classroom setting. A few examples are that were found were 'Brazil CNEN [Brazilian Nuclear Energy Commission, ed.]' project, a project at the Weizmann Institute in Israel and a project at La Sapienza University in Italy (Pereira, Rosa, Negretti, & Faletti, 1998). All the projects have received positive feedback, but no research has been done on the conceptual understanding of the projects. Furthermore, other than the practical researched here, the latter projects did not include a practical where students could work with ionizing radiation.

In the practical investigated in this research, earlier research has been done on students' conceptual understanding, investigating whether guided discovery (related to GIBL) would lead to an increase of conceptual learning compared to the 'step-by-step' or 'cookbook' (related to DI) approach (Janssen, 2016). The research concluded that there was no significant greater increase in conceptual understanding between open and closed IRP ($p = .595$). This conclusion however, can be questioned due to the small sample size ($N = 42$) and the use of pooled data. Because of pooling data, it is difficult to state any reliable and valid conclusions because the results of the one school may not be representative for the results of schools in general. Following the relative dearth of research in this field, it is useful to focus on the conceptual understanding that are learned in a more hands-on physics practical. More concretely, the learning methods DI and GIBL will be compared to investigate its effects on conceptual understanding in the physics practical.

Position of this research

The aim of this research is to get a better understanding on the effects of DI and GIBL on the conceptual understanding in a practical setting. Ideally, to overcome as many external factors as possible like school setting and teacher experience, an Education Design Research (EDR) (Van den Akker et a., 2010) should be used. In an EDR, the preparation time to the practical and the time spent on the practical should be the same in the practical based on DI and based on GIBL. Unfortunately, due to practical constrains, it is not possible to use EDR in this research, but the same practical as Janssen (2016) investigated will be used.

This research however, is not a mere repetition of Janssen's (2016) research. This research takes the next step towards controlling factors and therefore finding a more representative insight in the effects of either DI or GIBL on conceptual understanding in science practicals. Where possible, data will be collected through more ways than used in Janssen's research, the methods used will be evaluated in between where possible, after which research continues with revised versions of the methods used. Furthermore, more schools will be analyzed to reduce possible teacher or school biases and to find a general trend in the effects of DI and GIBL on the conceptual understanding in science practicals.

Research question

From the theory above, the research question that can be formulated is: In the case of a high school practical on ionizing radiation for grade 9 to 12 students, what are the differences in change of conceptual understanding between Guided Inquiry Based Learning and Direct Instruction? The research takes place in Dutch education, where grade 9 to 12 comply with 4-5 HAVO to 5-6 VWO.

Hypothesis

In this research, our research question is whether, in the case of a high school practical on ionizing radiation for grade 9 to 12 students, there are differences in change of conceptual understanding between Guided Inquiry Based Learning and Direct Instruction. What can be expected during this research when it comes to the conceptual understanding in this research? As also shown in Hattie (2008), educational research is subject to varying settings, teachers and contents. For our research, we focus on a physics practical on ionizing radiation.

Since students experiment with different setups, it is expected that students will gain a ‘hands-on’ experience, while using existing and gaining new theory about ionizing radiation. As mentioned earlier, the DI worksheets provide a ‘cookbook’ work plan for students how to execute the experiment. It therefore can be hypothesized that students can start without in depth conceptual knowledge of the physics involved.

The GIBL worksheets on the other hand, require students to overthink and prepare a research question, hypothesis and work plan before they can start experimenting. The preparation requires more intensive collaboration between the group members and is more time consuming than when using the DI worksheets. The more intensive preparation, leads us to hypothesize that a more intensive preparation will lead to higher conceptual understanding in the GIBL approach, compared to DI.

Methods

Setting

This research focused on a physics practical called the Ionizing Radiation Practical (IRP, Dutch: ‘Ioniserende Stralen Practicum’) of Utrecht University. The practical – started in 1972 – gives 15 to 18-year upper level secondary school students (grade 9-12) the opportunity to gain some hands-on experience dealing with radioactive materials, which otherwise stays an abstract topic in secondary schools.

IRP initially started with a DI-only approach. This shifted after the Institute for Curriculum Development (SLO; Dutch: Stichting Leerplan Ontwikkeling) called for a lesson approach which would be more open to input of students (Hulsbeek, Meijerink, Schurink, & Wiegman, 1999). From 2011 onwards, an option of both the DI and GIBL has been introduced, therefore making it a good setting to investigate the students’ conceptual understanding of the physics.

The IRP consists of 24 different experiments; the experiments are divided in four categories, half-life, absorption, x-rays and miscellaneous. Teachers often preassign the experiments to groups of two to three students, where the practical has space for up to 30 students working simultaneously. For every experiment, a worksheet is offered to support students through the experiment. In total, there are 24 worksheets based in DI, and another eight experiments also have a separate GIBL worksheets.

The DI worksheets provide a step-by-step guide how to conduct an experiment. In the case of IRP, the step by step approach directs students through the measurements and their analysis, leaving very little room for students' own initiative.

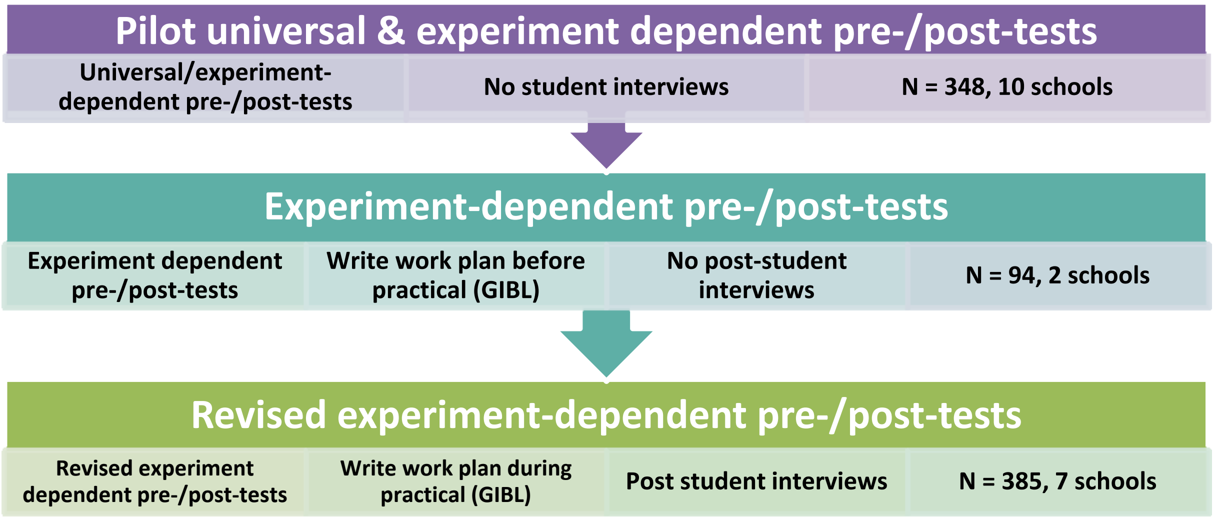
In the GIBL worksheets, students only receive guiding questions about the experiment. These questions should be used to make their own step-by-step work plans how to execute the experiment. As mentioned earlier, Guided Inquiry Based Learning (GIBL) is utilized in the GIBL worksheets in IRP, which is leaning towards student-initiated while guiding students through the experiment (Caps & Crawford, 2013).

For the actual practical, school classes go to Utrecht University to execute their experiments or IRP provides the practical at the secondary schools themselves in mobile labs. For schools, this means that they do not have to keep any radioactive materials their schools, which saves costs and safety precautions. In two to three hours, students generally perform three experiments, depending on the experiment and the teacher. In the case of the GIBL assignments, students must make a work plan how they want to do execute the experiment. Students can either write the work plan before the actual practical or during the practical itself; this is up to the teacher. For 90% of this research, students write the work plan during the IRP.

When the work plan has been checked by the teacher, the students can start. The GIBL worksheets take more time than the DI worksheets, because in GIBL students write their work plan before they can start experimenting. Therefore, it is often decided to use only one GIBL worksheet, after which DI worksheets follow (Utrecht University, 2018). Reporting all the data is partly carried out during the experiments, but is also carried out after the experiments (especially for the GIBL worksheets). After the IRP project, the worksheets can also be evaluated in class.

Research design

In this research, we investigated the main research question: In the case of a high school practical on ionizing radiation for grade 9 to 12 students, what are the differences in change of conceptual understanding between Guided Inquiry Based Learning and Direct Instruction? Since very little research has been done onto the topic of conceptual understanding between DI and GIBL in practical



settings, this was an exploratory research. The research section is separated in three phases, which are visualized in Figure 1 (below).

This research focused on investigating differences in conceptual understanding of ionizing radiation between DI and GIBL in a practical setting. The research took place through a mixed methods approach, using pre- and post-tests and semi-structured post interviews.

Since no existing pre- and posttest could be used for this research, new tests were designed and piloted. The pre- and post-test can either be universal or experiment-dependent. In the case of a universal pre- and post-test, there was one test for all 24 experiments available in the IRP. When designing an experiment-dependent pre- and post-test, one pair of tests was designed for every of the 24 available experiments, focusing on the specific topic of that experiment. If there were both a GIBL and DI worksheet available for one experiment, the pre- and post-tests for these experiments were the same.

Upon designing the pre- and post-tests, it was aimed to find the proper balance between the time the tests take in the practical and the level of the tests. All tests designed were aimed for use in school classes ranging from grade 9 to 12 (4 HAVO up until 6 VWO). Furthermore, the tests should take 5-10 minutes at maximum, due to the limited time available at the practical.

It was too idealistic to expect clear and singular results regarding the differences in conceptual understanding, because the results also are reliant on external factors such as the school culture and the teacher. The varying measurability of conceptual understanding is also argued by Hattie (2015). Another aspect that was kept in mind was the possibility that students would learn from the pre-post-tests themselves, rather than from the practical investigated; this is also known as the testing-effect. Research by Karpicke and Blunt (2011) and by McDaniel and Fisher (1991) showed that taking knowledge tests can help students with memorizing the concepts involved. Dirkx, Kester and Kirschner (2014) also showed in their research that a testing-effect can lead to better fact retention and that it positively affects deeper learning. And although these effects on itself can be beneficial to education, it may be unwanted when pre- and post-tests are compared.

At the same time, the learning effect (Hartley, 1973) could become a problem with pre-and post-tests in research, since it became less clear whether a student learned from the pretest or from the practical itself. To overcome as much learning effects as possible, the phrasing of questions was varied between pre-and post-tests.

Universal pre-/post-tests

The first pre- and post-test that was designed, was universal and therefore could be used for every experiment in IRP. It thus measured the level of conceptual understanding of ionizing radiation in general, not specific to a specific experiment in the practical. Upon designing the universal pre- and post-test there was a limited number of pre- and post-test on ionizing radiation that already existed. The pre- and post-test that was considered, was the test used by Janssen (2016). This test, however, would take too much time during the practical and therefore it was decided to design a new universal pre- and post-test.

The new universal pre- and post-test was designed based on the following:

- Universal pre- and post-test by Janssen (2016)
- Syllabi Physics 2018, as mentioned earlier in the meaning of conceptual understanding
- Discussion with four staff of the practical, all with a background in physics
- The worksheets students use in IRP

The combination of the four have led to a universal pre- and post-test of seven questions per test, which covered the contents of all the experiments in the IRP. Three different types of questions were added in the tests, that is four closed multiple choice questions, two semi-open questions and one multiple choice question with the possibility to elaborate on the answer given. In addition, students were asked for the first two letters of their first name, their date of birth, level and school, to enable us to combine the students' pre- and post-test. An example of the universal pre- and post-test is added in Appendix B.

Experiment-dependent pre-/post-tests

For the experiment-dependent pre/post-tests, no known tests were available from literature, therefore the tests needed to be piloted thoroughly. The experiment-dependent pre-/post-tests were designed based on the same basic elements as the universal tests. In the experiment-dependent tests however, the focus was on the contents of the specific experiment involved; the test was adjusted to the concrete contents of the experiment to maximize insight into the students' conceptual understanding.

Other than using seven questions like in the universal pre-/post-test, it was decided to use three closed multiple choice questions per test. It was hoped that this would enable students to finish the tests in less time, leaving more time for the practical. As in the universal tests, students were asked for their birth day, first two letters of their first name, school and level, to couple the pre- and post-test. An example of an experiment-dependent pre-/post-test is added in Appendix C.

Student interviews

To also get a qualitative understanding of the thoughts and conceptual understanding of students, a semi-open post student interview was added when using the revised experiment-dependent pre- and post-test. It was intended to organize group interviews of 3-5 students of every class that was investigated using the revised experiment-dependent pre- and post-tests. As far as found in literature, there was no existing interview scheme on the conceptual understanding of students in a physics practical on ionizing radiation. The core questions therefore were designed first, based on the contents of the IRP. Furthermore, the interview examined if students had different perceptions between the conceptual understanding in the DI experiments and in the GIBL experiments.

For the design of this interview, it was decided to leave three till four weeks between the IRP and the post interview. In this time, student could process the data collected in the IRP (as far as they did not do this during the practical) and most of the time students also finished the chapter on ionizing radiation in physics class; the latter was relevant to see if there is a possible relation between theoretical knowledge on ionizing radiation from the book and the score in the pre-and post-tests.

The interviews were conducted at the secondary school involved, mostly during physics class or shortly before or after. Students were randomly selected by the teacher, by for instance assigning a number to every student and randomly select numbers. After selection, students were free to refuse collaborating in the interview; another student would then be randomly selected. All students involved were asked if recording the interview (by using a voice recorder) is allowed, after which the interview would start.

Since the focus of this research was on the results of the pre- and post-tests, the interviews were not coded; yet typical quotes from the interview will be discussed. The core questions of the interview are added in Appendix D.

Timeline research

As mentioned earlier, students did three to four experiments in the IRP; this research only focused on the first experiment students encounter in the practical, as also showed in Figure 2 (below). This meant that the pre- and post-tests as shown will be taken shortly before and directly after the first experiment students execute. It depends on the teacher whether students were able to process all the data before they filled in the post-test and continued to the second experiment, or if students processed the measurements later. Focusing on the first experiment enabled us to explicitly compare students conceptual understanding between the DI and GIBL worksheets; there were only eight GIBL worksheets are available, forcing the remaining students to start with a DI worksheet. After approximately four weeks, three to five students of each group participated in a semi-structured group interview at the secondary school involved. However, only focusing on the first experime

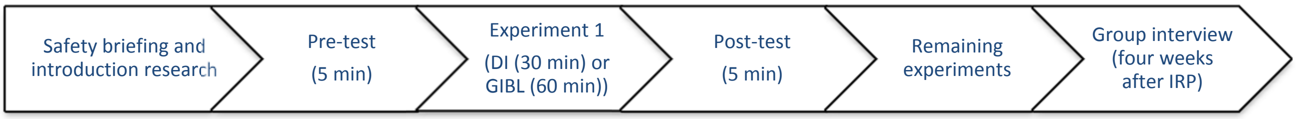


Figure 2: Timeline of the research on conceptual understanding in the IRP.

Statistical analysis

For the data analysis, a MANOVA was considered to compare the data that was collected, because the pre- and post-tests produced data that required a test that combined both a between-subjects and a within-subjects analysis. To use a MANOVA, it was required that the data set was normally distributed. Therefore, a Shapiro-Wilk test of normality was used to determine the normality. If the data would be normally distributed, the analysis would proceed with a MANOVA. The effect size Cohens d then was also calculated if the data was normally distributed.

If the data was not normally distributed, a Mann-Whitney U Test was used to determine if there was a difference in conceptual understanding between the DI worksheets and the GIBL worksheets. To use a Mann-Whitney U Test, the results from the pre- and post-test were combined into one $\Delta Score$, as shown in Figure 3. As a result, there now was one dataset for the DI worksheets and one dataset for the GIBL worksheets. Again, a Shapiro-Wilk test was used to check for normality. If the dataset would be parametric, analysis proceeded with an independent samples t-test. If the data was not normally distributed, analysis proceeded with a Mann-Whitney U Test.

$$Score_{post} - Score_{Pre} = \Delta Score$$

Figure 3: Formula used to calculate the difference of student scores between the pre-test and post-test.

For the $\Delta Score$ dataset, the effect size was measured using Cohens d for parametric datasets. For the nonparametric dataset, the effect size was calculated using the formula shown in Figure 4 (Rosenthal, 1994). Z is the standardized test statistic, N is defined as the total sample size of the dataset, leading to effect size r.

$$r = \frac{Z}{\sqrt(N)}$$

Figure 4: Formula used to calculate the effect size of nonparametric data sets

Results

The results section of the research will be divided in three main parts. First, the results of the pilots will be discussed qualitatively. Second, the results of the experiment-dependent pre-/post-test will be discussed in a quantitative fashion. The results of this part need to be analyzed separate from the last (revised) pre-/post-tests, since the questions in the pre- and post-tests differ. Finally, the results of the revised pre-/post-tests will be discussed with the most typical quotes from the group interviews.

Pilot universal pre-/post-tests

In the starting phase of this research, both the universal and experiment-dependent pre-/post-tests were piloted amongst secondary school students, participating in the Ionizing Radiation Practical. In the universal pre-/post-test, 159 secondary school students participated, all leveled grade 11 (HAVO 5). The students examined only used the DI worksheets, therefore the data could not be used for answering the research question. However, the data was useful for determining the final version of the pre-/post-test.

Upon using the universal pre-/and post-test in the IRP, it was observed that the questions in the universal tests covered more than just the contents of the experiments the students executed as their first experiment. This was also confirmed by verbal feedback of the teachers and written feedback by some students. Since the universal pre- and post-test was too broad to measure the change in conceptual understanding effectively, it was decided to terminate designing a universal pre-/post-test and continue designing the experiment-dependent-pre/post-test.

Pilot experiment-dependent pre-/post-tests

In the pilot of the experiment-dependent pre- and post-test, 121 secondary school students leveling between grade 9 to 12 participated in the pre- and post-test. The students examined, only used the DI worksheets in the experiments. Therefore, this data set could not be used for answering the research question. It however was expected that the students' starting level of conceptual understanding on ionizing radiation would be independent of the type of worksheet students use. Therefore, this data set would be useful for determining the appropriate level of the final tests.

From the data collected and the verbal feedback of teachers, staff involved and students that participated in the research, the main feedback was that the test connected well to the experiment for which it was designed for. Any mistakes in the tests were corrected before continuing to the next phase.

Experiment-dependent pre/post-tests

After processing the feedback in the pilot experiment-dependent pre- and post-test, the second phase was initiated. Two secondary schools were investigated (grade 9 to 11), in which 10 students used the GIBL worksheets, whereas another 84 used the DI worksheets. The pre- and post-tests were graded on a three-point scale, ranging from 0 to 3 points depending on the amount of multiple choice questions that students filled answered correct. It was intended to analyze all data with a Mixed ANOVA (MANOVA) in SPSS, enabling us to check for differences in conceptual understanding between the DI and GIBL worksheets. To do so, it was required that the data set would be normally distributed. A Shapiro-Wilk test of normality was executed, of which the results are combined with general descriptive data in Table 2 (below).

		Descriptives			*Test of Normality*		
	VersionIRP	n	M	SD	Statistic	df	Sig.
ScorePRE	DI	84	2.60	.623	.409	84	.000
	GIBL	10	2.80	.422	.482	10	.000
ScorePOST	DI	84	2.67	.584	.440	84	.000
	GIBL	10	2.53	.633	.370	10	.000

Table 2: General descriptives and Shapiro-Wilk normality test for results of experiment-

In addition, as shown in Table 2 and as an example visualized in Figure 5 (right), it becomes clear that none of the data sets collected are normally distributed. Therefore $\Delta Score$ was introduced, as explained in the Methods.

Upon introducing $\Delta Score$, it was required to again perform a Shapiro Wilk normality check, of which the results combined with the general descriptives are summarized in Table 3 (below) and a visualization of the distribution is shown in Figure 6 (below).

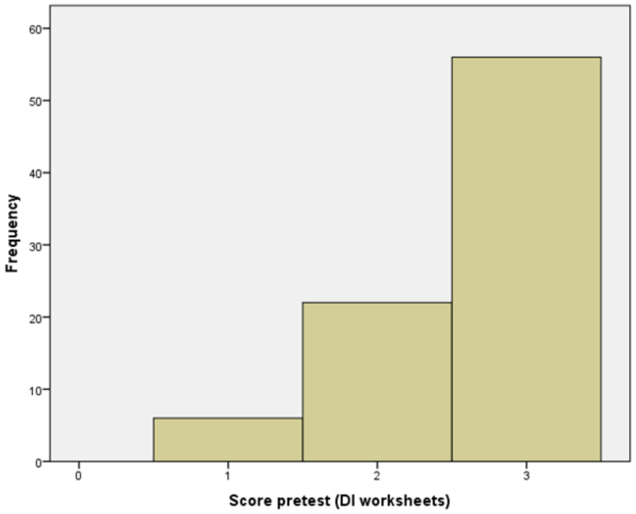


Figure 5: Frequency diagram pre-test (DI worksheets)

	VersionIRP	Descriptives			Test of Normality		
		n	M	SD	Statistic	df	Sig.
ΔScore	DI	84	.08	.843	.854	84	.000
	GIBL	10	-.27	.887	.902	10	.233

Table 3: General descriptives and Shapiro-Wilk normality test for results of $\Delta Score$ of the experiment-dependent pre- and post-test.

From the results in Table 3, it becomes clear that only the dataset of the GIBL worksheets is normally distributed (*p = .233*). Since only one out of two datasets was considered as normally distributed, it was decided to continue analyzing the data through a non-parametric test. An Independent Samples Mann-Whitney U Test was used to analyze $\Delta Score$. The null hypothesis was defined as “If the null hypothesis is true, there is no difference in change in conceptual understanding between students using the DI or GIBL worksheets in IRP.” The results show that there is no significant difference (*p = .223*) between the DI (*Mdn = 0*) and GIBL worksheets (*Mdn = 0*) with (*U = 330*), (*Z = -1.22*) and a small effect size of (*r = .13*).

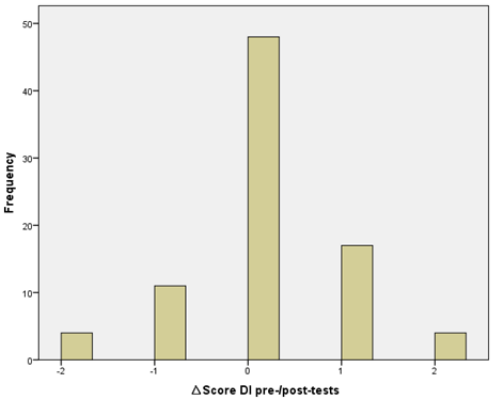


Figure 6: Distribution of $\Delta Score$ for the experiment-dependent pre-/post-tests.

Revised experiment-dependent pre/post-tests

From the data of the second phase, it became clear that the average score of the pre-tests already was predominantly towards the maximum score of 3 points. For this reason, it was decided to increase complexity of the questions asked in the pre- and post-tests. Since most questions changed, the revised experiment-dependent pre- and post-test (as shown Figure 1, also known as phase 3) were analysed separate from the previous data set.

In the third phase, 385 secondary school students of six schools participated, levelling between grade 11 to 12 (5 HAVO and 6 VWO). Of the 385 students, 188 used the DI worksheets, whereas the remaining 197 students used the GIBL worksheets. The revised pre- and post-test data was analysed in the same way as was done in the second phase, that is by a Mixed Anova. A Shapiro-Wilk test was executed to determine the normality of the data. The results are shown in Table 4 (below),

		Descriptives			*Test of Normality*		
	VersionIRP	n	M	SD	Statistic	df	Sig.
ScorePRE	DI	188	2.37	.731	.758	188	.000
	GIBL	197	2.32	.792	.772	197	.000
ScorePOST	DI	188	2.16	.780	.813	188	.000
	GIBL	197	2.18	.798	.799	197	.000

combined with the general descriptives.

In addition, as shown in Table 4 and as an example visualized in Figure 7 (right), it becomes clear that data is more advanced than in the previous pre- and post-test. Still, the data set was not normally distributed. This is also shown in an example of the distribution of the data used, visualized in Figure using Shapiro-Wilk test was executed on the dataset of $\Delta Score$ of the revised experiment-dependent pre- and post-tests and shown in Table 5 (below).

Table 4: General descriptives and Shapiro-Wilk norm revised experiment-dependent pre- ar

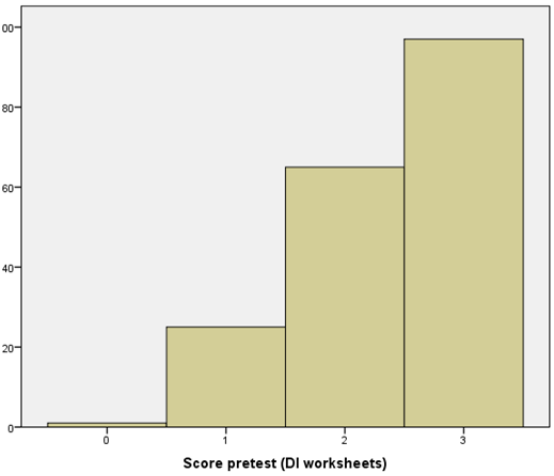


Figure 7: Frequency diagram revised pre-test (GIBL worksheets)

		Descriptives			*Test of Normality*		
	VersionIRP	n	M	SD	Statistic	df	Sig.
ΔScore	DI	188	-.21	1.031	.903	188	.000
	GIBL	197	-.14	.869	.875	197	.000

Since Table 5 shows that the dataset of $\Delta Score$ for the revised experiment-dependent pre- and post-test is n Mann-Whitn

Table 5: General descriptives and Shapiro-Wilk normality test for results of $\Delta Score$ of the revised experiment-dependent pre- and post-test.

null hypothesis is true, there is no difference in change in conceptual understanding between students using the DI or GIBL worksheets in IRP.” The results show that there is no significant difference (*p = .202*) between the DI (*Mdn = 0*) and GIBL worksheets (*Mdn = 0*), with (*U = 19,827*), (*Z = 1.28*) and a small effect size of (*r = .07*).

Group interviews

Beside the revised experiment-dependent pre- and post-tests, four groups of students were interviewed in two schools in a Dutch spoken semi-structured interview. The students that participated – from either grade 11 or 12 – were chosen randomly by the physics teacher. One group of three to five students per physics class was interviewed about four weeks after the IRP and students participated on a voluntary basis. The interview was focused on the thoughts of students regarding the practical in general, their perception on the IRP and their thoughts on the conceptual understanding students gained during the IRP.

One of the main findings from the interviews is that there may be very little to no change in conceptual understanding of the students. Furthermore, external factors – such as timing of the practical and experience with writing work plans – play an important role in students’ thoughts regarding conceptual understanding about ionizing radiation. In one of the interviews, 12th grade secondary school students were interviewed. The IRP of the practical was planned about one year after the theoretical lessons about Ionizing Radiation. Upon asking what students thought to have learned from the practical, L1 responded:

> I found it fun to do, it was okay; it is different than sitting in school all day. But I do not have the feeling that I really learned much new knowledge .

After which L2 added:

> When the practical was executed during the chapter [on ionizing radiation, ed], I think it would have really added something.

In another interview, the same question was asked, after which L3 responded:

> I now know better how to deal [with the machines involved, ed], gaining knowledge and, in particular, how to apply knowledge.

L4 also summarized this as:

> It is more like that you apply theory for real. And that was interesting, because when you read theory you think ‘whatever’, whereas if you see it happening you think ‘Oh, so it really is like this.’

The answers of L1, L3 and L4 indicate that students there is little to no change in conceptual understanding about ionizing radiation during the practical. This is in line with the results from the pre- and post-tests. Moreover, L3 and L4 mention that the practical was useful to apply the existing knowledge. The answer of L2 indicates that the timing of the practical is also essential; the practical should be planned when students spend time on ionizing radiation in physics class.

Upon asking whether there was a difference between the GIBL and DI-based experiments, L5 answered:

> In the open [GIBL, ed] version, you were required to think of the steps much more by yourself, whereas you could just follow the steps in the closed [DI, ed] version.

L2 responded to L5:

> In the closed [DI, ed] version, it really was like push the button and write down, push the button and write down, push the button and write down, calculate the average and then you know the answer.

L6 then answered about their DI-based experiment:

> Our experiment involved pulses or something. In the beginning, there was a text which I did not understand at all, but I have been able to make the complete assignment without knowing what I was doing. The assignments were not very instructive.

Although there seems to be little to no change in conceptual understanding during the IRP, the answers provided by L2, L5, L6 and answers provided by few other students indicate that students learn even less in the DI-based experiments. At the same time, most students tend to agree that the practical serves an opportunity to apply the theory on ionizing radiation in a practice. Again, the interviews should be investigated more extensively before drawing general conclusions.

Conclusion

This research investigated the effects of DI and GIBL on in conceptual understanding in a physics practical on ionizing for grade 9 to 12 students. In this practical, students used worksheets based on Direct Instruction or Guided Inquiry-Based Learning. Through experiment-dependent pre- and post-tests and group interviews, the effects conceptual understanding of students were measured. The research question central to this research was: In the case of a high school practical on ionizing radiation for grade 9 to 12 students, what are the differences in change of conceptual understanding between Guided Inquiry Based Learning and Direct Instruction?

In the first experiment-dependent pre-and post-test, 94 students participated (grade 9 to 11) of which 10 students used the GIBL worksheets. The Mann-Whitney U Test indicated there is no significant difference (*p = .223*) between the change in the conceptual understanding when using either the DI if GIBL-based worksheets.

For the revised experiment-dependent pre- and post-test 385 grade 11 & 12 students participated. Of 385 students, 197 used the GIBL worksheets, whereas 188 used the DI worksheets. As in the previous dataset, there is no significant difference (*p = .202*) between the change in the conceptual understanding when using either the DI if GIBL-based worksheets.

In the four interviews with three to five students, the typical quotes show answers that tend to be in line with the same conclusion as drawn from the pre-/and post-tests. The interviews indicate that the conceptual understanding is not the main learning outcome of the IRP, even less in DI than in GIBL. Moreover, the interviews show that students consider the practical as an opportunity to apply the theoretical knowledge. Students seem to apply the data better when using GIBL worksheets compared to DI worksheets, since students are more actively involved in the experiment during the GIBL experiment. Summarized, the first results from the interviews seem useful, but the interviews should be analyzed more in depth before any general conclusions can be drawn.

The research question was: *In the case of a high school practical on ionizing radiation for grade 9 to 12 students, what are the differences in change of conceptual understanding between Guided Inquiry Based Learning and Direct Instruction?* Based on the experiment-dependent pre- and post-tests, there is no significant difference in change, between a DI or a GIBL physics practical on ionizing radiation. The group interviews should first be analyzed more extensively, before general conclusions can be drawn from the interview data.

Discussion & implications

The practical investigated in this design, served several roles. In the practical, students were able to get acquainted with radioactive materials, gain experience in collaboration with other students and the practical served to create a deeper conceptual understanding of knowledge from concepts introduced in earlier physics classes. The different tests and interview questions were designed around this assumption to find changes in the conceptual understanding in DI and GIBL.

Investigating students' conceptual understanding

If we assume the change in conceptual understanding to be debatable, the methods are up to debate. Since no significant differences were found, it can be debated whether the experiment-dependent pre- and post-test is the proper way to measure conceptual understanding in a practical setting. In the method used, both the pre-test and post-test consisted of three questions per test. The number of questions was chosen, mainly because of time constraints during the practical, but may have been too little to investigate students' conceptual understanding properly.

The level of the questions was designed by researchers with a background in physics to cover all grade 9 to 12 students that could participate in the IRP. It may not have been realistic to design tests with a limited number of questions, covering all school levels while at the same time enabling research to demonstrate changes in conceptual understanding.

As mentioned earlier, the time between the pre-test and post-test varied between 30-60 minutes depending whether students executed a DI or GIBL worksheet. In both cases, there may have been a testing-effect, in which students learned from the pre-test itself rather than from the practical. Research on the testing-effect shows that taking tests positively influences the fact retention and deeper learning (Dirkx, Kester & Kirschner, 2014). For students, the pre- and post-tests therefore may have been beneficial, but for the actual research it may not. One of the considerations to reduce the testing-effect, is by leaving more time between the pre- and post-test. However, by increasing the amount of time between the tests, would have made it even more difficult to find change in conceptual understanding that can be attributed to either DI or GIBL in a practical setting.

Aim practical

It can be debated what the main aim of the practical is. According to verbal feedback by the teachers of the school classes involved, a teacher and school perspective often is that the practical serves to create new knowledge deeper conceptual understanding with the students. What kind of knowledge is meant? The teachers argue that this is conceptual knowledge on ionizing radiation and this is sometimes agreed by students in the group interviews. However, most of the students argue that they have not learned new knowledge, and that the practical served to apply the existing knowledge from previous regular classroom lessons. It is not likely that there was no change in conceptual understanding, but it may be argued that the aim of this research and of the teachers, do not comply with the aim students enter a practical.

The same question can be raised regarding the aim of either DI and GIBL worksheets. The latter was introduced in 2011 to enable students to make students more engaged and deepen their learning process by allowing them to write their own work plans and execute those plans. It was assumed that the level of conceptual understanding would not decrease, but until this study, this has been researched on a limited scale. If the level of inquiry would be utilized completely in a practical, this would mean that the student can design their experiment from scratch. However, safety constrains of the radioactive materials do require guidance for students. Still the level of inquiry is larger in the GIBL worksheets than in the DI worksheets, which should enable researchers to find differences in students' learning and conceptual understanding, as also argued in settings of discovery learning by De Jong and van Joolingen (1998).

Significance results

No significant difference was found between DI and GIBL worksheets in the practical, while the methods were intended for measuring the students' (deeper) conceptual understanding. The effect size that were found varied between ($r = .07$) and ($r = .13$). It can be argued that, by just looking at the values, the effect sizes seem to be relatively high for the time span of the intervention. The intervention time is only 30 to 60 minutes, whereas research in education sometimes struggles to achieve these effect sizes even in long-term research with more interventions (Hattie, 2008).

It can also be discussed whether a difference in conceptual understanding on ionizing radiation in this physics practical can be expected at all. Or in other words, is conceptual understanding the field where students will improve upon introducing more inquiry-based worksheets? Ideally, when introducing a new teaching method, the students' the results of students improve. However, in this research over 450 secondary school students participated, which may seem like a decent number to state reliable outcomes. And the fact there is no change in conceptual understanding can also be interpreted as positive.

Setting

In educational research, it often becomes clear that the results are subject to the setting where the research takes place (Hattie, 2015). This was also the case in this research. Here, the conceptual understanding in a physics practical is subject to students' secondary school and to their physics teacher. For example, in the student interviews, there were indications that some students lacked experience in writing work plans, making it more difficult to use the GIBL-based worksheets. For teachers, it therefore is important to also prepare students for inquiry, as is also in line with the trend of *Onderwijs2032* (also known as Curriculum.nu) in education (Curriculum.nu, 2018).

At the same time, students who did have experience could apply this in the IRP. To overcome as much of the setting dependence as possible in the future, it is important to do qualitative preliminary research in the schools involved. It should become clear what the students' starting level of inquiry is and to what extent the teacher uses inquiry in his/her regular school lessons. If sample sizes are big enough, this preliminary data can be used to investigate to what extent the level of inquiry in regular school classes influences the results in the physics practical.

Future research

Since the students' answers in the interviews give indications on the change in conceptual understanding of students, the focus should be on taking group interviews and code the data that is collected. At the same time, taking pre- and post-tests should still be considered. The tests can provide a way to confirm the data collected in interviews, creating more reliable outcomes. The current experiment-dependent pre- and post-test can be used, but the questions need be adjusted to fit the aim of the research. In addition, the length of the pre- and post-test could be adjusted, keeping logistical concerns in mind.

Upon asking students' opinion on the function of the practical, one of the themes mentioned was applying theory. In this research, it was assumed that gaining deeper conceptual understanding on ionizing radiation was the aim of the practical. It was also assumed that applying the conceptual understanding in practical setting was separate from conceptual understanding on itself. Therefore, the level of application of knowledge was not measured in this research, although the interview questions show signals that this would have been interesting.

Future research therefore should focus on practicals, examining the difference in long-term conceptual understanding between DI and GIBL and examining whether the GIBL practical enables students to better apply knowledge than when students execute a DI practical. Again, external factors as secondary school and teacher will be difficult to control, but larger sample sizes could show a general trend.

Another aspect to investigate is the relation between conceptual understanding and students' motivation in science practicals. Last year, Teun Nooijen (2017) investigated the relation between the same GIBL or DI practical and students' intrinsic motivation. Nooijen found a statistically significant difference in students' motivation, in favor of GIBL. If there is a proper way to investigate students' conceptual understanding on ionizing radiation, this could be connected to the motivation of students. It for instance can be investigated if in a science practical the conceptual understanding is higher amongst students with a high intrinsic motivation. The data from this future research could give directions in how to use intrinsic motivation to increase the conceptual understanding.

This research showed that – when it comes to conceptual understanding on ionizing radiation – there is no statistical difference between using GIBL or DI in a physics practical. However, this research does also show that there are many effects that remain unanswered when comparing Direct Instruction and Guided Inquiry-Based Learning in science practicals. The effects of DI and GIBL in science practicals on for instance students' autonomy, collaboration, students' motivation, students' existing level of inquiry before a practical and ability of students to apply existing knowledge remain largely unknown. Here future research offers a chance to answer these question, enabling teachers to better adjust education to the needs and wants of their students.

References

American National Research Council. (2000). Inquiry and the National Science Education Standards: A Guide for Teaching and Learning (p. 24). Washington, D.C.: National Academies Press. https://doi.org/10.17226/9596

Capps, D. K., & Crawford, B. A. (2013). Inquiry-Based Instruction and Teaching About Nature of Science: Are They Happening? *Journal of Science Teacher Education*, *24*(3), 497–526. https://doi.org/10.1007/s10972-012-9314-z

Carnine, D. (2000). *Why Education Experts Resist Effective Practices (And What It Would Take To Make Education More Like Medicine)*. Washington, DC: Thomas B. Fordham Foundation. Retrieved from https://files.eric.ed.gov/fulltext/ED442804.pdf

Chiappetta, E. L. (2008). Historical Development of Teaching Science as Inquiry. In *Science as inquiry in the secondary setting* (pp. 21–30). Arlington, Va: NSTA Press.

College voor Toetsen en Examens. (2016a). *Natuurkunde HAVO: syllabus centraal examen 2018*. College voor Toetsen en Examens. Retrieved from https://www.examenblad.nl/examenstof/syllabus-2018-natuurkunde-HAVO/2018/HAVO/f=/natuurkunde_HAVO_2_versie_2018.pdf

College voor Toetsen en Examens. (2016b). *Natuurkunde VWO: syllabus centraal examen 2018*. College voor Toetsen en Examens. Retrieved from https://www.examenblad.nl/examenstof/syllabus-2018-natuurkunde-VWO/2018/VWO/f=/natuurkunde_2_versie_VWO_2018.pdf

Curriculum.nu. (2018). Curriculum.nu. Retrieved from https://curriculum.nu/

Dalen, B. van, & Ottink, H. (2013). *Systematische natuurkunde. Basisboeken VWO 4 Basisboeken VWO 4*. Amersfoort: ThiemeMeulenhoff.

De Jong, T., & Van Joolingen, W. R. (1998). Scientific Discovery Learning with Computer Simulations of Conceptual Domains. *Review of Educational Research*, *68*(2), 179–201. https://doi.org/10.3102/00346543068002179

Dirkx, K. J. H., Kester, L., & Kirschner, P. A. (2014). The Testing Effect for Learning Principles and Procedures from Texts. *The Journal of Educational Research*, *107*(5), 357–364. https://doi.org/10.1080/00220671.2013.823370

Edelson, D. C., Gordin, D. N., & Pea, R. D. (1999). Addressing the Challenges of Inquiry-Based Learning Through Technology and Curriculum Design. *Journal of the Learning Sciences*, *8*(3–4), 391–450. https://doi.org/10.1080/10508406.1999.9672075

Flokstra, J., Groenewold, A., Hooyman, K., Kootwijk, C., Kortland, K., Over, P., … Wolters, J. (2015). *Newton: natuurkunde voor de bovenbouw. VWO. Basisboeken 6 VWO. Basisboeken 6*. Amersfoort: ThiemeMeulenhoff.

Furtak, E. M., Seidel, T., Iverson, H., & Briggs, D. C. (2012). Experimental and Quasi-Experimental Studies of Inquiry-Based Science Teaching: A Meta-Analysis. *Review of Educational Research*, *82*(3), 300–329. https://doi.org/10.3102/0034654312457206

Gibson, H. L., & Chase, C. (2002). Longitudinal impact of an inquiry-based science program on middle school students' attitudes toward science. *Science Education*, *86*(5), 693–705. https://doi.org/10.1002/sce.10039

Gooding, D. (1985). "In Nature's School": Faraday as an Experimentalist. In *Farday Rediscovered: Essays on the life and work of Michael Faraday, 1791-1867* (pp. 105–135). Macmillan Press.

Gormally, C., Brickman, P., Hallar, B., & Armstrong, N. (2009). Effects of Inquiry-based Learning on Students' Science Literacy Skills and Confidence. *International Journal for the Scholarship of Teaching and Learning*, *3*(2). https://doi.org/10.20429/ijsotl.2009.030216

Hänze, M., & Berger, R. (2007). Cooperative learning, motivational effects, and student characteristics: An experimental study comparing cooperative learning and direct instruction in 12th grade physics classes. *Learning and Instruction*, *17*(1), 29–41. https://doi.org/10.1016/j.learninstruc.2006.11.004

Hart, C., Mulhall, P., Berry, A., Loughran, J., & Gunstone, R. (2000). What is the purpose of this experiment? Or can students learn something from doing experiments? *Journal of Research in Science Teaching*, *37*(7), 655–675. https://doi.org/10.1002/1098-2736(200009)37:7<655::AID-TEA3>3.0.CO;2-E

Hartley, J. (1973). The effect of pre-testing on post-test performance. *Instructional Science*, *2*(2), 193–214. https://doi.org/10.1007/BF00139871

Hattie, J. (2009). *Visible learning: a synthesis of over 800 meta-analyses relating to achievement*. London ; New York: Routledge.

Herman, W. E., & Pinard, M. R. (2015). Critically Examining Inquiry-Based Learning: John Dewey in Theory, History, and Practice. In P. Blessinger & J. M. Carfora (Eds.), *Innovations in Higher EducationTeaching and Learning* (Vol. 3, pp. 43–62). Emerald Group Publishing Limited. https://doi.org/10.1108/S2055-364120150000003016

Hulsbeek, M., Meijerink, G., Schurink, M., & Wiegman, H. (1999). *Stralen of dralen? haalbaarheidsonderzoek naar het ioniserende stralen practicum (isp)*. SLO.

Janssen, B. (2016). Increased conceptual learning through a guided discovery approach of the ISP. Retrieved from https://dspace.library.uu.nl/handle/1874/335724

Karpicke, J. D., & Blunt, J. R. (2011). Retrieval Practice Produces More Learning than Elaborative Studying with Concept Mapping. *Science*, *331*(6018), 772–775. https://doi.org/10.1126/science.1199327

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*(2), 75–86. https://doi.org/10.1207/s15326985ep4102_1

Magliaro, S. G., Lockee, B. B., & Burton, J. K. (2005). Direct instruction revisited: A key model for instructional technology. *Educational Technology Research and Development*, *53*(4), 41–55. https://doi.org/10.1007/BF02504684

McDaniel, M. A., & Fisher, R. P. (1991). Tests and test feedback as learning sources. *Contemporary Educational Psychology*, *16*(2), 192–201. https://doi.org/10.1016/0361-476X(91)90037-L

Millar, R., & Abrahams, I. (2009). Practical work : Making it more effective. *School Science Review*, *91*(334), 59–64.

Minner, D. D., Levy, A. J., & Century, J. (2010). Inquiry-based science instruction-what is it and does it matter? Results from a research synthesis years 1984 to 2002. *Journal of Research in Science Teaching*, *47*(4), 474–496. https://doi.org/10.1002/tea.20347

Nooijen, T. (2017). *Improving Students' Intrinsic Motivation by Using an Inquiry-based Learning Practical*. Utrecht University. Retrieved from https://dspace.library.uu.nl/handle/1874/353621

Pedaste, M., Mäeots, M., Siiman, L. A., de Jong, T., van Riesen, S. A. N., Kamp, E. T., … Tsourlidaki, E. (2015). Phases of inquiry-based learning: Definitions and the inquiry cycle. *Educational Research Review*, *14*, 47–61. https://doi.org/10.1016/j.edurev.2015.02.003

Pereira, O. S., Rosa, M. D. S., Negretti, C., & Faletti, R. (1998). *Project for the formation of a nuclear consciousness*. International Atomic Energy Agency.

Rosenthal, R. (1994). Parametric measures of effect size. In H. Cooper & L. V. Hedges (Eds.), The handbook of research synthesis. (pp. 231-244). New York: Russell Sage Foundation.

Ryan, R. M., & Deci, E., L. (2000). Self-Determination Theory and the Facilitation of Intrinsic Motivation, Social Development, and Well-Being. *American Psychologist*, *55*(1), 68–78.

Ryder, R. J., Burton, J. L., & Silberg, A. (2006). Longitudinal Study of Direct Instruction Effects From First Through Third Grades. *The Journal of Educational Research*, *99*(3), 179–192. https://doi.org/10.3200/JOER.99.3.179-192

Savelsbergh, E. R., Prins, G. T., Rietbergen, C., Fechner, S., Vaessen, B. E., Draijer, J. M., & Bakker, A. (2016). Effects of innovative science and mathematics teaching on student attitudes and achievement: A meta-analytic study. *Educational Research Review*, *19*, 158–172. https://doi.org/10.1016/j.edurev.2016.07.003

Schreiner, C., & Sjøberg, S. (2007). *Science education and youth's identity construction -two incompatible projects?* Retrieved from http://folk.uio.no/sveinsj/Values-ROSE-Schreiner-Sjoberg.pdf

Shamos, M. H. (1995). *The myth of scientific literacy*. New Brunswick, N.J: Rutgers University Press.

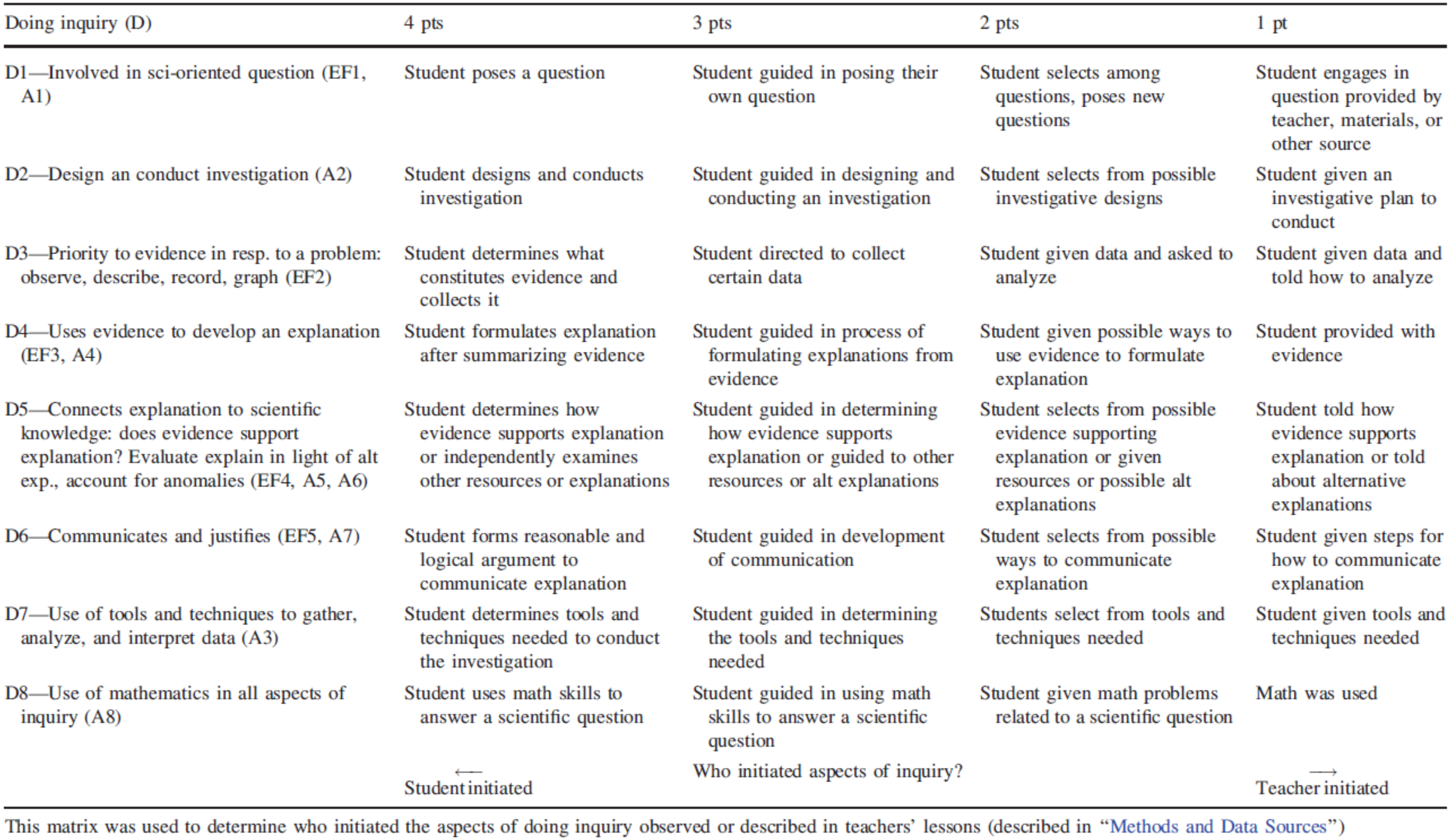
Traub, J. F. (1999). The Unknown and the Unknowable. *The Sciences*, *39*(1), 39–44. https://doi.org/10.1002/j.2326-1951.1999.tb03412.x

Utrecht University. (2018). Uitvoeringsvariant Ioniserende Stralen Practicum. Retrieved from https://stralenpracticum.nl/voor-docenten/uitvoeringsvariant/

Van den Akker, J., Bannan, B., Kelly, A. E., Nieveen, N., & Plomp, T. (2010). *An introduction to educational design research: proceedings of the seminar conducted at the East China Normal University, Shanghai (PR China), November 23-26, 2007*. Enschede: SLO. Retrieved from http://www.slo.nl/downloads/2009/Introduction_20to_20education_20design_20research.pdf/download

Appendices

Appendix A: Matrix level of Inquiry (Capps & Crawford, 2013)



Doing inquiry (D)	4 pts	3 pts	2 pts	1 pt
D1—Involved in sci-oriented question (EF1, A1)	Student poses a question	Student guided in posing their own question	Student selects among questions, poses new questions	Student engages in question provided by teacher, materials, or other source
D2—Design an conduct investigation (A2)	Student designs and conducts investigation	Student guided in designing and conducting an investigation	Student selects from possible investigative designs	Student given an investigative plan to conduct
D3—Priority to evidence in resp. to a problem: observe, describe, record, graph (EF2)	Student determines what constitutes evidence and collects it	Student directed to collect certain data	Student given data and asked to analyze	Student given data and told how to analyze
D4—Uses evidence to develop an explanation (EF3, A4)	Student formulates explanation after summarizing evidence	Student guided in process of formulating explanations from evidence	Student given possible ways to use evidence to formulate explanation	Student provided with evidence
D5—Connects explanation to scientific knowledge: does evidence support explanation? Evaluate explain in light of alt exp., account for anomalies (EF4, A5, A6)	Student determines how evidence supports explanation or independently examines other resources or explanations	Student guided in determining how evidence supports explanation or guided to other resources or alt explanations	Student selects from possible evidence supporting explanation or given resources or possible alt explanations	Student told how evidence supports explanation or told about alternative explanations
D6—Communicates and justifies (EF5, A7)	Student forms reasonable and logical argument to communicate explanation	Student guided in development of communication	Student selects from possible ways to communicate explanation	Student given steps for how to communicate explanation
D7—Use of tools and techniques to gather, analyze, and interpret data (A3)	Student determines tools and techniques needed to conduct the investigation	Student guided in determining the tools and techniques needed	Students select from tools and techniques needed	Student given tools and techniques needed
D8—Use of mathematics in all aspects of inquiry (A8)	Student uses math skills to answer a scientific question	Student guided in using math skills to answer a scientific question	Student given math problems related to a scientific question	Math was used
	⟵ Student initiated	Who initiated aspects of inquiry?		⟶ Teacher initiated

This matrix was used to determine who initiated the aspects of doing inquiry observed or described in teachers' lessons (described in "Methods and Data Sources")

Appendix B: Example universal pre- and post-test

1. *Wat is de betekenis van de 'activiteit' van een radioactief materiaal? Kies één optie uit de onderstaande antwoorden.*
   De 'activiteit' van een radioactief materiaal is…
   ○ het aantal instabiele kernen dat per jaar vervalt
   ○ het aantal instabiele kernen dat per seconde vervalt
   ○ de totale energie van de uitgezonden straling, uitgedrukt in Joule
   ○ de hoeveelheid geabsorbeerde energie, uitgedrukt in Joule/kg.

2. *Polonium-210 is een radioactief materiaal, dat vervalt onder uitzending van* α-straling. Waar bestaat α-straling uit? Kruis *één optie uit de onderstaande antwoorden.*
   ○ Alleen positronen
   ○ Alleen elektronen
   ○ Alleen heliumkernen
   ○ Alleen neutronen
3. *Thorium-233 is een radioactief materiaal, dat vervalt onder uitzending van* ß$^{-}$*-straling. Achter de vragen is een stuk van de tabel 25 uit het BINAS toegevoegd. Hieronder is een begin gemaakt aan de vervalreactie, vul deze verder aan.*

$$^{233}_{90}Th \rightarrow \quad \ldots \quad + \quad \ldots$$

4. *Welke van de onderstaande typen straling, is het schadelijkste bij besmetting? Kies één optie uit de onderstaande antwoorden.*
   ○ α straling
   ○ ß$^{-}$ straling
   ○ γ straling
   ○ Röntgenstraling

5. *Bismuth-210 heeft een halveringstijd van 5 dagen. Na hoeveel dagen is de activiteit afgenomen tot 6,25% van de beginactiviteit? Kies één optie uit de onderstaande antwoorden.*
   De activiteit is afgenomen tot 6,25% van de beginactiviteit na…
   ○ *5 dagen*
   ○ *10 dagen*
   ○ *15 dagen*
   ○ *20 dagen*

6. *Een radioactief materiaal kan* α, ß$^{-}$ en γ*-straling uitzenden. Sorteer deze typen straling* ***van hoogste naar laagste doordringend vermogen****.*
   Hoogste doordringend
   vermogen:……………………………………………………………………………………….
   Gemiddeld doordringend vermogen:………………………………………………………………………………
   Laagste doordringend vermogen:…………………………………………………………………………………

7. Lood wordt regelmatig als bescherming gebruikt, omdat het ioniserende straling absorbeert. Voordat de ioniserende straling de loden plaat ingaat, is de activiteit maximaal. Aan de achterzijde van de plaat is er nog 1/32e deel over van de maximale activiteit. Hoeveel halveringsdiktes gaan er in de loden plaat? Kruis hieronder één antwoord aan en leg uit hoe je tot je antwoord bent gekomen.
   ○ 2 halveringsdiktes
   ○ 3 halveringsdiktes
   ○ 4 halveringsdiktes
   ○ 5 halveringsdiktes

   Leg hier uit hoe je tot je antwoord bij vraag 7 bent gekomen:
   ………
   ………
   ………

Example questions in English

- *What is the meaning of 'activity'? Choose one option of the answers below.* The 'activity of a radioactive material is…
  ○ the number of unstable nuclei that decay per year
  ○ the number of unstable nuclei that decay per second
  ○ the total energy of the emitted radiation, expresses as Joule
  ○ the amount of energy absorbed, expressed in Joule / kg.

- *A radioactive material can emit* α*,* ß$^{-}$ *en* γ*-radiation. Sort these types from highest to lowest penetrating power.*
  Highest penetrating power:……………………………………………………………………………………….
  Average penetrating power:………………………………………………………………………………………
  Lowest penetrating power:…………………………………………………………………………………………

Appendix C: Example experiment-dependent pre- and post-test

Experiment 2A

1. *Een isotoop van bismuth heeft een halveringstijd van 7 dagen. In de hoeveelste week is de activiteit afgenomen onder 10% van de beginactiviteit? Kies één optie uit de onderstaande antwoorden. Wanneer het correcte weeknummer ontbreekt, kies dan de optie die de week het eerst opvolgt.*
   De activiteit is voor het eerst afgenomen onder 10% van de beginactiviteit in…
   ○ *Week 1*
   ○ *Week 2*
   ○ *Week 3*
   ○ *Week 4*

2. *Thorium-233 is een radioactief materiaal, dat vervalt onder uitzending van ß*$^-$*-straling. Achter de vragen is een stuk van de tabel 25 uit het BINAS toegevoegd. Hieronder is een begin gemaakt aan de vervalreactie, vul deze verder aan.*

$$^{233}_{90}\ldots \rightarrow \quad \ldots \; + \; \ldots$$

3. *Een radioactief gas wordt ingeademd door een volwassene. Deze stof zendt zowel α, ß*$^-$ *als γ straling uit en de volwassene ontvangt een dosis van 1 Gy (Gray) in de longen. Bij welk type straling ontvangt een volwassene de hoogste dosisequivalent bij inademing? Kies één optie uit de onderstaande antwoorden.*
   ○ α straling
   ○ ß$^-$ straling
   ○ γ straling
   ○ α, ß$^-$ als γ straling zijn even schadelijk bij inademing

Appendix D: Core question student group interview in Dutch and English

Leerling welkom heten. Alle antwoorden die je (leerling) geeft zijn goed en de antwoorden die je hier geeft hebben geen invloed op schoolcijfers. Doel van het onderzoek is om te onderzoeken wat je als leerling leert van het Ioniserende Stralen Practicum en we hopen het practicum te kunnen verbeteren met de data. Alle gegevens uit het onderzoek worden anoniem behandeld en alleen gebruikt voor mijn eindscriptie en de bijbehorende presentaties. We zijn ongeveer tien minuten bezig met het interview, maar je kan op ieder moment stoppen als je dat wilt. Heb je nog vragen vooraf?

1. Hebben jullie al eerder les gehad over het onderwerp ioniserende straling? Zo ja, wanneer?
    a. Hebben jullie al eerder een toets gehad over ioniserende straling/radioactieve materialen?
    b. Indien toets na ISP, heb je wat gehad aan het ISP in de voorbereiding op de toets? Zo ja/nee, waarom wel/niet?
    c. Was het door het ISP makkelijker/moeilijker om opdrachten uit het boek te maken?
        i. Voorbeelden?
2. Wat kan je je nog herinneren van het practicum?
    a. Hebben jullie een werkplan geschreven tijdens een van de experimenten?
        i. Zo ja: wat vond je hiervan?
        ii. Zo nee: door naar volgende vraag
3. Wat vond je van het practicum?
    a. Welke experimenten heb je uitgevoerd?
    b. Waren de experimenten ingedeeld of kon je zelf kiezen?
4. Hoeveel tijd heb je in de klas of thuis besteed aan het voorbereiden van het practicum?
    a. Indien open → Wat heb denk je geleerd te hebben tijdens het maken van het werkplan?
    b. Welk experiment vond je het leukste en minst leuk?
    c. Waren jullie ingedeeld in groepjes, of kon je zelf groepen vormen?

5. Wat denk je geleerd te hebben van de experimenten?
    a. Wat vond je moeilijk bij het uitvoeren van open/gesloten experiment?
    b. Hoe vaak is de docent bij jullie langs geweest om vragen te beantwoorden?
    c. Wat vond je makkelijk bij het uitvoeren van open/gesloten experiment?
        i. Hoe kwam het dat je makkelijk vond? Wist je bijv. de theorie al, al vooraf antwoorden gedeeld met andere leerlingen of iets anders?

Welcome all students. All answers that students give, do not influence the school grades. Aim of research is to investigate students' learning in the IRP. It is hoped that the practical can be improved with this research. All data of this research, will be treated anonymously and will only be used for my Master thesis and the relevant presentations. The interview will take about ten minutes, but one can stop at any moment. Does student have questions upon the start of the interview?

1. Did you have earlier physics classes on ionizing radiation? If so, when?
    a. Did you already take a test about ionizing radiation/radioactive materials.
    b. If the test is after IRP, do you think IRP was useful to prepare for the test? If (not) so, why?
    c. Was it because of IRP easier/more difficult to make the assignments in the book?
        i. Examples?
2. What do you recall of the practical?
    a. Did you write a workplan during one of the experiments?
        i. If yes: what is your opinion on this?
        ii. If no, proceed to next question.
3. What is your opinion on the practical?
    a. Which experiments did you execute?
    b. Were the experiments pre-assigned or were you able to choose your own?

4. How much time did you spend in class or at home to prepare for the practical?
    a. If student did do open/GIBL experiment → What do you think to have learned from writing the work plan?
    b. Which experiment did you fancy most and least?
    c. Were you grouped beforehand, or were you able to make your own groups?

5. What do you think to have learned from the experiments?
    a. What did you find difficult when executing the open/closed experiment?
    b. How often has the teacher helped you to answer questions?
    c. What did you find easy when executing the open/closed experiment?
        i. Because of what do think that is was easy? Did you for instance know theory beforehand, did you share answers with other students, something else?