Go-Lab

Global Online Science Labs for Inquiry Learning at School

Collaborative Project in European Union's Seventh Framework Programme

Grant Agreement no. 317601



Deliverable D8.3

First trial report

Editors	Evita Tasiopoulou
	Ton de Jong
Date	31-10-2015
Dissemination Level	Public
Status	Final



© 2015, Go-Lab consortium

The Go-Lab Consortium

Beneficiary Number	Beneficiary name	Beneficiary short name	Country
1	University Twente	UT	The Netherlands
2	Ellinogermaniki Agogi Scholi Panagea Savva AE	EA	Greece
3	Ecole Polytechnique Fédérale de Lausanne	EPFL	Switzerland
4	EUN Partnership AISBL	EUN	Belgium
5	IMC AG	IMC	Germany
6	Reseau Menon E.E.I.G.	MENON	Belgium
7	Universidad Nacional de Educación a Distancia	UNED	Spain
8	University of Leicester	ULEIC	United Kingdom
9	University of Cyprus	UCY	Cyprus
10	Universität Duisburg-Essen	UDE	Germany
11	Centre for Research and Technology Hellas	CERTH	Greece
12	Universidad de la Iglesia de Deusto	UDEUSTO	Spain
13	Fachhochschule Kärnten – Gemeinnützige Privatstiftung	CUAS	Austria
14	Tartu Ulikool	UTE	Estonia
15	European Organization for Nuclear Research	CERN	Switzerland
16	European Space Agency	ESA	France
17	University of South Wales	USW	United Kingdom
18	Institute of Accelerating Systems and Applications	IASA	Greece
19	Núcleo Interactivo de Astronomia	NUCLIO	Portugal
20	Cardiff University	CU	United Kingdom

Contributors

Name	Institution
Ton de Jong	UT
Siswa van Riesen	UT
Jakob Sikken	UT
Anjo Anjewierden	UT
Lars Bollen	UT
Ellen Wassink-Kamp	UT
Bas Kollöffel	UT
Hannie Gijlers	UT
Erik Bong	UT
Casper de Jong	UT
Mario Mäeots	UTE
Margus Pedaste	UTE
Leo Siiman	UTE
Urmas Heinaste	UTE
Zacharias Zacharia	UCY
Tasos Hovardas	UCY
Nikoletta Xenofontos	UCY
Anna Fiakkou	UCY
Evita Tasiopoulou	EUN
Agueda Gras	EUN
Gina Mihai	EUN
Fani Stylianidou	EA
Eleftheria Tsourlidaki	EA
Internal reviewers	
Sofoklis Sotiriou	EA
Danilo Garbi Zutin	CUAS

Legal Notices

The information in this document is subject to change without notice. The Members of the Go-Lab Consortium make no warranty of any kind with regard to this document, including, but not limited to, the implied warranties of merchantability and fitness for a particular purpose. The Members of the Go-Lab Consortium shall not be held liable for errors contained herein or direct, indirect, special, incidental or consequential damages in connection with the furnishing, performance, or use of this material. The information and views set out in this deliverable are those of the author(s) and do not necessarily reflect the official opinion of the European Union. Neither the European Union institutions and bodies nor any person acting on their behalf may be held responsible for the use which may be made of the information contained therein.

Executive Summary

This deliverable presents the results of a first set of evaluations in Go-Lab. The deliverable follows the research questions that were presented in Go-Lab deliverable D8.1. These research questions were divided into three clusters: questions aimed at students, teachers, and organisations. In each of these clusters the focus of the questions is on a specific Go-Lab intervention (e.g., ILSs or elements of ILSs, Golabz, and/or the Go-Lab authoring facilities) and measure different types of outcomes (mainly knowledge and inquiry skills). The three different clusters of research questions are presented in separated parts in this deliverable.

The studies that evaluated students covered various tools from the Go-Lab set of tools, different configurations of ILSs, different age groups, different labs and domains, and different assessment methods. All studies were conducted "in vivo" meaning that we have always conducted the studies in real classes under realistic circumstances. This sometimes gave specific challenges, such as hampering internet connections, and made that research conditions, in terms of time allowed for the study or number of subjects, were not always optimal. We should also consider that in all of the studies these were most probably the students' (and teachers') first encounters with inquiry learning. Despite this some general conclusions can be drawn.

First, in all of the studies in which knowledge was measured (and where no internet issues appeared) we have seen a significant increase in scores on knowledge tests. There have been no comparisons with other, more traditional, approaches (this will be done in Y4 of the project) but in any case offering online labs makes that students learn about the domain. In the case of inquiry skills such an increase was not always measured, which can be explained from the fact that for those skills to develop properly, we need more time and a prolonged training of the skills.

For a specific set of tools we could find direct effects on students' acquisition of knowledge and inquiry skills. The conclusion tool, the hypothesis scratchpad, and the experiment design tool all showed specific effects in some of the studies. In some case a comparison was made to a condition in which the tool was not offered, in other cases a comparison between a fully specified tool and a tool that was rather "empty" (such as a hypothesis scratchpad without pre-defined terms) was made. Studies on the hypothesis scratchpad showed that most probably offering pre-defined concepts in the tool was beneficial for learning compared to letting students configure these terms themselves. In the study on the concept map, no differences between including and not offering the concept map were found but in this case the concept map did not have any predefined terms in the pull down menu. If we extend the results of the studies with the hypothesis scratchpad to the concept map, we might expect better effects when these terms will be offered in the concept map's pull down menu. So, a second conclusion n might be that tools often support students but that they might need to be filled with domain terms in order to create an effect.

A third conclusion might be that there is no "one size fits all" solution. Several of our studies (especially the ones on the experiment design tool (EDT)) show that tools are specifically effective (differentially for knowledge and inquiry skills) for students with lower prior knowledge or for younger students. The studies with the EDT also show that in these cases there is an interaction with the difficulty of the domain involved, effectiveness of a tool for the students who need this might be more distinct when the domain gets more difficult.

Finally, also the specific configuration of a tool might matter. The studies with the data viewer, for example showed, that when the data viewer automatically incorporated data an improvement of inquiry skills was reached, whereas students who imported data themselves gained better conceptual knowledge.

When deciding which tools to include in an ILS and how these tools should be configured, it is important to consider students' age, level of education, and prior knowledge, as well as the difficulty of the domain, and whether the goal is for students to gain conceptual knowledge, to acquire inquiry skills or both.

The (Phase B) teacher evaluation has been based on the analysis of pre- and post-data of 130 teachers. A close look at those teachers teaching and technical skills reveals that our sample has been composed of advanced teachers who were not only very much interested in the use of online laboratories but had also quite developed pedagogical and technological skills. Their background knowledge in combination with their interests and the support mechanisms offered by Go-Lab, had an impact on those teachers' knowledge and motivation. What is particularly interesting is that although the majority of these teachers intended to mostly use ready-made ILSs and online laboratories, at the end of Pilot Phase B, most of them have started creating their own ILS.

When it comes to the evaluation of organisations, the analysis of teachers', headmasters and policy makers' interviews reveals that the applications and impact of Go-Lab expand to multiple levels. Within schools, STEM teachers are teaming up with colleagues from other disciplines and use Go-Lab to develop interdisciplinary activities (i.e., combining STEM topics with language learning, schools' collaboration, and special needs education). Headmasters begin to realise how the use of online laboratories can contribute to both their teachers' and students' development in an easy and cost effective way. At the same time, policy makers' understanding on the use of online laboratories is also strengthened which opens possibilities for more support regarding the Go-Lab implementations.

Table of Contents

PART I: STUDENT EVALUATION 11

1 INTRODUCTION; OVERVIEW OF GUIDANCE AND STUDIES WITH STUDENTS 13

1.1 GUIDANCE 13
1.1.1 Orientation: Concept mapper 13
1.1.2 Conceptualisation: Hypothesis and question scratchpad 14
1.1.3 Investigation: Experiment design tool 14
1.1.4 Investigation: Data viewer tool 15
1.1.5 Investigation: Experimental error tool 16
1.1.6 Conclusion: Conclusion tool 16
1.1.7 Discussion: Reflection tool 17
1.1.8 General: Quiz tool 18
1.2 STRUCTURE OF THE STUDIES 18
1.3 REFERENCES 19

2 SCAFFOLDING STUDENTS' ORIENTATION: CONCEPT MAPPER 21

2.1 ABSTRACT 21
2.2 INTRODUCTION 21
2.3 EXPERIMENT 1 22
2.3.1 Method 22
2.3.2 Results 24
2.3.3 Conclusion and discussion 25
2.4 EXPERIMENT 2 25
2.4.1 Method 26
2.4.2 Results 27
2.4.3 Conclusion and Discussion 27
2.5 REFERENCES 28

3 SCAFFOLDING STUDENTS' CONCEPTUALISATION WITH THE QUESTION AND HYPOTHESIS SCRATCHPAD 29

3.1 ABSTRACT 29
3.2 INTRODUCTION 29
3.3 METHOD 29
3.3.1 Participants 30
3.3.2 Materials 30
3.3.3 Assessment 32
3.4 PROCEDURE 32
3.5 RESULTS 32
3.6 DISCUSSION AND CONCLUSION 33

4 HOW MANY WORDS ARE ENOUGH? ASSESSING THREE DIFFERENT CONFIGURATIONS OF THE HYPOTHESIS SCRATCHPAD 35

4.1 ABSTRACT 35
4.2 INTRODUCTION 35
4.3 METHOD 36
4.3.1 Participants 37
4.3.2 Materials 37
4.4 PROCEDURE 40
4.5 RESULTS 41

4.5.1 Results in the UK *41*
4.5.2 Results in the Netherlands *43*
4.5.3 Results in Estonia *45*
4.5.4 Overview of results across all research settings *48*
4.6 DISCUSSION, CONCLUSIONS AND IMPLICATIONS FOR FUTURE RESEARCH 48
4.7 REFERENCES 49

5 SCAFFOLDING STUDENTS' EXPERIMENTATION 51

5.1 ABSTRACT 51
5.2 INTRODUCTION 51
5.3 METHOD 53
5.3.1 Participants *54*
5.3.2 Materials *54*
5.4 PROCEDURE 57
5.5 RESULTS 57
5.6 CONCLUSION AND DISCUSSION 60
5.7 REFERENCES 61

6 GUIDING STUDENTS IN DESIGNING PROPER EXPERIMENTS; DOES ABILITY LEVEL MAKE A DIFFERENCE? 65

6.1 ABSTRACT 65
6.2 INTRODUCTION 65
6.3 METHOD 66
6.3.1 Participants *67*
6.3.2 Materials *67*
6.3.3 Assessment *68*
6.4 PROCEDURE 68
6.5 RESULTS AND CONCLUSIONS 69
6.6 CONCLUSION AND DISCUSSION 70
6.7 REFERENCES 70

7 SUPPORTING PLANNING AND CONDUCTING EXPERIMENTS 73

7.1 ABSTRACT 73
7.2 INTRODUCTION 73
7.3 METHOD 76
7.3.1 Participants *77*
7.3.2 Materials *77*
7.3.3 Procedure *79*
7.4 RESULTS 80
7.5 CONCLUSION AND DISCUSSION 82
7.6 REFERENCES 82

8 SCAFFOLDING EXPERIMENTATION IN VIRTUAL LABS: THE CASE OF THE GO-LAB EXPERIMENT DESIGN TOOL 87

8.1 ABSTRACT 87
8.2 INTRODUCTION 87
8.3 METHOD 88
8.3.1 Participants *89*
8.3.2 Materials *89*
8.4 PROCEDURE 93
8.5 RESULTS 93
8.5.1 Content Knowledge test *93*

8.5.2 Inquiry skills test 93
8.5.3 Summary of results 94
8.6 DISCUSSION, CONCLUSIONS AND IMPLICATIONS FOR FUTURE RESEARCH 94
8.7 REFERENCES 95

9 DOES USING THE EXPERIMENTAL ERROR TOOL CHANGE STUDENTS' UNDERSTANDING ABOUT 'HOW SCIENCE WORKS'? 97

9.1 ABOUT THE STUDY 97
9.2 METHOD 98
9.2.1 Evaluation instrument 98
9.2.2 Participants 100
9.2.3 The lessons 100
9.2.4 Data Collection and Analysis 102
9.2.5 Discussion of findings 107
9.3 REFERENCES 108

10 THE EFFECT OF THE GO-LAB CONCLUSION TOOL ON STUDENTS' SCIENCE LEARNING 109

10.1 ABSTRACT 109
10.2 INTRODUCTION 109
10.3 METHODS 110
10.3.1 Participants 111
10.3.2 Materials 111
10.4 PROCEDURE 114
10.5 RESULTS 115
10.5.1 Content Knowledge test 115
10.5.2 Inquiry skills test 115
10.6 DISCUSSION, CONCLUSIONS AND IMPLICATIONS FOR FUTURE RESEARCH 116
10.7 REFERENCES 116

11 CONSTRUCTING GRAPHS BY MEANS OF THE DATA VIEWER: COMPARISON BETWEEN TWO DIFFERENT CONFIGURATIONS OF THE TOOL IN TERMS OF THEIR EFFECTS ON STUDENT KNOWLEDGE AND SKILLS 119

11.1 ABSTRACT 119
11.2 INTRODUCTION 119
11.3 METHOD 120
11.3.1 Participants 120
11.3.2 Materials 120
11.4 PROCEDURE 125
11.5 RESULTS 125
11.6 DISCUSSION, CONCLUSIONS AND IMPLICATIONS FOR FUTURE RESEARCH 126
11.7 REFERENCES 127

12 SCAFFOLDING STUDENTS' REFLECTION WITH THE REFLECTION TOOL 129

12.1 ABSTRACT 129
12.2 INTRODUCTION 129
12.3 METHOD 129
12.3.1 Participants 129
12.3.2 Procedure 130
12.3.3 Materials 130
12.3.4 Assessment 132
12.4 RESULTS AND DISCUSSION 132
12.5 CONCLUSION 134

12.6 REFERENCES 134

13 OVERALL CONCLUSIONS FROM THE STUDIES WITH STUDENTS 137

13.1 SET-UP AND RESULTS STUDIES 137
13.1.1 Concept map 137
13.1.2 Hypothesis and questioning scratchpad 137
13.1.3 Experiment Design Tool 138
13.1.4 Experimental Error Tool 138
13.1.5 Conclusion Tool 138
13.1.6 Data viewer 139
13.1.7 Reflection tool 139
13.2 OVERALL CONCLUSION 139

PART 2: TEACHER EVALUATION 141

14 TEACHER EVALUATION 143

14.1 INTRODUCTION 143
14.2 DATA SAMPLE 143
14.3 PRE-QUESTIONNAIRE EVALUATION DATA 144
14.3.1 How do teachers intend to use Go-Lab? 145
14.3.2 How experienced are teachers regarding their science teaching 145
14.3.3 How do teachers rate their technological skills 147
14.3.4 How do teachers understand IBSE scenarios 147
14.3.5 How often do teachers intend to use specific parts of Go-Lab 149
14.4 POST-QUESTIONNAIRE EVALUATION DATA 150
14.4.1 How have teachers used Go-Lab? 151
14.4.2 How experienced are teachers in teaching science? 151
14.4.3 How do teachers rate their technological skills? 152
14.4.4 How do teachers understand IBSE scenarios? 153
14.4.5 Teachers' views on inquiry based science 154
14.4.6 How often do teachers use specific parts of Go-Lab 155
14.4.7 Go-Lab use and usability 156
14.4.8 Which is the most useful component of the Go-Lab portal for science teachers and why? 156
14.4.9 Which is the least useful component of the Go-Lab portal for science teachers and why? 157
14.5 CONCLUSIONS 157

PART 3: EVALUATION OF THE ORGANISATION 159

15 ORGANIZATION EVALUATION 161

15.1 INTRODUCTION 161
15.2 INTERVIEWS 161
15.2.1 Teachers' interviews 161
15.2.2 Conclusions 187
15.2.3 Conclusions 195
15.3 CASE STUDIES 196
15.4 REFERENCES 196
APPENDIX 1 - GO-LAB CASE STUDY PROTOCOL TEMPLATE 197
15.5 CASE STUDY TEMPLATE 200

Part I: Student evaluation

1 Introduction; overview of guidance and studies with students

This part of the deliverable reports a number of studies on the use of Go-Lab scaffolds (or tools) by students. In this first phase of the Go-Lab evaluation we have focused on the effects of the Go-Lab scaffolds since they are designed specifically for the Go-Lab ILSs. Data we gather may help to assess their effectiveness and provide us with suggestions of how to improve them. In this, introductory, part we will first recapitulate an overview of the available scaffolds, then present the structure of the studies that have been conducted. After that each study will be presented individually.

1.1 Guidance

Characteristic for Go-Lab ILSs is that they provide students with integrated guidance. In deliverable D1.1 we have listed the different types of guidance applied in Go-Lab and in internal deliverable G1.3 we have presented more specific descriptions of one type of guidance, the *scaffolds*. The scaffolds are described from a more technical perspective in D5.5. In D1.1. we have defined scaffolds as: “Scaffolds are tools that help students perform a learning process by supporting the dynamics of the activities involved. Scaffolds often provide students with the components of the process and thus structure the process” (de Jong, 2013, p. 16). The scaffolds (or tools as we also call them) that we have introduced in Go-Lab and that are evaluated in the studies presented in this deliverable are summarized in the following sections.

1.1.1 Orientation: Concept mapper

To help students revive their prior knowledge and give them a structured start of the inquiry process the orientation phase may offer them a concept map. Overall, concept maps (Novak, 1990) are found to be good facilitators of the learning process in supporting higher order cognitive skills (e.g., Bramwell-Lalor & Rainford, 2014), to guide students through a domain (Hagemans, van der Meij, & de Jong, 2013) or to support collaboration between students (Gijlers & de Jong, 2013). In Go-Lab the concept map can include predefined concepts and relations and students. In the particular study reported in this deliverable the concept maps was generated by students themselves (with predefined concepts and relations present). Go-Lab also has the facility, however, to provide students with readymade (partial) concept maps which may also be a way to help students in the learning process (Chang, Sung, & Chen, 2001, 2002). Figure 1.1 shows an example of a Go-Lab concept map.

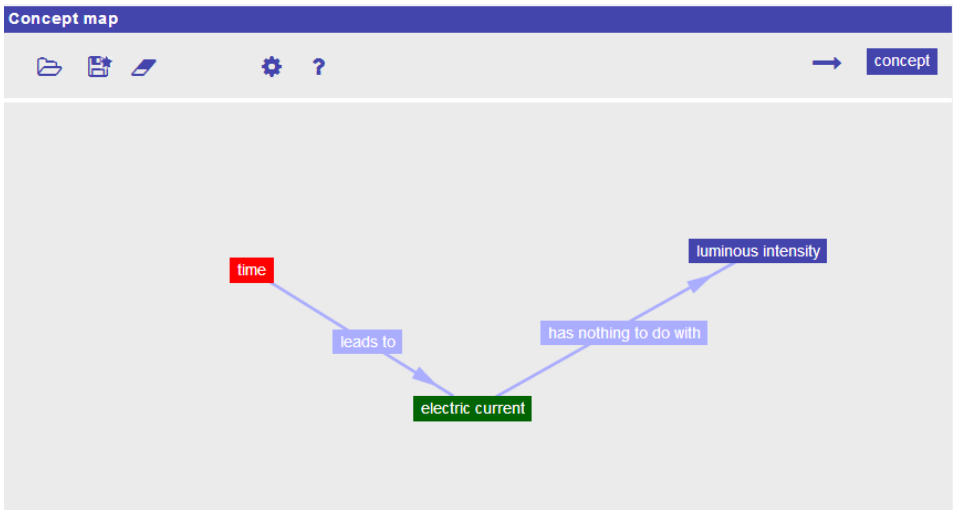


Figure 1.1. Example of a Go-Lab concept map.

1.1.2 Conceptualisation: Hypothesis and question scratchpad

The hypothesis scratchpad offers students a structured way to create hypotheses. The scratchpad offers (by the teacher) pre-defined elements and relations and students can add their own elements and relations as well. This scaffold is based on older work by van Joolingen and de Jong (1993). Teachers may also decide to give students a set of predefined completed hypotheses to start from (see e.g., Njoo & de Jong, 1993). Figure 1.2 shows the hypothesis scratchpad.

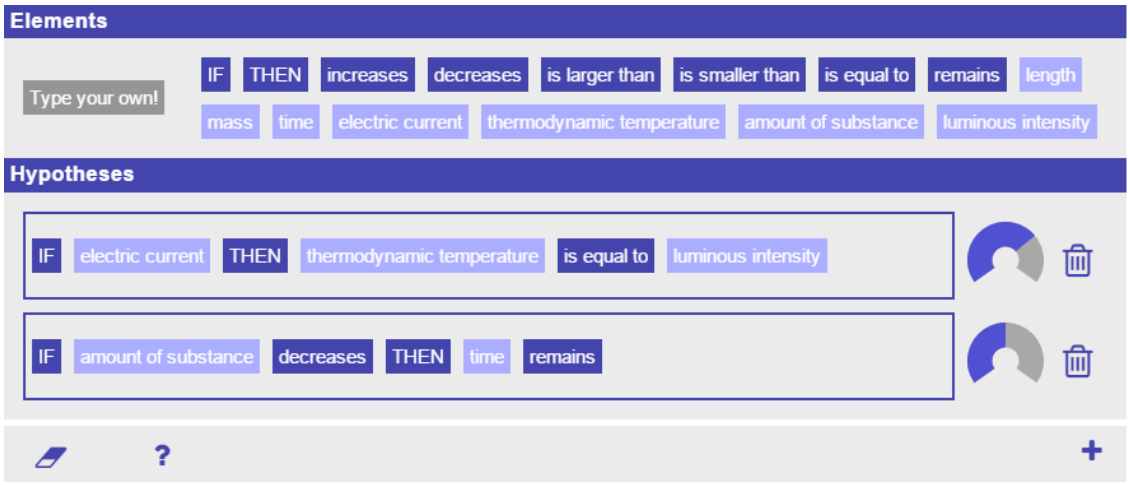


Figure 1.2. Example of a Go-Lab hypothesis scratchpad.

An alternative to the hypothesis scratchpad is the "question scratchpad". The question scratchpad offers a more open way of creating statements but still gives a structure by means of the predefined terms that are available for the students. This question scratchpad can be used when students have more open issues to explore. Figure 1.3 gives an example of the question scratchpad.

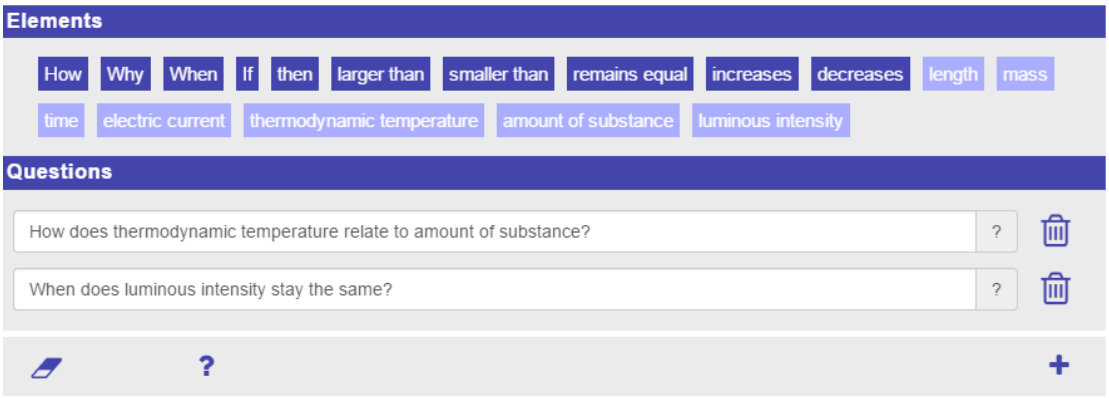


Figure 1.3. The Go-Lab question scratchpad.

1.1.3 Investigation: Experiment design tool

Creating an informative and unconfounded experiment is a challenge for many students (Arnold, Kremer, & Mayer, 2014; Klahr, Zimmerman, & Jirout, 2011; Morgan & Brooks, 2012). The experiment design tool (EDT) created for Go-Lab is a new type of tool that helps students by presenting them an overview of variables and a means to structure these in well-designed experiments. Figure 1.4 shows an example of the EDT as used in the "Guppies' domain.

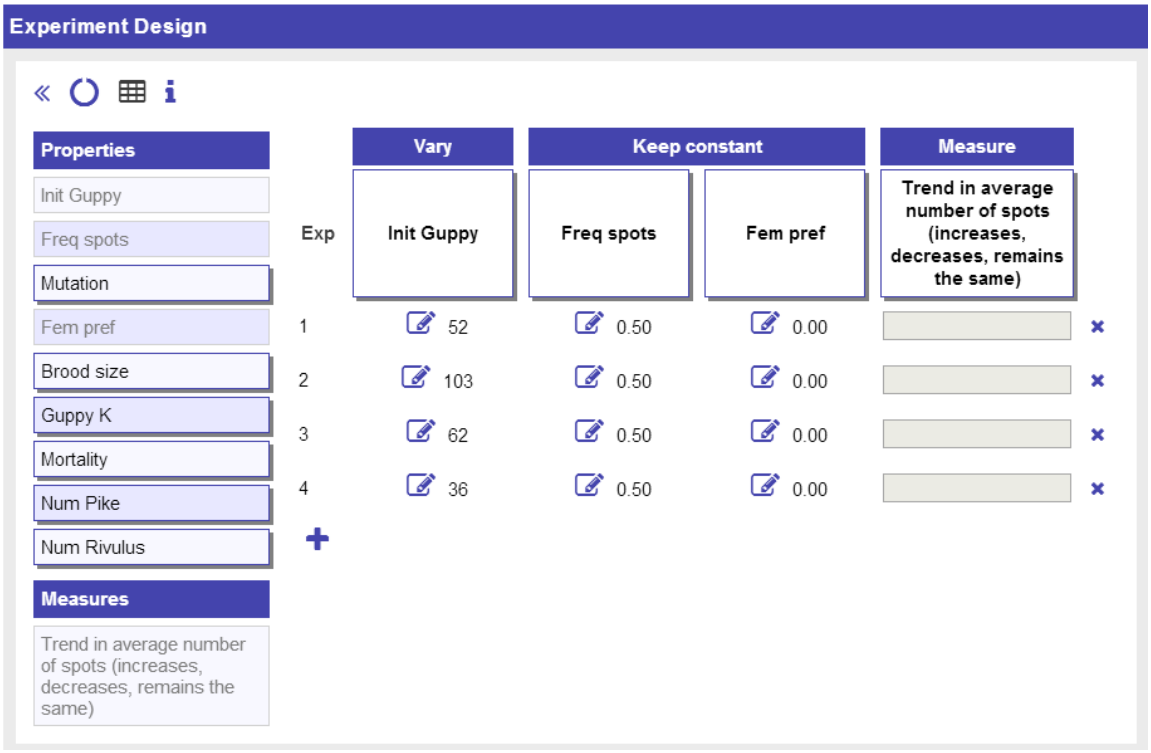


Figure 1.4. Example of the Go-Lab Experiment Design Tool (EDT).

1.1.4 Investigation: Data viewer tool

Interpreting data and finding trends in data is a skill that is very central for inquiry learning but often not well valued. It appears though that students often have real trouble making sense of data and seeing them in the right way (see e.g., Eckhardt, Urhahne, Conrad, & Harms, 2013). The data viewer tool allows students to plot variables from their experiments against each other and see the relations in different representations. Figure 1.5 shows an example of the data viewer tool, here the experiment had two variables that are plotted against each other and a bar chart has been selected to view the relation between the variables.

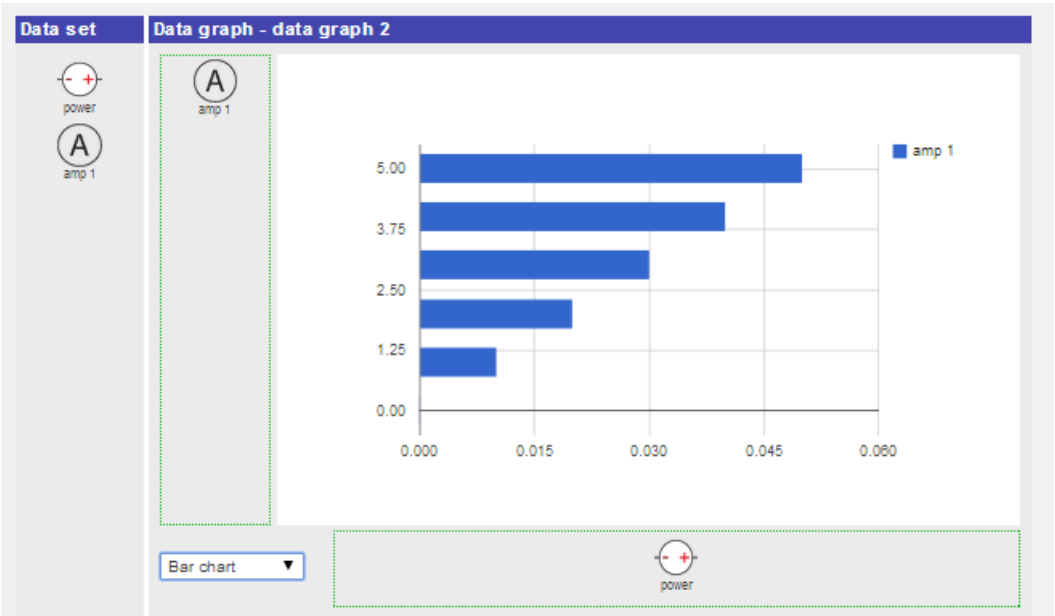


Figure 1.5. Example of the Go-Lab data viewer tool.

1.1.5 Investigation: Experimental error tool

Virtual experiments normally do produce neat data without any measurement errors but in real experiments (and thus also in remote experiments) measurement (random and systematic) errors play a crucial role. Research shows that students are not always aware of these issues (see e.g., Kanari & Millar, 2004). The Go-Lab experimental error tool helps students to get an idea of what plays a role in interpreting data that may be effected by measurement errors and helps them to make the appropriate calculations. The experimental error tool allows students to calculate experimental errors that stem from real experimental setups. Using this tool, students may learn about the different sources of error that occur when performing experiments and about the different types of errors that can be calculated so as to decide whether an experiment is precise and accurate.

Figure 1.6 shows an example of a part of the experimental error tool, the upper parts are more theoretical and explain the students the background of experimental errors.

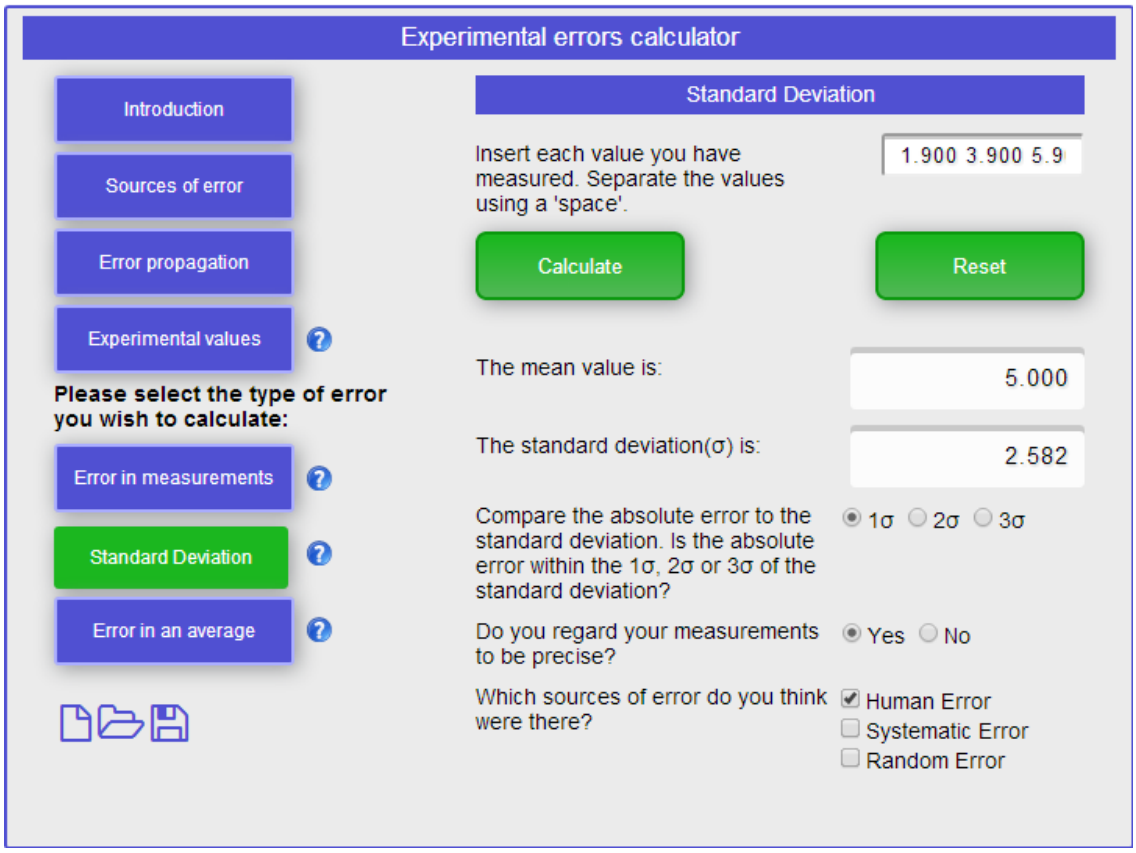


Figure 1.6. Screenshot of the Go-Lab experimental error tool.

1.1.6 Conclusion: Conclusion tool

After having collected and interpreted data these data have to come in touch with the original questions and/or hypotheses a student had. This is the process of drawing conclusions (de Jong, 2006) that is again a focal process in inquiry (Scanlon, Anastopoulou, Kerawalla, & Mulholland, 2011). The Go-Lab conclusion tool facilitates this process by letting students select a question from the set of questions and/or a hypothesis from the set of hypotheses and connect this question and/or hypothesis with student observations (as noted down in the observation tool) and/or saved explorations from the data viewer. The idea behind the conclusion tool is that students can systematically evaluate their hypotheses and questions. For hypotheses they can adapt their original confidence in each hypothesis in the conclusion tool. In case, students do not find any

data or observations related to the question or hypothesis they can go back to experiment to generate more data.

Figure 1.7 shows the Go-Lab conclusion tool in action.

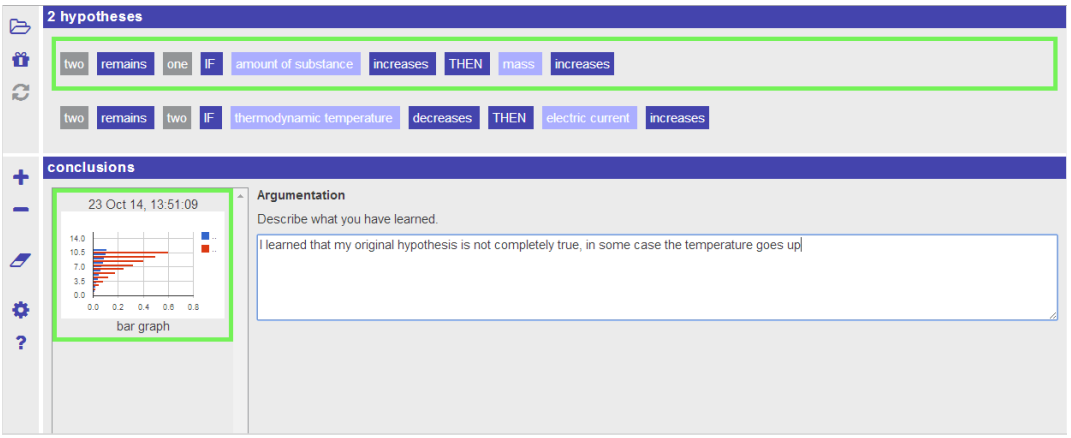


Figure 1.7. The Go-Lab conclusion tool.

1.1.7 Discussion: Reflection tool

Reflection is a key aspect of learning. By reflection learners try to think at a meta-level on the knowledge acquired or the process gone through (Land & Zembal-Saul, 2003). In Go-Lab we focus on a number of different aspects to reflect on both on the product and the process side. For example, a tool showing aggregated concept maps (a kind of average concept map of all students) that individual students can use to compare their own concept map to. A reflection tool that focuses on the process is a tool that shows students their own time spending in the different phases of an ILS and present a “norm” (indicated by the teacher). Students can then compare their own distribution of time against this norm. Figure 1.8 presents a screendump of this reflection tool.

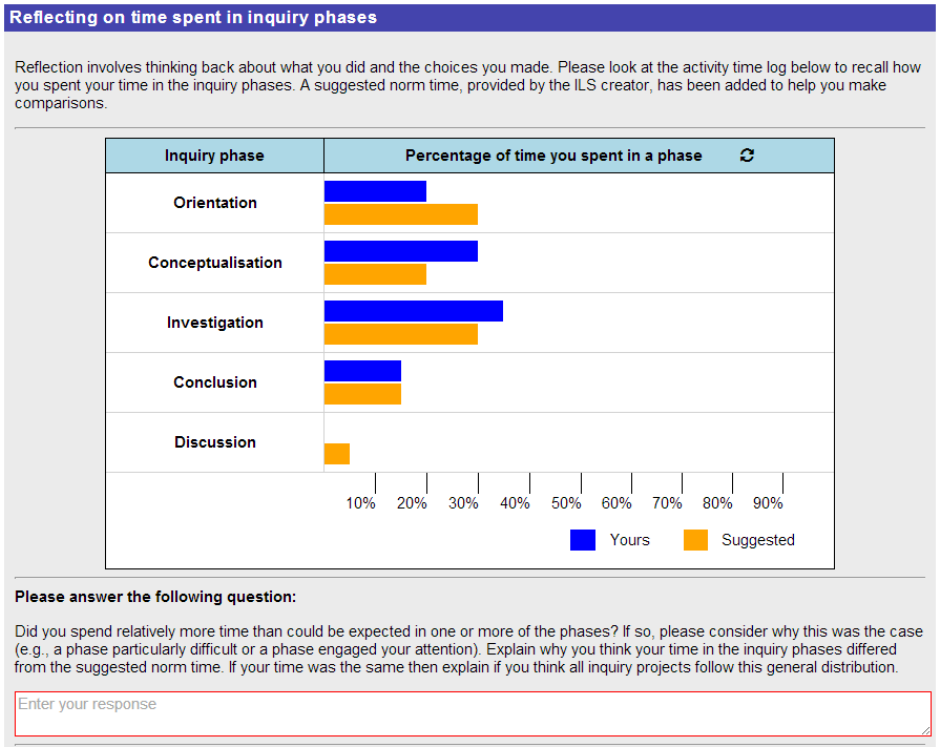


Figure 1.8. Example of the time spent reflection tool.

1.1.8 General: Quiz tool

Asking students for their knowledge and insights may have different functions in an inquiry process, it may help to revive prior knowledge or as a kind of formative testing of developing understanding. With the Go-Lab quiz tool a teacher can design a multiple choice test and define the feedback for each alternative. Figure 1.9 shows part of a Go-Lab quiz.

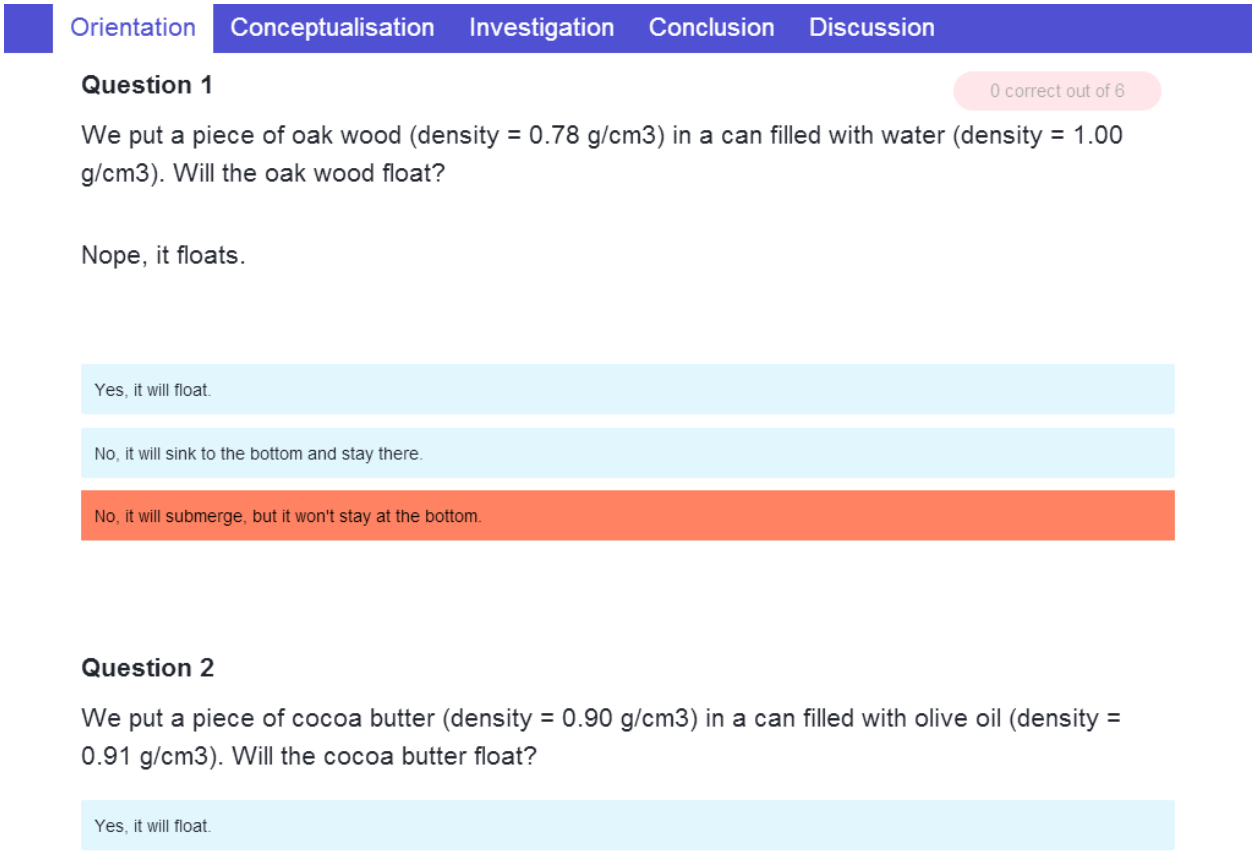


Figure 1.9. Part of a Go-Lab quiz.

1.2 Structure of the studies

All studies follow a classical experimental pre-test – intervention – post-test design with different experimental conditions. One experimental condition involves a lab/ILS with the scaffold(s) under study included, other conditions lack this/these scaffold(s) and in some cases include alternative forms of scaffolding or guidance. In almost all cases measurements of domain knowledge are taken, and, if relevant measures of inquiry skills, attitude, motivation, and knowledge of Nature of Science are included. Interventions were relatively brief so overall we did not expect a large effect on inquiry skills, for these to develop more prolonged interventions seem necessary.

Students were assigned randomly to conditions if possible, in other cases quasi-experimental designs were used. In one case (Chapter 4) we were able to perform what we have called a “concurrent replication” meaning that the same study was conducted at several places in a, somewhat, concurrent way. This gave us the opportunity to compare results over different contexts, labs and types of students.

Table 1.1 presents an overview of the studies that were conducted. The chapters to follow each report on one of these studies.

Table 1.1
Overview of Go-Lab experimental studies with student

Scaffold	Lab	N*	Type of student	Measurement	Partner	Chapter
Concept mapper	Splash	131	Pre-university track	Domain knowledge	UT	2
Hypothesis scratchpad and question tool	Splash, Guppies	43	Splash: Grade 7, Guppies: Grade 11,	Inquiry skills	UTE	3
Hypothesis scratchpad	Splash	385	Differing levels	Domain knowledge Inquiry skills	UT ULEIC UCY UTE	4
Experiment design tool	Splash	120	Pre-university track	Domain knowledge	UT	5
Experiment design tool	Splash	210	Pre-university track	Domain knowledge	UT	6
Experiment design tool	Splash	173	Pre-university track	Domain knowledge	UT	7
Experiment design tool	Splash	49	Grade 5 and Grade 8	Domain knowledge Inquiry skills	UCY	8
Experimental errors tool	Radioactivity	49	Grade 11	Inquiry skills	EA	9
Conclusion tool	Electricity	27	Grade 10	Domain knowledge, inquiry skills	UCY	10
Data viewer tool	Electricity	30	Grade 6 and Grade 10 or 11	Domain knowledge	UCY	11
Reflection tool	Acid based solutions and pH Scale	40	Middle school	Reflection content and level	UTE	12

* In some studies the number of students (N) in the final data analyses are smaller than the number of students who started in the study (which is the number displayed in the table) due to sickness, loss of data etc. The numbers include students who participated in control groups.

1.3 References

Arnold, J. C., Kremer, K., & Mayer, J. (2014). Understanding students' experiments—what kind of support do they need in inquiry tasks? *International Journal of Science Education, 36*, 2719-2749. doi: 10.1080/09500693.2014.930209

Bramwell-Lalor, S., & Rainford, M. (2014). The effects of using concept mapping for improving advanced level biology students' lower- and higher-order cognitive skills. *International Journal of Science Education, 36*, 839-864. doi: 10.1080/09500693.2013.829255

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

10.1002/tea.3660220208

Chang, K. E., Sung, Y. T., & Chen, S. F. (2001). Learning through computer-based concept mapping with scaffolding aid. *Journal of Computer Assisted Learning, 17*, 21-33. doi: 10.1111/j.1365-2729.2001.00156.x

Chang, K. E., Sung, Y. T., & Chen, S. F. (2002). The effect of concept mapping to enhance text comprehension and summarization. *The Journal of Experimental Education, 71*, 5-23. doi: 10.1080/00220970209602054

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

Dillashaw, F. G., & Okey, J. R. (1980). Test of the integrated science process skills for secondary science students. *Science Education, 64*, 601-608. doi: 10.1002/sce.3730640506

Eckhardt, M., Urhahne, D., Conrad, O., & Harms, U. (2013). How effective is instructional support for learning with computer simulations? *Instructional Science, 41*, 105-124. doi: 10.1007/s11251-012-9220-y

Gijlers, H., & de Jong, T. (2013). Using concept maps to facilitate collaborative simulation-based inquiry learning. *Journal of the Learning Sciences, 22*, 340-374. doi: 10.1080/10508406.2012.748664

Hagemans, M. G., van der Meij, H., & de Jong, T. (2013). The effects of a concept map-based support tool on simulation-based inquiry learning. *Journal of Educational Psychology, 105*, 1-24. doi: 10.1037/a0029433

Kanari, Z., & Millar, R. (2004). Reasoning from data: How students collect and interpret data in science investigations. *Journal of Research in Science Teaching, 41*, 748-769. doi: 10.1002/Tea.20020

Klahr, D., Zimmerman, C., & Jirout, J. (2011). Educational interventions to advance children's scientific thinking. *Science, 333*, 971-975. doi: 10.1126/science.1204528

Land, S. M., & Zembal-Saul, C. (2003). Scaffolding reflection and articulation of scientific explanations in a data-rich, project-based learning environment: An investigation of progress portfolio. *Educational Technology: Research & Development, 51*, 65-84. doi: 10.1007/Bf02504544

Morgan, K., & Brooks, D. W. (2012). Investigating a method of scaffolding student-designed experiments. *Journal of Science Education and Technology, 21*, 513-522. doi: 10.1007/s10956-011-9343-y

Njoo, M., & de Jong, T. (1993). Supporting exploratory learning by offering structured overviews of hypotheses. In D. Towne, T. de Jong, & H. Spada (Eds.), *Simulation-based experiential learning* (pp. 207-225). Berlin: Springer.

Novak, J. D. (1990). Concept mapping: A useful tool for science education. *Journal of Research in Science Teaching, 27*, 937-949. doi: 10.1002/tea.3660271003

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 inquiry cycle. *Educational Research Review, 14*, 47-61.

Scanlon, E., Anastopoulou, S., Kerawalla, L., & Mulholland, P. (2011). How technology resources can be used to represent personal inquiry and support students' understanding of it across contexts. *Journal of Computer Assisted Learning, 27*, 516-529. doi: 10.1111/j.1365-2729.2011.00414.x

van Joolingen, W., & de Jong, T. (1993). Exploring a domain with a computer simulation: Traversing variable and relation space with the help of a hypothesis scratchpad. In D. Towne, T. de Jong & H. Spada (Eds.), *Simulation-based experiential learning* (pp. 191-206). Berlin: Springer.

2 Scaffolding Students' Orientation: Concept Mapper

2.1 Abstract

In two studies the learning effects of using the concept mapper tool were examined. Different comparisons were made, including using the concept mapper at the beginning of the inquiry process (in the orientation phase), or at the end (in the conclusion phase), with or without pre-structuring, or learning without a concept mapper. The learning effects were examined in quasi-experimental studies that took place in real school settings. Between-subject pre-test post-test designs were used. In the first study, no differences were found regarding learning outcomes between a condition in which students constructed a concept map at the beginning of the inquiry task and a control condition in which students performed the same inquiry task but without a concept mapper tool. In the second study, significant learning effects were observed between the pre-test and the post-test but not between conditions.

2.2 Introduction

There is a general consensus that learning is successful when students achieve a deep understanding about a domain and are able to transfer their newly acquired knowledge and skills to new problems and situations. This is often called 'meaningful learning'. Several cognitive processes are thought to facilitate meaningful learning. Mayer (2002) summarizes these processes as: selecting, organizing, and integrating. *Selecting* refers to the processes in which students attend to the information and *select* those pieces of the information of which that they think are relevant and need to be processed further. *Organizing* refers to interpreting, connecting and organizing the selected pieces of information and to build a coherent mental model of the learning materials. *Integrating* refers to connecting the new information with prior knowledge. These processes of selecting, organizing, and integrating are thought to occur in an iterative rather than a linear fashion (Mayer, 2003) and is relevant for inquiry learning, where students are actively involved in constructing their own, meaningful knowledge. The question is however, whether students need some form of scaffolding helping them to engage in selecting, organizing, and integrating information. A scaffolding tool that seems particularly useful in this respect is the concept map.

A concept map is a kind of graphical organizer in which labelled nodes denote main concepts and links between the nodes denote the relationships between the concepts. These links can be directional or non-directional and labelled or unlabelled. Furthermore, concept maps can either be given to students, or, as a learning activity, students can be asked to build (or modify) concept maps themselves. Here we focus on the latter: students constructing concept maps.

According to Novak and Gowin (1984), the purpose of constructing a concept map is to assess the relevance of information. Shute and Zapata-Rivera (2008) argue that a concept map is an external representation of an internal mental model. A concept map helps to organize the knowledge of students by visualizing information in a form (Shute and Zapata-Rivera, 2008). According to Chmielewski and Dansereau (1998), creating a concept map requires cognitive processes in which knowledge is extracted from information. Graft and Kemmers (2007) argue that the skills associated with a scientific way of working, inter alia, promote the development of coherent concepts. Seeing the consistency and relationships between concepts is what is done when creating a concept map.

In a meta-analysis, Nesbit and Adesope (2006) found that concept mapping activities were more effective for attaining knowledge retention and transfer as compared to activities such as reading text passages, attending lectures, and participating in class discussions. Concept mapping was found to benefit learners across a broad range of educational levels, subject areas, and settings. Lambiotte and Dansereau (1992) argue that students with little prior knowledge benefit more from creating a concept map than students with more knowledge. Nesbit and Adesope (2006) indicate that students with more knowledge benefit from creating a concept map as well. However, this is significantly less than the students with little knowledge (Nesbit & Adesope, 2006). This may mean that students at the beginning of the lesson benefit more from the organization of knowledge. Getting some grip on the most important domain concepts and variables and having some overview of how they relate to each other is an important part of both the orientation and the conceptualization phase in the inquiry learning process and can form an important basis for designing and setting up experiments. On the other hand, it might also well be that getting such a domain overview in the early phases of inquiry is too much asked from the students and causes cognitive overload. Constructing a concept map can also be done at the end of the inquiry process. Then the mapping serves as a kind of summarizing. Externalizing one's mental model by constructing a concept map might be helpful to further integrate one's knowledge and understanding, but might also help to bring to light possible gaps in one's understanding (Cox, 1999).

Below two studies will be presented and discussed in which the effects of constructing concept maps was examined in Go-Lab. We start with a study that aimed to assess the learning effects of constructing a concept map during the orientation phase in the inquiry cycle.

2.3 Experiment 1

2.3.1 Method

2.3.1.1 Participants

The study was conducted with three classes consisting of Dutch students of secondary education. The students are in the second year of pre-university education (VWO). The study was conducted with 44 Dutch students of secondary education. The students were in the second year of pre-university education (VWO). The group consisted of 17 boys and 27 girls. Participants were on average 13.00 years old with a standard deviation of 0.482.

Design

Two conditions were compared in an experimental study with a pre-test post-test, between-subjects design. Participants in the experimental condition were asked to construct a concept map during the Orientation phase. Participants in the control condition followed the same instruction, except for the concept map.

2.3.1.2 Materials

Learning environment

Students in both conditions worked in an online Go-Lab learning environment on relative density. The learning environments in both conditions were identical, except for the concept mapper tool in the orientation phase in the experimental condition.

Virtual laboratory: Splash

Splash is a virtual laboratory in which students can conduct experiments to learn about buoyancy and Archimedes' principle. It consists of several sub labs that have different

levels of difficulty. The current study focused on one of these sub labs, namely the lab about relative density. In the lab containers filled with (variable) liquids are displayed. Students can drop balls in those containers. Students can choose the mass, volume and density of the balls by means of sliders. After the properties of the balls are chosen they can run experiments. The balls are dropped in the containers and students can observe whether the balls sink, drift or float in the liquid (see Figure 2.1).

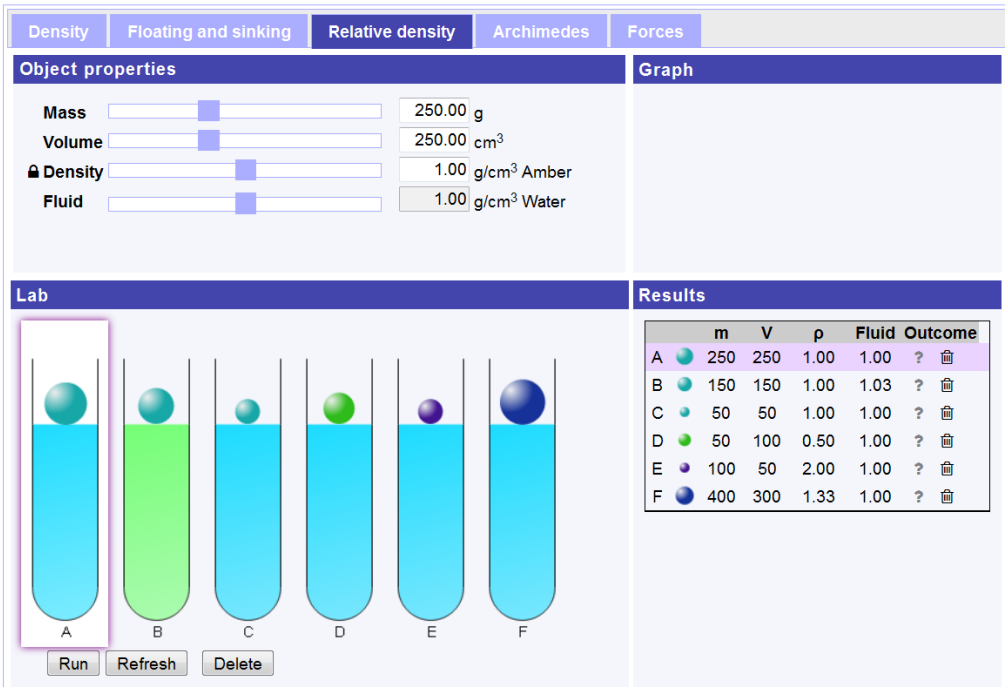


Figure 2.1. Splash sub lab on relative density.

The Concept Mapper Tool

The Concept Mapper Tool used in the experiment didn't provide domain related terminology. Participants could enter their own terms instead.

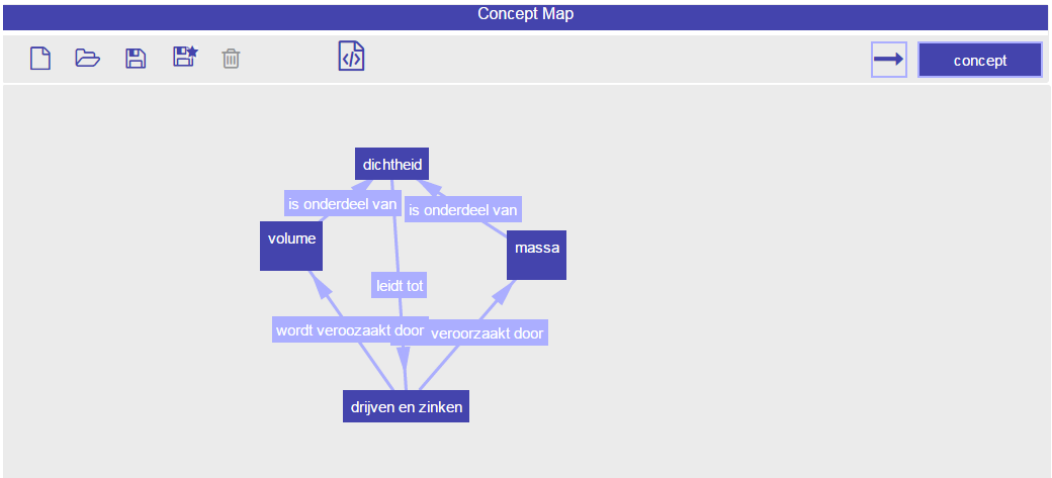


Figure 2.2. The Concept Mapper Tool.

2.3.1.3 Assessment

Students' conceptual knowledge was assessed both before and after the intervention with a parallel knowledge test that was designed for this study. The test consisted of two parts. The first part of the test (25 questions) concerns buoyancy and covers the topics of Splash:

floating, drifting and submerging, density, and relative density. The test measured students' understanding of the key concepts and principles. Students were asked to write down definitions of the key concepts and they had to apply this knowledge by providing the masses, volumes and densities of balls in different situations, the amount of displaced water, and/or forces that act upon the ball or the displaced water. For example, in the first question they were first asked to: "Give a definition of density. Include the terms volume and mass in your definition". In the second question, they had to: "Give the volume, the mass, and the density of three different balls". To do this they were required to understand that density equals mass in grams divided by volume in cm^3. Students received one point for each correct answer. In the example they could receive one point for the correct definition and one point for each correctly composed ball. The second part of the test contained one item that was intended to assess structural knowledge. Here, students were asked to note the concepts relevant to the domain of relative density and to indicate their relations. For this item, a maximum number of 12 points could be obtained. Points were assigned on the basis of a scoring rubric. The students' responses to this item were scored by a panel of four independent raters. In cases where scores were different, the raters discussed the scoring with each other until consensus was reached.

2.3.1.4 Procedure

During the first session, participants completed the pre-test. One week later, the actual experiment took place. The lesson started with a series of six quiz questions about floating and sinking of several exploratory questions on the subject. After taking the online quiz, the instruction began. At the end of the lesson, another online quiz was presented, basically similar to the first one, except for some surface features. After the participants finished the online lesson, they completed the post-test, which was a parallel version of the pre-test.

2.3.2 Results

On the first part of both the pre-test and the post-test, the students were able to achieve a total of 25 points. The average test scores are displayed in Table 2.1 Test scores (first part) of pre-test and post-test (max. score = 25).

Table 2.1. Test scores (first part) of pre-test and post-test (max. score = 25)

	Control (*n* = 21)		Experimental (*n* = 23)	
Test	M	SD	M	SD
Pre-test	15.57	5.11	16.09	6.95
Post-test	16.19	4.21	15.65	6.33

A repeated measures Anova was performed on the data in Table 2.1 Test scores (first part) of pre-test and post-test (max. score = 25), with time as dependent variable and condition as independent variable. The outcomes showed that there was no main effect of time (Wilks' Lambda $F(1, 42) = 0.03$, $p = .88$) and neither was there an interaction effect between time and condition (Wilks' Lambda $F(1, 42) = 0.18$, $p = .68$).

The second part of the pre- and post-test consisted of an overview of relevant domain concepts and relations between them. The scores on these items are displayed in Table 2.2 Test scores (second part) of pre-test and post-test (max. score = 12).

Table 2.2. Test scores (second part) of pre-test and post-test (max. score = 12)

	Control (*n* = 21)		Experimental (*n* = 23)	
Test	M	SD	M	SD
Pre-test	5.62	2.11	5.78	2.21
Post-test	5.67	1.32	5.91	2.15

A repeated measures Anova was performed on the data in Table 2.2 Test scores (second part) of pre-test and post-test (max. score = 12), with time as dependent variable and condition as independent variable. The outcomes showed that there was no main effect of time (Wilks' Lambda F (1, 42) = 0.05, p = .82) and neither was there an interaction effect between time and condition (Wilks' Lambda F (1, 42) = 0.01, p = .92).

2.3.3 Conclusion and discussion

In this study, we investigated whether making a concept map affects knowledge acquisition. The lesson the students have followed about relative density was based on the principles of inquiry learning. The experimental group had the opportunity to create a concept map during the orientation phase of the inquiry cycle. Participants in the control condition did not have this tool. In order to verify whether the students in the experimental condition indeed benefited from the making of a concept map, their pre-test post-test gains were compared with those of participants in the control condition. Before and after the class there is a test conducted to measure the knowledge. At the beginning and at the end of the lesson there was a quiz in the curriculum, but the results could not be retrieved and analysed. It was found that there is no significant difference between the experimental group and the control group.

For further research, it is advisable to analyse whether the concept maps have effects in the long term. As a study by Taskin and colleagues (2011) showed then concept maps have effect if these are made over a long period.

There are some limitations to this study. During the experiment, the internet did not work properly due to circumstances beyond the experimenters' control. Therefore, not all students could see the entire lesson at the start of the session. Because the concept map was also meant to be constructed at the beginning of the lesson, this may have prevented students from using the tool. The internet connectivity problems at the beginning of the session caused agitation in the classroom. The students therefore may not have been able to concentrate optimally. Finally, at the end of the lesson, a concert scheduled for the students started. The students indicated they would like to stand in front and it was only after they had gone through the entire lesson, this may have influenced the results of the post-test.

If we look at the distribution of the measurements before and after, we see that the average close to maximum points. Probably, there is a ceiling effect. This means that there are students for which the test was too easy.

2.4 Experiment 2

In order to foster a stronger alignment with other Go-Lab studies, a new experiment was set-up, in which fewer tools were used. Also a modified version of the knowledge test was used. This time we used the same test as our partners in Cyprus. In this new experiment,

the research focus was broader than in the previous experiment: this time, three research questions were examined:

1. Does constructing a concept map lead to better learning outcomes?
2. Does it matter (in terms of learning outcomes) when a concept map is constructed in the inquiry process (early or at the end)?
3. Is a pre-structured concept map more effective for learning than building a concept map from scratch?

2.4.1 Method

2.4.1.1 Participants

The study was conducted with four classes consisting of Dutch students of secondary education. The students are in the second year of high school (HAVO) or pre-university education (VWO). The number of respondents was 87 students (49 boys and 38 girls). Participants were on average 13.76 years old with a standard deviation of 0.573.

2.4.1.2 Design

Four conditions were compared in an experimental study with a pre-test post-test, between-subjects design. There were three experimental conditions:

- CM-Begin: condition in which students build the concept map during the orientation phase
- CM-End: condition in which students build their concept map at the end of the inquiry process, namely in the conclusion phase
- CM-End+Structure: similar to condition CM-End except that here the students are provided with a pre-structured concept map in which a structure of (empty) nodes and links is provided and students only have to fill in the concepts in the nodes.

Participants in the control condition followed the same instruction, except for the concept map.

2.4.1.3 Materials

Learning environment

Students in both conditions worked in an online Go-Lab learning environment on relative density. This time a lighter version of the learning environment was used. By "lighter" we mean that the ILS only contained the tools that were within the focus of the current study, in this case the concept mapper tool. Other tools were left out as much as possible. Where necessary, other solutions were used. For example instead of providing the students with a questioning tool they were provided with a given research question. Hypotheses were provided as well.

The learning environments in all conditions were identical, except for the concept mapper tool. In the CM-Begin condition this tool was provided in the orientation phase; in the CM-End en CM-End+Structure conditions, the tool was provided in the conclusion phase; and in the Control condition there was no concept mapper.

The Concept Mapper Tool

The Concept Mapper Tool used in the experiment didn't provide domain related terminology. Participants could enter their own terms instead.

2.4.1.4 Assessment

Students' conceptual knowledge was assessed both before and after the intervention with a knowledge test that was designed for this study. For the creation of the knowledge test,

Bloom's revised taxonomy of the cognitive domain has been followed (see also de Jong, 2014; Zervas, 2013). The corresponding levels were “Remember”, “Understand”, “Apply” and “Think critically and creatively”. Our knowledge test consisted of six items: one open-ended “Remember” item, 3 closed-ended items, of which, two referred to “Understand” and another one to “Apply”, and two open-ended “Think Critically and Creatively” items. All items focused on density, the relation between mass, volume and density, and the role of density to sinking and floating. In total, a maximum score of 12 point could be obtained.

2.4.1.5 Procedure

During the first session, participants completed the pre-test. One week later, the actual experiment took place. The lesson started with a series of six quiz questions about floating and sinking of several exploratory questions on the subject. After taking the online quiz, the instruction began. At the end of the lesson, another online quiz was presented, basically similar to the first one, except for some surface features. After the participants finished the online lesson, they completed the post-test, which was a parallel version of the pre-test.

2.4.2 Results

The average test scores on the pre-test and post-test are displayed in Table 2.3 Test scores of pre-test and post-test (max. score = 12).

Table 2.3. Test scores of pre-test and post-test (max. score = 12)

	Control (*n* = 22)		CM-Begin (*n* = 21)		CM-End (*n* = 21)		CM-End+Struct. (*n* = 23)	
Test	M	SD	M	SD	M	SD	M	SD
Pre-test	7.05	2.13	7.14	2.58	7.10	2.59	7.30	3.05
Post-test	7.82	1.92	7.38	2.58	7.52	2.44	7.96	3.10

A repeated measures Anova was performed on the data in Table 2.3 Test scores of pre-test and post-test (max. score = 12), with time as dependent variable and condition as independent variable. The outcomes showed that there was a main effect of time (Wilks’ Lambda $F(1, 83) = 6.89$, $p < .05$), but no interaction effect between time and condition (Wilks’ Lambda $F(1, 83) = 0.35$, $p = .79$).

2.4.3 Conclusion and Discussion

In this second experiment, significant increase of knowledge from the pre-test to the post-test was observed. There were no significant differences between conditions, in other words, students in all groups benefitted equally from the lesson, regardless of whether they constructed a concept map at the beginning of the inquiry task, at the end, a pre-structured concept map, or even did not construct a concept map at all. We can only speculate why we didn't find learning effects from concept mapping. An explanation could be that there are learning effects but our knowledge test is not suitable or sensitive enough to detect them. However, our knowledge test is not very different from tests used in regular tests in schools. So, if our test doesn't capture the effects of concept mapping, then it is not very likely that concept mapping will result in higher test scores in schools. It may also be the case that concept mapping does not lead to learning effects in our study. Nesbit and Adesope (2006) observed the strongest effects of concept mapping in domains that

are highly text-based, such as humanities, law, and social studies (average effect size 1.27) and only minor (but nonetheless) significant effects in physical science (average effect size 0.28). The domain in our study is part of physical science, our instruction does not contain much text and the relative density is a topic that can be characterized with relatively small number of concepts (if necessary with 3 or 4 concepts). So, it might be that our domain does not lend itself very well for concept mapping. It would be interesting if future research can further explore the effects of concept mapping in other domains, more text-rich and concept-rich domains. It would also worthwhile to explore possible side-effects of concept mapping. Our impression is that students enjoy making the concept maps and in that sense it can be a welcome and motivating variation when applied in combination with other learning activities.

2.5 References

Alverman, D. (1981). The compensatory effect of graphic organizers on descriptive text. The *Journal of Educational Research 75*, 44-48.

Chmielewski, T., & Dansereau, D. (1998). Enhancing the recall of text: knowledge mapping traning promotes implicit transfer. *Journal of Educational Psychology*, *90*, 407-413.

de Jong, T. (Ed.). (2014). *D1.3: Preliminary inquiry classroom scenarios and guidelines: Go-Lab project.*

Lambiotte, J., & Dansereau, D. (1992). Effects of knowledge maps and prior knowledge on recall of science lecture content. *Journal of Experimental Education 60*, 189-201.

Mayer, R. E. (2002). Rote versus meaningful learning. *Theory into Practice, 41*, 226-232.

Mayer, R. E. (2003). The promise of multimedia learning: Using the same instructional design methods across different media. *Learning and Instruction, 13*(2), 125-139.

Nesbit, J., & Adesope, O. (2006). Learning with concept and knowledge maps: A meta-analysis. *Review of Educational Research 76*, 413-448.

Shute, V., & Zapata-Rivera, D. (2008). Using an evidence-based approach to assess mental models. In D. Ifenthaler, P. Pirnay-Dummer, & J. Spector, *Understanding models for learning and instruction: Essays in honor of Norbert M. Seel* (pp. 23-41). New York: Springer.

Taskin, M., Pepe, H., Taskin, C., Gevat, C. & Taskin, H. (2011). The effect of concept maps in teaching sportive technique. *Procedia Social and Behavioral Sciences, 11*, 141-144. doi:10.1016/j.sbspro.2011.01.049

Zervas, P. (Ed.). (2013). D2.*1. The Go-Lab inventory and integration of online labs – labs offered by large scientific organisations: Go-Lab project.*

3 Scaffolding students' conceptualisation with the question and hypothesis scratchpad

3.1 Abstract

We assessed the impact of two Go-Lab scaffolds (the question and hypothesis scratchpad) on students' inquiry skills. Assessment mainly relied on pre- and post-tests to measure students' inquiry skills before and after a Go-Lab intervention. The Go-Lab intervention involved two conditions that differed in the level of support provided by the two scaffolds. In one condition the scaffolds included predefined terms to guide students to formulate research questions and hypotheses. In the other condition the scaffolds did not include these predefined terms. Students were randomly placed into one of the two conditions. We conducted studies with two different age groups: one study with secondary school students (ages 18-19) and the other study with basic school students (ages 13-14). We discuss the results of these studies and based on these results offer suggestions to improve the implementation of the Go-Lab interventions.

3.2 Introduction

Inquiry-based science education is an important pedagogical teaching practice because it helps increase children's and student's interest and attainment levels in science while at the same time stimulating teacher motivation. The inquiry approach gives greater opportunity for learners to inductively build their knowledge, and thus strongly emphasizes observation and experimentation. An important step in beginning inquiry investigations is formulating research questions and/or hypotheses. Pedaste et al. (2015), after conducting an extensive literature review of terms used to describe steps in the inquiry process, combined posing research questions and/or formulating hypotheses under the category of Conceptualization. The Conceptualization phase is a process where a learner begins to understand the concepts belonging to a problem and then raises questions that need to be answered and/or predictions that need to be tested.

In the Go-Lab learning environment the question and hypothesis scratchpad are scaffolds to help students pose research questions and formulate hypotheses. They present a structured way to create research questions and hypotheses by offering students a list of pre-defined list of elements that can be put together to create a full research question or hypothesis.

3.3 Method

In this study, students from two different age groups worked with two different Go-Lab Inquiry Learning Spaces. The younger students worked with a physics-based virtual laboratory called Splash, whereas the older students worked with a biology-based virtual laboratory called Sexual Selection in Guppies. The studies were performed during normal school hours when it was possible to arrange computer classroom time for the students. Two university researchers involved in the Go-Lab project helped the school teacher during implementation of the intervention at school.

The Go-Lab intervention involved two conditions: one where the scaffolds included predefined terms to guide students to formulate research questions and hypotheses, and the other where the scaffolds did not include these predefined terms. Students were randomly placed into one of the two conditions. It was expected that students in the predefined scaffolds condition would have the highest gains in inquiry skills.

3.3.1 Participants

Students from two age groups participated in these studies. A total of 24 students aged 13-14 (mean age 13.6) from a basic school and 19 students aged 18-19 (mean age 18.7) from a secondary school in Estonia participated. All participants took an inquiry skills pre-test on paper, were randomly assigned to one of the two Go-Lab intervention conditions to complete a Go-Lab ILS on the computer, and afterwards completed an inquiry skills post-test on paper.

3.3.2 Materials

3.3.2.1 Learning environment

Students in all conditions worked in an online learning environment that had a similar structure. The learning environments differed in the support they contained. In one condition students had question and hypothesis scratchpads with predefined terms, in the other case students worked with question and hypothesis scratchpads that did not have predefined terms. In both the Splash and Guppies ILSs there were additional Go-Lab scaffolds: the Concept Map Tool, the Experimental Design Tool, the Observation Tool and the Conclusion Tool. The Guppies ILS also included the Reflection Tool.

3.3.2.2 Virtual laboratory: Splash

As discussed already, the Splash lab is a virtual laboratory about buoyancy and Archimedes' principle. The concept of Archimedes' principle can be applied to explain why objects float. A floating object displaces a weight of fluid equal to its own weight. If an object displaces less fluid when submerged, then it sinks. In our Go-Lab Inquiry Learning Space students could manipulate the object mass, object volume, object density and fluid density to explore what happens when the object is submerged into the fluid. The natural outcome is that when the object density is less than the fluid density then the object floats. Conversely, when the object density is more than the fluid density then the object sinks.

3.3.2.3 Virtual laboratory: Sexual Selection in Guppies

The Sexual Selection in Guppies virtual laboratory is an online interactive simulation provided by the Virtual Biology Lab Project, http://virtualbiologylab.org/EndlersGuppies.htm, to allow for the recreation of classic experiments performed by the biologist John Endler when he first investigated the balance of natural and sexual selection in guppy fish in the 1970s. Natural selection affects traits related to survival and producing offspring, while sexual selection specifically affects traits related to obtaining mates. Often traits which increase an individual's access to mates, actually decreases their survival. The Sexual Selection in Guppies virtual laboratory deals with this evolutionary trade-off. In guppies, females prefer to mate with males that have lots of spots, but those males are more easily seen by predators. The virtual laboratory allows students to manipulate the strength of female preference and the number of predators to see the influence of environment in shaping different characteristics of guppies (see Figure 3.1).

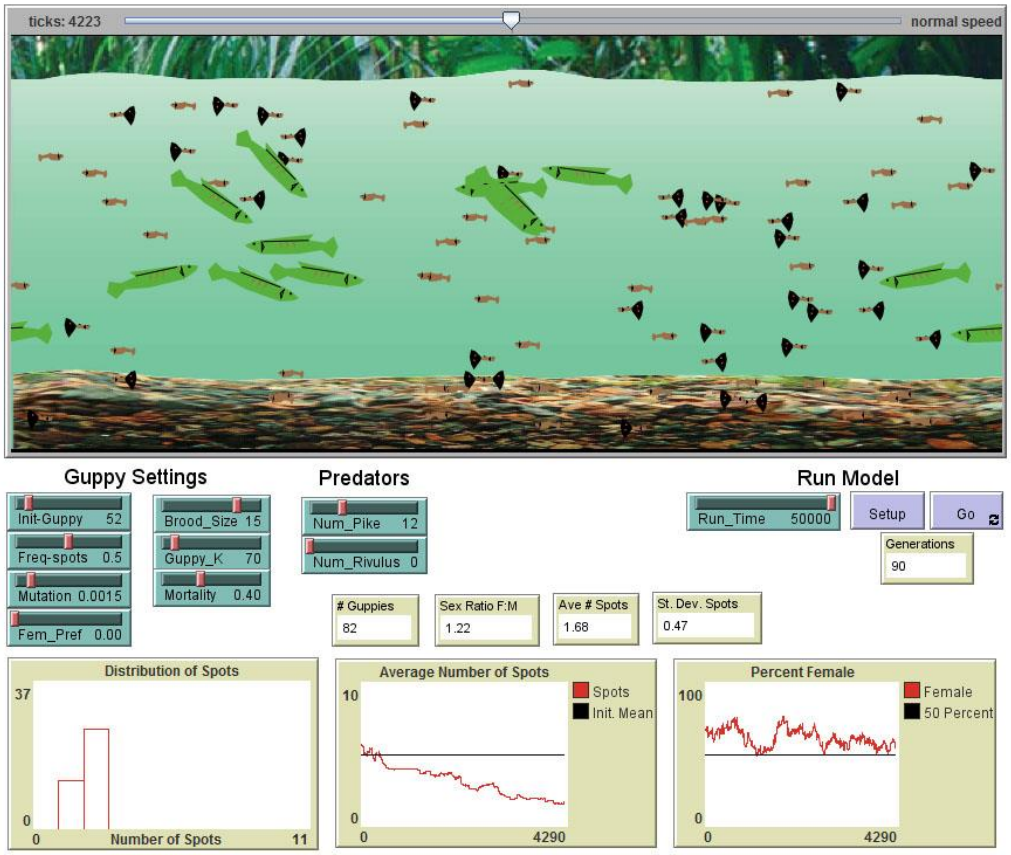


Figure 3.1. Guppies lab on natural and sexual selection in a guppy fish environment.

3.3.2.4 The Hypothesis and Question Scratchpad

This question and hypothesis scratchpads are based on work by van Joolingen and de Jong (1993). They offer students a structured way to pose research questions and formulate hypotheses. In the experiments the condition with predefined terms (i.e. Condition A) included 18 terms in the questioning scratchpad and 15 terms in the hypothesis scratchpad for the Splash lab study; and 13 terms in the questioning scratchpad and 9 terms in the hypothesis scratchpad for the Guppies lab study. The condition without predefined terms (i.e. Condition B) required students to write research questions and hypotheses in their own words (see Figure 3.2 for the scratchpad in both conditions).



Figure 3.2. Two conditions of the Questioning scratchpad scaffold in the Sexual Selection in Guppies study: (a) Condition A with predefined terms, (b) Condition B without predefined terms.

3.3.3 Assessment

Assessment of inquiry skills was conducted using the Test of the Integrated Science Process Skills (TIPS) developed by Dillashaw and Okey (1980) and the TIPS II (Burns, Okey, & Wise, 1985), which is an extended version of the original TIPS. TIPS was used as the pre-test and TIPS II as the inquiry skills post-test. For our study only items related to measuring two inquiry skills were selected: 1) identifying variables (i.e. dependent, independent, and controlled variables), and 2) identifying testable hypotheses. These selection criterions lead to a total of 21 multiple choice items, each with four alternatives. The test reliability, Cronbach's Alpha, is 0.86.

3.4 Procedure

In both age group studies the students were allotted 20 minutes to take the pre- and post-tests. In the study with 18-19 years-olds they had 75 minutes to complete the Sexual Selection in Guppies ILS. In the study with 13-14 years-olds they had 45 minutes to complete the Splash ILS. Two university researchers involved in the Go-Lab project led the Go-Lab intervention at the respective schools were the studies occurred. Their presence on-site during the intervention was accompanied by the taking of observer notes, in order to provide some insight on the effectiveness of the intervention.

3.5 Results

The results of the pre- and post-tests are presented in Table 3.1 and Table 3.2 for the two age groups studied. No statistically significant differences were found in the pre- and post-test data. An interesting observation between the two age groups is that younger students performed better on average on items related to identifying variables. On closer

examination it appears that a large majority of the older students confused the vocabulary terms control and independent variables on both the pre- and post-tests. The Go-Lab intervention did not specifically address this issue because vocabulary in the ILS used phrases like “variables that remain the same” for control variables and “variables that you change” for independent variables. It was assumed that students could implicitly transfer this knowledge.The older students did show better scores on average on items related to identifying testable hypotheses when compared to younger students.

Table 3.1. Pre- and post-test scores for the 13-14 year old student study

	Condition A (*n* = 12)		Condition B (*n* = 12)	
Inquiry Skills Test	M	SD	M	SD
Identifying testable hypotheses (max. score = 9)				
Pre-test	5.50	1.62	6.08	1.56
Post-test	5.75	2.60	6.00	1.41
Difference	0.25	2.30	-0.08	1.62
Identifying variables (max. score = 12)				
Pre-test	5.42	2.39	4.67	1.72
Post-test	5.00	2.83	3.83	1.70
Difference	-0.42	3.03	-0.83	1.50

Table 3.2. Pre- and post-test scores for the 18-19 year old student study

	Condition A (*n* = 12)		Condition B (*n* = 7)	
Inquiry Skills Test	M	SD	M	SD
Identifying testable hypotheses (max. score = 9)				
Pre-test	7.08	1.44	7.71	0.95
Post-test	7.42	0.90	7.86	1.86
Difference	0.33	1.78	0.14	1.86
Identifying variables (max. score = 12)				
Pre-test	4.50	1.93	4.86	1.57
Post-test	4.17	1.47	3.43	1.62
Difference	-0.33	1.92	-1.43	1.81

3.6 Discussion and Conclusion

During the Go-Lab intervention when students were working through the Go-Lab Inquiry Learning Spaces, the two researchers responsible for leading the session made observations about student behaviour in the Go-Lab learning environment. The researchers observed that students from both age groups would benefit enormously from a better understanding of how to budget their time while working in a Go-Lab ILS. It was observed that many students spent an excessive amount of time with the introductory material and thus did not have sufficient time to explore their research questions and hypotheses using the online virtual laboratory. For example, embedded videos placed in the Orientation phase consumed a lot of time (8½ minute running time for the Guppy ILS, about 4 minutes for the Splash ILS). Although 4 minutes running time is relatively short, because the videos were in English with sometimes English text appearing, the 13-14 year

old Estonian students would pause and rewind the video at a leisurely pace to understand it. This increased the video viewing time. This finding now has been transferred into a design tip at the support page of Golabz. Another factor that increased the working time for students was the inclusion of open-ended scaffolds. For example, the ILSs included the Concept Map Tool without a predefined structure and therefore students had to spend time thinking of concepts and how to connect them with each other. A predefined Concept Map with more concepts and relationships already diagrammed, and that asked students to add only two or three new concepts would have been a more efficient use of time. In general, ILS design should match the allocated time for an in-class intervention. A consequence of limited time is that some students are not motivated to seriously complete the ILS when they notice that time is running out. Thus, ensuring that students can work efficiently in the allotted classroom time is an important consideration for Go-Lab ILS design. Another observation was that students require some time to get accustomed to using Go-Lab scaffolds. This getting familiar with the Go-Lab environment and scaffolds could possibly be sped up if students are assigned a homework assignment to use a Go-Lab ILS. During the homework activity students could comfortably move from one inquiry phase to another and get a sense of the different tasks required in each phase. They could also practice moving the terms in the question and hypothesis scratchpads to formulate their responses. After this familiarization process the actual in-class Go-Lab intervention time could flow more efficiently since additional in-class time to learn the functionality of tools in the Go-Lab learning environment would be avoided.

In conclusion, these initial studies helped identify areas in the implementation of Go-Lab ILSs that were common for two differing student age groups and where improvements can be easily made to facilitate the Go-Lab learning experience for students. Further research on evaluation that takes into account these suggestions may better reveal differences in how students receiving alternate forms of scaffolding develop in their inquiry skills.

4 How many words are enough? Assessing three different configurations of the Hypothesis Scratchpad

4.1 Abstract

The objective of the study was to assess the impact of three different configurations of a tool designed to scaffold students in formulating hypotheses. The first configuration included all words necessary to be integrated in hypotheses. The second configuration offered students a subset of variables and conditionals needed to formulate hypotheses. Finally, the third configuration did not contain any term. Our exploratory study wished to examine which configuration of the tool would have the most marked effect on student learning, namely, student knowledge and inquiry skills. The same learning environment was used by secondary school students in the UK, the Netherlands, Estonia, and Cyprus. This environment focused on relative density and included the Hypothesis Scratchpad. The general trend that was revealed pointed towards a highly heterogeneous assembly of results. A commonality among research settings indicated that a fully-fledged guidance in hypotheses formulation might be more beneficial for separate dimensions of student knowledge and inquiry skills.

4.2 Introduction

Formulating hypotheses is considered among the most demanding tasks for students during educational interventions in science education (Chin & Brewer 1993; Germann et al., 1996; Hofstein et al., 2005; Kremer et al., 2014). Providing support to students, when they are about to generate hypotheses, might prove beneficial for student learning, overall. For instance, previous research has suggested that offering students a hypothesis menu, which includes variables necessary to carry out an experimentation, might improve student performance (Chang et al., 2008; de Jong, 2006a; Njoo & de Jong, 1993; van Joolingen & de Jong, 1991, 2003). In this direction, a hypothesis scratchpad might offer a twofold guidance (Zacharia et al., 2015). First, it would outline appropriate variables for the phenomenon under study, and additionally, it would showcase the proper structure of a hypothesis in the form of an if-then clause.

There are three main challenges to be addressed in designing and assessing scaffolds, including tools like the hypothesis scratchpad. First, task complexity or complexity in tool usage might increase cognitive load substantially and this might severely compromise the contribution of the tool in student learning (e.g., van Joolingen & de Jong, 1997). Therefore, there is a need to simplify learning tasks by structuring them accordingly.

However, a rigid tool structure might subtract freedom from students in enacting explorations (e.g., Chang et al., 2008; van Berkum & de Jong, 1991). This brings us to the second challenge in designing and assessing scaffolds. The tension between structuring tasks and “problematizing” students that has been underlined by Reiser (2004). Namely, there is a tension between removing undesirable cognitive load from students (offloading), on the one hand, and increasing “germane” or useful cognitive load, on the other (see also Kalyuga, 2007). A latent contradistinction between structuring tasks and “problematizing” students might be also present in the debate on optimizing inquiry learning approaches in terms of appropriate guidance or openness (e.g., Arnold et al., 2014; Koksal & Berberoglou, 2014).

The third challenge in designing and assessing scaffolds relates to adapting learning environments so that scaffolds might “fade in” or “fade out” in a timely and learner-tailored fashion, according to learning tasks and student needs (de Jong, 2006b; Kalyuga, 2007).

All three challenges mentioned above are crucial in designing and assessing a hypothesis scratchpad. How many words would be enough in a hypothesis scratchpad to simplify the task of formulating hypotheses, and, at the same time, “problematize” students? Could different configurations of a hypothesis scratchpad, including each different numbers of variables and conditionals, be employed in attempts to adapt a learning environment in changing student ability and needs?

A first effort to address these issues was attempted in the frame of a study that was implemented in four different countries. Our objective was to investigate the effect of a tool developed to support students in formulating hypotheses. Specifically, we examined three different configurations of this tool on student knowledge and inquiry skills. The first configuration included all words necessary to be integrated in the hypotheses students would form. In that case, students had to put all words in the right place to generate hypotheses. The second configuration offered students a subset of variables and conditionals needed to formulate hypotheses. Here students were prompted by words present and they had to complete hypotheses with the rest of variables and conditionals. The third configuration did not contain any term. In this condition, students would have to write down on their own all words necessary to form their hypotheses. Our exploratory study wished to examine which configuration of the tool would have the most marked effect on student learning, namely, student knowledge and inquiry skills.

4.3 Method

The study in each country involved three different configurations of the Hypothesis Scratchpad (Figure 4.1). In the first condition (Condition 1), the tool had all the conditionals and variables needed for formulating a hypothesis. The second condition (Condition 2) included a subset of words (e.g., selection of conditionals and variables). In the third condition (Condition 3), no word was provided. Before and after the intervention, students completed a knowledge test and an inquiry skills test (pre and post-tests). Further, after the completion of the pre-tests students got familiar with the Hypothesis Scratchpad tool by formulating hypotheses in a stand-alone fashion. After the completion of the post-tests, students used the tool in a stand-alone mode again, and formulated hypotheses in a new context.

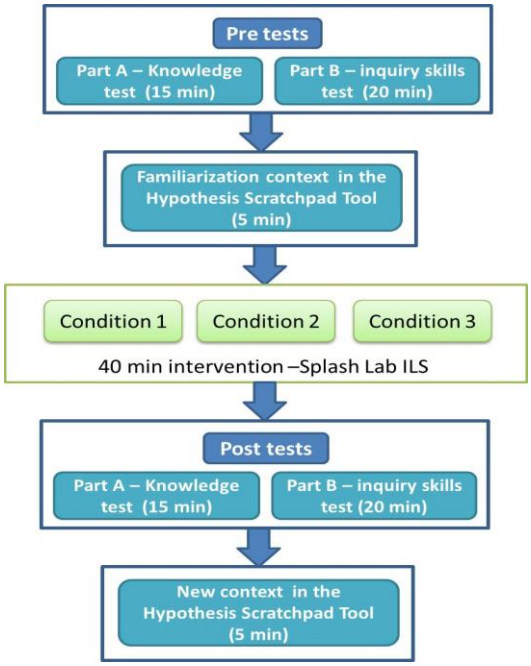


Figure 4.1. Experimental design of the study.

4.3.1 Participants

Overall, for a total of 312 students data were analysed (more students participated in the studies). Number, mean age, and gender of participants for each country are presented in Table 4.1. In each country, students were assigned on equal numbers to three different conditions, which corresponded to the three configurations of the Hypotheses Scratchpad.

Table 4.1. Sample characteristics.

	Number of participants	Mean age (years)	Gender (% female)
ULEIC	27	14.9	25.9
UT	173	12.0	44.5
UTE	39	13.0	59.0
UCY	73	14.5	58.9

4.3.2 Materials

4.3.2.1 Learning environment

Learning activities were undertaken in an online learning environment called Inquiry Learning Space (ILS), which was developed within the inquiry cycle framework (Pedaste et al., 2015) and by means of the Go-Lab authoring tool (de Jong, Sotiriou, & Gillet, 2014). The ILS focused on relative density. Students investigated whether an object sinks or floats by manipulating variables such as mass, volume and density of the object and density of the fluid. The ILS involved five inquiry phases, namely, the *Orientation*, the *Conceptualization*, the *Investigation*, the *Conclusion* and the *Discussion* phase.

In the *Orientation* phase, students watched a video that introduced them in the basic variables they would later encounter. In the *Conceptualization* phase, students explored the “Splash” Laboratory (“Splash: Virtual Buoyancy Laboratory”; http://www.golabz.eu/;

Figure 4.2).) to locate the four variables involved in the phenomenon under study (i.e., mass, volume and density of the object and density of the fluid). In that lab, students were able to manipulate the mass, volume and density of an object, and the kind of the fluid in a container. Then they run the experiment, and they observed it outcome, namely, if the object would sink or float. All manipulations and outcomes were presented in a Results´ table on the right side of the tool´s interface. A red arrow in the table meant that the object would sink, while a green star would denote that the object would float.

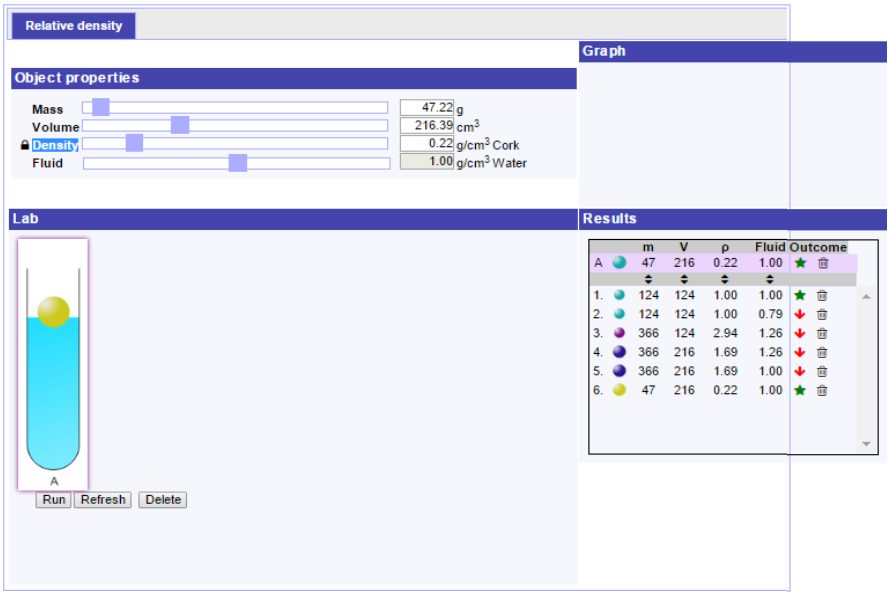


Figure 4.2. Splash: Virtual Buoyancy Laboratory.

After students became familiar with the Splash Laboratory, they formulated hypotheses using the Hypothesis Scratchpad. Overall, there were three different conditions (Figure 4.3).

In Condition 1, students were offered all words needed to formulate their hypotheses in the form of an “if…then” statement. Students could drag and drop these words to the space provided by the tool to generate their hypotheses. In Condition 2, students were given a subset of words, and in Condition 3, no word was given to students. Students could type their own words or phrases and place them in their hypothesis. This option was also valid in Condition 1. In all conditions, students could adjust their confidence level for each hypothesis, by changing the color of the “horseshoe” next to the space provided by the tool. If the “horseshoe” was blue, overall, that would indicate that a student was absolutely confident that his/her hypothesis was correct.



Figure 4.3. The three different conditions of the Hypothesis Scratchpad used in the study.

After student hypotheses had been formulated, students continued to the next phase, the *Investigation* phase, where they conducted their experiments to confirm or reject their hypotheses. Students were prompted to keep notes about ideas, thoughts and observations, by means of the Observation Tool (Figure 4.4).

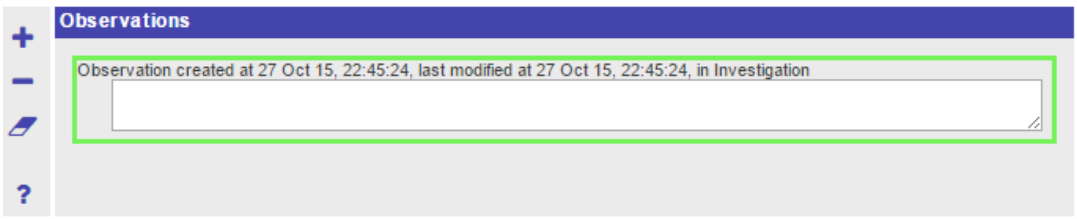


Figure 4.4. Observation Tool.

In the *Conclusion* phase, students used the Conclusion Tool (Figure 4.5) to retrieve their hypotheses and observations and, then, argue how their confidence for each hypothesis had changed after their investigation or not.

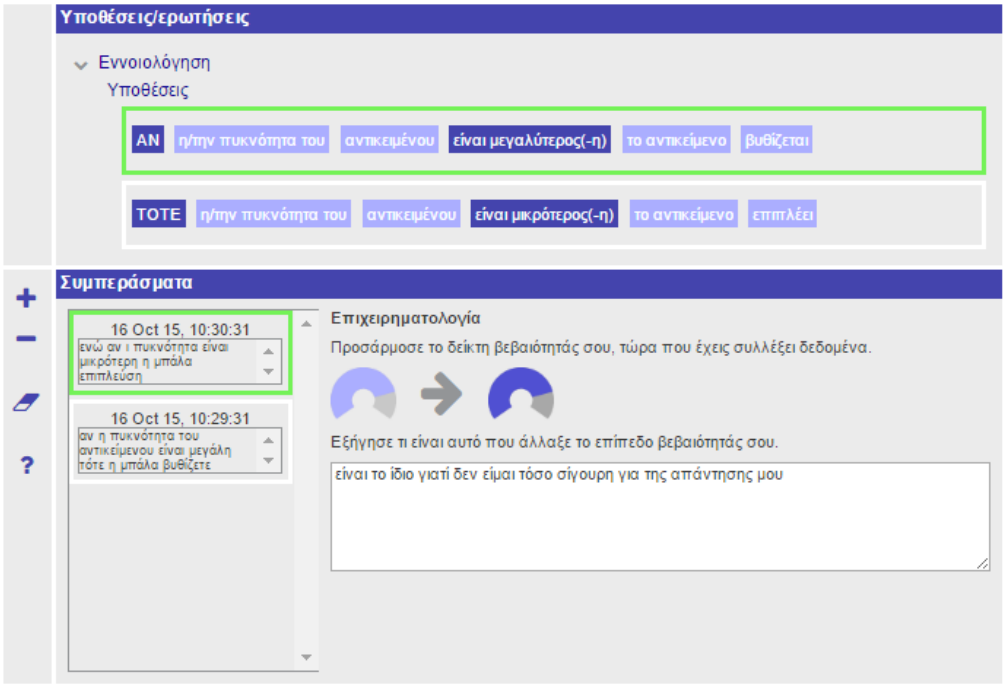


Figure 4.5. Conclusion Tool.

In the *Discussion* phase, students performed reflection activities by answering questions concerning the change in their confidence level for their hypotheses. They also described the most difficult phase during the whole inquiry cycle.

4.3.2.2 Assessment

Data collection involved two different tests, namely, a knowledge test and the inquiry skills test (Data collected through the knowledge test in Estonia were not included in data analyses due to many incomplete responses). For the creation of the knowledge test, a revised taxonomy of cognitive processes has been followed (for more details see de Jong, 2014; Zervas, 2013). Cognitive processes included in the test were "*Remember*" (1 item), "*Understand*" (1 item) and "*Apply*" (2 items). The items of the test focused on density, the relation between mass, volume and density, and the role of density tin sinking and floating.

For the inquiry skills test, items from the TIPSII instrument (Burns, Okey, & Wise, 1985) were selected and translated in the language of each country, where the study took place. The test consisted of 21 multiple-choice items, addressing "*Identifying variables*" (12 items) and "*Identifying and stating hypotheses"* (9 items). The number of items in each subscale was the same as in the initial TIPSII instrument.

For each test, one point was given to each correct response. Scores for knowledge and inquiry skills dimensions were rescaled to range between 0 and 1.

4.4 Procedure

The implementations were carried out by science teachers in each country. Students first completed pre-tests and got familiarized with the Hypothesis Scratchpad. Then, students used computers to access the ILS and it took them 30 to 40 minutes to go through the entire ILS. During the lesson, the only help students received from teachers concerned technical issues with regard to the use of tools and the lab. Whenever such technical issues appeared, they were solved without causing any considerable delay to the

completion of the lesson. After the ILS, students completed post-tests and used the Hypothesis Scratchpad in a new context.

4.5 Results

The results of the common study are first presented for each country separately. Then, an overview across research settings follows.

4.5.1 Results in the UK

Table 4.2. Student overall performance in the knowledge test

	Condition 1 (all words)	Condition 2 (some words)		Mann Whitney Test Z
Pre-test	0.54	0.32		ns
Post-test	0.57	0.42		-1.98*
Wilcoxon Signed Ranks Test Z	ns	ns		

Note: ns = non-significant* p < 0.05; ** p < 0.01; *** p < 0.001.

Table 4.3. Student performance in "Remember" (knowledge dimension)

	Condition 1 (all words)	Condition 2 (some words)		Mann Whitney Test Z
Pre-test	0.00	0.00		ns
Post-test	0.50	0.39		ns
Wilcoxon Signed Ranks Test Z	-2.65**	-2.24*		

Note: ns = non-significant* p < 0.05; ** p < 0.01; *** p < 0.001.

Table 4.4. Student performance in "Understand" (knowledge dimension)

	Condition 1 (all words)	Condition 2 (some words)		Mann Whitney Test Z
Pre-test	0.50	0.15		ns
Post-test	0.36	0.15		ns
Wilcoxon Signed Ranks Test Z	ns	ns		

Note: ns = non-significant* p < 0.05; ** p < 0.01; *** p < 0.001.

Table 4.5. Student performance in "Apply" (knowledge dimension)

	Condition 1 (all words)	Condition 2 (some words)		Mann Whitney Test Z
Pre-test	0.69	0.49		ns
Post-test	0.67	0.51		ns
Wilcoxon Signed Ranks Test Z	ns	ns		

Note: ns = non-significant* p < 0.05; ** p < 0.01; *** p < 0.001.

Table 4.6. Student overall performance in the inquiry skills test

	Condition 1 (all words)	Condition 2 (some words)		Mann Whitney Test Z
Pre-test	0.60	0.48		ns
Post-test	0.60	0.54		ns
Wilcoxon Signed Ranks Test Z	ns	ns		

Note: ns = non-significant* p < 0.05; ** p < 0.01; *** p < 0.001.

Table 4.7. Student performance in "Identifying variables" (skill dimension)

	Condition 1 (all words)	Condition 2 (some words)		Mann Whitney Test Z
Pre-test	0.57	0.49		ns
Post-test	0.57	0.56		ns
Wilcoxon Signed Ranks Test Z	ns	ns		

Note: ns = non-significant* p < 0.05; ** p < 0.01; *** p < 0.001.

Table 4.8. Student performance in "Stating hypotheses" (skill dimension)

	Condition 1 (all words)	Condition 2 (some words)		Mann Whitney Test Z
Pre-test	0.64	0.48		ns

Post-test	0.64	0.51		ns
Wilcoxon Signed Ranks Test Z	ns	ns		

Note: ns = non-significant* p < 0.05; ** p < 0.01; *** p < 0.001.

4.5.2 Results in the Netherlands

Table 4.9. Student overall performance in the knowledge test

	Condition 1 (all words)	Condition 2 (some words)	Condition 3 (no words)	Kruskal-Wallis Test Chi-square
Pre-test	0.14	0.18	0.20	ns
Post-test	0.32	0.26	0.27	ns
Wilcoxon Signed Ranks Test Z	-4.28***	-2.06*	ns	

Note: * p < 0.05; ** p < 0.01; *** p < 0.001.

Table 4.10. Student performance in "Remember" (knowledge dimension)

	Condition 1 (all words)	Condition 2 (some words)	Condition 3 (no words)	Kruskal-Wallis Test Chi-square
Pre-test	0.00	0.02	0.00	ns
Post-test	0.05	0.06	0.11	ns
Wilcoxon Signed Ranks Test Z	ns	ns	-2.24*	

Note: non-significant; * p < 0.05; ** p < 0.01; *** p < 0.001.

.

Table 4.11. Student performance in "Understand" (knowledge dimension)

	Condition 1 (all words)	Condition 2 (some words)	Condition 3 (no words)	Kruskal-Wallis Test Chi-square
Pre-test	0.05	0.19	0.27	8.47*
Post-test	0.29	0.23	0.24	ns
Wilcoxon Signed Ranks Test Z	-2.98**	ns	ns	

Note: non-significant; * p < 0.05; ** p < 0.01; *** p < 0.001.

Table 4.12. Student performance in "Apply" (knowledge dimension)

	Condition 1 (all words)	Condition 2 (some words)	Condition 3 (no words)	Kruskal-Wallis Test Chi-square
Pre-test	0.35	0.34	0.32	ns
Post-test	0.62	0.49	0.44	6.51*
Wilcoxon Signed Ranks Test Z	-3.51**	-2.38*	-2.04*	

Note: non-significant; * p < 0.05; ** p < 0.01; *** p < 0.001.

Table 4.13. Student overall performance in the inquiry skills test

	Condition 1 (all words)	Condition 2 (some words)	Condition 3 (no words)	Kruskal-Wallis Test Chi-square
Pre-test	0.42	0.39	0.39	ns
Post-test	0.42	0.40	0.39	ns
Wilcoxon Signed Ranks Test Z	ns	ns	ns	

Note: non-significant; * p < 0.05; ** p < 0.01; *** p < 0.001.

Table 4.14. Student performance in "Identifying variables" (skill dimension)

	Condition 1 (all words)	Condition 2 (some words)	Condition 3 (no words)	Kruskal-Wallis Test Chi-square
Pre-test	0.37	0.35	0.36	ns
Post-test	0.36	0.33	0.36	ns
Wilcoxon Signed Ranks Test Z	ns	ns	ns	

Note: non-significant; * p < 0.05; ** p < 0.01; *** p < 0.001.

Table 4.15. Student performance in "Stating hypotheses" (skill dimension)

	Condition 1 (all words)	Condition 2 (some words)	Condition 3 (no words)	Kruskal-Wallis Test Chi-square

Pre-test	0.47	0.44	0.42	ns
Post-test	0.48	0.47	0.42	ns
Wilcoxon Signed Ranks Test Z	ns	ns	ns	

Note: non-significant; * p < 0.05; ** p < 0.01; *** p < 0.001.

4.5.3 Results in Estonia

Table 4.16. Student overall performance in the inquiry skills test

	Condition 1 (all words)	Condition 2 (some words)	Condition 3 (no words)	Kruskal-Wallis Test Chi-square
Pre-test	0,51	0,52	0,49	ns
Post-test	0,62	0,54	0,54	ns
Wilcoxon Signed Ranks Test Z	ns	ns	ns	

Note: non-significant; * p < 0.05; ** p < 0.01; *** p < 0.001.

Table 4.17. Student performance in "Identifying variables" (skill dimension)

	Condition 1 (all words)	Condition 2 (some words)	Condition 3 (no words)	Kruskal-Wallis Test Chi-square
Pre-test	0.519	0.544	0.468	ns
Post-test	0.506	0.417	0.479	ns
Wilcoxon Signed Ranks Test Z	ns	ns	ns	

Note: non-significant; * p < 0.05; ** p < 0.01; *** p < 0.001.

Table 4.18. Student performance in "Stating hypotheses" (skill dimension)

	Condition 1 (all words)	Condition 2 (some words)	Condition 3 (no words)	Kruskal-Wallis Test Chi-square
Pre-test	0.504	0.488	0.509	ns
Post-test	0.735	0.659	0.602	ns
Wilcoxon Signed Ranks Test Z	ns	ns	ns	

Note: non-significant; * p < 0.05; ** p < 0.01; *** p < 0.001.

Table 4.19. Student overall performance in the knowledge test

	Condition 1 (all words)	Condition 2 (some words)	Condition 3 (no words)	Kruskal-Wallis Test Chi-square
Pre-test	0.31	0.27	0.28	ns
Post-test	0.33	0.38	0.35	ns
Wilcoxon Signed Ranks Test Z	ns	ns	ns	

Note: ns = non-significant; * p < 0.05; ** p < 0.01; *** p < 0.001.

Table 4.20. Student performance in "Remember" (knowledge dimension)

	Condition 1 (all words)	Condition 2 (some words)	Condition 3 (no words)	Kruskal-Wallis Test Chi-square
Pre-test	0.00	0.00	0.00	ns
Post-test	0.00	0.10	0.08	ns
Wilcoxon Signed Ranks Test Z	ns	ns	ns	

Note: ns = non-significant; * p < 0.05; ** p < 0.01; *** p < 0.001.

Table 4.21. Student performance in "Understand" (knowledge dimension)

	Condition 1 (all words)	Condition 2 (some words)	Condition 3 (no words)	Kruskal-Wallis Test Chi-square
Pre-test	0.58	0.32	0.42	ns
Post-test	0.50	0.50	0.42	ns
Wilcoxon Signed Ranks Test Z	ns	ns	ns	

Note: ns = non-significant; * p < 0.05; ** p < 0.01; *** p < 0.001.

Table 4.22. Student performance in "Apply" (knowledge dimension)

	Condition 1 (all words)	Condition 2 (some words)	Condition 3 (no words)	Kruskal-Wallis Test Chi-square
Pre-test	0.36	0.50	0.42	ns
Post-test	0.49	0.61	0.58	ns
Wilcoxon Signed Ranks Test Z	ns	ns	-2.98**	

Note: ns = non-significant; * p < 0.05; ** p < 0.01; *** p < 0.001.

Table 4.23. Student overall performance in the inquiry skills test

	Condition 1 (all words)	Condition 2 (some words)	Condition 3 (no words)	Kruskal-Wallis Test Chi-square
Pre-test	0.35	0.36	0.38	ns
Post-test	0.54	0.48	0.44	ns
Wilcoxon Signed Ranks Test Z	-4.29***	-3.67***	-3.02**	

Note: ns = non-significant; * p < 0.05; ** p < 0.01; *** p < 0.001.

Table 4.24. Student performance in "Identifying variables" (skill dimension)

	Condition 1 (all words)	Condition 2 (some words)	Condition 3 (no words)	Kruskal-Wallis Test Chi-square
Pre-test	0.34	0.39	0.35	ns
Post-test	0.57	0.50	0.49	ns
Wilcoxon Signed Ranks Test Z	-4.30***	-2.63**	-3.62***	

Note: ns = non-significant; * p < 0.05; ** p < 0.01; *** p < 0.001.

Table 4.25. Student performance in "Stating hypotheses" (skill dimension)

	Condition 1 (all words)	Condition 2 (some words)	Condition 3 (no words)	Kruskal-Wallis Test Chi-square
Pre-test	0.35	0.32	0.39	ns
Post-test	0.50	0.45	0.38	6.01*
Wilcoxon Signed Ranks Test Z	-3.67***	-3.44***	ns	

Note: ns = non-significant; * p < 0.05; ** p < 0.01; *** p < 0.001.

4.5.4 Overview of results across all research settings

Table 4.26 presents an overview of significant improvements in student performance (i.e., knowledge and inquiry skills) across research settings.

Table 4.26. Improvement of student knowledge and inquiry skills across research settings

	UK	The Netherlands	Estonia	Cyprus
Overall performance in the knowledge test		Con 1; 2		
"Remember"	Con 1; 2	Con 3		
"Understand"		Con 1		
"Apply"		Con 1 > 2; 3		Con 3
Overall performance in the inquiry skills test				Con 1; 2; 3
"Identifying variables"				Con 1; 2; 3
"Stating hypotheses"				Con 1 > 2

Note: "Con" = Conditions; Conditions presented revealed significant improvement after the intervention; ">" denotes increased improvement over conditions.

4.6 Discussion, conclusions and implications for future research

It is obvious there was quite a heterogeneous picture of results. Such a diversity might imply that the use of the Hypothesis Scratchpad, in any configuration that it had been offered, might have been highly dependent upon the context, even if the learning activity sequence had been aligned among countries. The exact impact of the tool on student knowledge and inquiry skills might not be readily identifiable and each implementation might have been too idiosyncratic to locate common trends across research settings.

Given the heterogeneity in the results of the study, we might still discern two commonalities. First, we can observe that, when improvement differed significantly among conditions, then Condition 1 (all words) scored higher ("Apply" in the Netherlands; "Stating

hypotheses" in Cyprus). This appeared twice, in both knowledge and inquiry skill dimensions. The distinct improvement of Condition 1 might indicate that offering a fully-fledged guidance to students might be the most beneficial configuration of the tool, at least for separate dimensions of knowledge and inquiry skills. In this case, structure should be preferred over problematizing students while they formulate their hypotheses.

The second commonality refers to Condition 3, which did not offer any word to students. In two cases engaging knowledge dimensions, there was an improvement in this condition only ("Remember" in the Netherlands; "Apply" in Cyprus). This finding indicates that a configuration of the Hypothesis Scratchpad leaning towards "problematisation" might prove beneficial for cognitive processes.

Since the results of the present study cannot be conclusive, future research might investigate for more trends among different configurations of the Hypothesis Scratchpad. Such research that could engage novice learners as compared to more experienced learners, as well as students with relatively low prior knowledge as compared to students with relatively high prior knowledge. In these cases, Condition 1 of the tool might be contrasted to Condition 3. Finally, these two configurations of the tool could be employed in subsequent inquiry cycles to investigate whether they could be used to facilitate fading out of guidance in hypothesis formulation.

4.7 References

Arnold, J. C., Kremer, K., & Mayer, J. (2014). Understanding students' experiments - What kind of support do they need in inquiry tasks? *International Journal of Science Education*, 36, 2719-2749. doi: 10.1080/09500693.2014.930209

Burns, J. C., Okey, J. R., & Wise, K. C. (1985). Development of an integrated process skill test: TIPS II. *Journal of Research in Science Teaching*, 22(2), 169-177. doi: 10.1002/tea.3660220208

Chang, K. E., Chen, Y. L., Lin, H. Y., & Sung, Y. T. (2008). Effects of learning support in simulation-based physics learning. *Computers & Education*, 51(4), 1486-1498. doi: 10.1016/j.compedu.2008.01.007

Chin, C. A., & Brewer, W. F. (1993). The role of anomalous data in knowledge acquisition: A theoretical framework and implications for science instruction. *Review of Educational Research*, 63, 1-49.

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

de Jong, T. (2006b). Computer simulations – Technological advances in inquiry learning. *Science*, 312, 532–533.

de Jong, T. (Ed.). (2014). D1.3: Preliminary inquiry classroom scenarios and guidelines: Go-Lab project.

de Jong, T., Sotiriou, S., & Gillet, D. (2014). Innovations in STEM education: the Go-Lab federation of online labs. *Smart Learning Environments*, 1(1), 1-16. doi: 10.1186/s40561-014-0003-6

Germann, P. J., Aram, A., & Burke, G. (1996). Identifying patterns and relationships among the responses of seventh-grade students to the science process skill of designing experiments. *Journal of Research in Science Teaching*, 33, 79–99.

Hofstein, A., Navon, O., Kipnis, M., & Mamlok-Naaman, R. (2005). Developing students' ability to ask more and better questions resulting from inquiry-type chemistry laboratories. *Journal of Research in Science Teaching*, 42, 791–806.

Kalyuga, S. (2007). Expertise reversal effect and its implications for learner-tailored instruction. *Educational Psychology Review*, 19, 509-539. doi:10.1007/s10648-007-9054-3

Koksal, E. A., & Berberoglou, G. (2014). The effect of guided inquiry instruction on 6th grade Turkish students' achievement, science process skills, and attitudes toward science. *International Journal of Science Education*, 36, 66-78. doi: 10.1080/09500693.2012.721942

Kremer, K., Specht, C., Urhahne, D., & Mayer, J. (2014). The relationship in biology between the nature of science and scientific inquiry. *Journal of Biological Education*, 48, 1–8.

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.

Reiser, B. J. (2004). Scaffolding complex learning: The mechanisms of structuring and problematizing student work. *The Journal of the Learning Sciences*, 13(3), 273-304. doi:10.1207/s15327809jls1303_2

van Berkum, J. J. A., & de Jong, T. (1991). Instructional environments for simulations. *Education & Computing*, 6, 305-358.

van Joolingen, W. R., & de Jong, T. (1991). Supporitng hypothesis generation by learners exploring an interactive computer simulation. *Instructional Science*, 20, 389-404.

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

van Joolingen, W. R., & de Jong, T. (2003). SimQuest: authoring educational simulations. In T. Murray, S. Blessing, & S. Ainsworth (Eds.), *Authoring tools for advanced technology educational software: Toward cost-effective production of adaptive, interactive, and intelligent educational software* (pp. 1–31). Dordrecht: Kluwer Academic Publishers.

Zacharia, Z.C., Manoli, C., Xenofontos, N., de Jong, T., Pedaste, M., van Riesen, S., Kamp, E., Mäeots, M., Siiman. L., Tsourlidaki, E. (2015). Identifying potential types of guidance for supporting student inquiry when using virtual and remote labs: A literature review. *Educational Technology Research and Development*, 63, 257-302. doi: 10.1007/s11423-015-9370-0

Zervas, P. (Ed.). (2013). D2.1. *The Go-Lab inventory and integration of online labs – labs offered by large scientific organisations*: Go-Lab project.

5 Scaffolding Students' Experimentation

5.1 Abstract

Inquiry learning is an educational approach in which learners actively construct their own knowledge by conducting inquiry. Experimentation is a core element in science inquiry learning. Learners need guidance in order to design and conduct successful experiments. We combined the successful elements from previous work by others on forms of support into an Experiment Design Tool (EDT), which supports students in planning and conducting experiments in an online lab. In this study we compared the gain of conceptual knowledge for students using the EDT with two other conditions, in which students were guided by either main research questions or main research questions combined with more specific guiding questions. Students in all conditions worked with an online learning environment in which they had to plan and conduct experiments to answer research questions that were provided to them in the learning environment. All learning environments consisted of research questions, an online virtual laboratory about buoyancy and Archimedes' principle, and a conclusion text box. In addition to this, the EDT was included in the learning environment of the EDT condition, but not in the other two conditions. Students' conceptual knowledge was measured using a parallel knowledge test. Results showed a significant learning gain in all conditions but no significant differences between the conditions. However, a significant difference in learning gain between conditions in favour of the EDT was found for students with low prior knowledge, which indicates that low prior knowledge students do benefit from additional support in the form of the EDT.

5.2 Introduction

Inquiry learning has been widely recognized as an important teaching approach in science education (e.g. Minner, Levy, & Century, 2010). In science inquiry learning, learners actively construct their knowledge using methods and practices similar to those used by professional scientists (Keselman, 2003). They engage in multiple phases of inquiry by familiarizing themselves with the topic of interest, formulating research questions or hypotheses, planning and conducting experiments, drawing conclusions, reflecting upon their inquiry processes and results, and communicating their findings to others (de Jong, 2006; Pedaste et al., 2015; White & Frederiksen, 1998). The effectiveness of inquiry learning has been demonstrated in many studies, so long as learners are guided in their inquiry process (e.g., Alfieri, Brooks, Aldrich, & Tenenbaum, 2011; Furtak, Seidel, Iverson, & Briggs, 2012; Minner et al., 2010). Guided inquiry learning has also been found to promote a positive attitude towards learning (Hwang, Sung, & Chang, 2011), to foster critical thinking and high-level processing (Carnesi & DiGiorgio, 2009), and to motivate learners to acquire, incorporate, and apply new knowledge (Edelson, Gordin, & Pea, 1999).

One of the core elements of inquiry learning is the actual investigation (Osborne, Collins, Ratcliffe, Millar, & Duschl, 2003) during which learners design and conduct experiments. Designing experiments involves a number of processes and strategies. Learners first need to identify the variables associated with answering their research question or testing their hypothesis. They need to determine what variables to measure or observe (dependent variables), what variable to manipulate in order to test its effect on the dependent variable (independent variable), and what variables to control for because they may affect the results in unintended ways (controlled variables) (Arnold, Kremer, & Mayer, 2014). The second step in designing experiments is to assign values to the independent and

controlled variables. Different values are assigned to the independent variable across experimental trials, allowing the learner to investigate their effects on the dependent variable. Variables that are not manipulated, that is, controlled variables, have the same value across experimental trials, creating similar background conditions that allow the learner to compare results (Schunn & Anderson, 1999; Tschirgi, 1980).

Well-designed experiments serve as a bridge between the research question or hypothesis and the data analysis (Arnold et al., 2014), and provide the learner with adequate information to answer research questions or test hypotheses (de Jong & van Joolingen, 1998). A strategy often applied in well-designed experiments is the Control of Variables Strategy (CVS) in which learners vary only the variable of interest and keep all other (potentially causal) variables constant (Klahr & Nigam, 2004). CVS allows the learner to conclude that any effect that occurs is due to the one variable that was varied.

However, learners find it difficult to set up well-designed experiments (de Jong, 2006). They often design experiments that do not support answering their research question (de Jong & van Joolingen, 1998; Lawson, 2002). For example, they tend to manipulate variables that have nothing to do with the research question (de Jong & van Joolingen, 1998; van Joolingen & de Jong, 1991) or fail to identify the manipulable and observable variables within their research question or hypothesis. If the research question or hypothesis does not contain a directly manipulable variable, learners are often unable to convert abstract or theoretical variables into variables they can measure or observe (Lawson, 2002). Another difficulty is that learners sometimes vary too many variables, making it challenging to draw conclusions. When varying too many variables, one cannot tell which variable is responsible for an observed effect (Glaser, Schauble, Raghavan, & Zeitz, 1992).

To overcome these difficulties, learners can be guided in their inquiry. The majority of guidance that is used to support the investigational element in inquiry learning is presented as heuristics – rules of thumbs in the form of hints and suggestions about how to carry out certain actions, assignments, or learning processes (Zacharia et al., 2015). Heuristics are especially useful for novice learners who still need to learn about effective strategies that can be used to set up an investigation or for whom this process is not yet automated (Veermans, van Joolingen, & de Jong, 2006). Examples of heuristics for designing unconfounded experiments are ‘vary one thing at a time’, ‘assign simple values to the independent variable’, and ‘control all other variables by using the same value across experimental trials’ (Veermans et al., 2006). Another form of support can be found in scaffolds – tools that change the task or take over part of it for the learner, to help learners accomplish the learning task and gain higher order understandings they could not achieve on their own (de Jong, 2006; Reiser, 2004; Simons & Klein, 2007). One example of such a scaffold for science inquiry learning in an online environment is a monitoring tool (Veermans, de Jong, & van Joolingen, 2000) that allows learners to store experiments and presents the values and variables in table format. These variables can be sorted in ascending or descending order to easily compare different experiments, and experiments can be replayed. A second example is the SCY Experimental Design Tool, which presents an overview of the experimental process, that is, the research question, hypothesis, principle of manipulation, materials, and manipulation and data treatment. Learners can select an inquiry process and describe their experiment. They can evaluate their experimental setup with a checklist and they can retrieve additional information in the form of text that describes the selected process and gives instructions on how to perform the task (Lazonder, 2014).

In this study, we sought to combine heuristics and successful scaffolding elements into one combined scaffold called the Experiment Design Tool (EDT) that is intended to support learners in both planning and conducting experiments, as will be further explained in the method section. In the current study, the general effectiveness of the EDT was assessed.

The EDT was integrated into an online learning environment containing a virtual laboratory. Virtual laboratories are software simulation programs in which users carry out experiments on their computer (de Jong, Sotiriou, & Gillet, 2014). These laboratories have the advantage that variables can take on many values; they are also accurate, time- and cost-effective, and experiments can be repeated easily (Balamuralithara & Woods, 2009; de Jong, Linn, & Zacharia, 2013; Gomes & Bogosyan, 2009; Schiffhauer et al., 2012). The virtual laboratory used for the current study covered buoyancy and Archimedes' principle. Buoyancy plays an important role in science education and in everyday life, and it can be challenging for learners of all ages. It requires a conceptual understanding of density, floating, and sinking. Understanding the concept of density requires the simultaneous consideration of the two dimensions of mass and volume (Hardy, Jonen, Möller, & Stern, 2006). Whether an object sinks or floats depends on the density of the object as well as the density of the fluid in which the object is placed. The ideas that primary and secondary students have about this domain are often misconceptions. For example, they think that light objects float and heavy objects sink regardless of other features such as volume or material, that a true floating object must be entirely above the fluid, and that the volume of the fluid influences the flotation of an object (Biddulph & Osborne, 1984). Even adolescents and adults experience difficulties understanding the role of density in sinking and floating; they often predict that the weight of an object determines its flotability, they fail to recognize the ratio between mass and volume in materials, or they focus on specific features of objects such as holes (Driver, Squires, Rushworth, & Wood-Robinson, 1994, in Loverude, 2009; McKinnon & Renner, 1971, in Loverude, Kautz, & Heron, 2003).

Archimedes' principle can be explained in terms of fluid displacement, or in terms of forces. It entails that the mass of fluid displaced by a floating or submerged object is equal to the mass of the object, and the volume of fluid displaced by a sunken or submerged object is equal to the volume of the object (Hughes, 2005). Archimedes' principle states "an object fully or partially immersed in a fluid is buoyed up by a force equal to the weight of the fluid that the object displaces" (Halliday, Resnick, & Walker, 1997, in Hughes, 2005). The current study examined the effect of the Experiment Design Tool (EDT) on students' gain of conceptual knowledge about Archimedes' principle. Heron et al. (2003) found that it was helpful to provide learners with laboratory experience in buoyancy prior to introducing Archimedes' principle, in order to address intuitive ideas and misconceptions. To understand Archimedes' principle, students need a certain level of understanding of buoyancy. The students in the current study had already learned about buoyancy in previous science courses. Therefore, we chose buoyancy as the topic for students' introduction to the learning environment, which would also activate the prior knowledge necessary for understanding Archimedes' principle. The focus of the current study was on students' gain of conceptual knowledge about Archimedes' principle. The main research question was: Does using the EDT have a positive effect on students' gain of conceptual knowledge about Archimedes' principle?

5.3 Method

Students worked in an online learning environment where they received research questions and then had to plan and conduct experiments to answer them in a virtual laboratory called Splash. Three conditions were compared, with progressive levels of

support. It was expected that students who received the highest level of support would show the highest gain in conceptual knowledge about Archimedes' principle and that students with the lowest level of support would show the lowest gain.

5.3.1 Participants

A total of 120 third year pre-university students from four classes at two secondary schools in the Netherlands participated in this study. Within their own class, students were randomly assigned to one of the three conditions. Two classes involved students who participated in a special science interest program. Regular students and students in the science program were evenly divided over the conditions.

After eliminating students who missed a session or who conducted fewer than four experimental trials in the Archimedes parts of the Splash lab, the data from a total of 86 students were taken into account in the analyses.

5.3.2 Materials

5.3.2.1 Learning environment

Students in all conditions worked in an online learning environment with a similar structure. Students first received instructions explaining that they had to plan experiments and conduct them in Splash in order to answer a set of research questions. The learning environment itself incorporated three main elements: a research question (Figure 5.1, A), the Splash virtual lab in which experiments could be conducted (Figure 5.1, B), and a text field in which conclusions from the experiments could be entered (Figure 5.1, C). After students entered their conclusion for the research question, they received a new research question to investigate.

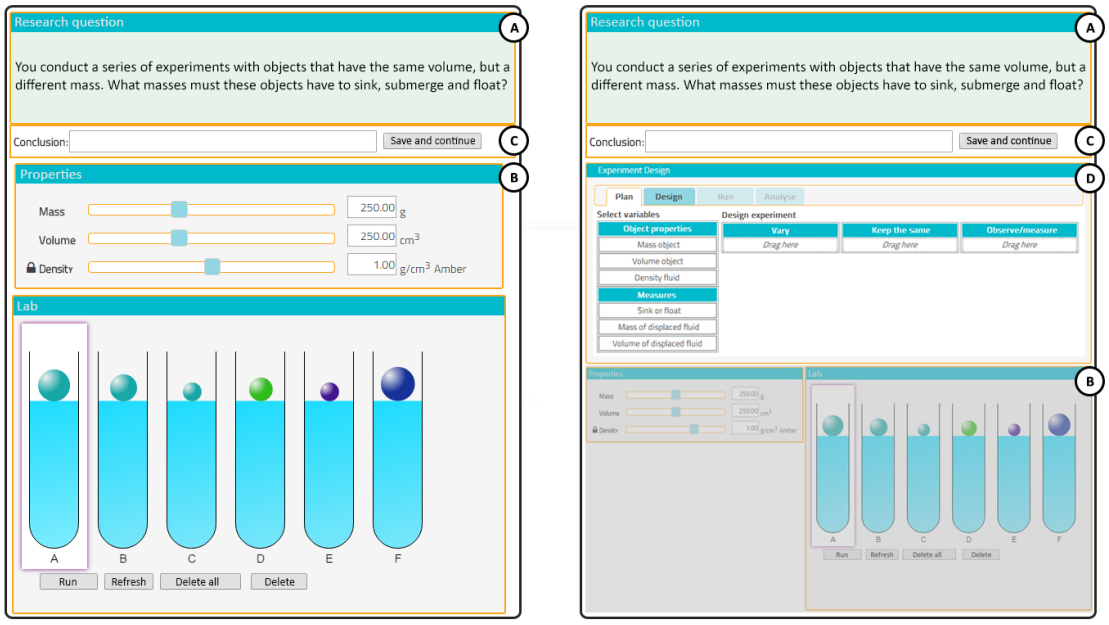


Figure 5.1. The learning environment. On the left is the interface seen in the main questions and guiding questions conditions (with a sub-question from the guiding questions condition in the Research question field), and the interface for the EDT condition (in the Plan step) is on the right.

The three learning environments differed in the support they provided. Students in the main questions condition worked with a version of the learning environment that gave them one main research question for each of the five topics, for example, "How do the mass

and volume of an object influence whether it sinks, submerges, or floats?" Students in the guiding questions condition worked in essentially the same learning environment. The only difference was that they were given thirteen sub-questions to conduct experiments about; this would enable them to answer the main questions. Introducing sub-questions breaks the task down into smaller pieces to guide students in the right direction. An example of such a question is: "You conduct a series of experiments with objects that have the same volume, but a different mass. What masses must these objects have to sink, submerge and float?" Students in the EDT condition received the same sub-questions. In addition, they were provided with the EDT, in which they planned their experiments and entered the results they obtained (Figure 5.1, D). Splash and the EDT are further explained in the next section.

5.3.2.2 Virtual laboratory: Splash

Splash is a virtual laboratory consisting of five sub-labs (Figure 5.2) in which students can conduct an unlimited amount of experiments to learn about buoyancy and Archimedes' principle. The first sub-lab displays containers filled with water in which students can place balls. Students can choose the mass, volume and density of the balls by means of sliders. After the properties of the balls are chosen, they can run experiments. The designed balls are dropped in the containers and students can observe whether the balls sink fully, submerge or float in water. The other sub-labs elaborate on this; in the third buoyancy sub-lab students can additionally adjust the fluid (density), and in the two Archimedes' principle labs students can also measure the volume and mass of displaced water, and measure forces within the buoyancy domain.

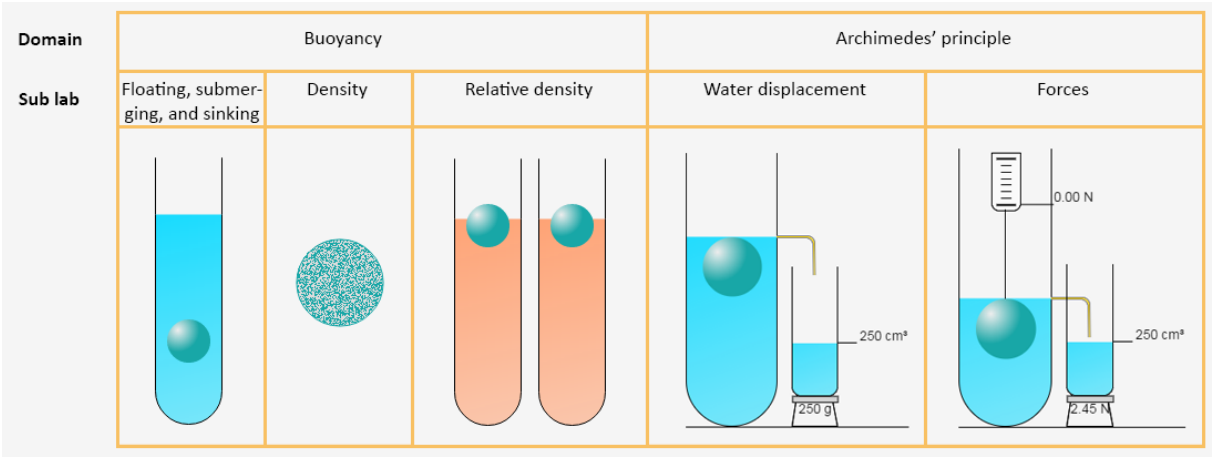


Figure 5.2. Examples of the five sub-labs in Splash.

5.3.2.3 The Experiment Design Tool

The Experiment Design Tool (Figure 5.3) breaks down the process of designing and conducting an experiment into several steps; 1) choosing the independent, controlled, and dependent variables, 2) assigning values to the variables, 3) conducting the experiment, and 4) analysing the results.

In the first step, students are given a list of predefined variables. For each variable, they decide whether they want to vary it across experimental trials (independent variable), keep it the same (controlled variable) or measure/observe it (dependent variable). They receive feedback based on their actions by means of a pop-up screen. For example, if they indicate that they want to vary a variable, they receive feedback that this means that they want to study the effect of the chosen variable on the variable they want to measure or observe. In the second step, students specify the number of experimental trials that together make

up one experiment and students choose the values of the controlled and independent variables. They assign one value per experimental trial to each independent variable (e.g., in the first trial they experiment with a mass of 300 grams and in the second trial they use a different mass of 400 grams), and a value to each controlled variable that remains the same over experimental trials within an experiment (e.g., volume = 200 cm^3 in every trial for this experiment). In the third step, students run their experiment. The trials they designed in the EDT were automatically transferred to Splash. After observing the results in Splash they document their observation or measurement of the dependent variable in the tool. Finally, students analyse their results. They can sort their data in ascending or descending order by variable, which provides them with some structure. This makes it easier to compare results and to decide if they can now draw conclusions based on their data or if they need to plan and conduct more trials or even more experiments to answer the research question.

5.3.2.4 Assessment

Students' conceptual knowledge was assessed both before and after the intervention with a parallel knowledge test that was designed especially for this study. The pre- and post-test contained the same questions, but the numbers within questions, as well as the order of questions, were changed. The tests consisted of two parts that each addressed what students would encounter in a session with the learning environment in the current study – buoyancy and Archimedes' principle – and contained open-ended questions for which students could obtain a maximum of 35 points. The tests had two parts because of the structure of this study, in which the initial session with the learning environment related to buoyancy served as an introduction to familiarise students with the learning environment and to encourage their retrieval of prior knowledge about this topic, while Archimedes' principle was the new topic to learn about in the second session through inquiry. The first part of the tests (25 points could be obtained) concerned buoyancy and covered the first three Splash sub-lab topics: floating, submerging and sinking; density; and relative density. The second part of the tests (10 points could be obtained) concerned Archimedes' principle and covered the final two Splash sub-lab topics: water displacement and forces. After students' tests were scored, one item related to Archimedes' principle was removed from the analysis because students interpreted that question very differently than was intended, which left a total of nine points to be obtained for the second part of the test.

The tests measured students' understanding of the key concepts and principles. Students were asked to write down definitions of the key concepts and they had to apply this knowledge by providing the masses, volumes and densities of balls in different situations, the amount of displaced water, and/or forces that act upon the ball or the displaced water. For example, in the first question they were asked to: "Give a definition of **density**. Include the terms **volume** and **mass** in your definition". In the second question, they were asked to: "Give the **volume**, the **mass**, and the **density** of three different balls". To do this, they had to understand that density equals mass in grams divided by volume in cm^3. Students obtained one point for each correct answer. In the example, they could obtain one point for the correct definition and one point for each correctly specified ball.

Because the test consisted of open questions that were scored using a coding scheme, a second researcher also used the coding scheme to score the post-tests from one of the four classes (n = 30). Agreement between the two researchers reached a Kappa = 0.943 with $p < 0.001$, which is considered to be a very high agreement.

To determine the reliability of the tests, separate Cronbach alpha's for both parts of the pre-test and the post-test were determined, based on the participants whose data were

taken into account for this study (n = 86). The first part of the pre-test (about buoyancy) had a Cronbach's alpha of .92 (25 points) and the second part of the pre-test (about Archimedes' principle) a Cronbach's alpha of .84 (9 points). The post-test had a Cronbach's alpha of .91 for the buoyancy part (25 points) and a Cronbach's alpha of .88 for the Archimedes' principle part (9 points).

5.4 Procedure

The study took place during four lessons of 50-60 minutes each, over a period of two and a half weeks. During the first session, students' prior conceptual knowledge was measured using the pen-and-paper pre-test as described above. They could use the entire lesson to complete the test, but all students finished within half an hour. The intervention began in the second session, where students worked with their assigned type of learning environment to familiarize themselves and to activate their prior knowledge about floating, sinking and submerging, density, and relative density. Students were placed in one of three conditions and were told that they were going to do experiments about floating, sinking and submerging, density, and relative density by individually designing and conducting experiments on the computer to answer the provided research questions**.** All of the required information was presented to them in the learning environment. They had one lesson to complete this introductory task. In the third session students worked in the same learning environment as in the second session, but they learned about different topics – water displacement and forces. During the fourth session students' conceptual knowledge was measured with the post-test, for which they could again use the entire lesson.

5.5 Results

In this study, the three conditions that were compared differed in the support students received for carrying out inquiry activities in an online learning environment. Because the data were not normally distributed, an independent samples Kruskal-Wallis test was conducted to check if the three experimental groups were comparable. No significant differences were found between the conditions regarding physics grade, $H(2) = 1.97$, $p = .374$, math grade, $H(2) = 2.24$, $p = .327$, and pre-test scores on both buoyancy, $H(2) = 1.26$, $p = .534$ and Archimedes' principle, $H(2) = 1.70$, $p = .428$.

First, we explored whether students did gain knowledge about buoyancy and Archimedes' principle. Wilcoxon signed-rank tests showed significant increases in scores from pre- to post-test for buoyancy ($Z = 2.414$, $p = .016$, $d_{Cohen} = 0.33$) and for Archimedes' principle ($Z = 6.126$, $p < .001$, $d_{Cohen} = 0.99$).

Separate analyses per condition were also performed to explore students' learning gain by condition. Scores for buoyancy significantly increased from pre- to post-test only in the extra questions condition ($Z = 1.992$, $p = .046$, $d_{Cohen} = 0.41$), with no significant increase in score from pre- to post-test found in the EDT condition ($Z = 1.732$, $p = .083$, $d_{Cohen} = 0.43$) or the main questions condition ($Z = 0.458$, $p = .647$, $d_{Cohen} = 0.10$). It should be noted that, as was expected, students scored very high on buoyancy in the pre-test, leaving little room for improvement. Students' scores regarding Archimedes' principle significantly increased from pre- to post-test in all conditions (main questions condition: $Z = 3.044$, $p = .002$, $d_{Cohen} = 0.92$, extra questions condition: $Z = 3.576$, $p < .001$, $d_{Cohen} = 0.83$; EDT condition: $Z = 4.133$, $p < .001$, $d_{Cohen} = 1.33$), indicating a learning effect in all conditions. Table 5.1 shows the means and SDs of the pre- and post-test scores for all conditions, as well as the difference scores between pre- and post-test.

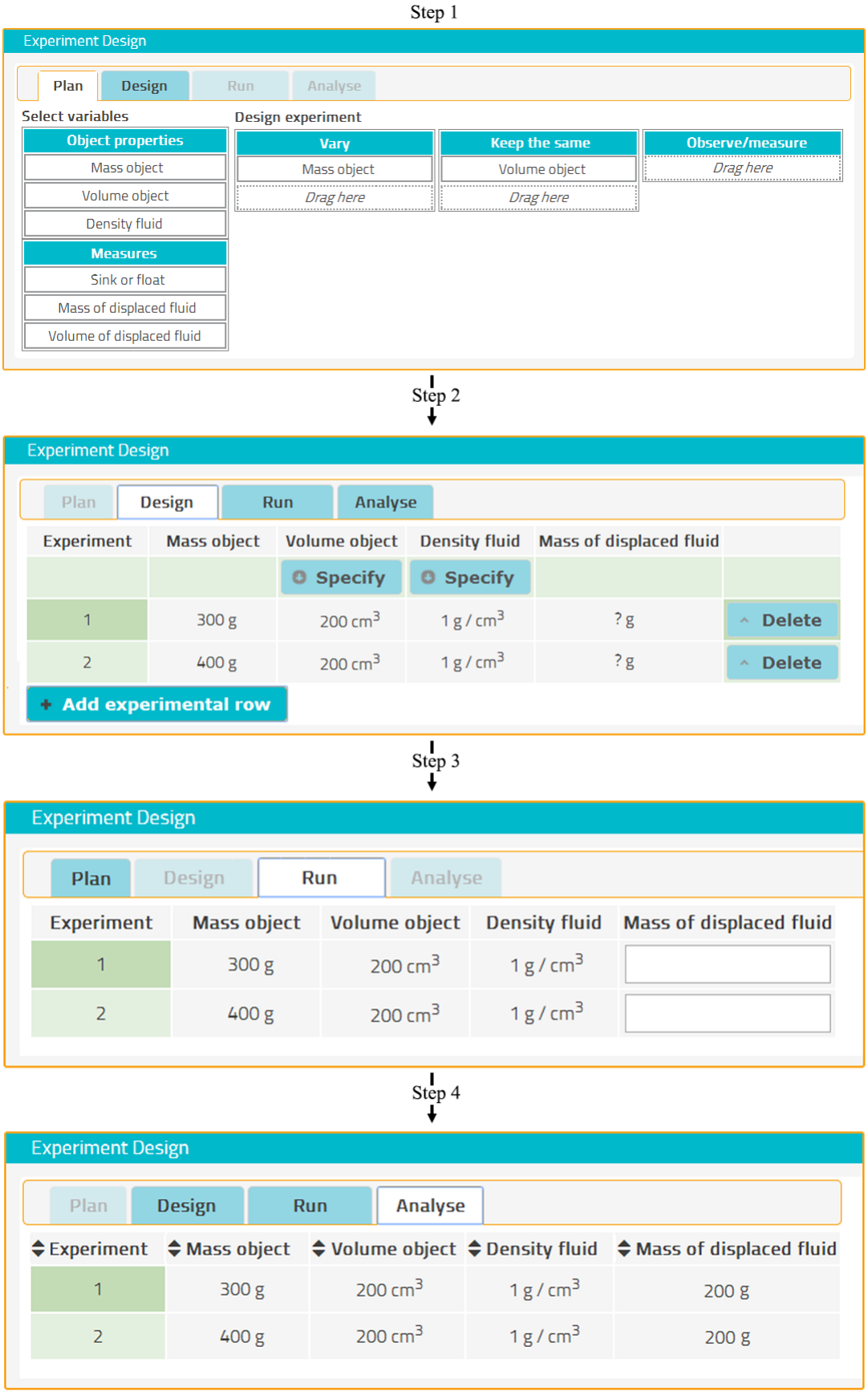


Figure 5.3. The Experiment Design Tool.

Table 5.1. Test Scores per Condition for All Students

	Main questions (*n* = 28)	Extra questions (*n* = 32)	EDT (*n* = 26)
Test	M (SD)	M (SD)	M (SD)
Buoyancy (Max = 25)			
Pre-test	21.61 (4.33)	20.16 (6.03)	19.96 (6.10)
Post-test	22.07 (4.54)	22.28 (4.25)	22.27 (4.41)
Difference score	0.46 (5.14)	2.13 (5.88)	2.31 (7.23)
Archimedes' principle (Max = 9)			
Pre-test	1.43 (2.06)	1.59 (1.78)	2.31 (2.72)
Post-test	3.89 (3.18)	3.72 (3.12)	5.65 (2.30)
Difference score	2.46 (3.61)	2.13 (2.66)	3.35 (2.58)

To determine whether students who received additional support gained more conceptual knowledge than students who did not receive the additional support, an independent samples Kruskal-Wallis test was performed. No significant differences were found between the conditions for both parts of the test (buoyancy: $H(2) = 1.51$, $p = .470$, and Archimedes' principle: $H(2) = 2.96$, $p = .228$.

However, the literature suggests that students with low prior knowledge benefit more from additional guidance than students with high prior knowledge (Tuovinen & Sweller, 1999). Therefore, the data from the current study was also analysed, looking at only low prior knowledge students. A total of 63 students were classified as low prior knowledge students because they had a maximum of two out of nine correct answers on the Archimedes' principle part of the pre-test, which is less than 25% of the maximum score. An independent samples Kruskal-Wallis test showed no significant differences in learning gain regarding buoyancy, $H(2) = 1.03$, $p = .60$, but did demonstrate a significant difference on the Archimedes' principle part of the test, $H(2) = 6.54$, $p = .038$. Follow-up Mann-Whitney analyses showed significantly higher learning gains regarding Archimedes' principle only for low prior knowledge students in the EDT condition compared to low prior knowledge students in the extra questions condition, $U = 115.50$, $z = 2.438$, $p = .015$, $r = .16$. Table 5.2 presents the means and SDs of pre- and post-test scores of low prior knowledge students, by condition. These findings suggest that low prior knowledge students do benefit from additional support in the form of an Experiment Design Tool.

Table 5.2. Test Scores per Condition for Low Prior Knowledge Students

	Main questions (*n* = 22)	Extra questions (*n* = 23)	EDT (*n* = 18)
Test	M (SD)	M (SD)	M (SD)
Buoyancy (Max = 25)			
Pre-test	21.68 (4.44)	18.74 (6.51)	19.28 (6.65)
Post-test	22.77 (3.28)	21.43 (4.74)	22.17 (4.54)
Difference score	1.09 (5.23)	2.70 (6.81)	2.89 (8.03)
Archimedes' principle (Max = 9)			
Pre-test	0.50 (0.86)	0.70 (0.93)	0.72 (0.89)
Post-test	3.64 (3.20)	2.74 (3.08)	5.17 (2.33)
Difference score	3.14 (3.27)	2.04 (2.99)	4.44 (2.23)

5.6 Conclusion and discussion

The goal of the study was to assess the effectiveness of the Experiment Design Tool that guides students in planning and conducting experiments in an online lab to answer research questions. Three conditions were compared that differed in the level of support students received in carrying out their inquiries. Students were guided by 1) main questions, 2) main questions together with guiding sub-questions, or 3) main questions, guiding sub-questions and the Experiment Design Tool. It was expected that students in the EDT condition would gain the most conceptual knowledge about Archimedes' principle.

Results indicated that students in all conditions showed a significant increase in conceptual knowledge about Archimedes' principle, which is in line with previous studies demonstrating the effectiveness of guided inquiry learning (Alfieri, Brooks, Aldrich, & Tenenbaum, 2011; Furtak, Seidel, Iverson, & Briggs, 2012; Minner et al., 2010). No significant differences in learning gain were found between the conditions. However, when the analysis focused on only low prior knowledge students, there was a significant difference between the conditions in favour of the EDT. Providing learners with additional support in the form of the EDT thus seems to be especially valuable for learners with low prior knowledge, which is in line with the literature. For instance, Tuovinen and Sweller (1999) suggest that novice learners benefit from higher levels of guidance, because guidance acts as a substitute for missing schemas. Experienced learners, on the other hand, are less likely to need additional instructional guidance, as they already possess schemas that support the construction of mental representations. Kalyuga (2007) refers to this phenomenon as the "expertise reversal effect", where guidance becomes redundant and thus, ineffective for experienced learners. The current study suggests that guidance indeed becomes redundant for learners with more prior knowledge.

5.7 References

Alfieri, L., Brooks, P. J., Aldrich, N. J., & Tenenbaum, H. R. (2011). Does discovery-based instruction enhance learning? *Journal of Educational Psychology, 103*(1), 1-18. doi: 10.1037/A0021017

Arnold, J. C., Kremer, K., & Mayer, J. (2014). Understanding students' experiments - What kind of support do they need in inquiry tasks? *International Journal of Science Education, 36*, 2719-2749. doi: 10.1080/09500693.2014.930209

Balamuralithara, B., & Woods, P. C. (2009). Virtual laboratories in engineering education: The simulation lab and remote lab. *Computer Applications in Engineering Education, 17*(1), 108-118. doi: 10.1002/cae.20186

Biddulph, F., & Osborne, R. (1984). Pupils' ideas about floating and sinking. *Research in Science Education, 14*, 114-124.

Blanchard, M. R., Southerland, S. A., Osborne, J. W., Sampson, V. D., Annetta, L. A., & Granger, E. M. (2010). Is inquiry possible in light of accountability? A quantitative comparison of the relative effectiveness of guided inquiry and verification laboratory instruction. *Science Education, 94*(4), 577-616. doi: 10.1002/sce.20390

Carnesi, S., & DiGiorgio, K. (2009). Teaching the inquiry process to 21st century learners. *Library Media Connection, 27*, 32-36.

de Jong, T. (2006). Computer simulations - Technological advances in inquiry learning. *Science, 312*(5773), 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*(2), 179-201. doi: 10.2307/1170753

de Jong, T., van Joolingen, W. R., Giemza, A., Girault, I., Hoppe, U., Kindermann, J., . . . van der Zanden, M. (2010). Learning by creating and exchanging objects: The SCY experience. *British Journal of Educational Technology, 41*(6), 909-921. doi: 10.1111/j.1467-8535.2010.01121.x

de Jong, T., Linn, M. C., & Zacharia, Z. C. (2013). Physical and virtual laboratories in science and engineering education. *Science, 340*(6130), 305-308. doi: 10.1126/science.1230579

de Jong, T., Sotiriou, S., & Gillet, D. (2014). Innovations in STEM education: The Go-Lab federation of online labs. *Smart Learning Environments, 1*, 1-16. doi: 10.1186/s40561-014-0003-6

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. doi: 10.1207/s15327809jls0803&4_3

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. doi: 10.3102/0034654312457206

Glaser, R., Schauble, L., Raghavan, K., & Zeitz, C. (1992). Scientific reasoning across different domains. In E. de Corte, M. Linn, H. Mandl, & L. Verschaffel (Eds.), *Computer-based learning environments and problem solving* (pp. 345-373). Berlin, Germany: Springer-Verlag.

Gomes, L., & Bogosyan, S. (2009). Current trends in remote laboratories. *IEEE Transactions on Industrial Electronics, 56*, 4744-4756. doi: 10.1109/TIE.2009.2033293

Hardy, I., Jonen, A., Möller, K., & Stern, E. (2006). Effects of instructional support within constructivist learning environments for elementary school students' understanding of "floating and sinking". *Journal of Educational Psychology, 98*(2), 307-326. doi: 10.1037/0022-0663.98.2.307

Heron, P. R. L., Loverude, M. E., Shaffer, P. S., & McDermott, L. C. (2003). Helping students develop an understanding of Archimedes' principle. II. Development of research-based instructional materials. *American Journal of Physics, 71*(11), 1188. doi: 10.1119/1.1607337

Hughes, S. W. (2005). Archimedes revisited: A faster, better, cheaper method of accurately measuring the volume of small objects. *Physics Education, 40*, 468-474.

Hwang, G., Sung, H., & Chang, H. (2011). *The effect of integrating STS strategy to online inquiry-based learning on students' learning performance*. Paper presented at the IEEE International Conferences on Internet of Things, and Cyber, Physical and Social Computing.

Kalyuga, S. (2007). Expertise reversal effect and its implications for learner-tailored instruction. *Educational Psychology Review, 19*, 509-539. doi: 10.1007/s10648-007-9054-3

Keselman, A. (2003). Supporting inquiry learning by promoting normative understanding of multivariable causality. *Journal of Research in Science Teaching, 40*, 898-921. doi: 10.1002/Tea.10115

Klahr, D., & Nigam, M. (2004). The equivalence of learning paths in early science instruction: Effect of direct instruction and discovery learning. *Psychological Science, 15*(10), 661-667. doi: 10.1111/j.0956-7976.2004.00737.x

Lawson, A. E. (2002). Sound and faulty arguments generated by preservice biology teachers when testing hypotheses involving unobservable entities. *Journal of Research in Science Teaching, 39*, 237-252. doi: 10.1002/Tea.10019

Lazonder, A. W. (2014). Inquiry learning. In J. M. Spector, M. D. Merrill, J. Elen, & M. J. Bishop (Eds.), *Handbook of research on educational communications and technology* (pp. 453-464). New York: Springer Science + Business Media.

Loverude, M. E. (2009). A research-based interactive lecture demonstration on sinking and floating. *American Journal of Physics, 77*(10), 897. doi: 10.1119/1.3191688

Loverude, M. E., Kautz, C. H., & Heron, P. R. L. (2003). Helping students develop an understanding of Archimedes' principle. I. Research on student understanding. *American Journal of Physics, 71*(11), 1178. doi: 10.1119/1.1607335

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. doi: 10.1002/Tea.20347

Osborne, J., Collins, S., Ratcliffe, M., Millar, R., & Duschl, R. (2003). What "ideas-about-science" should be taught in school science? - A Delphi study of the expert community. *Journal of Research in Science Teaching, 40*, 692-720. doi: 10.1002/Tea.10105

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. doi: 10.1016/j.edurev.2015.02.003

Reiser, B. J. (2004). Scaffolding complex learning: The mechanisms of structuring and problematizing student work. *Journal of the Learning Sciences, 13*(3), 273-304. doi: 10.1207/s15327809jls1303_2

Schiffhauer, S., Gößling, J., Wirth, J., Bergs, M., Walpuski, M., & Sumfleth, E. (2012). *Fostering experimental skills by a combination of hands-on and computer-based learning-environments*. Paper presented at the Annual Meeting of the American Educational Research Association (AERA), Vancouver, BC, Canada.

Schunn, C. D., & Anderson, J. R. (1999). The generality/specificity of expertise in scientific reasoning. *Cognitive Science, 23*, 337-370. doi: 10.1207/s15516709cog2303_3

Simons, K. D., & Klein, J. D. (2007). The impact of scaffolding and student achievement levels in a problem-based learning environment. *Instructional Science, 35*, 41-72. doi: 10.1007/s11251-006-9002-5

Tschirgi, J. E. (1980). Sensible reasoning: A hypothesis about hypotheses. *Child Development, 51*, 1-10. doi: 10.2307/1129583

Tuovinen, J. E., & Sweller, J. (1999). A comparison of cognitive load associated with discovery learning and worked examples. *Journal of Educational Psychology, 91*, 334-341. doi: 10.1037/0022-0663.91.2.334

van Joolingen, W., & de Jong, T. (1991). A hypothesis scratchpad as a supportive instrument in simulation learning environments.

Veermans, K., de Jong, T., & van Joolingen, W. R. (2000). Promoting self-directed learning in simulation-based discovery learning environments through intelligent support. *Interactive Learning Environments, 8*, 229-255. doi: 10.1076/1049-4820(200012)8:3;1-D;FT229

Veermans, K., van Joolingen, W. R., & de Jong, T. (2006). Use of heuristics to facilitate scientific discovery learning in a simulation learning environment in a physics domain. *International Journal of Science Education, 28*(4), 341-361. doi: 10.1080/09500690500277615

White, B. Y., & Frederiksen, J. R. (1998). Inquiry, modeling, and metacognition: Making science accessible to all students. *Cognition and Instruction, 16*(1), 3-118.

Zacharia, Z. C., Manoli, C., Xenofontos, N., de Jong, T., Pedaste, M., van Riesen, S. A. N., . . . Tsourlidaki, E. (2015). Identifying potential types of guidance for supporting student inquiry when using virtual and remote labs in science: A literature review. *Educational Technology Research and Development, 63*, 257-302. doi: 10.1007/s11423-015-9370-0

6 Guiding students in designing proper experiments; does ability level make a difference?

6.1 Abstract

In inquiry learning learners design and conduct experiments to build their own knowledge. However, inquiry learning has found to be difficult by students and previous research indicate that students with different levels of prior knowledge should receive other guidance. In the current study students received research questions they had to answer by conducting experiments in an online learning environment. Two versions were compared in which students either received direct instructions telling them what variable they had to vary and what they had to keep constant, or they had to decide this by themselves. The learning environments were used by students in the first and third year of secondary education, who followed either a lower or higher level of education. Results showed that students from both years of the higher level education performed equally well in the two learning environments, whereas both first and third year lower level students significantly benefited from direct instructions that let them know what to vary and what to keep constant in their experiments. These results indicate that especially students' level of education influences their need for additional support in inquiry learning.

6.2 Introduction

Science plays an important part in society and education. Learners are taught important principles in several science disciplines through direct instructions and other educational approaches. An approach that has received an increasing amount of attention is inquiry learning. During inquiry learning learners build their own knowledge by conducting inquiries instead of passively absorbing knowledge presented to them by the teacher or through other media (Carnesi & DiGiorgio, 2009). Learners follow (part of the) inquiry cycle; i.e., they formulate hypotheses and/or research questions, orient on the subject matter, set up investigations, draw conclusions, and reflect upon their process and results (Pedaste et al., 2015). Research shows that inquiry learning is an effective educational approach, as long as learners receive appropriate guidance to help them design meaningful inquiries (Alfieri, Brooks, Aldrich, & Tenenbaum, 2011). Inquiry learning has found to motivate learners to acquire, integrate and apply new knowledge and it promotes a positive attitude towards learning (Edelson, Gordin, & Pea, 1999; Hwang, Sung, & Chang, 2011).

One of the core phases of inquiry learning is the actual investigation phase, during which learners set up and conduct experiments to test their hypothesis or answer their research question, and gain (new) knowledge (Osborne, Collins, Ratcliffe, Millar, & Duschl, 2003). In order to gain conceptual knowledge it is essential for learners to set up valuable experiments that allow them to draw conclusions (van Riesen, Gijlers, Anjewierden, & de Jong, submitted). However, setting up useful experiments has found to be difficult for learners of all ages, and levels of education and experience (de Jong & van Joolingen, 1998). A mistake often made is that the designed experiment does not match the hypothesis or research question, i.e., learners incorporate variables in their experiment design that have nothing to do with their hypothesis or research question (de Jong & van Joolingen, 1998; Lawson, 2002). Another common mistake is that learners vary too many variables, making it impossible to draw sound conclusions based on unconfounded experiments (Glaser, Schauble, Raghavan, & Zeitz, 1992).

Learners' prior domain knowledge has found to influence the selected and implemented strategies they use in designing experiments, and inquiry learning in general. Learners with low prior knowledge encounter more problems and make use of less sophisticated strategies to induce knowledge than high prior knowledge learners, and they also execute those strategies less effectively (Schauble, Glaser, Raghavan, & Reiner, 1991). Low prior knowledge learners often have unsystematic ways of conducting inquiries, whereas high prior knowledge learners show well-structured, goal-oriented inquiry behaviour (Alexander & Judy, 1988; Hmelo, Nagarajan, & Roger, 2000; Schauble et al., 1991).

Studies have been inconclusive about the type and amount of support that should be offered to learners with different levels of prior knowledge and experience. One strand of studies suggests that novice learners should receive more guidance, as opposed to more experienced learners, because this can act as substitute for their missing schemas (Tuovinen & Sweller, 1999). Others conclude that additional guidance may be difficult to interpret and understand by novice learners and impedes on their cognitive capacities (van Joolingen & de Jong, 1991; Roll, Baker, Aleven, & Koedinger, 2014; Vollmeyer, Burns, & Holyoak, 1996). In order to gain better understanding about the need for guidance by learners with different levels of education and experience, we conducted a study in which we compared two online inquiry learning environments that differed in the guidance provided to learners of distinct years and levels of education that were presumed to have different prior knowledge and experiences. In both learning environments learners received research questions for which they had to design and conduct experiments; in one condition they received direct information about what to vary and keep the same, and in the other condition they did not.

6.3 Method

In the current study secondary students worked in an online learning environment in which they learned about buoyancy and Archimedes' Principle by planning and conducting experiments in a virtual laboratory called Splash. All students received research questions they had to answer based on the results of their experiments. Two learning environments were compared that differed in the level of support students received. In one condition students received research questions that did not contain any specific directions or instructions that could lead them towards an answer (non-instructed questions condition) and in the other condition students received research questions and additional instructions that informed them which variable to vary and which variables to keep the same in order to be able to answer the research question (instructed questions condition). First and third year senior general secondary education students (HAVO)[1] and first and third year pre-university students (VWO) took part in this study. It was expected that first year HAVO level students would benefit most from the additional support and that third year VWO level students would benefit least from the additional support.

[1] In the Netherlands students usually enter secondary school at the age of 12 (first year in this article) and attend secondary education according to their abilities. Three main levels of education are distinguished: 1) Pre-vocational secondary education (MAVO) is the lowest level and takes four years to prepare students for further vocational training, 2) Senior general secondary education (HAVO) is the middle level and takes five years to prepare students for applied universities, and 3) Pre-university education (VWO) is the highest level and takes six years to prepares students for university. After secondary education students can attend further education up to and including the highest level they are qualified for.

6.3.1 Participants

A total of 210 students from nine classes of one secondary school in the Netherlands participated in this study, of which 183 remained after eliminating outliers and students that had not participated in all sessions. All participants were first or third year HAVO- or VWO-students. Within their year and level of education they were assigned to one of the two conditions based on their pre-test scores so that the conditions were comparable in pre-test scores.

6.3.2 Materials

6.3.2.1 Learning environment

Students in both conditions worked in an online learning environment that had the same overall structure. Students first received instructions explaining that they had to design experiments and conduct those in Splash in order to answer research questions. The online learning environment itself consists of five elements: a research question (Figure 6.1, 1), a help button that contains domain information (Figure 6.1, 2), the Experiment Design Tool that scaffolds the design of experiments (Figure 6.1, 3), an online laboratory called Splash in which experiments can be conducted (Figure 6.1, 4), and a conclusion text box that appears when the experiments have been conducted. After students enter the conclusion in the text box a new research question appears.

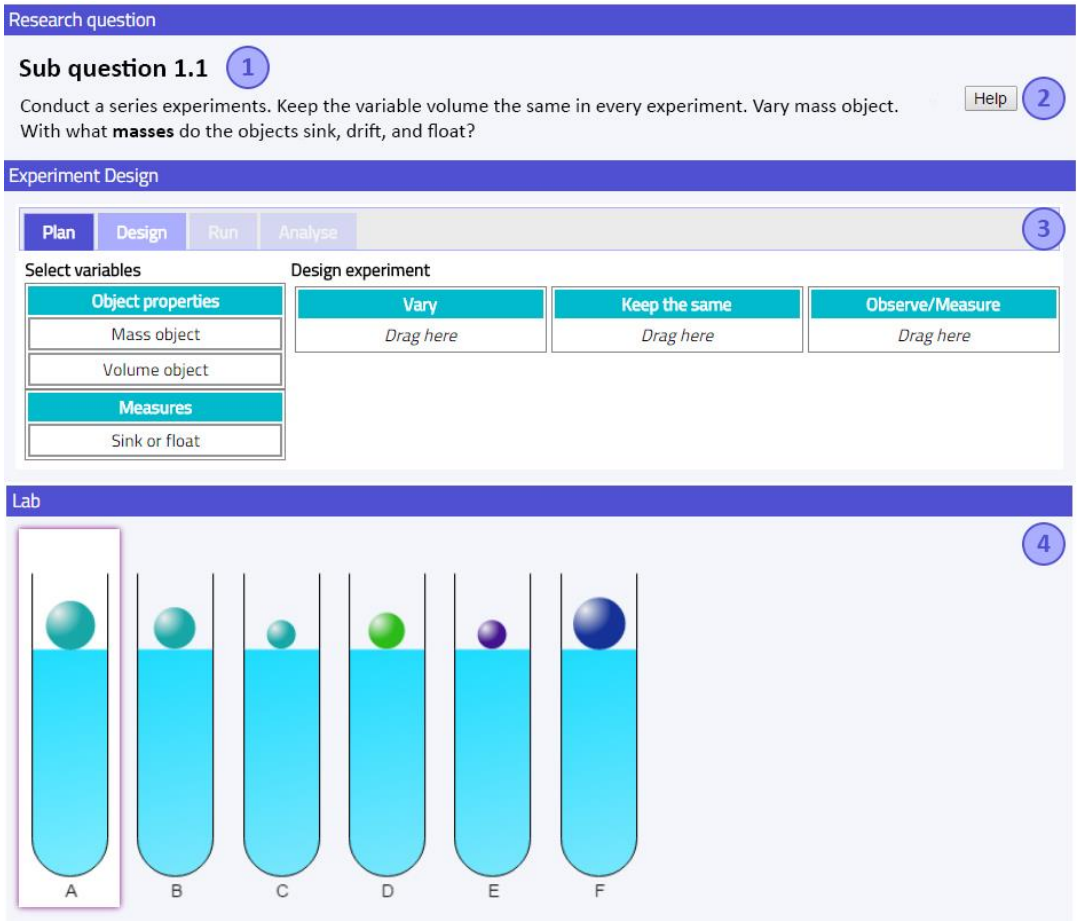


Figure 6.1. The online learning environment.

6.3.2.2 Virtual laboratory: Splash

Splash is a virtual laboratory about buoyancy and Archimedes' principle. Five sub labs are distinguished that each are about different topics: floating, sinking and drifting, density,

relative density, water displacement, and forces. Students can manipulate certain properties of the objects, as well as the density of the fluid, and observe what happens if they drop a ball into a fluid-filled container.

6.3.2.3 Experiment Design Tool

The Experiment Design Tool (EDT) scaffolds students' design of experiments by providing structure and segmenting the experimentation process. The tool breaks down the processes of designing and conducting experiments into smaller steps; 1) choose the control, independent, and dependent variables for the experiment, 2) assign values to these variables to create several experimental trials, 3) run the experiment, 4) and analyse the results.

6.3.2.4 Research questions

The learning environments differed in the research questions students had to answer. In the non-instructed questions condition students received research questions without direct instructions telling them what to do. An example of such a question is, "what is the influence of **mass** on the sinking, drifting, or floating of objects?". In the instructed questions condition students received similar questions, but additionally were instructed which variable to vary and which variables to keep the same. For example, "conduct a series of experiments. Keep the variable volume the same in every experiment. Vary mass object. With what **masses** do the objects sink, drift or float?"

6.3.2.5 Help button

The learning environment contained a help-button students could use if they were unfamiliar with a domain-specific term. For example, the term "mass" is explained as "the weight of the material of an object (in grams)".

6.3.3 Assessment

Students' knowledge of the topics in Splash was assessed both before and after the intervention by a parallel pre- and post-test taken from a previous study conducted by van Riesen and colleagues (submitted). The test consisted of 35 open questions that measured students' understanding of the key concepts and key principles of the topics in Splash. Students had to write down definitions and apply their knowledge by providing the mass, volume, and density of balls and fluids in different situations, the amount of water displaced by the ball, and/or forces that are present in the provided situations. To determine the reliability of the tests, separate Cronbach Alpha's for both the pre-test and the post-test were determined based on the participants that were taken into account for this study (N=183). The pre-test showed a Cronbach's Alpha of .954 (35 items) and the post-test showed a Cronbach's Alpha's of .958 (35 items), both of which demonstrate very high reliabilities.

6.4 Procedure

The study was performed during three sessions of 50 minutes each, over a period of two weeks. In the first session students received a brief introduction in which the advantages of computer-supported inquiry learning were presented. After this, they had 20 minutes to complete the pre-test as described above. Finally, they were informed about the activities in the next sessions, relevant concepts were explained, and they were given a demonstration of the learning environment. In the second session students were placed in one of the two conditions and worked individually with the corresponding learning environment to design and conduct experiments to answer research questions about the topics floating, sinking and drifting, density, and relative density. In the third session

students spent 30 minutes working with the learning environment about water displacement and forces. In the final 20 minutes of the third session students took the post-test.

6.5 Results and conclusions

In this study students had to design and conduct experiment to answer research questions they received. Two conditions that differed in the formulation of the research questions and instructions were compared in 1HAVO, 1VWO, 3HAVO, and 3VWO.

First, we explored if students (per group and condition) had learned from the intervention. Wilcoxon signed-rank tests showed significant increases in scores from pre-test to post-test for students in all groups and conditions (1HAVO instructed questions: $Z = 3.415$, $p = .001$; 1HAVO non-instructed questions: $Z = 2.613$, $p = .009$; 1VWO instructed questions: $Z = 4.170$, $p < .001$; 1VWO non-instructed questions: $Z = 3.830$, $p < .001$; 3HAVO instructed questions: $Z = 4.851$, $p < .001$; 3HAVO non-instructed questions: $Z = 3.594$, $p < .001$; 3VWO instructed questions: $Z = 3.072$, $p = .002$; 3VWO non-instructed questions: $Z = 3.470$, $p = .001$). Table 6.1 shows the means and SDs of the pre- and post-test scores of all groups.

Table 6.1. Means and SDs of pre-and post-test scores for each group separately

Group	*N*	Pre-test scores (Max.=35)		Post-test scores (Max.=35)	
		Mean	SD	Mean	SD
1HAVO instructed	20	0.25	0.64	4.05	3.46
1HAVO non-instructed	25	1.36	1.73	3.12	3.19
1VWO instructed	23	1.30	1.92	11.78	6.73
1VWO non-instructed	20	1.05	1.70	10.15	6.03
3HAVO instructed	37	10.76	7.76	20.00	7.64
3HAVO non-instructed	30	12.63	8.18	17.83	8.90
3VWO instructed	12	17.33	6.39	29.50	1.57
3VWO non- instructed	16	18.88	5.84	29.19	2.95

The main aim of the current study was to find out if students that differed in prior knowledge and experience should receive the same support for inquiry learning or not. We investigated if students gained more conceptual knowledge in the instructed research questions condition or the non-instructed research questions condition. First, separate analyses were done for all groups to analyse the difference within each group regarding the two conditions: 1HAVO, 1VWO, 3HAVO and 3VWO. Mann-Whitney tests showed significant differences between the two conditions in 1HAVO ($U = 156.50$, $z = -2.17$, $p = .030$, $r = -.32$) and 3HAVO ($U = 375.00$, $z = -2.27$, $p = .023$, $r = -.28$). No significant differences were found between the two conditions in 1VWO ($U = 201.00$, $z = -0.71$, $p = .479$, $r = -.11$) and 3VWO ($U = 80.00$, $z = -0.75$, $p = .478$, $r = -.14$). These results indicate that HAVO students (medium level education) benefit more from instructed questions than from non-instructed questions, whereas VWO students (highest level education) do not. There seems to be no difference between younger and older students of the same level of education regarding the type of research questions they learn more from.

6.6 Conclusion and discussion

In this study we investigated if students of different years and levels of education should receive different forms of support for their inquiry learning processes. We compared two online learning environments that differed in the support offered to students, and tested their effect on the learning gain of first and third year students from the HAVO and VWO track. Results showed that students from both years of the VWO track performed equally well on the two learning environments, whereas both first and third year HAVO students significantly benefited from direct instructions that let them know what to vary and what to keep the same in their experiments. These results indicate that especially students' level of education influences their need for additional scaffolding in inquiry learning. In a critical review of studies about inquiry learning and other educational approaches, Kirschner and colleagues (2006) argued that "minimally guided instruction is less effective and less efficient than instructional approaches that place a strong emphasis on guidance of the student learning process". Findings of the current study seem to be in line with these critical notes. HAVO students who received direct instructions telling them what to do performed better than students who had to find out how to set up valuable experiments themselves, whereas VWO students did equally well when they did or did not receive these direct instructions. The current study thus is in line with research that demonstrate that different learners need different levels of guidance; i.e. lower level students benefit from direct instructions in inquiry learning. It also became apparent that age did not influence the need for direct instruction in inquiry learning.

6.7 References

Alexander, P. A., & Judy, J. E. (1988). The interaction of domain-specific and strategic knowledge in academic performance. *Review of Educational Research, 58*, 375-404.

Alfieri, L., Brooks, P. J., Aldrich, N. J., & Tenenbaum, H. R. (2011). Does discovery-based instruction enhance learning? *Journal of Educational Psychology, 103*, 1-18. doi: 10.1037/a0021017

Carnesi, S., & DiGiorgio, K. (2009). Teaching the inquiry process to 21st century learners. *Library Media Connection, 27*, 32-36.

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*, 391-450.

Glaser, R., Schauble, L., Raghavan, K., & Zeitz, C. (1992). Scientific reasoning across different domains. In E. de Corte, M. Linn, H. Mandl, & L. Verschaffel (Eds.), *Computer-based learning environments and problem solving* (pp. 345-373). Berlin, Germany: Springer-Verlag.

Hmelo, C. E., Nagarajan, A., & Roger, S. (2000). Effects of high and low prior knowledge on construction of a joint problem space. *Journal of Experimental Education, 69*, 36-56.

Hwang, G., Sung, H., & Chang, H. (2011). *The effect of integrating STS strategy to online inquiry-based learning on students' learning performance*. Paper presented at the IEEE International Conferences on Internet of Things, and Cyber, Physical and Social Computing.

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: 10.2307/1170753

de Jong, T., van Joolingen, W. R., Giemza, A., Girault, I., Hoppe, U., Kindermann, J., . . . van der Zanden, M. (2010). Learning by creating and exchanging objects: The SCY

experience. *British Journal of Educational Technology, 41*(6), 909-921. doi: 10.1111/j.1467-8535.2010.01121.x

van Joolingen, W., & de Jong, T. (1991). An hypothesis scratchpad as a supportive instrument in simulation learning environments.

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

Lawson, A. E. (2002). Sound and faulty arguments generated by preservice biology teachers when testing hypotheses involving unobservable entities. *Journal of Research in Science Teaching, 39*, 237-252. doi: 10.1002/Tea.10019

Osborne, J., Collins, S., Ratcliffe, M., Millar, R., & Duschl, R. (2003). What "ideas-about-science" should be taught in school science? A Delphi study of the expert community. *Journal of Research in Science Teaching, 40*, 692-720. doi: 10.1002/Tea.10105

Pedaste, M., Maeots, 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. doi: 10.1016/j.edurev.2015.02.003

van Riesen, S. A. N., Gijlers, A. H., Anjewierden, A. A., & de Jong, T. (submitted). Scaffolding students' experimentation. *International Journal of Science Education*.

Roll, I., Baker, R. S. J. d., Aleven, V., & Koedinger, K. R. (2014). On the benefits of seeking (and avoiding) help in online problem solving environment. *Journal of the Learning Sciences*. doi: 10.1080/10508406.2014.883977

Schauble, L., Glaser, R., Raghavan, K., & Reiner, M. (1991). Causal models and experimentation strategies in scientific reasoning. *The Journal of the Learning Sciences, 1*, 201-238.

Tuovinen, J. E., & Sweller, J. (1999). A comparison of cognitive load associated with discovery learning and worked examples. *Journal of Educational Psychology, 91*, 334-341. doi: 10.1037/0022- 0663.91.2.334

Vollmeyer, R., Burns, B. D., & Holyoak, K. J. (1996). The impact of goal specificity on strategy use and the acquisition of problem structure. *Cognitive Science: A Multidisciplinary Journal, 20*, 75-100. doi: 10.1207/s15516709cog2001_3

7 Supporting planning and conducting experiments

7.1 Abstract

In inquiry learning learners are stimulated to construct their own knowledge by means of designing and conducting experiments from which conclusions can be derived. However, students show difficulties designing experiments based on which they can draw conclusions to answer their research questions. Scaffolds have shown to help students design useful experiments. We created two versions of a scaffold to guide students in designing experiments. One version of the scaffold (Experiment Design Tool - EDT) provided students with a set structure where they had to apply the CVS strategy – vary only one variable, keep all other variables constant, and measure at least one variable – and where they had to prepare at least three experimental trials before they could start conducting their experiment. The other version (Experiment Tool – ET) had a more exploratory character and provided students with the freedom to conduct experiments without the obligation to apply the CVS strategy and without having to plan multiple experimental trials before they could start conducting their experiment. In the current study students learned about buoyancy and Archimedes' principle by means of conducting experiments in an online lab. Three conditions were compared in terms of students' gain of conceptual knowledge, in which students' experiment designs were either guided by one of the versions of the scaffold, or in which they did not work with a scaffold. Results showed that lower prior knowledge students benefited significantly from additional support in the form of the ET to learn about buoyancy, compared to the other conditions. No significant differences were found between the conditions for higher prior knowledge students.

7.2 Introduction

Educational objectives are shifting in current society in which technology is now playing a central role and learners need to know and understand how and where to find information rather than memorising information. Learners are increasingly expected to take control of their own learning. An educational approach that anticipates on this development is inquiry learning, an approach that has found to be effective for learning as long as learners are guided in the processes involved (Alfieri, Brooks, Aldrich, & Tenenbaum, 2011). Inquiry learning stimulates learners to actively construct their own knowledge by means of conducting inquiry, which allows learners to gain higher-order understandings, instead of passively absorbing information presented to them (Carnesi & DiGiorgio, 2009). Learners often follow (part of the) inquiry cycle that constitutes orienting on the subject of interest, formulating hypotheses and/or research questions, setting up and conducting experiments, drawing conclusions, and reflecting upon their inquiry (Pedaste et al., 2015). Moreover, inquiry learning promotes learners' autonomous working attitudes and gain of inquiry skills, both of which are important educational objectives in current curricula worldwide; it also promotes a positive attitude towards learning (Hwang, Sung, & Chang, 2011), and it motivates them to acquire, integrate, and apply new knowledge (Edelson, Gordin, & Pea, 1999).

A core phase of inquiry learning is the investigation phase (Osborne, Collins, Ratcliffe, Millar, & Duschl, 2003; van Riesen, Gijlers, Anjewierden, & de Jong, submitted). In this phase learners design and conduct experiments to test a hypothesis or answer a research question. Based on results from their experiments they analyse their data and draw conclusions accordingly. The experimentation phase thus builds a bridge between the hypothesis or research question and data analysis (Arnold, Kremer, & Mayer, 2014).

However, designing valuable experiments has found to be difficult (de Jong & van Joolingen, 1998). It involves several processes, requires understanding of inquiry, and learners need to possess or gain inquiry skills. They need to understand that they have to design experiments that comply with their hypothesis or research question, i.e. that their experiment must be designed in such a way that it can test their hypothesis or answer their research question. A mistake often made by learners is that they design experiments with which they cannot test their hypothesis or answer their research question (de Jong & van Joolingen, 1998; Lawson, 2002). For example, they design experiments in which they include variables that have nothing to do with their hypothesis or research question (de Jong & van Joolingen, 1998; van Joolingen & de Jong, 1991). To be able to test the hypothesis or answer the research questions, learners need to select relevant variables to test the hypothesis or answer the research question.

After selecting relevant variables, learners need to specify these variables; they have to determine what they want to measure (dependent variable), vary (independent variable) and control for (controlled variable), and assign values to the independent and controlled variables (Arnold et al., 2014). Learners often make the mistake of varying too many variables, which makes it difficult to draw correct conclusions because any effect that occurs may be due to a variety of influences (Glaser, Schauble, Raghavan, & Zeitz, 1992). An effective strategy often applied by professional researchers is the Control of Variable Strategy (CVS) in which only the variable of interest is varied and all other variables are kept constant (Klahr & Nigam, 2004). CVS allows learners to draw conclusions from unconfounded experiments. Any occurring effect on the dependent variable can then be ascribed to the independent variable of interest.

After learners have specified all variables, the experiment can be conducted, which can take place in field settings or more controlled settings like hands-on laboratories, remote laboratories, or virtual laboratories. Hands-on laboratories are traditional physical labs in which learners need to be present to set up and conduct the experiment (Reuter, 2009). It involves gathering and preparing all materials, and learners should make sure that the variables of interest can be manipulated with the materials that are available to them. Depending on the subject matter, hands-on labs can involve certain risks, both for learners and the utilised equipment (Corter, Esche, Chassapis, Ma, & Nickerson, 2011). An example is that certain materials can be toxic or can explode if the lab is not operated correctly. Online labs eliminate those kind of risk factors for the learner, because these labs are operated through a medium like a computer. Two kinds of online labs can be distinguished, remote and virtual labs. Remote labs are physical labs that can be operated through a medium (Almarshoud, 2011). The lab owner prepares the remote lab by setting up the experiment and making sure it can be operated remotely, which is a one-time procedure making it cost- and time effective (Almarshoud, 2011; Corter et al., 2011). The lab owner decides the actions learners can perform, and restricts other operations. This builds in some limitations for learners because they are unable to explore everything, but it also eliminates risks for both the learner and the apparatus through built-in safety mechanisms, providing learners with the opportunity to work with (advanced) equipment they would be prohibited to use otherwise (Cooper, 2005; Gomes & Bogosyan, 2009). Remote labs thus allow learners to operate physical equipment through a medium, within a restricted experimentation space. Virtual labs are also operated through a medium, but are described as simulations of reality (Sancristobal et al., 2012). Like in remote labs, the lab designer decides what variables can be manipulated and what values can be assigned to those variables. This provides learners with visual constrains which can lead them in the correct direction to build knowledge (Toth, Ludvico, & Morrow, 2014). An important

advantage of virtual labs is that they allow variables to take on many values, and learners can conduct an unlimited amount of experiments that consume less time than experiments conducted in hands-on labs, providing excellent opportunities to gain theoretical understandings (Almarshoud, 2011). In a recent study of Toth, Ludvico, and Morrow (2014) in which hands-on labs were compared with virtual labs it was found that virtual labs had significant advantages for gaining conceptual knowledge, and learning was more deep and purposeful than learning with hands-on labs. Virtual labs thus are very suitable to explore theoretical foundations of a domain. For our current study as described in this chapter, we chose to include a virtual lab in the online learning environment, because the focus of our study is on students' gain of theoretical knowledge.

Designing experiments requires self-regulation (de Jong & Njoo, 1992; Zacharia et al., 2015). An essential regulatory process in inquiry learning, found to predict declarative knowledge gains, is planning (de Jong & Njoo, 1992; Schraw, Crippen, & Hartley, 2006). During planning learners set goals, implement selected strategies to achieve the set goals, activate relevant prior knowledge, and budget their time (Schraw et al., 2006; Schunk, 1996). However, research indicates that learners have the tendency not to analyse a task or problem they have to solve, but to act immediately, without engaging in spontaneous or serious planning (Manlove, Lazonder, & de Jong, 2006; Veenman, Elshout, & Meijer, 1997). If learners do engage in planning, they often use unsystematic ways of planning, which may cause them to struggle with the task (de Jong & van Joolingen, 1998; Veenman et al., 1997).

Guiding learners in planning and designing experiments helps them to conduct useful and systematic experiments from which they can derive knowledge (Zacharia et al., 2015). In computer supported inquiry learning environments guidance can be provided in the form of, for example, heuristics, scaffolds, process constraints, performance dashboards, prompts, or information directly presented to the learner (de Jong & Lazonder, 2014). Heuristics – hints or suggestions about how to carry out certain assignments, actions, or learning processes – are among the most often used forms of guidance and are especially useful for novice learners who still have to become familiar with effective strategies to set up an investigation or who have not yet automated the processes involved (Veermans, van Joolingen, & de Jong, 2006; Zacharia et al., 2015). Examples of heuristics to direct learners to apply the CVS strategy are 'vary one thing at a time (VOTAT)', and 'control all other variables by using the same value across experimental trials' (Veermans et al., 2006). A second form of support often used in computer assisted inquiry learning are scaffolds – tools that transform or take over part of a task and thereby help learners accomplish tasks they would not have been able to do on their own (de Jong, 2006; Reiser, 2004; Simons & Klein, 2007). An example of such a scaffold is the monitoring tool in which experiments – described as a set of values assigned to input and output variables – are stored (Veermans, de Jong, & van Joolingen, 2000). The rationale behind this tool is that learners can focus on important relationships within the domain of interest because the tool takes over part of the task by providing learners with some sort of external memory where their conducted experimental trials are automatically stored. Learners can replay conducted trials, and rearrange them in ascending or descending order to be better able to compare results. The monitoring tool eliminates the difficulty of remembering the conducted experimental trials, interpreting the results, while simultaneously thinking of appropriate follow-up trials to conduct. A second example of a scaffold is the SCY Experimental Design Tool in which learners can write and evaluate their experimental setup by means of a checklist (Lazonder, 2014). The scaffold incorporates an overview and explanations of experimental processes, including the research question, hypothesis,

principle of manipulation, materials, and data treatment. Moreover, learners receive instructions on how to perform the task.

Based on successful heuristics and scaffolding elements, like the examples above, an Experiment Design Tool (EDT) was developed by van Riesen and colleagues (submitted) to help learners plan and conduct experiments. The EDT breaks down the process of planning and conducting an experiment into smaller steps in order not to overwhelm learners. Learners first select the variables they want to include in their experiment from a predefined list. Second, they specify those variables and determine their values (within predefined ranges) for each experimental trial. Third, they conduct the prepared experimental trials in the lab and document the results in the EDT. Finally, they analyse results and draw conclusions. The EDT is meant to encourage learners to plan their experiment before conducting it within a constrained learning environment, thereby allowing them to gain theoretical understandings. Results from a recent study with the EDT showed that low prior knowledge learners significantly benefited from the EDT (van Riesen et al., submitted). In the current study, presented in this chapter, the EDT was further developed based on observations and findings of the abovementioned study; two revised versions of the EDT called the Experiment Tool (ET) and the Experiment Design Tool (EDT) were compared in terms of effectiveness regarding students' learning gain about buoyancy and Archimedes' principle. The two revised versions of the EDT have the same overall structure, look and feel. However, in the EDT-version learners are obliged to apply the CVS-strategy and to plan multiple trials before conducting their experiment, whereas in the ET-version they are free to conduct their designed trials whenever they want and they are not obliged to apply the CVS-strategy.

The domains learners in the current study had to learn about concerned buoyancy and Archimedes' principle. Buoyancy was chosen because it plays an important role in science education, it can be challenging for learners of all ages, and its understanding is a prerequisite for understanding Archimedes' principle (van Riesen et al., submitted). Buoyancy requires a conceptual understanding of density (mass divided by volume), floating and sinking; objects placed in a fluid float when the density of the object is lower than the density of the fluid, and they sink when the density of the object is higher than the density of the fluid (Hardy, Jonen, Möller, & Stern, 2006). Learners of all ages experience the challenge of understanding the relationship between density and floating and sinking; they often think the floatability of an object is determined by its weight, they fail to recognize the relationship between mass and volume in materials, or they focus on specific features of objects such as holes (Driver, Squires, Rushworth, & Wood-Robinson, 1994, in Loverude, 2009; McKinnon & Renner, 1971, in Loverude, Kautz, & Heron, 2003). Archimedes' principle is related to buoyancy and is often used as extra subject matter in Dutch education. Archimedes' principle can be explained in terms of water displacement or forces (van Riesen et al., submitted). Floating objects have the same mass as the fluid they displace, sinking objects have the same volume as the displaced fluid, and submerged objects have the same mass and volume as the fluid they displace (Hughes, 2005). When explained in terms of forces, Archimedes' principle states that "an object fully or partially immersed in a fluid is buoyed up by a force equal to the weight of the fluid that the object displaces" (Halliday, Resnick, & Walker, 1997, in Hughes, 2005).

7.3 Method

In the current study students planned and conducted experiments in an online learning environment about buoyancy and Archimedes' Principle. The learning environments provided students with research questions they were to answer based on the results of

their experiments. Three learning environments were compared that contained different levels of support for planning and conducting experiments, but that were the same in all other aspects. In two of the learning environments students received additional support for planning and conducting experiments by means of one of the two revised versions of the EDT that was designed by van Riesen and colleagues (submitted). In the third learning environment students were not guided by an additional scaffold. It was expected that students who were guided by a scaffold would gain more conceptual knowledge from pre- to post-test than students who were not guided by a scaffold.

7.3.1 Participants

Three secondary schools in the Netherlands participated in the current study with a total of 173 students from six third grade classes, of which 159 students were in the pre-university track and 14 in the senior general secondary education track. The focus of the study was on pre-university students but one combined class with students from both tracks participated. Only pre-university students were taken into account in the analyses. After eliminating outliers and students that missed a session the data of a total of 104 students remained for analyses.

7.3.2 Materials

7.3.2.1 Learning environment

The learning environments in which students worked were all structured in similar ways. When students went to the learning environment they first received instructions explaining that they had to design experiments and conduct those in a virtual laboratory called Splash in order to answer research questions. After students continued they reached the actual learning environment (Figure 7.1) that consisted of a research question they had to answer, a virtual lab called Splash, a mechanism to prepare experiments, a help button to retrieve domain information, and a conclusion text box in which students could type the answer to the research question. Once students entered the conclusion in the text box a new research question to investigate appears.

The learning environments differed in the support offered to students. Students in the control condition worked with a learning environment that did not contain a scaffold to help them plan their experiments. Students prepared and conducted their experiments directly in Splash by means of sliders to adjust the values of the variables in the experiment. In the Experiment Tool (ET) condition students were offered a scaffold, as explained in the next section, with which they planned and prepared their experiments. In the Experiment Design Tool (EDT) condition students worked with the EDT to help them conduct systematic experiments, as explained later.

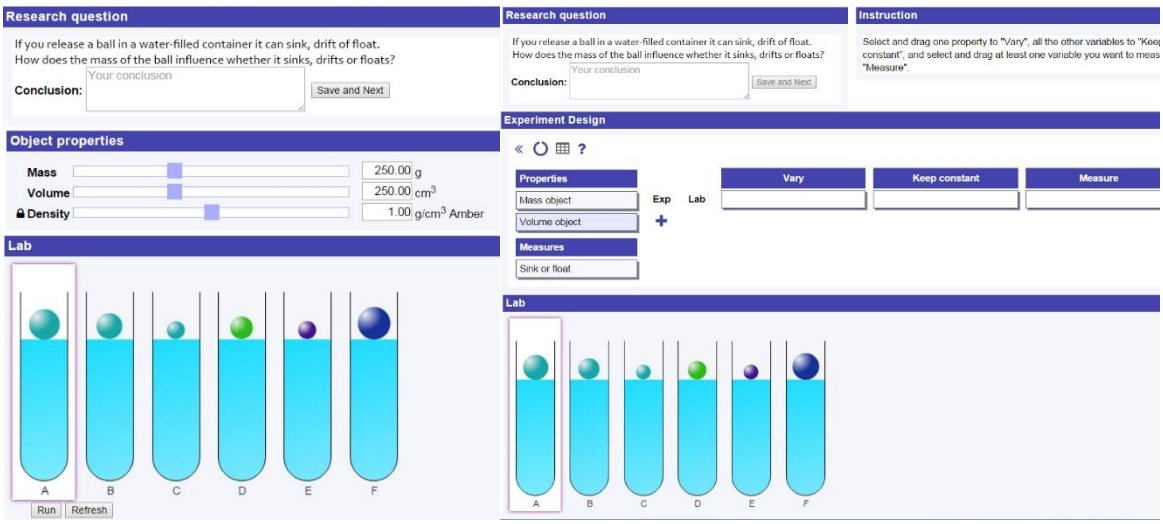


Figure 7.1. The learning environments with left the learning environment for the control group and right the learning environment for students in the ET and EDT condition.

7.3.2.2 Experiment Tool

The Experiment Tool (ET) guides students in planning and conducting their experiments, Figure 7.2. It provides students with structure in the form of a table that includes and emphasises three types of variables important in experimentation (independent, controlled and dependent). Students can select variables from a box and decide per variable if they want to vary it across experimental trials, keep it constant, or measure it. After students have specified the variables, they can assign values – within a restricted range –to the independent and controlled variables by means of sliders. They can choose to prepare one experimental trial or several trials in the ET before they conduct those in Splash. If they choose to prepare more than one trial they have to assign different values to all independent variables, and one value to each controlled variable. When students provide a value for a controlled variable the ET automatically assigns that same value to all trials within the experiment. Students can only enter results for the dependent variable after the trials are conducted.

In addition to preparing the experiments, the ET also offers a table in which all of the previously conducted trials are presented. This table provides students with the possibility to sort variable values ascending or descending, making it easier to reach conclusions or decide if more trials or even experiments are required to answer the research question. At all times, students can see instructions about how to operate the ET. The instructions students receive are just-in-time instructions that are provided to them based on their actions. For example, if students start planning they receive instructions to drag and drop all properties to the box vary or keep constant, and to drag at least one variable they want to measure to measure.

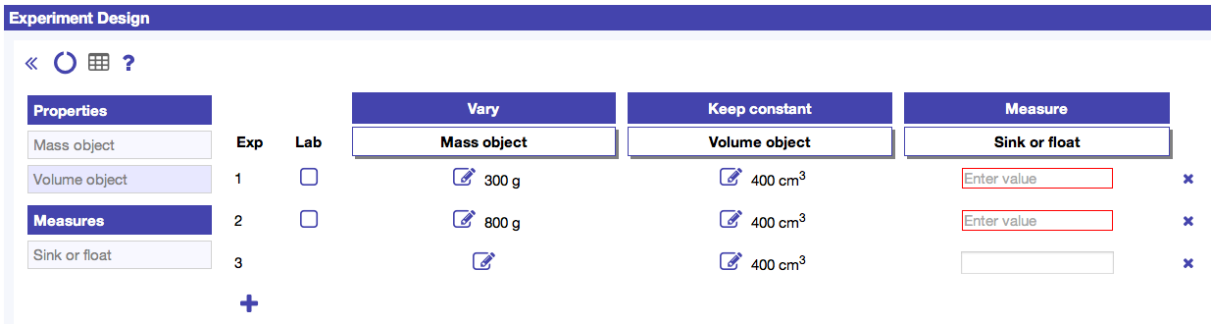


Figure 7.2. The Experiment Tool (ET).

7.3.2.3 Experiment Design Tool

The Experiment Design Tool (EDT) scaffolds students' design of experiments. The EDT is comparable to the ET in many ways but has two unique additions. In the EDT, as opposed to the ET, students have to apply the CVS strategy, i.e., they are only allowed to vary one variable at a time and have to keep all other variables constant. This strategy allows students to draw conclusions from unconfounded experiments. If students' experiment design in the EDT does not meet this criterion, they cannot proceed to the next step of conducting their experiment. The second feature of the EDT, is that students must first plan at least three experimental trials before they are able to conduct their experiment in Splash. As explained in the introduction, planning is an important regulatory process in inquiry learning. The EDT stimulates students to plan their experiment and to apply the CVS strategy.

7.3.2.4 Assessment

Students' knowledge of buoyancy and Archimedes' Principle was assessed both before and after the intervention with a parallel pre- and post-test. The test that was used was created by van Riesen and colleagues (submitted), but the part regarding Archimedes' principle was adapted to include more concepts within the domain. The adjusted version of the test consists of 58 open questions that measured students' understanding of the key concepts and principles of the topics in Splash. Students had to write down definitions and apply their knowledge by providing the mass, volume, and density of balls and fluids in different situations, the amount of water that was displaced, and/or forces that are present in the provided situations. To determine the reliability of the tests, separate Cronbach Alpha's for both parts (buoyancy and Archimedes' principle) of the pre-test (n = 118) and the post-test (n = 108) were determined based on the participants that completed the test and participated in the current study. The first part of the pre-test (about buoyancy) showed a Cronbach's Alpha of .934 (25 items), the second part of the pre-test (about Archimedes' principle) a Cronbach's Alpha of .880 (33 items). Cronbach's Alpha's of .899 were found for the buoyancy part of the post-test (25 items) and of .880 for the Archimedes' principle part of the post-test (33 items), all of which demonstrate very high reliabilities.

7.3.3 Procedure

The study was performed during four sessions of 50 minutes each, over a period of two and a half weeks. During the first session students were first briefly introduced to the researcher and they were explained what they were going to do. Thereafter they had half an hour to complete the pre-test for which they were allowed a calculator and a pen. Finally, they were placed into one of the three conditions and were given instructions about the tasks they were going to perform in the following lessons, and a demonstration was given on how to operate the learning environment. They could also ask any questions they

(still) had. During the second session students received a booklet matching the condition they were assigned to. All booklets contained instructions about the tasks they were going to perform, and the research questions they had to answer during the lesson. In addition, the booklets of students in the control group contained specific areas in which students could write down anything they wanted during the session in order for them to be able to plan experiments and write down results if desired. Students individually worked with the learning environment behind a computer. All the instructions were already provided to them in the first session and were additionally present in the learning environment and on paper, so they could immediately start planning and conducting experiments to learn about buoyancy. The third session was similar to the second session; students also worked with the learning environment but the topic of investigation was Archimedes' principle instead of buoyancy. During the fourth session students took the post-test and were told about the purpose of the study.

7.4 Results

In the current study three conditions were compared with different support for planning and conducting experiments in an online learning environment.

First, we explored if students in all conditions gained knowledge about buoyancy and Archimedes' principle. Wilcoxon signed-rank tests showed a significant increase in score for buoyancy from pre- to post-test in all conditions (Control condition: $n = 34$, $Z = 3.226$, p = .001; ET condition: $n = 33$, $Z = 3.302$, p = .001; EDT condition: $n = 37$, $Z = 3.015$, p = .003). A significant increase in score from pre- to post-test using Wilcoxon signed-rank tests was also found for all conditions for Archimedes' principle (Control condition: $n = 34$, $Z = 3.554$, p < .001; ET condition: $n = 33$, $Z = 2.943$, p = .003; EDT condition: $n = 37$, $Z = 2.757$, p = .006). These results indicate a learning effect for both topics in all conditions. Table 7.1 shows the means and SDs of the pre- and post-test scores of all conditions, as well as the difference scores.

Table 7.1. Test Scores per Condition for all students

	Control (n = 34)		ET (n = 33)		EDT (n = 37)	
Test	M	SD	M	SD	M	SD
			Buoyancy (Max = 25)			
Pre-test	14.82	8.12	17.06	6.80	15.19	6.41
Post-test	18.59	6.09	21.67	3.46	18.57	5.86
Difference score	3.76	5.89	4.61	6.80	3.38	6.23
			Archimedes' principle (Max = 33)			
Pre-test	4.50	4.50	4.52	4.28	4.76	4.02
Post-test	7.88	6.74	7.36	4.42	6.76	3.85
Difference score	3.38	4.65	2.85	4.99	2.00	4.31

With the current study we were especially interested in discovering if students who were guided by the ET or the EDT gained more conceptual knowledge than students who did not receive the additional support. Independent-Samples Kruskal-Wallis Tests showed no significant differences between the conditions for both parts of the test (buoyancy: H (2) = .253, p = .881; Archimedes' principle: H (2) = .651, p = .722).

However, a previous study with a similar learning environment by van Riesen and colleagues (submitted) showed that students with low prior knowledge do benefit from additional support in the form of the (previous version of the) Experiment Design Tool as opposed to students who did not receive this support. Therefore, for analyses purposes we divided students in two groups based on their pre-test scores; one group included students with the 50% lowest scores and the other group included students with the 50% highest scores. Students were divided based on their prior knowledge per domain, which means that a student who is considered a lower prior knowledge student for Archimedes' principle can have high prior knowledge about buoyancy and vice versa. Also, we divided the students in lower and higher knowledge groups without distinguishing between conditions, making it possible that the number of students with lower and higher prior knowledge is not necessarily equally divided between nor within conditions. Table 7.2 and Table 7.3 show the means and SDs of the pre- and post-test scores of all conditions, as well as the difference scores, split out for lower and higher prior knowledge students.

Table 7.2. Test Scores per Condition for Lower Prior Knowledge Students

Lower prior knowledge group Buoyancy (Max = 25)						
	Control (*n* = 18)		ET (*n* = 14)		EDT (*n* = 20)	
Test	M	SD	M	SD	M	SD
Pre-test	8.00	4.52	10.36	4.70	10.65	5.04
Post-test	15.17	6.04	21.29	3.45	16.40	5.99
Difference score	7.17	5.68	10.93	5.28	5.75	6.58
Lower prior knowledge group Archimedes' principle (Max = 33)						
	Control (*n* = 19)		ET (*n* = 17)		EDT (*n* = 19)	
Test	M	SD	M	SD	M	SD
Pre-test	1.00	1.73	0.82	1.59	1.26	1.85
Post-test	4.58	4.51	6.29	4.57	5.53	3.88
Difference score	3.58	4.49	5.47	4.46	4.26	3.86

An Independent-Samples Kruskal-Wallis Test demonstrated a significant difference between conditions in terms of gain of conceptual knowledge for lower prior knowledge students on the buoyancy part of the test between conditions, H (2) = 6.17, p = .046, in favour of the ET condition. No significant differences were found between the conditions for lower prior knowledge students on the Archimedes' principle part. For higher prior knowledge students no significant differences were found for both parts of the test.

Table 7.3. Test Scores per Condition for Higher Prior Knowledge Students

Higher prior knowledge group Buoyancy (Max. = 25)						
	Control (*n* = 16)		ET (*n* = 19)		EDT (*n* = 17)	
Test	M	SD	M	SD	M	SD
Pre-test	22.50	2.39	22.00	2.33	20.53	2.48
Post-test	22.44	3.22	21.95	3.54	21.12	4.69
Difference score	-0.06	3.19	-0.05	2.84	0.59	4.54
Higher prior knowledge group Archimedes' principle (Max. = 33)						
	Control (*n* = 15)		ET (*n* = 16)		EDT (*n* = 18)	
Test	M	SD	M	SD	M	SD
Pre-test	8.93	2.49	8.44	2.10	8.44	1.62
Post-test	12.07	6.88	8.50	4.10	8.06	3.47
Difference score	3.13	4.98	0.06	3.99	-0.39	3.43

7.5 Conclusion and discussion

The current study showed that only some students benefited from using tools to guide them in designing and conducting experiments. Only lower prior knowledge students benefited from additional support in the form of the ET to learn about buoyancy. The ET provided them with structure for their experiment design but also gave them enough space to explore the domain without the restriction to vary one thing at a time and having to plan at least three experiments before being able to conduct those in the lab. Interestingly, whereas lower prior knowledge students guided by the ET showed the highest knowledge gain regarding buoyancy, students who worked with the EDT had the lowest gain. This difference may be due to the restricted character of the EDT as opposed to the exploring character of the ET. Research indicates that learners perform better when they have the opportunity to first explore the domain of interest, rather than immediately starting with systematic ways of designing experiments.

No significant differences between conditions were found for higher prior knowledge students, which is in line with other research. Schauble and colleagues (1991) found that higher prior knowledge learners use more sophisticated strategies to induce knowledge and encounter less problems than lower prior knowledge students. Moreover, higher prior knowledge students often demonstrate more well-structured, goal-oriented inquiry behaviour by themselves (Alexander & Judy, 1988; Hmelo, Nagarajan, & Roger, 2000; Schauble et al., 1991). Additional support for these students has often found to be redundant, because they already have sufficient knowledge that helps them construct mental representations (Kalyuga, 2007).

7.6 References

Alexander, P. A., & Judy, J. E. (1988). The interaction of domain-specific and strategic knowledge in academic performance. *Review of Educational Research, 58*, 375-404. doi:10.3102/00346543058004375

Alfieri, L., Brooks, P. J., Aldrich, N. J., & Tenenbaum, H. R. (2011). Does discovery-based instruction enhance learning? *Journal of Educational Psychology, 103*, 1-18. doi:10.1037/A0021017

Almarshoud, A. F. (2011). The advancement in using remote laboratories in electrical engineering education: A review. *European Journal of Engineering Education, 36*, 425-433. doi:10.1080/03043797.2011.604125

Arnold, J. C., Kremer, K., & Mayer, J. (2014). Understanding students' experiments - What kind of support do they need in inquiry tasks? *International Journal of Science Education, 36*, 2719-2749. doi:10.1080/09500693.2014.930209

Carnesi, S., & DiGiorgio, K. (2009). Teaching the inquiry process to 21st century learners. *Library Media Connection, 27*, 32-36.

Cooper, M. (2005). Remote laboratories in teaching and learning-issues impinging on widespread adoption in science and engineering education. *International Journal of Online Engineering (iJOE), 1*.

Corter, J. E., Esche, S. K., Chassapis, C., Ma, J., & Nickerson, J. V. (2011). Process and learning outcomes from remotely-operated, simulated, and hands-on student laboratories. *Computers & Education, 57*, 2054-2067. doi:10.1016/j.compedu.2011.04.009

de Jong, T. (2006). Computer simulations - Technological advances in inquiry learning. *Science, 312*, 532-533. doi:10.1126/science.1127750

de Jong, T., & Lazonder, A. W. (2014). The guided discovery principle in multimedia learning. In R. E. Mayer (Ed.), *The Cambridge handbook of multimedia learning* (2nd ed., pp. 371-390). Cambridge: Cambridge University Press.

de Jong, T., & Njoo, M. (1992). Learning and instruction with computer simulations: Learning processes involved *Computer-based learning environments and problem solving* (pp. 411-427): Springer.

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:10.2307/1170753

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*, 391-450. doi:10.1207/s15327809jls0803&4_3

Glaser, R., Schauble, L., Raghavan, K., & Zeitz, C. (1992). *Scientific reasoning across different domains*: Springer.

Gomes, L., & Bogosyan, S. (2009). Current trends in remote laboratories. *IEEE Transactions on Industrial Electronics, 56*, 4744-4756. doi:10.1109/TIE.2009.2033293

Hardy, I., Jonen, A., Möller, K., & Stern, E. (2006). Effects of instructional support within constructivist learning environments for elementary school students' understanding of "floating and sinking". *Journal of Educational Psychology, 98*, 307-326. doi:10.1037/0022-0663.98.2.307

Hmelo, C. E., Nagarajan, A., & Roger, S. (2000). Effects of high and low prior knowledge on construction of a joint problem space. *Journal of Experimental Education, 69*, 36-56. doi:10.1080/00220970009600648

Hughes, S. W. (2005). Archimedes revisited: A faster, better, cheaper method of accurately measuring the volume of small objects. *Physics Education, 40*, 468-474.

Hwang, G., Sung, H., & Chang, H. (2011). *The effect of integrating STS strategy to online inquiry-based learning on students' learning performance*. Paper presented at the IEEE International Conferences on Internet of Things, and Cyber, Physical and Social Computing.

Kalyuga, S. (2007). Expertise reversal effect and its implications for learner-tailored instruction. *Educational Psychology Review, 19*, 509-539. doi:10.1007/s10648-007-9054-3

Klahr, D., & Nigam, M. (2004). The equivalence of learning paths in early science instruction: effect of direct instruction and discovery learning. *Psychol Sci, 15*, 661-667. doi:10.1111/j.0956-7976.2004.00737.x

Lawson, A. E. (2002). Sound and faulty arguments generated by preservice biology teachers when testing hypotheses involving unobservable entities. *Journal of Research in Science Teaching, 39*, 237-252. doi:10.1002/Tea.10019

Lazonder, A. W. (2014). Inquiry learning. In J. M. Spector, M. D. Merrill, J. Elen, & M. J. Bishop (Eds.), *Handbook of research on educational communications and technology* (pp. 453-464). New York: Springer Science + Business Media.

Loverude, M. E. (2009). A research-based interactive lecture demonstration on sinking and floating. *American Journal of Physics, 77*, 897-901. doi:10.1119/1.3191688

Loverude, M. E., Kautz, C. H., & Heron, P. R. L. (2003). Helping students develop an understanding of Archimedes' principle. I. Research on student understanding. *American Journal of Physics, 71*, 1178-1187. doi:10.1119/1.1607335

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

Osborne, J., Collins, S., Ratcliffe, M., Millar, R., & Duschl, R. (2003). What "ideas-about-science" should be taught in school science? A Delphi study of the expert community. *Journal of Research in Science Teaching, 40*, 692-720. doi:10.1002/Tea.10105

Pedaste, M., Maeots, 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. doi:10.1016/j.edurev.2015.02.003

Reiser, B. J. (2004). Scaffolding complex learning: The mechanisms of structuring and problematizing student work. *Journal of the Learning Sciences, 13*, 273-304. doi:10.1207/s15327809jls1303_2

Reuter, R. (2009). Online versus in the classroom: Student success in a hands-on lab class. *American Journal of Distance Education, 23*, 151-162. doi:10.1080/08923640903080620

Sancristobal, E., Martín, S., Gil, R., Orduña, P., Tawfik, M., Pesquera, A., . . . Castro, M. (2012). *State of art, initiatives and new challenges for virtual and remote labs*. Paper presented at the IEEE International Conference on Advanced Learning Technologies.

Schauble, L., Glaser, R., Raghavan, K., & Reiner, M. (1991). Causal models and experimentation strategies in scientific reasoning. *The Journal of the Learning Sciences, 1*, 201-238. doi:10.1207/s15327809jls0102_3

Schraw, G., Crippen, K. J., & Hartley, K. (2006). Promoting self-regulation in science education: Metacognition as part of a broader perspective on learning. *Research in Science Education, 36*, 111–139. doi:10.1007/s11165-005-3917-8

Schunk, D. H. (1996). *Learning theories*. New Jersey: Printice Hall Inc.

Simons, K. D., & Klein, J. D. (2007). The impact of scaffolding and student achievement levels in a problem-based learning environment. *Instructional Science, 35*, 41-72. doi:10.1007/s11251-006-9002-5

Toth, E. E., Ludvico, L. R., & Morrow, B. L. (2014). Blended inquiry with hands-on and virtual laboratories: The role of perceptual features during knowledge construction.

Interactive Learning Environments, 22, 614-630. doi:10.1080/10494820.2012.693102

van Joolingen, W., & de Jong, T. (1991). An hypothesis scratchpad as a supportive instrument in simulation learning environments.

van Riesen, S. A. N., Gijlers, A. H., Anjewierden, A. A., & de Jong, T. (submitted). Scaffolding students' experimentation. *Journal of Computer Assisted Learning*.

Veenman, M. V. J., Elshout, J. J., & Meijer, J. (1997). The generality vs domain-specificity of metacognitive skills in novice learning across domains. *Learning and Instruction, 7*, 187-209.

Veermans, K., de Jong, T., & van Joolingen, W. R. (2000). Promoting self-directed learning in simulation-based discovery learning environments through intelligent support. *Interactive Learning Environments, 8*, 229-255. doi:10.1076/1049-4820(200012)8:3;1-D;FT229

Veermans, K., van Joolingen, W. R., & de Jong, T. (2006). Use of heuristics to facilitate scientific discovery learning in a simulation learning environment in a physics domain. *International Journal of Science Education, 28*, 341-361. doi:10.1080/09500690500277615

Zacharia, Z. C., Manoli, C., Xenofontos, N., de Jong, T., Pedaste, M., van Riesen, S. A. N., . . . Tsourlidaki, E. (2015). Identifying potential types of guidance for supporting student inquiry when using virtual and remote labs in science: A literature review. *Etr&D-Educational Technology Research and Development, 63*(2), 257-302. doi:10.1007/s11423-015-9370-0

8 Scaffolding experimentation in virtual labs: The case of the Go-Lab Experiment Design Tool

8.1 Abstract

The purpose of this study was to assess the effect of a newly developed Experiment Design Tool, in the context of the Go-Lab project, on students' science learning and inquiry skills. Two studies were conducted. The first study was designed for primary students and the second one for secondary students. Two conditions were compared, namely the control condition, where the Experiment Design Tool had not been integrated in the learning environment, and the experimental condition, where the Experiment Design Tool had been included in the learning environment. Both before and after each implementation, students were administrated content knowledge and inquiry skills tests. For each study, pre- and post- tests were the same. The integration of the Experiment Design Tool in the learning environment seems to have different implications for students of different age. Namely, the tool has enhanced students designing investigations skills among younger students, while it has facilitated the advancement of content knowledge among older students. This latter finding might be attributed to the fact that the knowledge test for older students required more details, which referred to more comprehensive understanding of the phenomenon under study. Less experienced, younger students might benefit more from using the Experiment Design Tool for designing their experiments as compared to more experienced, older students, who might not find it that difficult to handle variables in their experimental design without using the tool. Future research should clarify the potential effect of former experience in experimental design on the gains derived from the use of Experiment Design Tool.

8.2 Introduction

Experimentation has been identified as an integral part of an inquiry enactment in science (Pedaste et al., 2015). It is usually used after a hypothesis has been formed in order to produce data/evidence concerning the variables at task. This particular data are then used for interpretation purposes, such as revealing if there is a relationship between the variables under study, in order to reach to conclusions (van Joolingen & Zacharia, 2009).

The process of experimentation involves a number of practices: (a) identifying the two variables under study, namely the independent and the dependent variable, (b) indentifying and setting up the materials/equipment needed for the experiment, (c) varying the independent variable and measuring the dependent variable, while keeping constant the rest of the related to the phenomenon variables, (d) repeating the experiment for a range of values of the independent variables in order to produce a number of values concerning the dependent variable, (e) organizing the data produced, usually in tables, and (f) producing meaningful representations from the produced data, usually graphs. Given this relatively big number of practices, as well as the demanding tasks that each of these practices involve during experimentation, someone could reasonably argue that experimentation might be a rather complex process for K-12 students. Research findings have shown instances when K-12 students fail to enact experimentation successfully (Hofstein & Lunetta, 2004). A closer examination of these studies reveals problems that are related to the aforementioned practices (for details see the review by Hofstein & Lunetta, 2004). The latter also explains all the research conducted so far for developing and using guidance tools (e.g., scaffolds, prompts, heuristics) to support students when designing and enacting an experiment (Zacharia et al., 2015).

A number of different types of guidance, such as prompts, performance dashboards, heuristics and scaffolds, have been developed over the years to provide support to the students during experimentation (Zacharia et al., 2015). For instance, van Joolingen and de Jong (2003) developed a monitoring performance dashboard tool in the SimQuest computer-based learning environment, which involved the affordance of storing an experiment and allowing students to later review and replay them if needed. Chang et al. (2008) have designed and developed an experiment prompting tool for guiding students through prompts on how to conduct an experiment (e.g., varying one variable at a time). Lin and Lehman (1999) have also used an experiment prompting tool that prompted students to reflect on their strategies when varying and controlling variables. A number of researchers have also advocated in favour of using heuristics for guiding experimentation. For example, Klahr et al. (1993) used heuristics for supporting students when observing in the context of an experiment, designing an experiment and focusing on one dimension of an experiment. Veermans et al. (2006) have also identified through a literature review a number of similar heuristics, such as the VOTAT (varying one thing at a time) heuristic, the equal increments heuristic, the extreme values heuristic and the make a graph heuristic. The literature of the heuristics domain also reveals ways of when and how to use heuristics for experimentation. Marschner et al. (2012) have shown that students using adaptive heuristics, accompanied by an explanation as to why they were given, differ in terms of experimentation behavior from students using a non-adaptive progression of heuristics.

In addition to the heuristics, a number of scaffolds were developed for guiding experimentation. For example, Veermans (2003) described scaffolding tools aiming to support data interpretation during experimentation and van Joolingen et al. (2011) described the development of an Experiment Design Tool (EDT), namely the SCYED tool, in which students can create a step by step procedure when designing an experiment by choosing the variables at task and their values and by describing how to conduct the experiment. Unfortunately, no thorough studies exist on examining the learning value of existing EDTs, as well as no empirical underpinnings exist about how an EDT should look like.

In this study we aimed at focusing on these issues. In particular, the purpose of this study was to assess the effect of a newly developed EDT, in the context of the Go-Lab project, on students' science learning and inquiry skills. In doing so, we used the Go-Lab EDT with students of different ages.

8.3 Method

Two studies were conducted. The first study was designed for primary students and the second one for secondary students. The experimental design followed in both studies is presented in Figure 8.1. Two conditions were compared, namely the control condition, where the EDT had not been integrated in the learning environment, and the experimental condition, where the EDT had been included in the learning environment. Both before and after each implementation, students were administered content knowledge and inquiry skills tests. For each study, pre- and post- tests were the same.

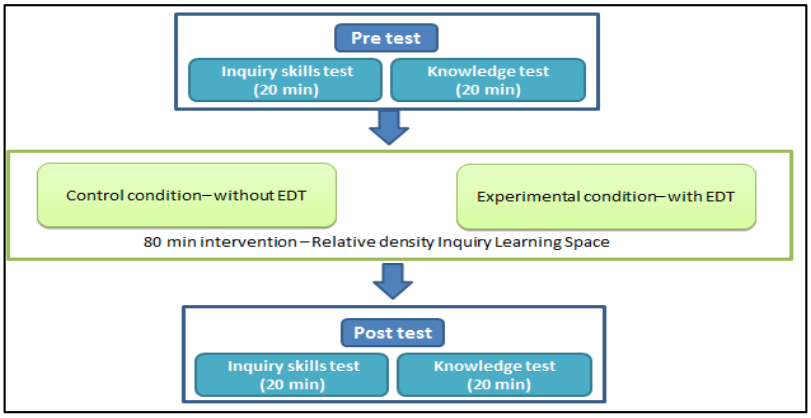


Figure 8.1. Experimental design for studies 1 and 2.

8.3.1 Participants

Study 1

Participants in study 1 were 26 fifth-graders (11-12 years old) from two classes (N_{class1} = 14 and N_{class2} = 12) of a public primary school in Limassol, Cyprus. The first class (5 boys and 9 girls) was assigned to the experimental condition and the second class (7 boys and 5 girls) to the control condition. The selection of the school was based on the teacher's interest to participate in the Go-Lab project and the selection of the two classes was made because both contained students of mixed learning abilities. In addition, the students in both classes were taught about sinking and floating in the age of nine and they had basic computer skills.

8.3.1.1 Study 2

Participants in study 2 were 23 eighth-graders (14-15 years old) from two classes (N_{class1} = 11 and N_{class2} = 12) of a public high school in Nicosia, Cyprus. The first class (5 boys and 6 girls) was assigned to the experimental condition and the second class (5 boys and 7 girls) to the control condition. The selection of both the school and the classes were based on the same criteria as in study 1. Students in both classes had learned about density in the sixth grade and had advanced computer skills.

8.3.2 Materials

8.3.2.1 Learning environment

For both studies, an Inquiry Learning Environment (ILS) was created. This was an online learning environment, based on the inquiry cycle design framework (Pedaste et al., 2015), which was configured through the Go-Lab authoring tool (de Jong, Sotiriou, & Gillet, 2014). Each ILS comprised five inquiry phases, namely, the *Orientation*, the *Conceptualization*, the *Investigation*, the *Conclusion* and the *Discussion* phase. The learning subject of both ILSs was relative density. Specifically, students learned about the relation of mass, volume and density and how density of an object influenced whether this object would sink or float in different fluids. Students conducted experiments in an online laboratory termed "Splash: Virtual Buoyancy Laboratory" (Figure 8.2), which is available in the Go-Labz repository (http://www.golabz.eu/). Below, more details about the ILSs used in each study are provided.

8.3.2.1.1 Study 1

In the *Orientation* phase, students predicted if some objects would sink or float by responding to a quiz. Then, they watched a video that contained information about the

basic concepts of the lesson. In the *Conceptualization* phase, students explored the environment of the Splash Laboratory, in order to identify the variables that were involved in the phenomenon under investigation. In the Splash Laboratory (Figure 8.2), students could manipulate mass, volume and density of an object, and the kind of the fluid in a container. They could add up to six containers and manipulate each one of them. When they run the experiment, they could see what would happen in each container. All manipulations and outcomes were presented in a Results´ table in the right side of the tool´s interface.

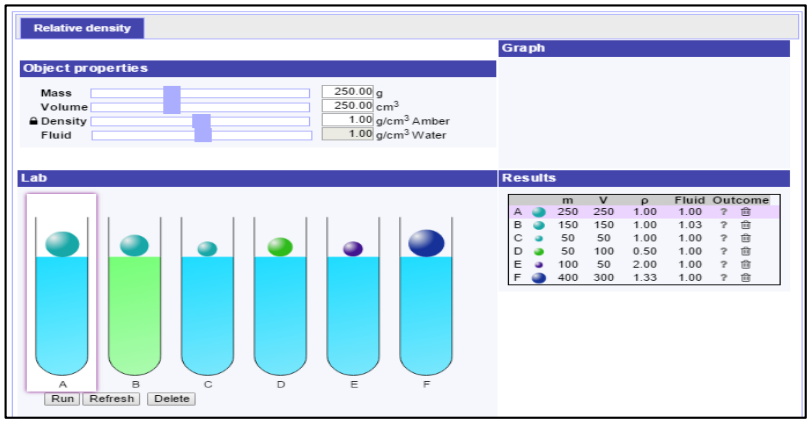


Figure 8.2. Splash: Virtual Buoyancy Laboratory.

After the students became familiar with the Splash Laboratory (5-7 min) they continued in the *Conceptualization* phase, where they had to formulate their hypothesis/hypotheses using the Hypothesis Scratchpad (Figure 8.3). In this tool, predefined conditionals and concepts were provided to students who could drag and drop them in the space provided by the tool to create a hypothesis in the form of an "if…then" statement. Additionally, they had the opportunity to adjust their confidence level for each hypothesis, by changing the colour of the "horseshoe" placed next to the hypothesis. If the "horseshoe" was blue, overall, that would have meant that a student was absolutely sure that his/her hypothesis was correct.

After formulating their hypotheses, students moved to the *Investigation* phase, where they had to conduct their experiments in order to confirm or reject their hypothesis/hypotheses. Before using the Splash Laboratory, students used the EDT to design their experiments. That was the configuration for the experimental condition, whereas in the control condition, the EDT was replaced by a note-taking tool and students were free to keep records of anything they could have considered as noteworthy before, during or after their experiments.

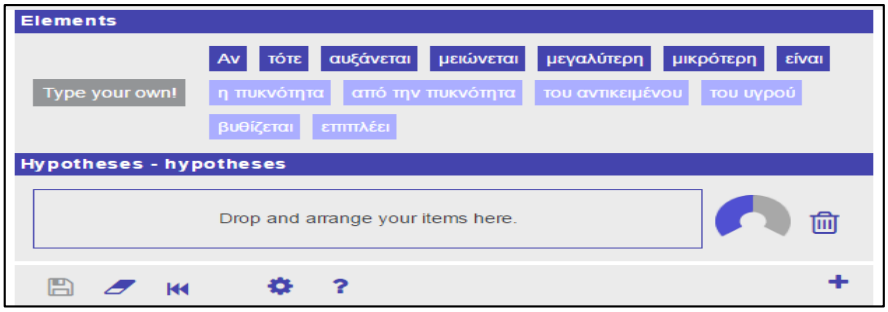


Figure 8.3. Hypothesis Scratchpad.

The EDT (Figure 8.4) allowed students to determine the independent ("Vary"), dependent variable ("Measure"), and control variables ("Keep constant") in their experiments. To do so, they dragged properties from the left side of the tool´s interface and dropped them in the proper column. Once a design was completed, the next step was to specify the values of each variable and add experimental trials. When an experimental trial was conducted in the laboratory, students returned to EDT to enter their measurements. They had to do this for each trial.



Figure 8.4. Experiment Design Tool.

In the *Conclusion* phase, students wrote their general conclusion about how density of an object influenced whether it would sink or float in different fluids. In order to confirm their rule about the relative density, they responded to a relevant quiz. In the *Discussion* phase, students discussed with their teacher and peers about their investigations and conclusions. They tried to explain what would have happened if a human tried to float in the Dead Sea. The last part of the *Discussion* phase involved reflection activities.

8.3.2.1.2 Study 2

The ILS used in study 2 was similar to the ILS described above, but small adaptations were made because it was to be used by older students. Specifically, some aspects of the content was more advanced (e.g. orientation video, texts) and more tools were added. In the *Investigation* phase, the Observation Tool was added (Figure 8.5) to let students keep notes about ideas, thoughts and observations right after they had run their experiment in the laboratory. Students were prompted to write their observations, because they would need to retrieve them in the next phase.

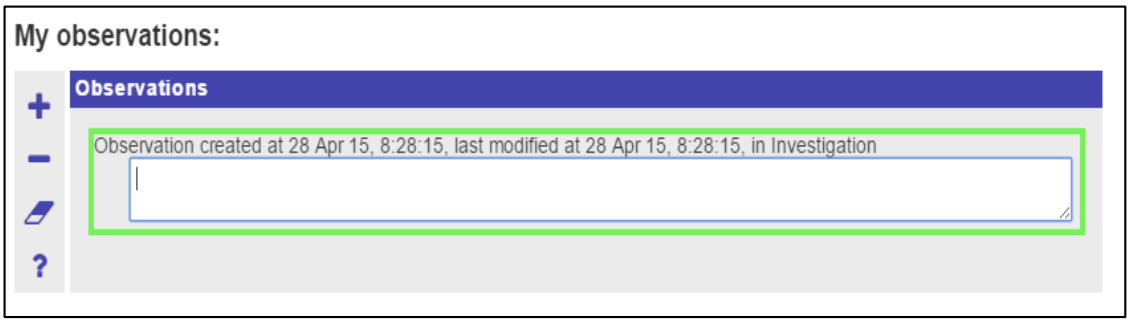


Figure 8.5. Observation Tool.

In the *Conclusion* phase, the Conclusion tool (Figure 8.6) was added, where students retrieved their hypothesis/hypotheses and observations in order to argue about how their

confidence levels for each hypothesis had changed after their investigation. In order to do so, they had to refer to observations from their experiments.

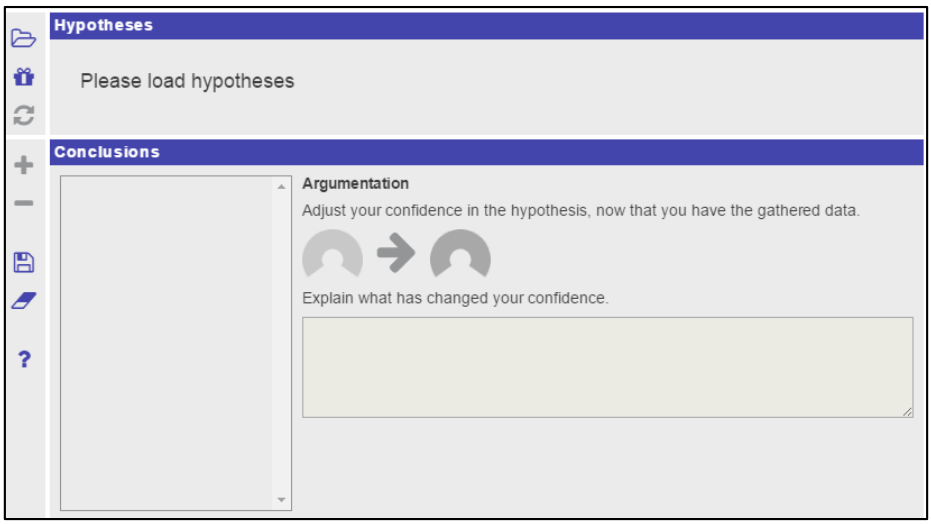


Figure 8.6. Conclusion Tool.

In the *Discussion* phase, the video about the Dead Sea was removed. The students in this phase discussed with the teacher and peer about their experiments and conclusions and performed reflection activities.

8.3.2.2 Assessment

Assessment of student knowledge and skills for both studies involved two different instruments, namely, a knowledge test and an inquiry skills test. For the creation of the knowledge test, a revised taxonomy of the levels of the cognitive domain of knowledge has been followed (more details in de Jong, 2014; Zervas, 2013). The corresponding levels were "*Remember*", "*Understand*", "*Apply*" and "*Think Critically and Creatively*". For the inquiry skills test, we selected and translated items included in TIPSII (Burns, Okey, & Wise, 1985).

All tests were scored blind to the condition in which each student had been placed. For open ended items in the knowledge test, a rubric was used, which specified scoring criteria. Inter-rater reliability was calculated based on ratings of two independent coders who had reviewed 20% of all data available. Inter-rater agreement was acceptable in both studies (Cohen's $Kappa_{study1}$= 0.92; Cohen's $Kappa_{study2}$= 0.95).

8.3.2.2.1 Study 1

The knowledge test consisted of four items, one for each level of knowledge. "*Remember*", and "*Think Critically and Creatively*" items were open-ended, while "*Understand*" and "Apply" items were closed-ended. All items focused on density and its role to sinking and floating.

The inquiry skills test consisted of 12 multiple-choice items. Of them, 4 addressed "*Identifying variables"*, 3 "*Identifying and stating hypotheses"*, 3 "O*perationally defining"* and, 2 "*Designing investigations"*. The criteria for the selection of TIPSII items were appropriateness of wording and content in respect to students' age.

8.3.2.2.2 Study 2

The knowledge test in study 2 was divided into two Parts, A and B. Part A consisted of 4 items, one open-ended "*Remember*" item and 3 closed-ended items, of which, two referred to "*Understand*" and another one to "*Apply*". Part B of the knowledge test consisted of two

open-ended "*Think Critically and Creatively*" items. Items in both parts of the instrument focused on density, the relation between mass, volume and density, and the role of density to sinking and floating.

The inquiry skills test consisted of 24 multiple-choice items. Of them, 12 addressed "*Identifying variables"*, 9 "*Identifying and stating hypotheses"* and 3 "*Designing investigations"*. The number of items in each subscale was the same as in the initial TIPSII instrument.

8.4 Procedure

The procedures for both studies were the same. The implementations were carried out by the science teacher of each class and they had a duration of four class meetings of 40 minutes each. Before the implementation, teachers participated in a face to face preparatory meeting. During the meeting, the teacher became familiar with the ILS, the Splash Laboratory and the tools in each phase of the ILS. In addition, some procedural issues were discussed and the role of the teacher during the lesson was clarified.

During the first class meeting (40 minutes), students completed the pre-tests (i.e., the knowledge test and the inquiry skills test). The next two class meetings (80 minutes) took place in the computer lab of each school, and each student used a computer in order to undertake activities of the ILS. At the beginning of the meeting, the teacher provided general guidance to his/her students and explained them how they would act in order to complete the lesson. During the lesson, the teacher provided help only for technical issues concerning the use of tools and the laboratory. Whenever technical issues arose, they were solved without causing any delay to the completion of the meeting. Finally, one day after the intervention, a last session was done (40 minutes) for the completion of post-tests (i.e., knowledge test and inquiry skills test).

8.5 Results

8.5.1 Content Knowledge test

8.5.1.1 Primary school students

The Wilcoxon test showed that both conditions improved primary school students' conceptual understanding for the ILS as a whole ($p < .05$). The Mann Whitney U test revealed that the two conditions did not differ in terms of their pre-test scores. Further, the Mann Whitney U test revealed no statistically significant differences between the two conditions scores in the post-test.

8.5.1.2 Secondary school students

The Wilcoxon test showed that both conditions improved secondary school students' conceptual understanding for the ILS as a whole ($p < .05$). The Mann Whitney U test did not reveal any statistical differences between the two conditions in terms of their pre-test scores ($p = .439$). However, the Mann Whitney U test revealed that students' post-test scores in the experimental condition were significantly higher than those of the students in the control condition ($Z = -2.89$; $p < .01$).

8.5.2 Inquiry skills test

8.5.2.1 Primary school students

The Wilcoxon test showed that both conditions improved primary school students' designing investigations skills ($p<.05$). No statistical significant differences were identified within either condition for either the identifying variables skills and the identifying and

stating hypothesis skills. It should be noted that there was an improvement in both cases, but it was marginally not found to be statistically significant.

The Mann Whitney U test revealed statistically significant differences in the case of students' designing investigations skills between the two conditions ($Z = -2.18$; $p < .05$). The students of the experimental condition appeared to have better developed their designing investigations skills during the study's intervention than the students of the control condition.

8.5.2.2 Secondary school students

The Wilcoxon test showed no statistical significant differences between the pre- and post-test scores for all the inquiry skills under study, namely identifying variables skills, identifying and stating hypothesis skills, and designing investigations skills ($p > .05$). As in the case of the primary schools students, there was an improvement in both cases, but it was marginally not found to be statistically significant across all skills.

The Mann Whitney U test revealed no statistically significant differences between the two conditions across all the inquiry skills under study ($p > .05$).

8.5.3 Summary of results

A summary of the results of both studies for both instruments is presented in Table 8.1.

Table 8.1. Comparisons in knowledge and skills scores for Study 1 and 2

	Knowledge scores	Skill scores
Study 1 (primary school students)		
Pre- vs. post-test comparison, overall	Improved scores in the post-test	Improved scores in the post-test
Control vs. experimental group; pre-test	Non-significant differences	Non-significant differences
Control vs. experimental group; post-test	Non-significant differences	Higher scores for "*Designing investigations*" in the experimental group
Study 2 (secondary school students)		
Pre- vs. post-test comparison, overall	Improved scores in the post-test	Improved scores in the post-test
Control vs. experimental group; pre-test	Non-significant differences	Non-significant differences
Control vs. experimental group; post-test	Higher scores in the experimental group	Non-significant differences

8.6 Discussion, conclusions and implications for future research

The integration of the EDT in the learning environment seems to have different implications for students of different age. Namely, the tool has enhanced students designing investigations skills among younger students, while it has facilitated the advancement of

content knowledge among older students. This latter finding might be attributed to the fact that the knowledge test for older students required more details, which referred to more comprehensive understanding of the phenomenon under study. Another advantage for the secondary school students is the fact that they have been explicitly taught about density at the early high-school grades, whereas this is not the case for the primary school students. Needless to say, the maturity level of the secondary school students might have enhanced their science learning more than the learning of the primary students.

Less experienced, younger students might benefit more from using the EDT for designing their experiments as compared to more experienced, older students, who might not find it that difficult to handle variables in their experimental design, at least for experiments such as the ones used in this study, without using the tool.

Future research should clarify the potential effect of former experience in experimental design on the gains derived from the use of EDT. This would necessitate two new studies that would control for varying experience in designing experiments among both primary school students as well as secondary school students.

8.7 References

Burns, J. C., Okey, J. R., & Wise, K. C. (1985). Development of an integrated process skill test: TIPS II. *Journal of Research in Science Teaching, 22*(2), 169-177. doi: 10.1002/tea.3660220208

Chang K. E., Chen Y. L., Lin H. Y., & Sung Y. T. (2008). Effects of learning support in simulation-based physics learning. *Computers & Education, 51*, 1486-1498. doi: 10.1016/j.compedu.2008.01.007

de Jong, T. (Ed.). (2014). *D1.3: Preliminary inquiry classroom scenarios and guidelines: Go-Lab project.*

de Jong, T., Sotiriou, S., & Gillet, D. (2014). Innovations in STEM education: The Go-Lab federation of online labs. *Smart Learning Environments, 1*(1), 1-16. doi: 10.1186/s40561-014-0003-6

Hofstein, A., & Lunetta, V.N., (2004). The laboratory in science education: foundation for the 21st century. *Science Education, 88*, 28-54. doi: 10.1002/sce.10106

Klahr, D., Fay, A., & Dunbar, K. (1993). Heuristics for specific experimentation: A developmental study. *Cognitive Psychology, 25*, 111-146. doi: 10.1006/cogp.1993.1003

Lin, X., & Lehman, D.J. (1999). Supporting learning of variable control in a computer-based biology environment: effects of prompting college students to reflect on their own thinking. *Journal of Research in Science Teaching, 36*, 837-858. doi: 10.1002/(SICI)1098-2736(199909)36:7<837::AID-TEA6>3.0.CO;2-U

Marschner, J., Thillmann, H., Wirth, J., & Leutner, D. (2012). How can the use of strategies for experimentation be fostered? *Zeitschrift Fur Erziehungswissenschaft, 15*, 77-93. doi: 10.1007/s11618-012-0260-5

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. doi: 10.1016/j.edurev.2015.02.003

Tschirgi, J. E. (1980). Sensible reasoning: a hypothesis about hypotheses. *Child Development, 51*, 1-10. doi: 10.2307/1129583

van Joolingen, W. R., & de Jong, T. (2003). SimQuest: Authoring educational simulations. In T. Murray, S. Blessing, & S. Ainsworth (Eds.), *Authoring tools for advanced technology educational software: Toward cost-effective production of adaptive,*

interactive, and intelligent educational software (pp. 1-31). Dordrecht: Kluwer Academic Publishers.

van Joolingen, W. R., Giemza, A., Bollen, L., Bodin, M., Manske, S., Engler, J., & Halik, K. (2011). *SCY cognitive scaffolds and tools* (DII.2). Twente: SCY consortium.

van Joolingen, W. R., & Zacharia, Z. C. (2009). Developments in inquiry learning. In N. Balacheff, S. Ludvigsen, T. de Jong, & S. Barnes (Eds), *Technology-enhanced learning* (pp. 21-37). Dordrecht, The Netherlands: Springer. doi: 10.1007/978-1-4020-9827-7_2

Veermans, K. H. (2003). *Intelligent support for discovery learning*. Ph.D. thesis, University of Twente.

Veermans, K. H., van Joolingen, W., & de Jong, T. (2006). Use of heuristics to facilitate scientific discovery learning in a simulation learning environment in a physics domain. *International Journal of Science Education, 28*, 341-361. doi: 10.1080/09500690500277615

Zacharia, Z. C., Manoli, C., Xenofontos, N., de Jong, T., Pedaste, M., van Riesen, S., Tsourlidaki, E. (2015). Identifying potential types of guidance for supporting student inquiry when using virtual and remote labs: A literature review. *Educational Technology Research & Development, 63*, 257-302. doi: 10.1007/s11423-015-9370-0

Zervas, P. (Ed.). (2013). D2.1. *The Go-Lab inventory and integration of online labs – labs offered by large scientific organisations: Go-Lab project*.

9 Does using the Experimental Error Tool change students' understanding about 'How Science Works'?

9.1 About the study

This study conducted in Greece, was concerned with testing the scaffold 'Experimental Error Tool' to see first its usability with students and second its potential impact on students' understanding about inquiry and the nature of science (called 'How Science Works' below).

The scaffold was developed by for the purpose of being included in the Go-Lab Apps (http://www.golabz.eu/apps/experimental-error-calculator) to allow students to calculate experimental errors that stem from real experimental setups. Using this tool, students may learn about the different sources of error that occur when performing experiments and about the different types of errors that can be calculated so as to decide whether an experiment is precise and accurate (for more on the scaffold see Section 1.2.5).

Due to the nature of the tool it was decided that it would be appropriate to use it in conjunction with the Radioactivity Lab, which is a remote lab at the University of Queensland (http://www.golabz.eu/lab/radioactivity-lab). The Radioactivity Lab examines the intensity of radiation over distance, demonstrating the effects of the inverse square law. Students investigate the intensity of radiation being emitted from a radioactive Strontium-90 source, by setting distances at which a Geiger counter measures radiation, and collecting radioactive particle counts at each of these distances. A table holding the radioactive source is rotated until the source is aligned with a hole in a thick lead plate. The head holding the Geiger-Muller tube is moved to these distances from the radioactive source. Students receive data in the form of a .csv file to be analysed in Microsoft Excel, or a similar data analysis tool.

The ILS used for this purpose was the "Is Radioactivity always harmful for humans?" (http://www.golabz.eu/spaces/radioactivity-always-harmful-humans) adapted and translated in Greek (http://graasp.eu/ils/54b8d81c51830bd46a666965?lang=el). The aim of the ILS is for students to understand the different types of radioactivity particles and the basics about radioactivity decay and use remotely a Geiger-Muller to study how radioactivity's absorption is affected by factors such as distance from radioactivity source, exposure time and presence of absorbers (see Figure 9.1).



Figure 9.1. Golabz page for the ILS "Is Radioactivity always harmful for humans?" (Greek version).

9.2 Method

9.2.1 Evaluation instrument

As mentioned above our hypothesis was that the scaffold selected might have an impact on students' understanding about inquiry and 'How Science Works'. To measure this impact we chose out of the evaluation instruments selected by the Go-Lab team and described in D8.1 to use a reduced version of the test developed by (P. M. Kind, 2013). Quoting from D8.1, "this test contains combinations of multiple-choice (mc), open, and ordered mc questions. It focuses on three major phases in inquiry learning: hypothesizing, experimenting, and evidence evaluation. The test recognizes that different types of knowledge are involved in scientific reasoning, such as science content knowledge, procedural knowledge, and epistemic knowledge. What is quite rare about this test, is that it also allows assessment of the depth of inquiry skills and to explain skill progression. In the reduced version of the questionnaire that we used, we chose items that focus mainly on the inquiry phases of 'experimenting' and 'evaluating evidence'. The former intended to identify procedural knowledge students use in experimentation and the latter epistemic knowledge they use in evidence evaluation, and both to establish understanding of progression in their corresponding areas. In particular from the original questionnaire of 10 sections (or 'testlets' as these are called by its author) we chose four: testlet 1 ('School experiment on dissolving sugar' or 'Experiment 1'), testlet 3 ('Evidence for how plants

increase in weight' or 'Evidence 1'), testlet 4 ('School experiment on friction' or 'Experiment 2') and testlet 5 ('Evidence for health risk' or 'Evidence 2'). Testlets 1 and 4 ('Experiment 1' and 'Experiment 2') tested students' procedural knowledge and testlets 3 and 5 (or 'Evidence 1' and 'Evidence 2') students' epistemic knowledge about inquiry.

More precisely, testlets 'School experiment on dissolving sugar' (also 'Experiment 1') and 'School experiment on friction' ('Experiment 2') tested *"students' knowledge about measurement of a single variable:*

- *Understanding that a single measurement has uncertainty and that repeated measurements therefore are likely to vary.*
- *Understanding strategies for making an accurate measurement, such as doing repeated measurements and finding the mean.*
- *Knowing about anomalous data and having a strategy for handling these."*

(Kind, ibid, p542)

With regard to progression in procedural knowledge in experimenting, three different levels of proficiency are distinguished. At the lowest level, students may understand measurements as direct observations of "true" values. They believe that a measurement will yield the true value, making repeated measurements unnecessary. In this case, students might still understand the cause-effect relationship between variables and be able to carry out "fair testing" strategies. At the second level, the student may believe that true values are attainable, but require repeated measurements. In this case, the student thinks that true value can be attained by repeating measurements until two identical values are observed. Kind (ibid.) argues that at the most advanced level, students understand a true value is unattainable. *"At this level student may have strategies for handling uncertainty in evaluating a series of measurements, such as averaging several measurements and looking for anomalous data."*

On the other hand, testlets 'Evidence for how plants increase in weight' (also 'Evidence 1') and 'Evidence for health risk' (also 'Evidence 2') tested students' understanding of:

- *The claim-evidence relationship in science*
- *The criteria for evaluating and using evidence*
- *The role of personal and social contexts in evidence evaluation and coordination.*

(Kind, ibid, p546)

With regard to progression in understanding and evaluating evidence in science, three different levels of proficiency are distinguished. At the lowest level students think that 'true' claims emerge automatically from science data. At the next level, students accept multiple explanations, have some understanding about using evidence to support claims and are more likely to understand that evidence has uncertainty. At the highest level, the students understand the complicated relationship between claims and evidence.

The items in the four testlets chosen used a combination of multiple-choice questions and open-ended questions and were organised around topics and situations, with several items on the same science experiment. They were chosen so that they cover domain knowledge from the three major school science subjects, chemistry, biology and physics, at a level lower than what one would expect the students of our study, all in Grade 11 (i.e. 16-17-year olds) to have. In this way, we hoped to get, if possible, less context-dependent

results about students' understanding of 'How Science Works'. The reduced version of the questionnaire was translated in Greek.

9.2.2 Participants

The trials took place in 2 schools, a private and a public school in the suburbs of Athens, The former has classes of mixed-ability students, whereas the latter has classes of more able students selected through exams. In both schools addressed students of Grade 11 (i.e. 16-17 years of age).

In the first school two similar in terms of student abilities, groups of 12 students each were selected to participate. The control group did the Radioactivity ILS without the experimental error tool and the experimental group with it. The reduced 'How Science Works' questionnaire was administered before and several weeks after the lesson, either online or using the hard copy version to both groups of students.

In the second school the approach was slightly different. The class was divided in two and during the same lesson 11 students (control group) did the Radioactivity ILS without using the experimental error tool, while 14 students (experimental group) did the same ILS using the scaffold tool. Once again, the same evaluation questionnaire was completed before and after the intervention by both groups of students.

9.2.3 The lessons

In both trials the lessons lasted two consecutive school sessions, that is 1.5 hours and were taught by a Go-Lab expert, with the help of another Go-Lab team member and the teacher (in the second trial). Students worked in pairs or teams of three sharing one computer. The procedure followed was the same and varied only as to whether students used or not the experimental error tool.

In the Orientation phase students were shown two videos: one about the Hiroshima atomic bomb and its consequences and one explaining what radioactivity is and its different kinds, using narration and animation. Both videos were in the English language.

Students were asked to answer in writing, using the Input text app, what they learned about radioactivity, if they consider it always harmful and if they knew of any way that one could be protected from it. In other words, after being sensitised about the topic students were asked to express their previous knowledge and beliefs about question in focus.

In the 'Conceptualisation / Hypothesis Generation' phase students discussed the factors that may determine how harmful radiation can be. They mentioned time of exposure, distance from the source, radiation barriers (absorbers) and energy of radiation. The teacher introduced the concepts of wave length and frequency, *"the smallest the wave length, the largest the penetration"*. She then announced that they would focus their experimentation on how time, distance and barriers (absorbers) affect the intensity of radiation and asked them to create relevant hypotheses about each of these factors. Due to limitations of time and technical problems the Hypothesis app was not used.

In the next phase called 'Experimental Set up' the experimental set up (Figure 9.2) was explained and the students were told that they will measure the intensity of radiation being emitted from a radioactive Strontium-90 source, using a remote lab in the University of Queensland.

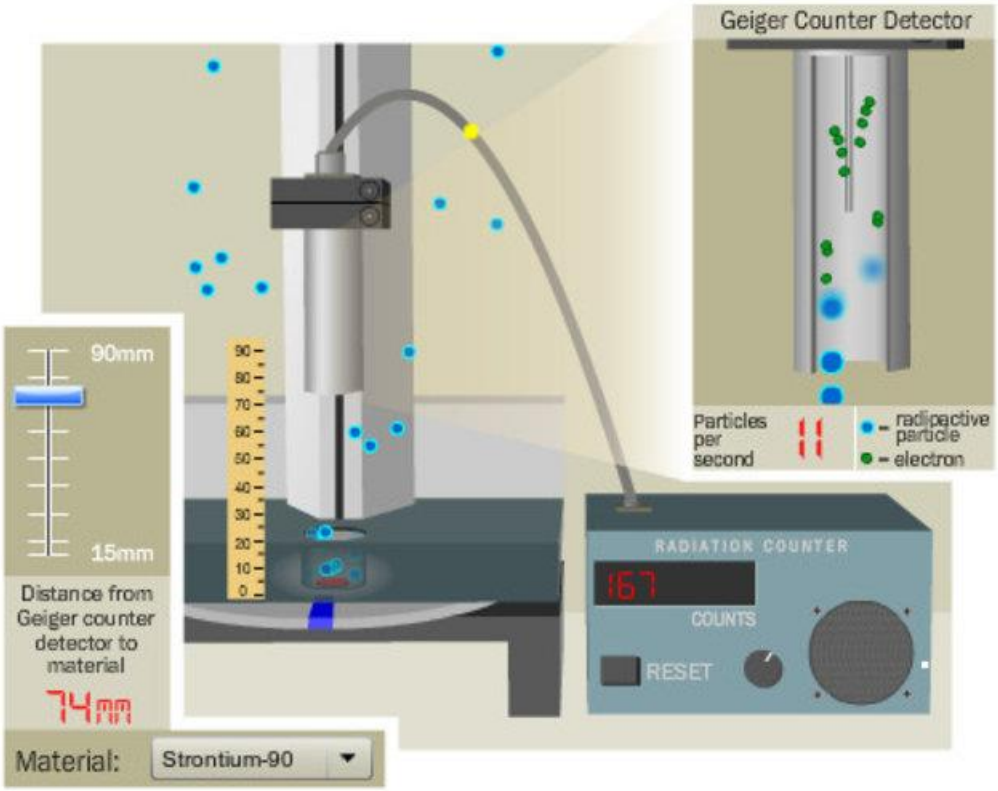


Figure 9.2. The experimental set-up (Source: Northwestern University's Office of STEM Education Partnerships, and Chicago-area high school science and math teachers).

Unfortunately, in neither trial was it possible for students to actually do the experiment remotely (in the 'Investigation' phase) for various technical reasons, including limited network bandwidth in the schools and problems with the remote lab involved. However, the teacher had already collected some real sets of measurements for each of the three experiments and asked the students to download the Excel file which contained them. A substantial discussion ensued about why the measurements differed, which could be considered as 'normal' and which one could be used to create the corresponding graph for each of the three experiments. Students' suggestions included:

- To calculate the average;
- To consider the largest and smallest values and find the value in-between;
- To consider only the largest, since radiation is dangerous for our health;
- To consider the values that appear more frequently.

The teacher pointed out that the anomalous values were usually discarded.

In the 'Data Analysis' phase the 'Experimental Error Tool' (see section 1.2.5) was introduced to the students (in the experimental groups) and the teacher asked them to explore its environment and come up with some answers about which values they could use in their graphs. The tool's language was English and the teacher had to help students understand the concepts of systematic error, maximum probable error and standard deviation and explained to them the difference between a precise and an accurate experiment, using the information provided by the tool. The students used the tool efficiently to calculate the mean value and corresponding errors.

The students in the control group on the other hand, used the Excel to calculate the average and did not discuss about errors. The same research design was applied in both school trials.

The corresponding graphs were drawn in Excel and there was a substantial discussion about which of the three factors one would choose to protect against in case of emergency ('Conclusions' phase).

Finally, in the 'Discussion' phase students watched a video with interesting facts about Chernobyl and reviewed their ideas about whether radioactivity is always harmful for humans.

9.2.4 Data Collection and Analysis

The data collected in the two trials can be seen in Table 9.1.

Table 9.1. Experimental set-up and number of subjects

		Experimental Group N of students	Control Group N of students
First trial	Pre-test	12	11
	Post-test	10 (7)	7 (5)
Second trial	Pre-test	8	9
	Post-test	6 (5)	9 (7)

Note: In parentheses one can see the number of students that participated in both pre- and post- tests.

In addition, there were also some students who did not want to identify themselves or cared to give a pseudonym in the questionnaires and as a result a number of questionnaires were not included in the above count for comparison purposes.

Student's responses for each test item were coded according to the scoring scale provided by the author of the test, and were analysed quantitatively. Progression in students' ability was reflected in the scale by considering both the difficulty of the items and the difficulty of each response alternative/score category. The analysis took place at different levels: from the whole test level to the testlet and item levels. The score of each testlet comprised of the sum of the scores for all the items of the testlet; the scores for the 'experimenting' and 'evaluating evidence' inquiry phases were calculated as the sum of the corresponding two testlets; and finally the total test score was the sum of the scores for both inquiry phases.

For each of the trials a number of statistical tests were conducted.

Table 9.2 and Table 9.3 show the descriptive statistics for each of the trials normalised (to a maximum score of 10), to facilitate possible comparisons. The non-highlighted rows refer to item scores; the highlighted yellow rows refer to testlet scores and the highlighted orange rows refer to the scores for each inquiry phase and the total test score. Only exception to this is the 'Experiment 1', which is both an item and a testlet.

The relationship of the original testlets/items of the test with the variables in the tables below is:

- Testlet 1 -> 'Experiment 1'

- Testlet 3 -> 'Evidence 1sum' (with items 'Evidence 1A', 'Evidence 1B' and 'Evidence 1C')
- Testlet 4 -> 'Experiment 2sum' (with items 'Experiment 2A', 'Experiment 2B' and 'Experiment 2C')
- Testlet 5 -> 'Evidence 2sum' (with items 'Evidence 2A', 'Evidence 2B' and 'Evidence 2C')

Table 9.2. Data from the first trial

FIRST TRIAL								
	PRE-TEST				POST-TEST			
	Experimental Group		Control Group		Experimental Group		Control Group	
	Mean	Stand. Dev.	Mean	Stand. Dev.	Mean	Stand. Dev.	Mean	Stand. Dev.
Experiment 1	4,2	5,1	1,8	4,0	1,0	3,2	1,4	3,8
Evidence 1A	5,6	4,6	4,2	3,4	7,7	4,2	5,7	4,6
Evidence 1B	4,2	4,1	3,9	2,9	7,0	2,9	3,8	4,9
Evidence 1C	2,5	2,9	1,5	2,3	3,0	1,1	1,0	1,6
Evidence 1sum	4,1	2,9	3,2	2,1	5,9	1,9	3,5	3,2
Experiment 2A	5,8	5,1	5,5	5,2	7,0	4,8	8,6	3,8
Experiment 2B	5,0	4,8	7,7	3,4	6,5	4,1	7,9	3,9
Experiment 2C	2,1	2,6	5,5	4,2	5,0	4,7	6,4	3,8
Experiment 2sum	4,0	2,6	6,4	3,2	6,0	3,3	7,4	2,8
Evidence 2A	2,5	4,5	3,6	5,0	2,0	4,2	2,9	4,9
Evidence 2B	3,8	2,3	3,2	3,4	5,0	4,7	6,4	2,4
Evidence 2C	2,1	4,0	3,2	4,0	1,0	2,1	2,1	3,9
Evidence 2sum	2,8	2,2	3,3	2,9	2,8	2,1	4,0	2,0
TotScore_Exp	4,0	2,2	5,6	2,8	5,2	2,5	6,4	2,6
TotScore_Evid	3,6	2,1	3,2	2,1	4,8	1,3	3,7	2,4
Totscore_total	3,8	1,4	4,0	2,1	4,9	0,8	4,5	2,3

Table 9.3. Data from the second trial

SECOND TRIAL								
	PRE-TEST				POST-TEST			
	Experimental Group		Control Group		Experimental Group		Control Group	
	Mean	Stand. Dev.	Mean	Stand. Dev.	Mean	Stand. Dev.	Mean	Stand. Dev.
Experiment 1	7,5	4,6	3,3	5,0	8,3	4,1	4,4	5,3
Evidence 1A	4,2	5,0	5,2	3,4	6,1	4,9	5,2	5,0
Evidence 1B	2,9	4,5	2,2	2,9	4,4	5,0	3,3	4,1
Evidence 1C	0,8	1,5	2,6	3,2	0,6	1,4	2,6	3,2
Evidence 1sum	2,6	2,7	3,3	2,4	3,7	3,3	3,7	2,0
Experiment 2A	10,0	0,0	7,8	4,4	10,0	0,0	10,0	0,0
Experiment 2B	6,3	3,5	4,4	3,9	6,7	2,6	5,6	3,0
Experiment 2C	5,0	4,6	3,9	3,3	4,2	2,0	5,0	4,3
Experiment 2sum	6,5	3,2	4,9	2,8	6,3	1,5	6,2	2,3
Evidence 2A	3,8	5,2	5,6	5,3	1,7	4,1	3,3	5,0
Evidence 2B	3,1	2,6	3,3	2,5	3,3	4,1	3,9	3,3
Evidence 2C	5,6	5,0	4,4	4,6	0,8	2,0	1,7	3,5
Evidence 2sum	4,3	2,9	4,2	2,7	2,0	2,2	2,9	2,3
TotScore_Exp	6,7	2,8	4,6	2,6	6,7	1,1	5,9	2,6
TotScore_Evid	3,2	2,1	3,7	2,0	3,1	2,3	3,4	1,6
Totscore_total	4,3	1,8	3,9	1,6	4,2	1,6	4,2	1,4

Looking at the scores, it is evident that in both trials students scored quite low (under 5 out of 10) overall in the test, with large variation in the scores. This indicated from the start that any comparison between or even within the experimental and control groups would not likely be statistically significant. Having said this, the total scores in the experimental

testlets are in both school trials above 5 and in the 2nd school trial close to 7 for the experimental group students.

We nevertheless conducted a series of additional statistical tests using SPSS, to compare the scores at different levels, as previously mentioned.

1. Considering all the sample together (experimental and control groups), for each trial school, we carried out:
   a) Paired samples t-test for the total pre- and post-test scores (var. 'TotScore_total'). We found no statistical difference ($p<0.05$) for either school trial.
   b) Paired samples t-test for the total pre- and post-test scores of all 'experimenting' testlets (var. 'TotScore_Exp'). We found no statistical difference ($p<0.05$) for either school trial.
   c) Paired samples t-test for the total pre- and post-test scores of all 'evaluating evidence' testlets (var. 'TotScore_Evid'). We found no statistical difference ($p<0.05$) for either school trial.
   d) Paired samples t-test for the total pre- and post-test scores of each testlet (var. 'Experiment 1', 'Evidence 1sum', 'Experiment 2sum' Evidence 2sum'). We found no statistical difference ($p<0.05$) for either school trial.
   e) Paired samples t-test for the pre- and post-test scores of each item (var. 'Experiment 1', 'Evidence 1A', 'Evidence 1B', 'Evidence 1C', 'Experiment 2A', 'Experiment 2B', 'Experiment 2C', 'Evidence 2A', 'Evidence 2B', 'Evidence 2C'). We found no statistical difference ($p<0.05$) for all items for either school trial, but the 'Evidence 2B' (1st trial school sig. 0.054) and the 'Evidence 2C' (2nd trial school sig 0.026).
2. Considering the sample of the experimental and control groups separately and for each trial school, we carried out:
   a) Paired samples t-test for the total pre- and post-test scores (var.'TotScore_total'). We found no statistical differences ($p<0.05$) for either group or school trial.
   b) Paired samples t-test for the total pre- and post-test scores of all 'experimenting' testlets ('TotScore_Exp'). We found no statistical differences but for the control group of the 2nd trial school sig. 0.045 ($p<0.05$).
   c) Paired samples t-test for the total pre- and post-test scores of all 'evaluating evidence' testlets ('TotScore_Evid'). We found no statistical differences ($p<0.05$) for either group or school trial.
   d) Paired samples t-test for the total pre- and post-test scores of each testlet (var. 'Experiment 1', 'Evidence 1sum', 'Experiment 2sum', 'Evidence 2sum'). We found no statistical differences but for 'Experiment 2sum' of the 2nd trial's control group, sig. 0.045 ($p<0.05$).
   e) Paired samples t-test for the pre- and post-test scores of each item (var. 'Experiment 1', 'Evidence 1A', 'Evidence 1B', 'Evidence 1C', 'Experiment 2A', 'Experiment 2B', 'Experiment 2C', 'Evidence 2A', 'Evidence 2B', 'Evidence 2C'). We found no statistical differences but for 'Experiment 2B' of the 2nd trial's control group, sig. 0.030 ($p<0.05$).
   f) Paired samples t-test between the pre-test scores of the 'experimenting' ('TotScore_Exp') and 'evaluating evidence' testlets ('TotScore_Evid'). It was statistically significant for the 1st trial's control group, sig. 0.008 ($p<0.05$).

g) Paired samples t-test between the post-test scores of the 'experimenting' ('TotScore_Exp') and 'evaluating evidence' testlets ('TotScore_Evid'). It was statistically significant for both trials' control groups, 1st trial sig. 0.006 ($p<0.05$), 2nd trial sig. 0.04 ($p<0.05$).

h) Paired samples t-test between the pre-test scores of the 'experimenting' ('TotScore_Exp') and 'evaluating evidence' testlets ('TotScore_Evid'). It was statistically significant for the 2nd trial's experimental group, sig. 0.023 ($p<0.05$).

i) Paired samples t-test between the post-test scores of the 'experimenting' ('TotScore_Exp') and 'evaluating evidence' testlets ('TotScore_Evid'). It was statistically significant for the 2nd trial's experimental group, sig. 0.027 ($p<0.05$).

To complement the paired samples t-test, we also conducted the nonparametric equivalent Wilcoxon Signed Ranks test:

j) Wilcoxon Signed Ranks test for: the total pre- and post-test scores (var.'TotScore_total'); the total pre- and post-test scores of all 'experimenting' testlets ('TotScore_Exp'); the total pre- and post-test scores of all 'evaluating evidence' testlets ('TotScore_Evid'), for each trial separately and each group (control – experimental) separately. We found no statistical differences ($p<0.05$).

AND

a) Independent samples t-test for the total post-test scores (var.'TotScore_total') between the experimental and control groups. We found no statistical difference ($p<0.05$) between the two groups for either school trial.

b) Independent samples t-test for the total post-test scores of all 'experimenting' testlets (var. 'TotScore_Exp') between the experimental and control groups. We found no statistical difference ($p<0.05$) between the two groups for either school trial.

c) Independent samples t-test for the total post-test scores of all 'evaluating evidence' testlets (var. 'TotScore_Evid') between the experimental and control groups. We found no statistical difference ($p<0.05$) between the two groups for either school trial.

d) Independent samples t-test for the post-test scores of each testlet (var. 'Experiment 1', 'Evidence 1sum', 'Experiment 2sum' Evidence 2sum') between the experimental and control groups. We found no statistical difference ($p<0.05$) between the two groups for either school trial.

e) Independent samples t-test for the post-test scores of each item (var. 'Experiment 1', 'Evidence 1A', 'Evidence 1B', 'Evidence 1C', 'Experiment 2A', 'Experiment 2B', 'Experiment 2C', 'Evidence 2A', 'Evidence 2B', 'Evidence 2C') between the experimental and control groups. We found no statistical difference ($p<0.05$) for all items, but the 'Evidence 1C' (1st trial school sig. 0.006, $p<0.05$).

To complement the independent samples t-test, we also conducted the nonparametric equivalent Mann-Whitney U test:

f) Mann-Whitney U tests to determine if there were differences in the pre-test scores of each testlet (var. 'Experiment 1', 'Evidence 1sum', 'Experiment 2sum' Evidence 2sum'); in the total pre-test scores of all 'experimenting' testlets (var. 'TotScore_Exp') and all 'evaluating evidence' testlets (var. 'TotScore_Evid'); and in the total pre-test scores (var.'TotScore_total') between the experimental and

control groups (for each trial). We found no statistical difference ($p<0.05$) between the two groups for either school trial.

g) Mann-Whitney U tests to determine if there were differences in the post-test scores of each testlet (var. 'Experiment 1', 'Evidence 1sum', 'Experiment 2sum' Evidence 2sum'); in the total post-test scores of all 'experimenting' testlets (var. 'TotScore_Exp') and all 'evaluating evidence' testlets (var. 'TotScore_Evid'); and in the total post-test scores (var.'TotScore_total') between the experimental and control groups (for each trial). We found no statistical difference ($p<0.05$) between the two groups for either school trial.

To sum up the above:

- In comparing the scores of the pre- vs post- tests for all the sample together (experimental and control groups) for each trial school, we found statistically significant results for the test items/variables:

'Evidence 2B': "What evidence supports that small amounts of phthalates should not be dangerous to humans?" We found an increase in the mean value from 35% to 65% in the 1st trial school.

'Evidence 2C': 'Which is the most correct conclusion based on the information above?' We found a decrease in the mean value from 40% to 15% in the 2nd trial school.

These findings may support the claim that overall the ILS on Radioactivity may have had a positive effect on students' understanding of the claim-evidence relationship in science, The increase in the score of the 1st trial students suggest that by the end of the intervention students accept that there may be multiple explanations about a phenomenon, and have some understanding about using evidence to support claims and that evidence has uncertainty. The results however from the 2nd trial students suggest that they have still considerable ground to cover to understand the complicated relationship between claims and evidence.

- In comparing the scores of pre- vs post- tests for the experimental and control groups separately for each trial school, we found statistically significant results for the following variables:

'TotScore_Exp': An increase from 46.3% to 59.3% for the control group in the 2nd trial.

'Experiment 2sum' ('Experiment on friction'): An increase from 48.9% to 62.2% for the control group in the 2nd trial.

'Experiment 2B' 'How should they decide which results they should use?": An increase from 44.4% to 55.6% for the control group in the 2nd trial.

These findings may suggest that the students in the 2nd trial control group may have secured and increased their procedural knowledge, in terms of understanding about the need for repeated measurements of a single variable, though they still lack in their knowledge about anomalous data and do not have a strategy for handling these.

- In comparing the scores in the 'Experimenting' vs 'Evaluating evidence' inquiry phases, for the experimental and control groups separately, we found statistically significant results:

  Before the intervention: scores of 56.1% vs 32.5% for the 1st trial's control group; and 66.7% vs 32.1% for the 2nd trial's experimental group

After the intervention: scores of 64.3% vs 36.7% for the 1st trial's control group; 66.7% vs 31% for the 2nd trial's control group; and 59.3% vs 34.1% for the 2nd trial's experimental group

These findings seem to suggest that students overall have more procedural than epistemic knowledge and thus a greater understanding of the 'Experimenting' than the 'Evaluating evidence' inquiry phase in both trials, and both Before and After the ILS 'Radioactivity' intervention. The difference in scores however between the two kinds of knowledge seems to strengthen and become more apparent After the intervention, both in the control and experimental groups, suggesting a possible effect of the ILS 'Radioactivity' on students' procedural knowledge.

- In comparing the post-tests scores between the Experimental and Control groups we found statistically significant difference for the variable:

  'Evidence 1C' 'If you weigh a growing pot plant you will find that the weight increases more than the water you add. Whose views this evidence supports or contradicts?': The 1st trial's Experimental group had a much greater post-test score than the equivalent Control group (30% vs 9.5%)

This finding extends the premises of our finding, that the intervention had some positive effect on the 1st trial students' epistemic knowledge about the claim-evidence relationship in science and show that this effect is greater for the experimental group that used the 'Experimental Error Tool' when using evidence for how plants increase in weight to support or contradict claims.

3. Finally, given that the same instructional strategy was used in both school trials, we used the Mann-Whitney U Test to examine if there were differences between the two experimental groups or between the two control groups among trials. There was no statistical difference in their total pre-test scores (var.'TotScore_total'). We thus proceeded in merging the two control groups and the two experimental groups into one and we proceeded in conducting the following tests:
   - Wilcoxon Signed Ranks Test between the pre- and post-test scores in all the aggregated variables, separately for the amalgamated control and experimental groups. No statistical significance was found, but for the control group in the aggregated 'experimenting' testlets (var. 'TotScore_Exp'). This is consistent with what we found in test 2b) above.

9.2.5 Discussion of findings

Given the above findings, it is safe to recognise that there was not any significant effect of the use of the 'Experimental Error Tool' on students' understanding about 'How Science Works'. A number of reasons can be identified for this:

a) Student's understanding about 'How Science Works', and in this particular case, students' procedural and epistemic knowledge are unlikely to change in the course of only one experiment sustained and long-term intervention is needed for this to occur.

b) The length of use of the Tool was not enough for this to have an effect, especially given that the students' were totally unfamiliar with it.

c) The 'Experimental Error Tool' is in the English language, and this may counteract its purpose of use as a scaffolding tool, at least for non-native English speaking students.

d) The ‘Experimental Error Tool’ included significant mathematics content itself, which needed to be understood before it could be used seamlessly to treat experimental values. The students may not have had enough time to acquire this understanding and thus never managed to appreciate fully its intended potential for analysis of experimental values and their certainty.

Further, longer and repeated investigations are needed to assert on the learning value of using the ‘Experimental Error Tool’.

Having said these, there were some indications that the intervention using the ILS ‘Radioactivity’ itself may have had some positive effects on student’s knowledge, especially on their procedural knowledge and more particularly on students’ understanding of the need for repeated measurements of a of a single variable. Again, additional, longer and repeated investigations are needed to prove this more solidly.

9.3 References

Kind, P. M. (2013). Establishing assessment scales using a novel disciplinary rationale for scientific reasoning. *Journal of Research in Science Teaching, 50*, 530-560. doi: 10.1002/tea.21086

10 The effect of the Go-Lab conclusion tool on students' science learning

10.1 Abstract

The purpose of this study was to assess the effect of the Go-Lab conclusion tool on students' science learning and inquiry skills. Two conditions were compared, the control condition, where the Conclusion Tool was not included in the learning environment and the experimental group, where the Conclusion Tool had been integrated in the learning environment. Before and after the intervention, students completed the same content knowledge test and the same inquiry skills test. Students' post-test content knowledge scores in the experimental condition were marginally higher than those of the students in the control condition. Additionally, statistical significant differences between the pre- and post-test scores were found for "*Identifying variables*" and "*Identifying and stating hypothesis*" for both conditions. The Conclusion Tool appears to enhance the content knowledge of students. Since the tool allows students to gather all necessary information they would need to formulate their conclusions at one place, this might make it easier for them to be able to reach a valid conclusion. Further, the tool seems to catalyse a reflection function on former learning products, which might imply a substantial meta-cognitive step before they can combine these products and formulate their conclusions. Future research should track student actions along learning activities in the *Conclusion* phase to investigate if learner pathways differ between experimental and control conditions.

10.2 Introduction

Conclusion has been identified as one of the phases involved in inquiry (Pedaste et al., 2015). In this phase, students reach to conclusive statements concerning research questions or hypotheses, which are usually formulated at the beginning of an inquiry enactment, after some sort of investigation (e.g., exploration, experimentation) that results in the collection of proper data/evidence (Scanlon et al., 2011). In the case of an open research question, the Conclusion phase leads to specifying the relationship between the variables under study, whereas in the case of a hypothesis the Conclusion phase requires a conclusive remark, accompanied by the necessary evidence, as far as the hypothesis is accepted or rejected.

As most of the inquiry phases, the Conclusion phase is also a challenging task to complete (van Joolingen & Zacharia, 2009; Zacharia et al., 2015). This arises because of the different factors that a student needs to consider before reaching to conclusions. Specifically, for reaching to conclusions a student needs to consider (a) the research question or hypothesis stated, (b) the data/evidence that emerged during the Investigation phase, usually through an experiment, and (c) the interpretations of the data, after they have been organized/represented and analysed (Pedaste et al., 2015). Research has highlighted a number of difficulties that students face during the Conclusion phase. For instance, students fail to consider all the aforementioned together to synthesize their conclusions, or fail to consider all the evidence collected (for more details see Zacharia et al., 2015).

One way to overcome the problems related to the Conclusion phase, researchers have advocated in favour of the provision of guidance, especially through the use of computer supported inquiry learning environments, considering how difficult it is for a teacher to provide individual based feedback and support to each student separately (Cho & Jonassen, 2012; Demetriadis et al., 2008; McNeill et al., 2006; Reiser et al., 2001;

Veermans et al., 2006; Woolf et al., 2002; Zumbach, 2009). The literature of the domain shows that researchers have designed and developed different types of guidance, such as performance dashboards, prompts, heuristics and scaffolds, to support student learning when enacting inquiry in computer supported learning environments. For instance, Woolf et al. (2002) developed the Final Case Review Tool to support students at the conclusion phase. This particular tool allowed students to have access to a review of all their observations and hypotheses tested when creating their final report. McNeill et al. (2006) developed prompts in the form of statements (e.g., "Write a sentence that connects your evidence to your claim that…") to support students when writing conclusive scientific explanations following the structure of claim-evidence-reasoning, whereas Demetriadis et al. (2008) developed prompts in the form of questions, namely the observe prompt, the recall prompt and the conclude prompt, in order to support the students to spot the necessary information and accompany them with proper reasoning for reaching to solid conclusions. Both sets of prompts were found to have a positive effect on the quality of students' conclusions. Veermans et al. (2006) have referred to the Present evidence heuristic developed by Schoenfeld (1985) to support the Conclusion phase. This particular heuristic reminded students when stating a conclusion about a certain hypothesis, to present evidence to support that conclusion. However, no empirical evidence was provided about the effectiveness of this particular tool.

In addition to prompts and heuristics, a number of scaffolds were developed to support students when formulating conclusions. For example, Reiser et al. (2001) used in their BGulLe computer supported learning environment a scaffold, namely the ExplanationConstructor tool of the investigation journal, that requested from students to directly connect their data and their explanations. Zumbach (2009) developed and used a scaffold which allowed students to represent their arguments (at a conclusion stage) with a text editor tool. In the text editor tool case, students were asked to classify pro and con arguments, whereas in the case of the graphical mind mapping tool, students were asked to connect these arguments and mark them either with "+" or "-", respectively. The latter scaffold was also found to enhance students' acquisition of knowledge.

Given all these conclusion related tools and their positive impact on students' formation of conclusions and learning, and given that the studies examining conclusion related tools effect on student learning are limited in number, we developed a conclusion tool in the context of the Go-Lab platform and aimed at examining its effectiveness. In particular, the purpose of this study was to assess the effect of the Go-Lab conclusion tool on secondary school students' science learning and inquiry skills.

10.3 Methods

The experimental design we have followed is presented in Figure 10.1. Two conditions were compared, the control condition, where the Conclusion Tool was not included in the learning environment and the experimental group, where the Conclusion Tool had been integrated in the learning environment. Before and after the intervention, students completed the same content knowledge test and the same inquiry skills test.

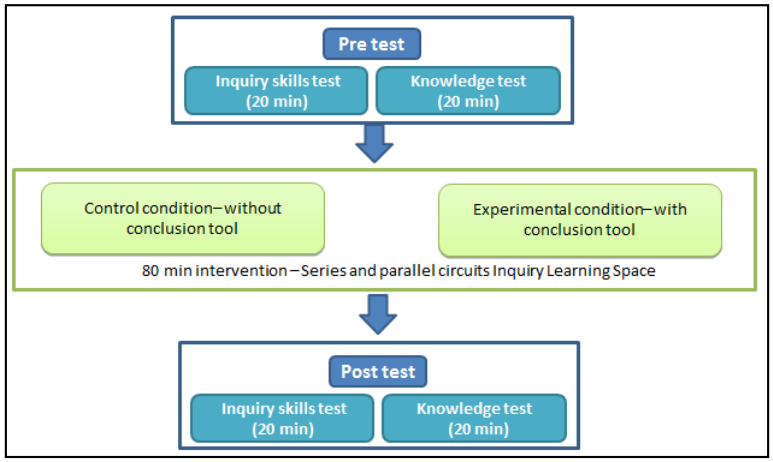


Figure 10.1. Experimental design.

10.3.1 Participants

The study involved 27 tenth graders (16-17 years old) from two classes (N_{class1} = 12 and N_{class2} = 15) of two public senior high schools in Larnaca, Cyprus. The students in class1 (11 boys and 1 girl) were assigned to the control condition and the students in class 2 (4 boys and 11 girls) to the experimental condition. The selection of the two classes was made based on the teachers interest to participate in the Go-Lab project and because students were at the same age, with the same prior knowledge and skills, concerning content knowledge and skills, respectively. In addition, students in both classes had advanced computer skills.

10.3.2 Materials

10.3.2.1 Learning environment

The students worked with an online learning environment termed Inquiry Learning Space (ILS), which was created based on the inquiry cycle design framework (Pedaste et al., 2015) through the Go-Lab authoring tool (de Jong, Sotiriou, & Gillet, 2014). Through the ILS students learned about the simple electric circuit and investigated differences between the series and parallel circuits. The ILS comprised five inquiry phases, namely, the *Orientation*, the *Conceptualization*, the *Investigation*, the *Conclusion* and the *Discussion* phase. For their investigations students used the Electric Circuit Laboratory Figure 10.2), which is available in the Golabz repository (http://www.golabz.eu/).

In the *Orientation* phase, students gathered information about the simple electric circuit and the series and parallel circuits through videos and text. In the *Conceptualization* phase, students first explored the Electrical Circuit Laboratory, in order to become familiar with its elements, meters and functions. In this laboratory students could create their circuits by dragging elements from the left side of the tool´s interface and drop them in the circuit board. In addition, students could connect several meters (e.g. ammeter, voltmeter) to their circuit in order to take measurements.

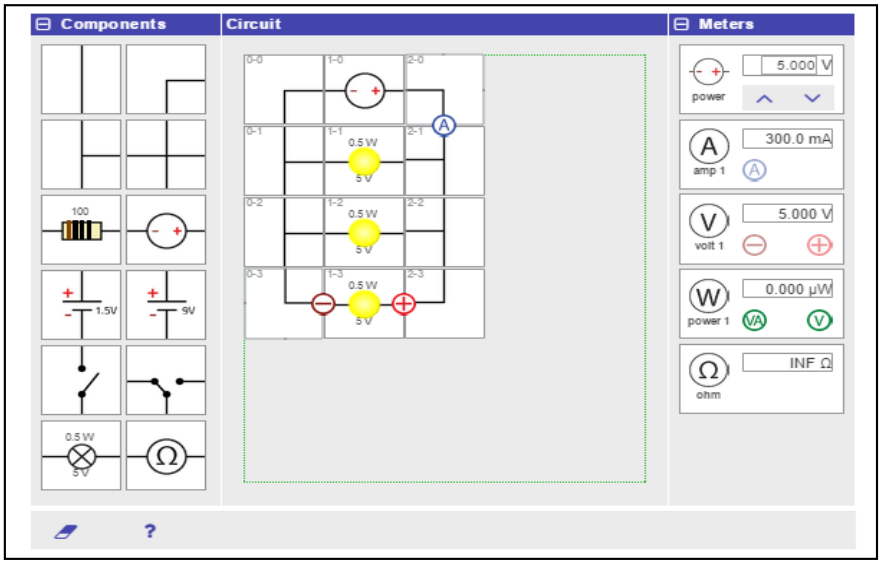


Figure 10.2. The Electrical circuit lab.

After students had played around with the Electrical Circuit Laboratory (5-7 min), they proceeded to the hypothesis/hypotheses formulation in the Hypothesis Scratchpad (Figure 10.3). In this tool, predefined conditionals and concepts were provided to students who could drag and drop them to create a hypothesis in the form of an "if…then" statement. Additionally, they could adjust their confidence level for each hypothesis, by changing the colour of the "horseshoe" placed next to the space provided by the toll to students to formulate their hypothesis. If the "horseshoe" was blue overall, that would mean that a student was absolutely sure that his/her hypothesis was correct.

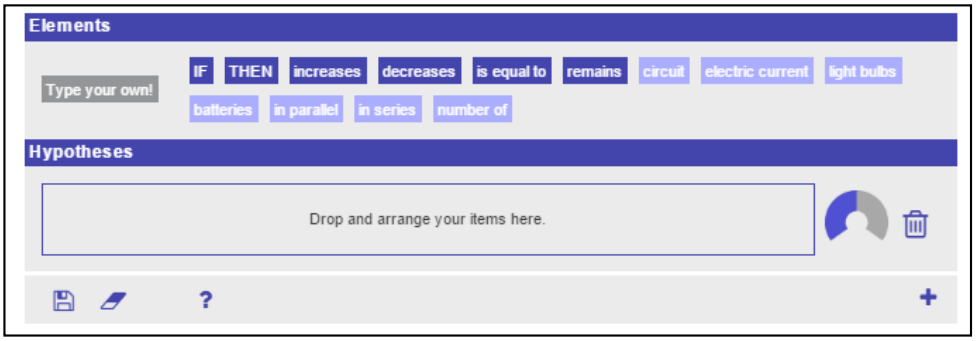


Figure 10.3. Hypothesis Scratchpad.

After the hypotheses formulation had been completed, students moved to the next phase, the *Investigation* phase, where they conducted their experiments to confirm or reject their hypothesis/hypotheses. Before running their experiments, students used the Experiment Design Tool (Figure 10.4). This tool allowed students to determine the independent variable ("Vary"), the dependent variable ("Measure") and the control variables ("Keep constant") in their experiments. To do so, they dragged the properties from the left side of the tool´s interface and dropped them in the proper column. Once a design had been completed, the next step was to specify the values of each variable and add experimental trials. After students had conducted an experimental trial in the lab, they returned to the Experiment Design Tool to enter the value of the independent variable. They had to do this for each one of their trials.

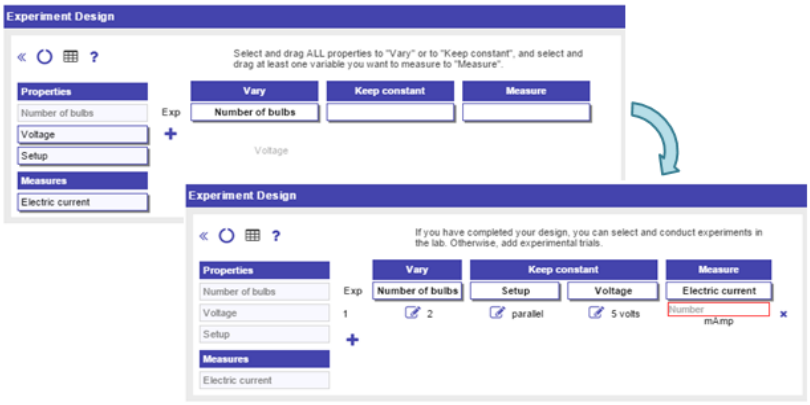


Figure 10.4. Experiment Design Tool.

During experimentation, students were prompted to keep notes about ideas, thoughts and observations, in the Observation Tool (Figure 10.5).

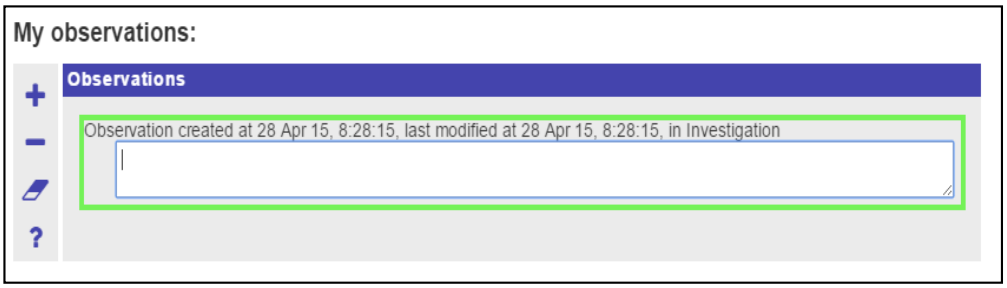


Figure 10.5. Observation Tool.

For data interpretation, students retrieved their data from the Experiment Design Tool and drew graphs in the Data Viewer Figure 10.6). For the creation of a graph students had to drag a set of data from the data set container into the data graph space. In addition, they had to respond to several questions that would help them interpret the relation between the variables depicted in the graph.

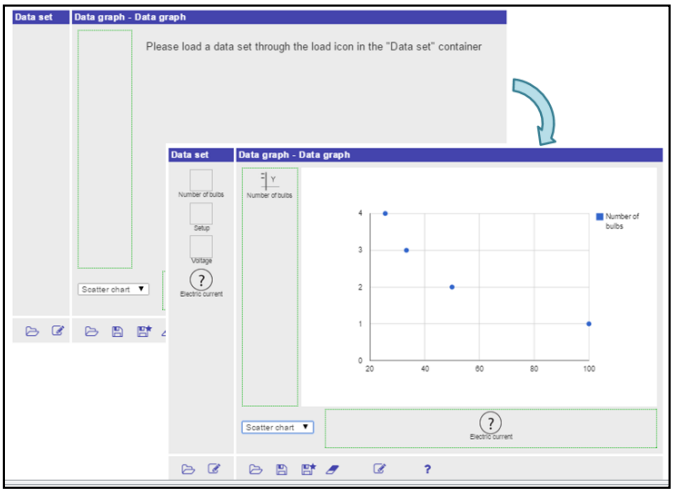


Figure 10.6. The Data viewer tool.

In the *Conclusion* phase, students used the Conclusion Tool (Figure 10.7) to retrieve their hypothesis/hypotheses, observations and graphs in order to argue how their confidence

for each hypothesis had changed after their investigation. That was the case in the experimental condition, whereas in the control condition, students proceeded directly to their final conclusion, which involved a reply to the initial problem on how the light fixtures in a house were connected.



Figure 10.7. Conclusion Tool.

In the *Discussion* phase, students performed reflection activities and discussed with the teacher and peers about their experiments and conclusions.

10.3.2.2 Assessment

Data collection involved two different tests, namely, the knowledge test and the inquiry skills test. For the creation of the knowledge test, a revised taxonomy of the levels of the cognitive domain of knowledge has been followed (more details in de Jong, 2014; Zervas, 2013). The corresponding levels were "*Remember*", "*Understand*", "*Apply*" and "*Think critically and creatively*". Specifically, the knowledge test consisted of seven items. Of them, one addressed "*Remember*", two "*Understand*", two "*Apply*" and two "*Think critically and creatively*". Three items were open-ended and four closed-ended. The items focused on the definition of the simple electric circuit and the differences between the two types of circuit set up (i.e., in series and in parallel), in terms of the brightness of the bulb and the electric current that will flow through a circuit.

For the inquiry skills test, items from the TIPSII instrument (Burns, Okey, & Wise, 1985) were selected and translated in Greek. The test consisted of 24 multiple-choice items. Of them, 12 addressed "*Identifying variables*", 9 "*Identifying and stating hypotheses"* and 3 "*Designing investigations"*. The number of items in each subscale was the same as in the initial TIPSII instrument.

After the implementation, both tests were scored blind to the condition in which each student had been placed. For the open ended items in the knowledge test, a rubric was used, that specified scoring criteria for each item. Inter-rater agreement between two independent coders who reviewed 20% of the data was found to be acceptable (Cohen's $Kappa_{study1}$= 0.94).

10.4 Procedure

The implementations were carried out by the science teacher of each class and they covered four class meetings of 40 minutes each. Before the implementation, teachers

participated in a face to face preparatory meeting to become familiar with the ILS, the Electrical Circuit Laboratory and the tools in each phase of the ILS. In addition, some procedural issues were discussed and the role of the teacher during the lesson was clarified.

In the first class meeting students completed pre-tests (i.e., the knowledge and inquiry skills tests). The next two meetings took place in the computer lab of each school, in order for each student to work in a computer to complete the activities of the ILS. At the beginning of the lesson, the teacher provided general guidance to the students, mainly explaining the way students were supposed to work in order to complete all the activities of the lesson. During the lesson, the only help students received from their teacher concerned some technical issues during the use of the several tools and the lab. Whenever technical issues appeared, they were solved without causing any delay to the completion of the lesson. Finally, the day after the intervention, the last session was done for the completion of the post-tests (i.e., the knowledge and inquiry skills tests).

10.5 Results

10.5.1 Content Knowledge test

The Wilcoxon test showed that both conditions improved secondary school students' conceptual understanding for the ILS as a whole ($p < .05$). The two conditions did not differ in terms of their pre-test scores. The Mann Whitney U test revealed that students' post-test scores in the experimental condition were marginally higher than those of the students in the control condition ($Z = -2.00$; $p = .048$).

10.5.2 Inquiry skills test

The Wilcoxon test showed statistical significant differences between the pre- and post-test scores for "*Identifying variables*" and "*Identifying and stating hypothesis*" for both conditions (Table 10.1), whereas no differences were found for "*Designing investigations*" ($p > .05$). The Mann Whitney U test revealed no statistically significant differences between the two conditions across all the inquiry skills under study ($p > .05$).

Table 10.1. Differences between pre- and post-test scores for "*Identifying variables*" and "*Identifying and stating hypothesis*" for both control and experiment<k

	Identifying variables	Identifying and stating hypothesis	Z
Control group			
Pre-test	4.58	2.92	-3.53***
Post-test	5.42	4.00	-3.25**
Experimental group			
Pre-test	3.89	2.94	-2.27*
Post-test	6.72	5.22	-2.57*

Note: * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$.

10.6 Discussion, conclusions and implications for future research

The Conclusion Tool appears to enhance the content knowledge of students. Since the tool allows students to gather all necessary information they would need to formulate their conclusions at one place, this might make it easier for them to be able to reach a valid conclusion. Further, the tool seems to catalyse a reflection function on former learning products, which might imply a substantial meta-cognitive step before they can combine these products and formulate their conclusions.

Future research should track student actions along learning activities in the *Conclusion* phase to investigate if learner pathways differ between experimental and control conditions. Moreover, time spent on writing conclusions should be counted and compared between conditions. It could be that time devoted to formulating conclusions might be less for students in the control condition if they had to return to earlier stages on the learning trajectory to look for necessary information. A last research question could refer to the quality of the conclusions per se, and if this quality varied between conditions. Namely, the ability of students to back up their claims by evidence and connect claims to evidence when expressing their reasoning.

10.7 References

Burns, J. C., Okey, J. R., & Wise, K. C. (1985). Development of an integrated process skill test: TIPS II. *Journal of Research in Science Teaching, 22*(2), 169-177.

Cho, Y., & Jonassen, D. (2012). Learning by self-explaining causal diagrams in high-school biology. *Asia Pacific Education Review, 13*, 171-184. doi: 10.1007/s12564-011-9187-4

de Jong, T. (Ed.). (2014). D1.3: *Preliminary inquiry classroom scenarios and guidelines: Go-Lab project*.

de Jong, T., Sotiriou, S., & Gillet, D. (2014). Innovations in STEM education: The Go-Lab federation of online labs. *Smart Learning Environments, 1*(1), 1-16.

Demetriadis, N. S., Papadopoulos, M. P., Stamelos, G. I., & Fischer, F. (2008). The effect of scaffolding students' context-generating cognitive activity in technology-enhanced case-based learning. *Computer & Education, 51*, 939-954. doi: 10.1016/j.compedu.2007.09.012

McNeill, K. L., Lizotte, D. J., Krajcik, J., & Marx, R. W. (2006). Supporting students' construction of scientific explanations by fading scaffolds in instructional materials. *Journal of the Learning Sciences, 15,* 153-191. doi: 10.1207/s15327809jls1502_1

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.

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

Scanlon, E., Anastopoulou, S., Kerawalla, L., & Mulholland, P. (2011). How technology resources can be used to represent personal inquiry and support students' understanding of it across contexts. *Journal of Computer Assisted Learning, 27*, 516-529. doi: 10.1111/j.1365-2729.2011.00414.x

Schoenfeld, A. (1985). *Mathematical problem solving*. New York: Academies Press.

van Joolingen, W. R., & Zacharia, Z. C. (2009). Developments in inquiry learning. In N. Balacheff, S. Ludvigsen, T. de Jong, & S. Barnes (Eds), *Technology-enhanced*

learning (pp. 21-37). Dordrecht, The Netherlands: Springer. doi: 10.1007/978-1-4020-9827-7_2

Veermans, K. H., van Joolingen, W., & de Jong, T. (2006). Use of heuristics to facilitate scientific discovery learning in a simulation learning environment in a physics domain. *International Journal of Science Education, 28*, 341-361. doi: 10.1080/09500690500277615

Woolf, B., Reid, J., Stillings, N., Bruno, M., Murray, D., Reese, P., .. Rath, K. (June, 2002). A general platform for inquiry learning. *Proceedings of the International Conference on Intelligent Tutoring Systems.* Biarritz, France. doi: 10.1007/3-540-47987-2_69, Retrieved January 24, 2012 from http://link.springer.com/chapter/10.1007%2F3-540-47987-2_69?LI=true

Zacharia, Z. C., Manoli, C., Xenofontos, N., de Jong, T., Pedaste, M., van Riesen, S., . . . Tsourlidaki, E. (2015). Identifying potential types of guidance for supporting student inquiry when using virtual and remote labs: A literature review. *ETR&D, 63*, 257-302

Zervas, P. (Ed.). (2013). D2.1. *The Go-Lab inventory and integration of online labs – labs offered by large scientific organisations: Go-Lab project.*

Zumbach, J. (2009). The role of graphical and text based argumentation tools in hypermedia learning. *Computers & Education, 25*, 811-817. doi: 10.1016/j.chb.2008.07.005.

11 Constructing graphs by means of the Data Viewer: Comparison between two different configurations of the tool in terms of their effects on student knowledge and skills

11.1 Abstract

The purpose of the current study was to assess a newly design graphing tool, the Data Viewer, which was developed within the context of the Go-Lab project. We compared two versions of the Data Viewer, which differed in some core functionalities. The first configuration (DataV1) was based on a connection between the Electrical Circuit Lab and the Data Viewer, which made available to students only one variable they had already handled/examined during their experimental trials. This configuration demanded from students to reflect on previous learning activities and be able to retrieve variables and their values from notes taken previously. The second configuration of the tool (DataV2) retrieved all variables handled by students in their experimental trials automatically. Students had to choose among four variables, which two to use, in order to construct their graph. The results showed that the two configurations of the tool had different implications. When data had to be loaded by students themselves (DataV1), knowledge gains were significant, while inquiry skills did not improve at all. When all data were loaded by the tool (DataV2), there were not significant knowledge gains for students but inquiry skills did improve substantially. Implications for “structuring” and “problematizing” student work are discussed.

11.2 Introduction

Data interpretation is considered as an essential task for inquiry learning in science. Pedaste et al. (2015) located data interpretation as a procedure that concludes the “Investigation” phase. During data interpretation, students need to make sense of data collected during their investigations or experimentations so that valid conclusions can be drawn. By interpreting their data, students address their hypotheses and decide how to proceed with their inquiry (Chang, Chen, Lin & Sung, 2008).

Since responding to hypotheses requires an examination of data and of any relation between variables handled during investigations and experimentations (van Joolingen & Zacharia, 2009), appropriate support might be often need to guide students through these complex tasks (Zacharia et al., 2015).

Students can be scaffolded during data interpretation by means of visualizations. Among the first attempts in that direction in computer supported learning environments, Mokros and Tinker (1987) aimed at supporting graph construction. More recently, technological advancement offered the opportunity of accompanying graph construction and interpretation with automatic feedback (Chang, Chen, Lin & Sung, 2008).

The use of graphing tools has also been attempted, where students might be requested to create data graphs themselves. Since this might be a time consuming and error prone process, automatic generation of graphs allows students to reallocate time from directly handling data collected to data interpretation (van Joolingen, de Jong, Lazonder, Savelsbergh & Manlove, 2005).

However, refocusing student concentration and time devoted to a task might have crucial implications in terms of their knowledge or inquiry skills. For instance, there might be different effects between letting students select among a pre-define set of variables or letting them create the set of variables to use in order to construct a graph. The purpose of the current study was to examine this issue by assessing a newly design graphing tool,

the Data Viewer, which was developed within the context of the Go-Lab project. Two different versions of this tool have been compared, based on their effect on student knowledge and inquiry skills.

11.3 Method

In the current study, we compared two versions of the Data Viewer, which differed in some core functionalities. The tool had been integrated in a specific learning activity of the Go-Lab online platform, namely an Inquiry Learning Space (ILS) on electric circuits (for details on ILSs see Pedaste et al. (2015)), and students used it right after they ran their experiments and collected data. Before and after the intervention, students in each condition (i.e., using each version of the Data Viewer) completed a knowledge test and an inquiry skills test.

11.3.1 Participants

The study involved 30 tenth graders (16-17 years old) from two classes (N_{class1} = 12 and N_{class2} = 18) of two public senior high schools (Lyceums) in Larnaca, Cyprus. Students in class1 (11 boys and 1 girl) worked with version 1 of the Data Viewer (DataV1), while students in class 2 (4 boys and 14 girls) worked with version 2 of the Data Viewer (DataV2). The selection of the two classes was based on the teachers´ interest to participate in the Go-Lab project. Additional selection criteria involved age alignment between classes, as well as alignment in terms of prior knowledge and skills. Further, students in both classes had advanced computer skills.

11.3.2 Materials

11.3.2.1 Learning environment

For the purpose of this study, an ILS was created through the use of the Go-Lab authoring tool (de Jong, Sotiriou, & Gillet, 2014). The ILS focused on the simple electric circuit and circuits connected in series and in parallel. Specifically, students learned about the differences between the two types of setup (i.e., series vs. parallel circuits). The ILS consisted of five inquiry phases, namely, the *Orientation*, the *Conceptualization*, the *Investigation*, the *Conclusion* and the *Discussion* phase. For their investigations, students conducted experiments in an online laboratory, the Electrical Circuit Lab (Figure 11.1), which is available in the Go-Lab online lab repository (http://www.golabz.eu/).

In the *Orientation* phase, students gathered information about the simple electric circuit and the series and parallel circuits through videos and text. In the *Conceptualization* phase, students first explored the Electrical Circuit Lab, in order to become familiar with its elements, meters and functions. In this laboratory, students could create their circuits by dragging elements from the left side of the tool´s interface and dropping them in the circuit board (Figure 11.1). In addition, students could connect several meters (e.g. ammeter, voltmeter) to their circuit in order to take measurements.

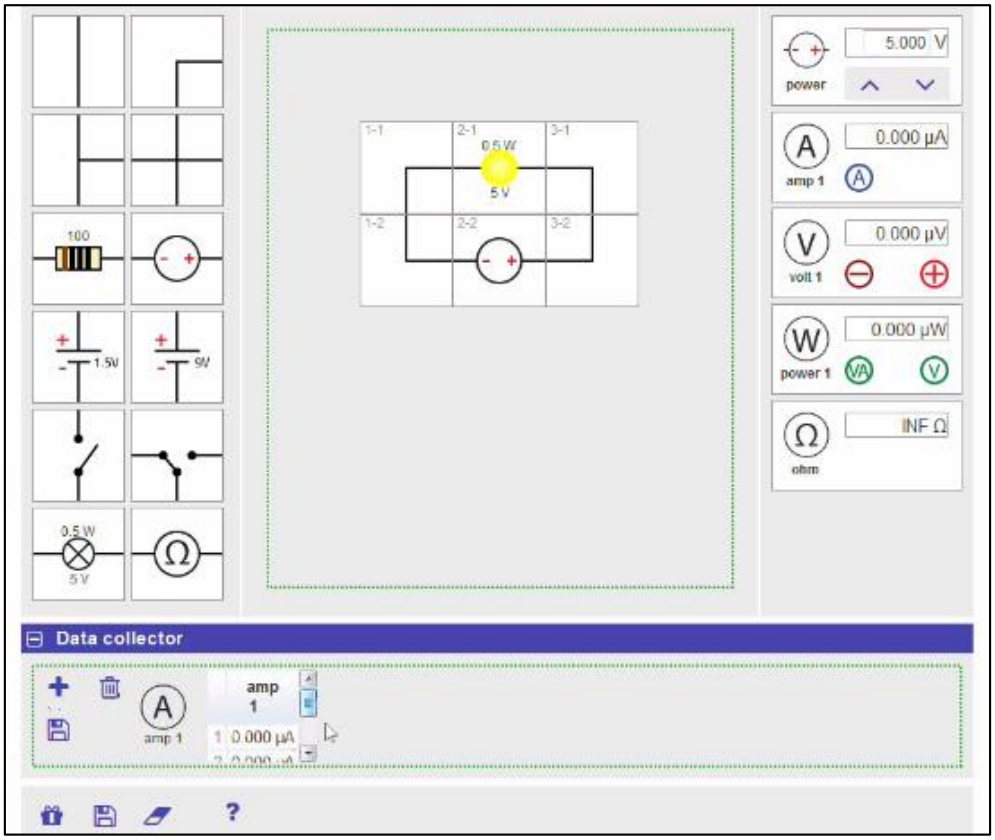


Figure 11.1. The Electrical Circuit lab.

After students had played around with the Electrical Circuit Lab for 5-7 minutes, they proceeded to the hypothesis/hypotheses formulation by means of the Hypothesis Scratchpad (Figure 11.2). In this tool, predefined conditionals and variables were provided to students who could drag and drop them to create a hypothesis in the form of an "if…then" statement. Additionally, they could adjust their confidence level for each hypothesis, by changing the colour of the "horseshoe" to the right side of the space, where the hypothesis had been formulated. If the "horseshoe" was blue, overall, that would mean that a student was absolutely sure that his/her hypothesis was correct.

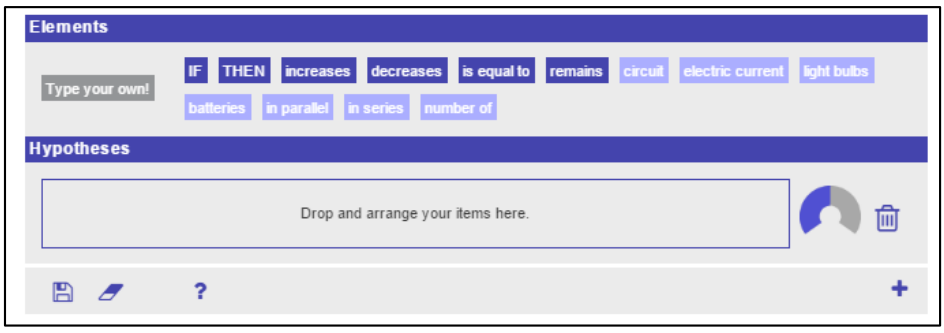


Figure 11.2. Hypothesis Scratchpad.

After the hypotheses formulation had been completed, students moved to the next phase, the *Investigation* phase, where they conducted their experiments to confirm or reject their hypothesis/hypotheses. Before running their experiments, students used the Experiment Design Tool (Figure 11.3). This tool allowed students to determine the independent variable ("Vary"), the dependent variable ("Measure") and the control variables ("Keep constant") in their experimental trials. To do so, they dragged variables from the left side

of the tool´s interface and dropped them in the proper column. Once a design had been completed, the next step was to specify the values of each variable and add experimental trials. After students had conducted an experimental trial in the lab, they returned to the Experiment Design Tool to enter the value of the dependent variable. They had to do this for each one of their trials.



Figure 11.3. Experiment Design Tool.

During experimentation, students were prompted to keep notes about ideas, thoughts and observations, by means of the Observation Tool (Figure 11.4).

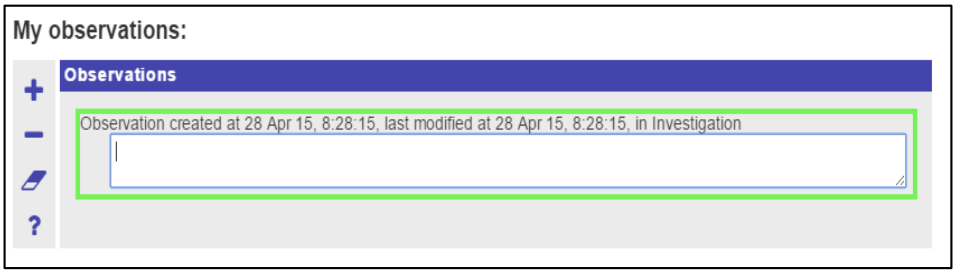


Figure 11.4. Observation Tool.

For data interpretation, students could retrieve data from their experiments and construct graphs by means of the Data Viewer. This tool was the only difference between the two conditions in the study. In the first configuration of the Data Viewer (DataV1), students recorded the electric current in the experimental trials they conducted and these data were collected and stored through the Data collector, which had been integrated in the lab (see Figure 11.1). When students were ready to use the Data Viewer, data stored in the Electrical Circuit lab were available to them through the ammeter, which was depicted in the Data set container of the Data Viewer (Figure 11.5). Since there was only one variable available, students had to create a new variable in the Data set container to be able to construct the graph and plot the values they had recorded in the Electrical Circuit lab. This was done by them manually and had to be based on notes taken previously. After students created the second variable, they had to drag both variables from the Data set container to the graph space to construct their graph.

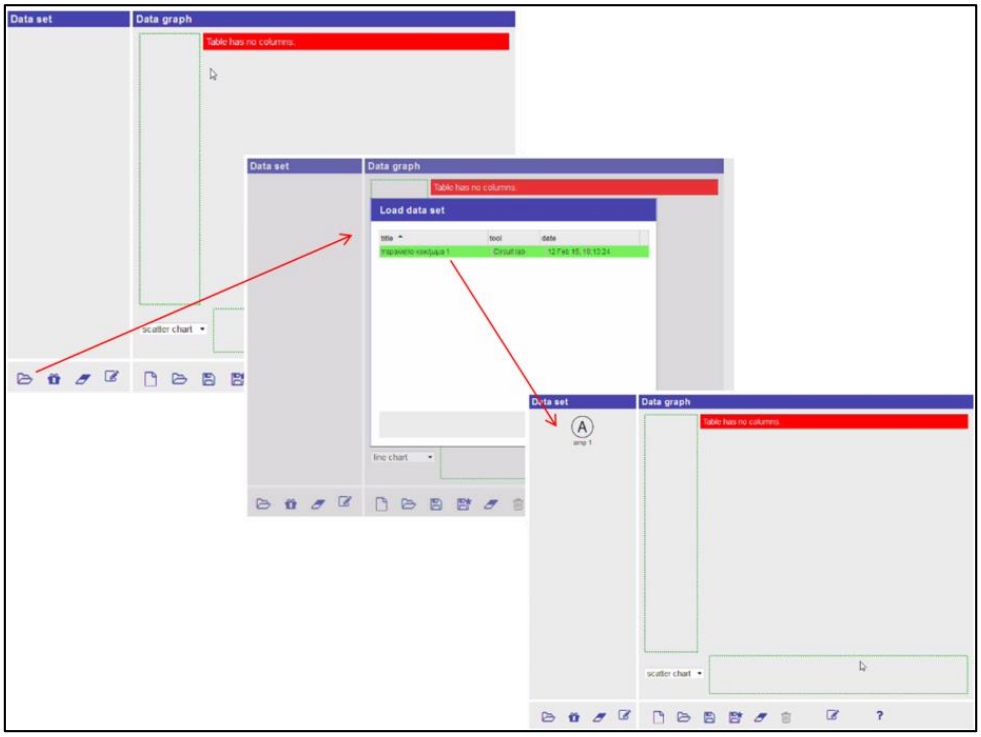


Figure 11.5. Data loading in DataV1.

In the second configuration of the Data Viewer (DataV2), all variables included in the Experiment Design Tool were automatically transferred to the Data set container of the Data Viewer (Figure 11.6). These variables were “number of bulbs”, “setup”, “voltage”, and “electric current”. Since values for all variables were saved by students after each experimental trial in the Experiment Design Tool, all these values were automatically retrieved and available for students. Whereas DataV1 was based on a link between the Electrical Circuit lab and the Data Viewer and could only transfer to the Data set container of the latter one variable (i.e., “electric current”), DataV2 was configured on a link between the Experiment Design Tool and the Data Viewer and could transfer all four variables handled by students in their experimental designs. In DataV2, students had to choose 2 out of 4 variables in the Data set container of the Data Viewer and drag them in the data graph space to construct their graph.

For both conditions, namely, for both configurations of the tool, students completed the data graphing activity and then responded to several questions that would help them interpret the relation between the two variables as this was depicted in their graph.

In the *Conclusion* phase, students used the Conclusion Tool (Figure 11.7) to retrieve their hypothesis/hypotheses, observations and graphs in order to argue how their confidence for each hypothesis had changed after their investigation.

In the *Discussion* phase, students performed reflection activities and discussed with the teacher and peers about their experiments and conclusions.

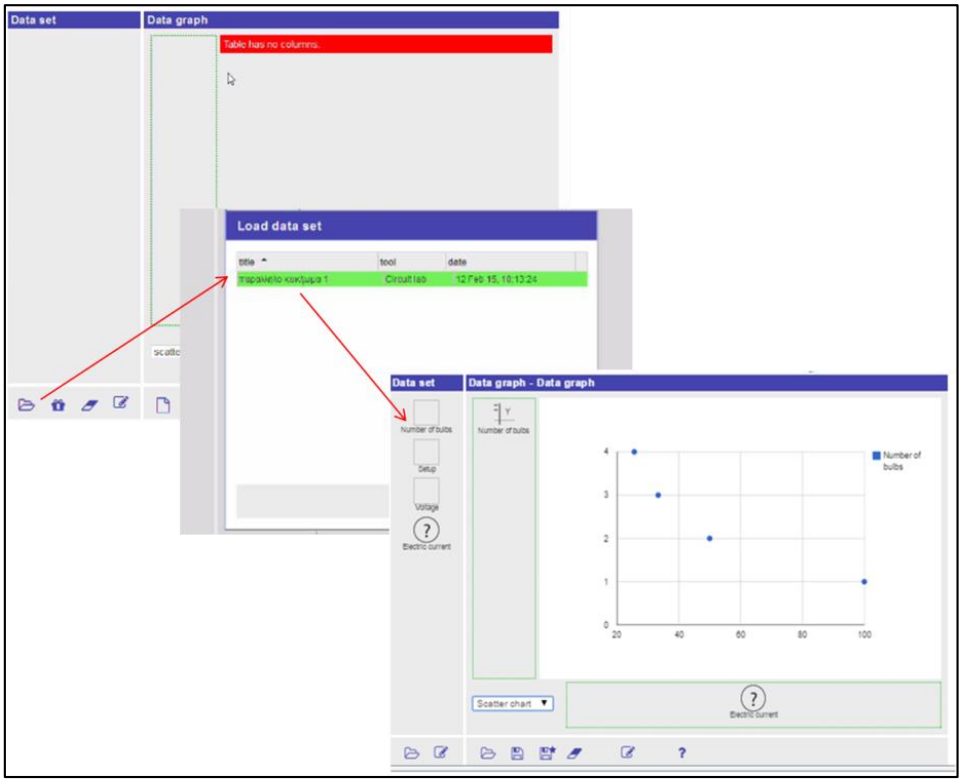


Figure 11.6. Data loading in the DataV2 condition.



Figure 11.7. Conclusion Tool.

11.3.2.2 Assessment

Data collection involved two different tests, namely, the knowledge test and the inquiry skills test. For the creation of the knowledge test, a revised taxonomy of the levels of the cognitive domain of knowledge has been followed (more details in de Jong et al. 2014; Zervas, 2013). The corresponding levels were "*Remember*" (one item), "*Understand*" (two items), "*Apply*" (two items) and "*Think critically and creatively*" (two items). Overall, the knowledge test consisted of six items. Three items were open-ended and four closed-ended. The items focused on the definition of the simple electric circuit and the differences between the two types of circuit setup (i.e., in series and in parallel), in terms of the brightness of the bulb and the electric current that would flow through a circuit.

For the inquiry skills test, items from the TIPSII instrument (Burns, Okey & Wise, 1985) were selected and translated in Greek. The test consisted of 30 multiple-choice items. Of them, 12 addressed "*Identifying variables*", 9 referred to "*Identifying and stating hypotheses"*, 3 corresponded to "*Designing investigations",* and another 6 addressed "*Graphing and interpreting*". The number of items in each subscale was the same as in the initial TIPSII instrument.

After the implementation, both tests were scored blind to the condition in which each student had been placed. For the open ended items in the knowledge test, a rubric was used, that specified scoring criteria for each item. Inter-rater agreement between two independent coders who reviewed 20% of the data was found to be acceptable (Cohen's Kappa= 0.93).

11.4 Procedure

The implementations were carried out by the science teacher of each class and they covered four class meetings of 40 minutes each. Before the implementation, teachers participated in a face to face preparatory meeting and became familiar with the ILS, the Electrical Circuit Lab, and the tools in each phase of the ILS. In addition, some procedural issues were discussed and the role of the teacher during the lesson was clarified.

In the first class meeting, students completed pre-tests (i.e., the knowledge and inquiry skills tests). The next two meetings took place in the computer lab of each school, in order for each student to work in a computer and complete the activities of the ILS. At the beginning of the lesson, the teacher provided general guidance to the students, mainly explaining the way students were supposed to work in order to complete all the activities of the lesson. During the lesson, the only help students received from their teacher concerned some technical issues during the use of the several tools and the lab. Whenever technical issues appeared, they were solved without causing any delay to the completion of the lesson. Finally, the day after the intervention, the last session was done for the completion of the post-tests (i.e., the knowledge and inquiry skills tests).

11.5 Results

Conditions (DataV1 and DataV2) did not differ significantly neither in knowledge pre-tests nor in inquiry skills pre-tests (Table 11.1 and Table 11.2, respectively; student performance has been weighed to range from 0 to 1, where 1 represents maximum knowledge or inquiry skill score). However, post-tests revealed significant differences between conditions in terms of both knowledge (Table 11.1; Mann-Whitney Test $Z = -2.58$; $p < 0.05$) and inquiry skills (Table 11.2; Mann-Whitney Test $Z = -2.05$; $p < 0.05$). In the case of the knowledge improvement, DataV1 was found to have higher rankings than DataV2 and in the case of the inquiry skills enhancement, DataV2 was found to have higher rankings than DataV1.

Table 11.1. Student overall performance in the knowledge test

	Data loaded by students (DataV1)	Data loaded by the tool (DataV2)	Mann-Whitney Test Z
Pre-test	0.37	0.32	-0.52
Post-test	0.62	0.44	-2.58*
Wilcoxon Signed Ranks Test Z	-2.86**	-1.95	

Note: * $p < 0.05$; ** $p < 0.01$.

Indeed, it seems that the two configurations of the tool had different implications for the two conditions. When data had to be loaded by students themselves, then knowledge gains were significant (Table 11.1, DataV1; Wilcoxon Signed Ranks Test Z = -2,86; p < 0,01), while inquiry skills did not improve at all (Table 11.2, DataV1). When all data were loaded by the tool, there were not significant knowledge gains for students (Table 11.1, DataV2) but inquiry skills did improve substantially (Table 11.2, DataV2; Wilcoxon Signed Ranks Test Z = -3,18; p < 0,01).

Table 11.2. Student overall performance in the inquiry skills test

	Data loaded by students (DataV1)	Data loaded by the tool (DataV2)	Mann-Whitney Test Z
Pre-test	0.42	0.36	-1.54
Post-test	0.42	0.56	-2.05*
Wilcoxon Signed Ranks Test Z	-0.05	-3.18**	

Note: * p < 0.05; ** p < 0.01.

11.6 Discussion, conclusions and implications for future research

The two different configurations of the Data Viewer had quite divergent effects on student knowledge and inquiry skills. The first configuration (DataV1) was based on a connection between the Electricity Circuit Lab and the Data Viewer, which made available to students only one variable they had already handled during their experimental trials. This configuration demanded from students to reflect on previous learning activities and be able to retrieve variables and their values from notes taken previously. This might have eventuated in triggering a metacognitive task that accumulated in knowledge gains for the students of this condition.

The second configuration of the tool (DataV2) retrieved all variables handled by students in their experimental trials automatically. Students had to choose among four variables, which two to use, in order to construct their graph. This configuration of the Data Viewer involved a selection task that proved beneficial for student inquiry skills. Indeed, selecting variables obviously presupposed identifying variables and interrelating the dependent variable with at least one independent variable to construct a graph. All these steps touched dimensions controlled by the items included in the inquiry skills test used in the current study.

Our results need to be discussed in terms of “structuring” and “problematizing” student work (Reiser, 2004). Indeed, these two mechanisms might often be in tension in design, namely, as structuring increases, problematizing might decrease and vice versa. DataV1 was different from DataV2 in that the former might have initiated a problematizing mechanism, while the latter might have structured their task on the Data Viewer better. Problematizing promoted by DataV1 might have prompted students to reason retrospectively and engage with the content of the ILS. On the other hand, structuring facilitated by DataV2 might have rendered variable selection a content independent assignment for students.

11.7 References

Burns, J. C., Okey, J. R., & Wise, K. C. (1985). Development of an integrated process skill test: TIPS II. *Journal of Research in Science Teaching*, 22(2), 169-177. doi: 10.1002/tea.3660220208

Chang, K. E., Chen, Y. L., Lin, H. Y., & Sung, Y. T. (2008). Effects of learning support in simulation-based physics learning. *Computers & Education*, 51(4), 1486-1498. doi: 10.1016/j.compedu.2008.01.007

de Jong, T. (Ed.). (2014). *D1.3: Preliminary inquiry classroom scenarios and guidelines: Go-Lab project.*

de Jong, T., Sotiriou, S., & Gillet, D. (2014). Innovations in STEM education: The Go-Lab federation of online labs. *Smart Learning Environments*, *1*(1), 1-16. doi: 10.1186/s40561-014-0003-6

Mokros, J. R., & Tinker, R. F. (1987). The impact of microcomputer-based labs on children's ability to interpret graphs. *Journal of Research in Science Teaching*, *24*(4), 369-383. doi: 10.1002/tea.3660240408

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. doi: 10.1016/j.edurev.2015.02.003

Reiser, B. J. (2004). Scaffolding complex learning: The mechanisms of structuring and problematizing student work. *The Journal of the Learning Sciences*, *13*(3), 273-304. doi:10.1207/s15327809jls1303_2

van Joolingen, W. R., & Zacharia, Z. C. (2009). Developments in inquiry learning. In N. Balacheff, S. Ludvigsen, T. de Jong, & S. Barnes (Eds), *Technology-enhanced learning* (pp. 21-37). Dordrecht, The Netherlands: Springer. doi: 10.1007/978-1-4020-9827-7_2

van Joolingen, W. R., de Jong, T., Lazonder, A. W., Savelsbergh, E. R., & Manlove, S. (2005). Co-Lab: research and development of an online learning environment for collaborative scientific discovery learning. *Computers in Human Behavior*, *21*(4), 671-688. doi: 10.1016/j.chb.2004.10.039

Zacharia, Z. C., Manoli, C., Xenofontos, N., de Jong, T., Pedaste, M., van Riesen, S., . . . Tsourlidaki, E. (2015). Identifying potential types of guidance for supporting student inquiry when using virtual and remote labs: A literature review. *Educational Technology Research and Development*, 63, 257-302. doi: 10.1007/s11423-015-9370-0

Zervas, P. (Ed.). (2013). D2.1. *The Go-Lab inventory and integration of online labs – labs offered by large scientific organisations*: Go-Lab project.

12 Scaffolding students' reflection with the reflection tool

12.1 Abstract

We studied the impact of the reflection tool scaffold on students' reflections after they had completed a Go-Lab learning activity. The reflection tool gives feedback to students about their use of time in an Inquiry Learning Space (ILS). The reflection tool displays the percentage of time a student spends in the various inquiry phases and compares this time to a norm set by the ILS owner. Students are prompted to reflect on their ILS use with open-ended questions. The experimental intervention involved two conditions. In one condition the reflection tool was included and in the other condition it was not. In our study assessment of the reflection tool relied on coding the response to an open-ended reflective question. Two aspects were assessed in students' reflections: content (technical, situational and sensitising) and reflection level (description, justification, critique, dialogue and transfer). We conducted a Mann-Whitney U-test and the results showed that there was a statistically significant difference in favour of the condition without the reflection tool. We discuss possible interpretations of this result.

12.2 Introduction

Research indicates that time spent on learning tasks (i.e., time-on-task) is one of the most important factors that can improve the quality of learning (Bloom, 1974; Karweit and Slavin, 1982; Stallings, 1980). Computer-based learning environments enable a relatively easy estimation of time-on-task based on trace data. Usually the collected data is used by learning analytics researchers and practitioners to evaluate students and the learning process. However, this data can also be potentially used as proactive feedback to students in order for them to reflect on their learning. Reflection is a cognitive process carried out in order to learn from experiences through individual inquiry and collaboration with others (Benammar, 2004; Dewey, 1933; Mezirow, 1991; Moon, 2004; Schön, 1983). Kori, Pedaste, Leijen, & Mäeots (2014) reviewed studies from 2007–2012 that used different support types to guide reflection in technology-enhanced learning. They identified three types of reflection support: technical tool, technical tool with predefined guidance, and technical tool with human interaction guidance. Examples of the second type (technical tool with predefined guidance) included prompts and guiding questions. In the Go-Lab learning environment the reflection tool is an example of this type of guidance. The focus of the current study was to assess whether the reflection tool has an effect on the responses students give to an open-ended reflection question. The research question addressed whether the reflection tool helps to identify students' reflection levels.

12.3 Method

In this study students worked with a Go-Lab Inquiry Learning Space about the chemistry concepts of acids, bases and pH. The ILS included two virtual laboratories from the PhET project (Acid-Base Solutions and pH Scale: Basics). The study was performed at school during a normal school lesson by students using Wi-Fi enabled tablets. A university researcher involved in the Go-Lab project helped the school teacher implement this intervention.

12.3.1 Participants

The Go-Lab intervention was conducted twice with two different classes at the same public school in Estonia. Complete data was collected from a total of 43 students aged 15-16 (mean age 15.0).

12.3.2 Procedure

Students had 45 minutes to solve the inquiry tasks presented in the Inquiry Learning Space. Each student was provided with a Wi-Fi enabled tablet to access the ILS and complete the inquiry tasks. A university researcher involved in the Go-Lab project led the intervention at the school as well as took notes about the intervention.

The Go-Lab intervention involved two conditions. In the first condition (i.e. Condition 1) the reflection tool was included and in the second condition (i.e. Condition 2) it was not. Condition 2 without the reflection tool included instead an input box tool which allowed students to answer the open-ended reflection question but without the additional visual feedback about use of time that the reflection tool provides. Students were randomly placed into one of the two conditions.

12.3.3 Materials

12.3.3.1 Learning environment

Students in both conditions worked in an online learning environment that had a similar structure. The learning environments differed in the support they contained. In one condition students had the reflection tool, in the other case students did not have the reflection tool and were presented only with open-ended reflection questions after finishing their inquiry learning activity. The Go-Lab environment included other scaffolds such as the hypothesis scratchpad, observation tool and conclusion tool.

12.3.3.2 Virtual laboratories: Acid-Base Solutions and pH Scale: Basics

The Acid-Base Solutions and pH Scale: Basics virtual laboratories are online interactive simulations provided by the PhET project (http://phet.colorado.edu). Figure 12.1 and Figure 12.2 show how these virtual laboratories appear.

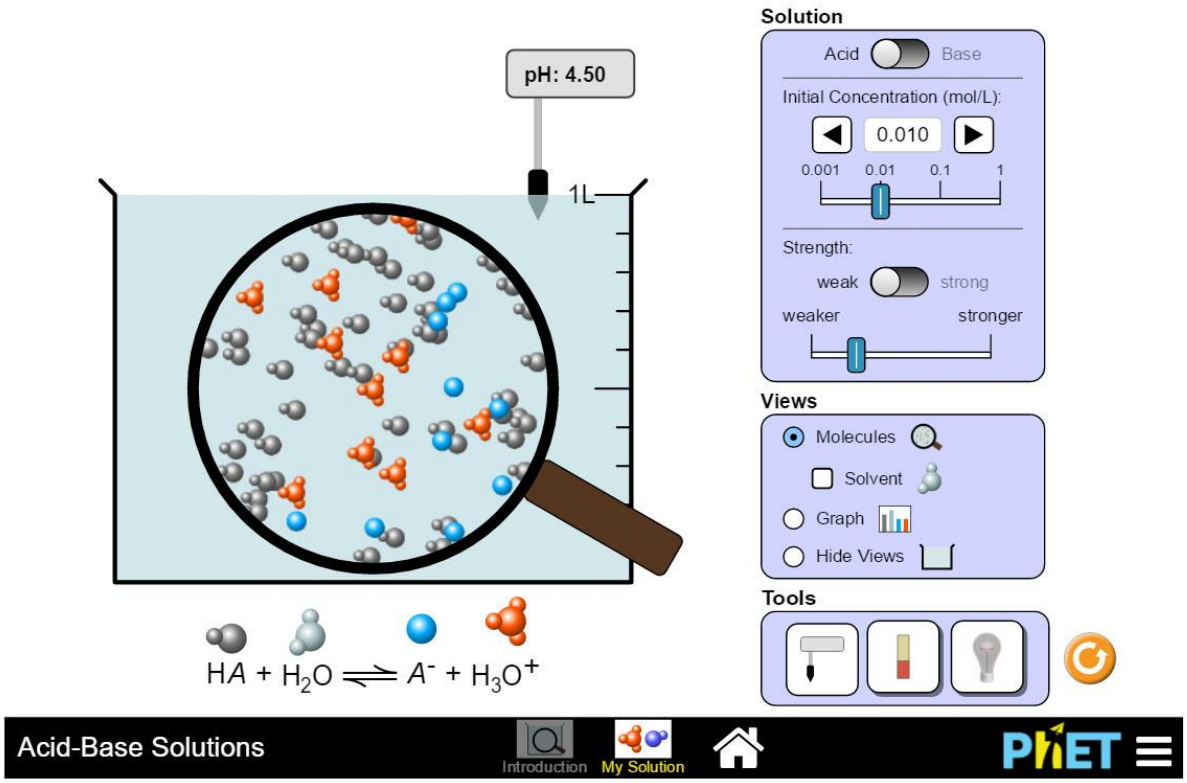


Figure 12.1 Screenshot of the Acid-Base Solutions virtual lab.

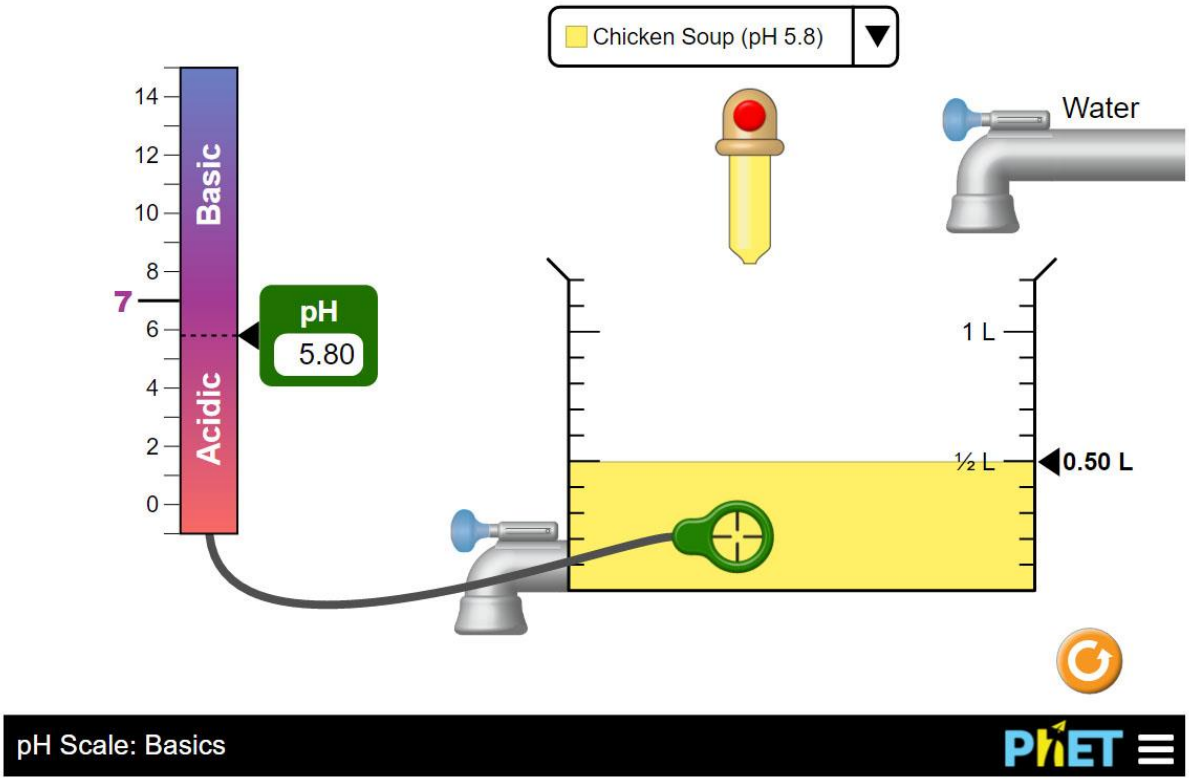


Figure 12.2 Screenshot of the pH Scale: Basics virtual lab.

PhET offers a library of science and math simulations that have been translated by volunteers into 76 languages, and in 2014 included over 75 million uses across more than 200 countries and territories. A valuable advantage of these two laboratories is that they have been created using HTML5 and are thus able to be used on tablet devices.

12.3.3.3 The Reflection Tool

The reflection tool gives feedback to students about their use of time in an Inquiry Learning Space. The reflection tool displays the percentage of time a student spends in the various inquiry phases and compares this time to a norm set by the ILS owner. Students are prompted to reflect on their ILS use with open-ended questions. An example of the data presented by the reflection tool is presented in Figure 12.3. The question used in our study in the reflection tool was “What was the most difficult phase during your inquiry activities? Why?”.

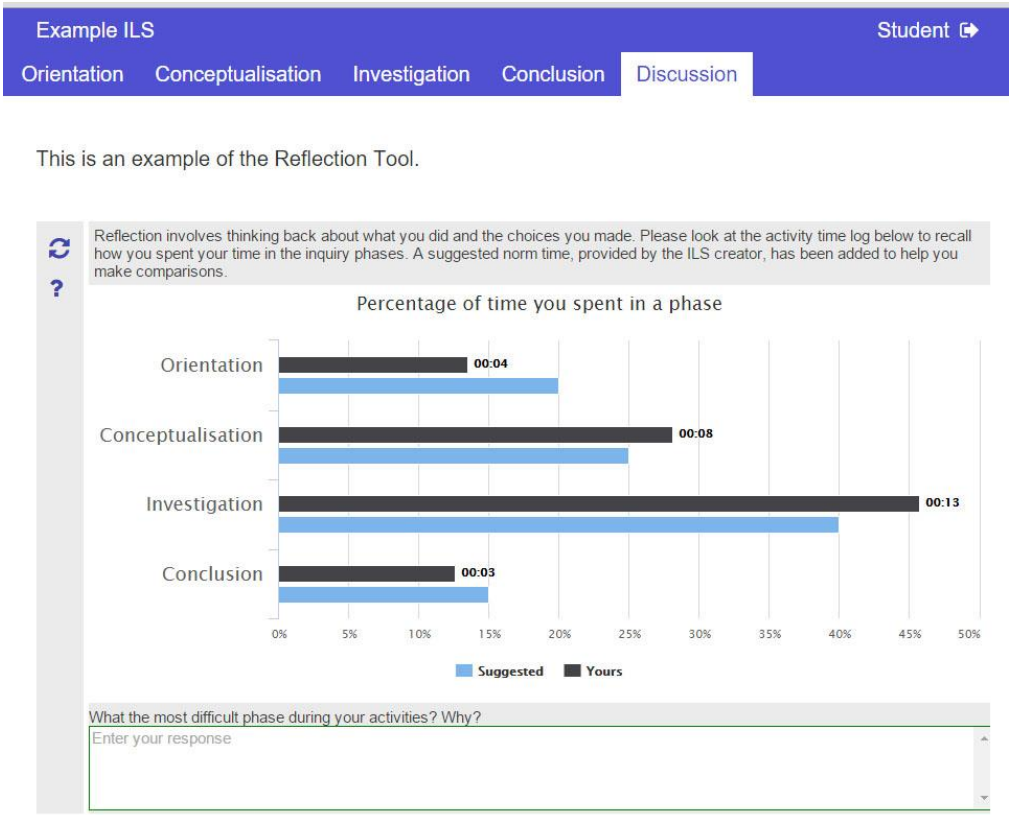


Figure 12.3. Screenshot of the reflection tool.

12.3.4 Assessment

Two aspects were assessed in students' reflections: content and level. The following three different contents were distinguished in students' reflection: technical, situational and sensitising (Leijen et al., 2012; Poldner et al, 2014). Reflections were categorized as technical if students pointed out technical difficulties that they had with the task. Reflections were categorized as situational if students pointed out difficulties that were related to the content of the task. And reflections were categorized as sensitising if students pointed out wider social, moral, ethical, or political aspects of the task.

Reflection levels were assessed based on the reflection levels developed by Poldner et al. (2014). Poldner et al. (2014) further elaborated the four reflection levels that were developed by Leijen et al (2012) and created the following five levels: description (descriptions of the difficulties that the student had), justification (rationale or logical explanation for the difficulties), critique (explanation and evaluation of the difficulties), dialogue (critical review of different solutions or alternative methods), and transfer (how the next action becomes different or better than the previous action was). Description is the lowest level and every higher level also contains information from lower levels (Leijen et al., 2012).

Two researchers assessed the reflection levels and contents and inter-rater reliability was calculated. In case of reflection levels Cohen's Kappa value was 0.715 and in case of reflection contents 0.6717.

12.4 Results and Discussion

Our study focused on whether the reflection tool helps to identify students' reflection content and their reflection levels.

The results addressing students' reflection content were analysed and compared between the two conditions (Condition 1 with the reflection tool and Condition 2 without the reflection tool). Figure 12.4 shows the distribution of students in the three different reflection content categories for the two conditions.

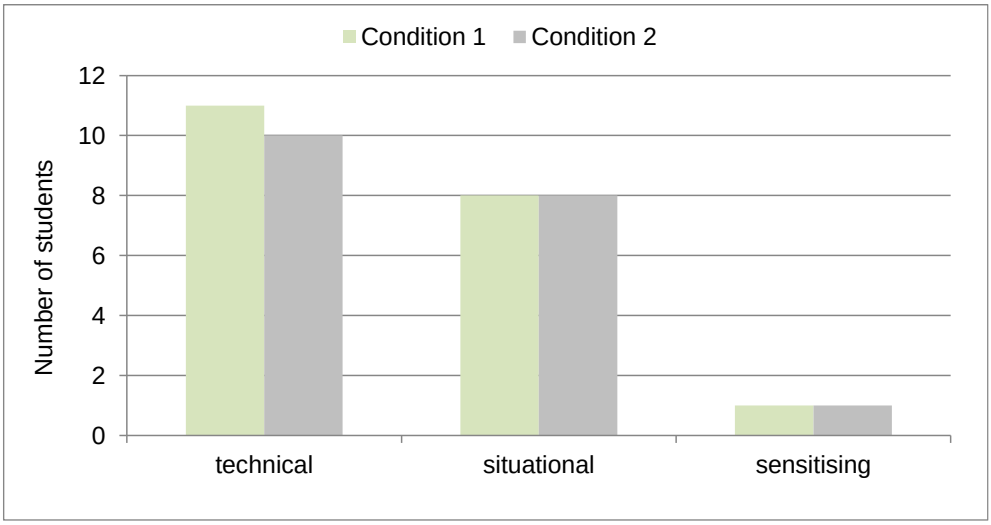


Figure 12.4. Number of students categorized by reflection content for Condition 1 (n=20) and Condition 2 (n=19).

As expected there was no statistically significant difference between the two conditions because the content category brings to attention factors that are not much related to time spent in inquiry phases. For example, a technical reflection content was "My iPad crashed all the time.", a situational reflection content was related to lack of general inquiry knowledge such as "I don't how to formulate hypotheses."; and a sensitising reflection content was related to personal issues such as "I feel that I am too tired to do this learning activity." In both conditions technical and situational reflection content was the most categorized responses. Sensitising reflection content was only grouped once in both conditions.

The results addressing students' reflection levels were analysed and compared between the two conditions. Figure 12.5 shows the distribution of students in the five different reflection level categories for the two conditions. In both conditions there were students that could be categorized into each of the five categories. In Condition 1 most students were categorized into justification reflection level whereas in Condition 2 most students were categorized into the transfer level (i.e. the highest level). If we compare both conditions then there are more students in the highest two levels for Condition 2 than Condition 1 and if we conduct a Mann-Whitney U-test then there is a statistically significant difference ($Z= -2.4304$; $p<0.05$) in favour of Condition 2 (the condition without the reflection tool). This difference may be explained from the researcher observations made during the intervention.

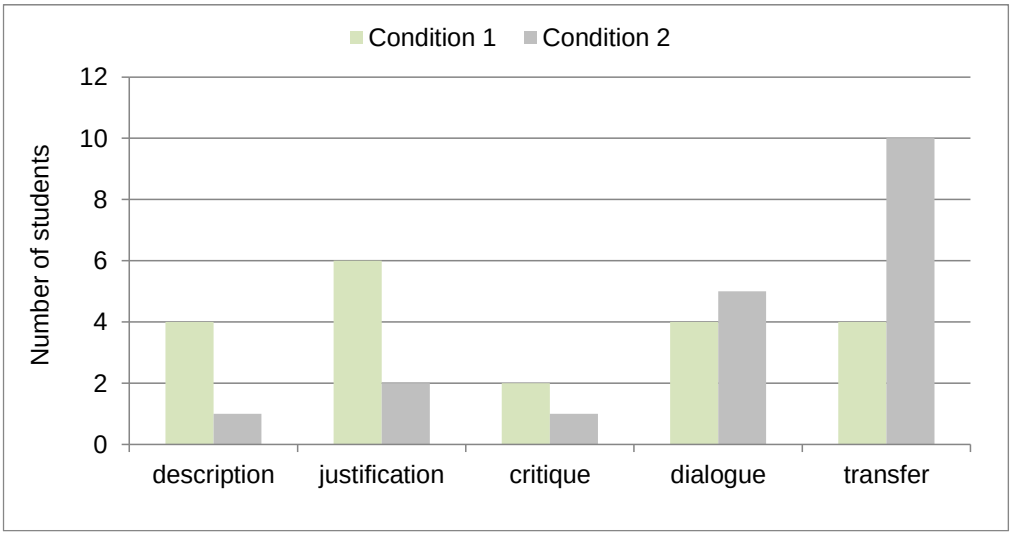


Figure 12.5. Number of students categorized by reflection level for Condition 1 (n=20) and Condition 2 (n=19).

Observations made by the researcher who led the intervention noted that students who were grouped into Condition 1 met more technical issues using the reflection tool than students who used the input box tool in Condition 2 to answer the reflection question. The reflection tool took longer to load on the tablets because of Wi-Fi internet connection speeds at the school. In addition, the size of the tablet screen makes text in the reflection tool appear smaller or harder to notice and requires scrolling. Student motivation to answer reflection questions was decreased according to the researcher and led to incomplete reflections.

12.5 Conclusion

The current study evaluated an initial version of the reflection tool using tablet computers after students finished solving an inquiry activity in a Go-Lab ILS. The results revealed a statistically significant difference in favour of the condition without the reflection tool. However, a majority of students' reflections (when including students from both conditions) could be categorized into the highest two reflection levels. This shows that students in general are capable of reflecting on their learning at a sophisticated level. Nevertheless, the Inquiry Learning Space can be further improved in order to avoid technical obstacles and demotivating issues that can affect students' reflections.

12.6 References

Benammar, K. (2004). *Conscious action through conscious thinking - reflection tools in experiential learning.* Public seminar. Amsterdam: Amsterdam University Press.

Bloom, B. S. (1974). Time and learning. *American Psychologist, 29*(9), 682–688.

Dewey, J. (1933). *How we think*. Buffalo, NY: Prometheus Books.

Karweit, N. & Slavin, R. E. (1982). Time-on-task: Issues of timing, sampling, and definition. *Journal of Educational Psychology, 74*(6), 844-851.

Kori, K., Pedaste, M., Leijen, Ä., & Mäeots, M. (2014). Supporting reflection in technology-enhanced learning. *Educational Research Review, 11*, 45-55.

Leijen, Ä., Valtna, K., Leijen, D. A., & Pedaste, M. (2012). How to determine the quality of students' reflections? *Studies in Higher Education, 37*(2), 203-217.

Mezirow, J. (1991). Transformative dimensions of adult learning. San Francisco, CA: Jossey-Bass.

Moon, J. A. (2004). *Reflection in learning and professional development*. New York: Routledge.

Poldner, E., Simons, P. R. J., Wijngaards, G., & Van der Schaaf, M. F. (2012). Quantitative content analysis procedures to analyse students' reflective essays: A methodological review of psychometric and edumetric aspects. *Educational Research Review, 7*(1), 19-37.

Schön, D. (1983). *The reflective practitioner: How professionals think in action*. New York: Basic Books.

Stallings, J. (1980). Allocated academic learning time revisited, or beyond time on task. *Educational Researcher, 9*(11), 11–16.

13 Overall conclusions from the studies with students

This section contains the conclusions of the conducted studies. First, the overall set-up of the studies is briefly described, followed by a short description of the studies and results per tool, and finally overall conclusions are presented.

13.1 Set-up and results studies

The studies in Section 1 of this deliverable focussed on the evaluation of tools that were designed for Go-Lab to guide students' inquiry learning processes. All studies had a pre-post-test between subjects design; in an number of cases students were assigned randomly to conditions, in other cases a quasi-experimental set-up was used.. First, students had to complete a conceptual knowledge and/or inquiry skills test, then they worked on an inquiry learning task, and after the intervention their conceptual knowledge and/or inquiry skills were measured again to determine the difference in knowledge and/or skills of before and after the intervention. The tests that were used to measure students' content knowledge depended on students' age and the domain of experimentation, and the inquiry skills tests were all based on the TIPS or TIPS II tests (see Deliverable D8.1). The TIPS tests consist of several constructs related to a specific skill; in the conducted studies only those constructs were used that were relevant for the specific study.

13.1.1 Concept map

Two studies were conducted to evaluate the concept map in terms of conceptual knowledge gain. Both studies focused on 13 year old students who had to learn about relative density. In the first study the Internet connection was lost; no significant results were found. In the second study four conditions were compared in which students 1) created a concept map before experimentation, 2) created a concept map after experimentation, 3) filled in concepts in a pre-structured concept map after experimentation, or 4) did not create a concept map at all. None of the provide concept maps in the first three conditions contained predefined terms, so students had to find the terms themselves. In all conditions, students learned significantly with the ILS, but there were no differences between conditions.

13.1.2 Hypothesis and questioning scratchpad

Three studies focused on the hypothesis and/or questioning scratchpad. In two of the conducted studies – one with 13 year olds and one with 17 year olds – two versions of the tools were compared in terms of their effect on students' inquiry skills. In one version students were provided with pre-defined terms they could use to formulate their research question/hypothesis, and in the other version they did not receive these terms. No learning effects were found in both studies, nor was there a difference between conditions. The third study, which was conducted in four countries with a total of 385 students, compared three different configurations of the hypothesis scratchpad. In the first version all concepts necessary for creating hypotheses were given in the second, version students were provided with a subset of the terms, and in the third version students had to create all the terms themselves. Effects were measured on knowledge and inquiry skills. ILSs and teste were offered in the local languages (Dutch, Greek, English, and Estonian). Data need to be further analysed in detail but an overall conclusion can be that offering more terms in the hypothesis scratchpad is beneficial for students in terms of knowledge and inquiry skills.

13.1.3 Experiment Design Tool

Four studies focused on the effect of the Experiment Design Tool (EDT) on students' learning. In one study students' conceptual knowledge gain was compared for students who designed and conducted experiments to answer 1) main research questions, 2) main research questions while guided by additional questions, 3) main research questions while guided by the same additional questions, and the EDT. Significant effects for learning were found in all conditions, but no differences were found between conditions. However, analyses on lower prior knowledge students showed significant differences in learning gain in favour of the EDT. In a second study two versions of the tool were compared regarding conceptual knowledge about buoyancy and Archimedes' principle. One version of the EDT had a more exploratory character and the other version had a structured CVS character in which students could vary just one thing at a time. A third condition was also included in the study in which no tool was used by students. Results showed significant learning effects in all conditions, but no overall difference between conditions. However, when analysing the data of students with lower prior knowledge, a significant difference was found for buoyancy in favour of the more exploratory version of the tool. In a third and fourth study students' gain of conceptual knowledge and inquiry skills were compared between students using the EDT and students not using the EDT. Participants in one study were eleven years old and in the other study students were fourteen years old. In all conditions students significantly gained conceptual knowledge, but only eleven year olds significantly increased in inquiry skill scores – in both conditions. Interestingly, in both studies significant differences were found in favour of the EDT, but for younger students this difference was found only for inquiry skills and not for conceptual knowledge gain, whereas for older students the difference was found for conceptual knowledge gain and not for inquiry skills. A fifth study that included the EDT, but that did not focus on the effect of the EDT, focused on the formulation of research questions for students to investigate. Students aged 13 and 15, with a middle or high level of secondary education, had to design and conduct experiments and received either direct instructions as to which variables they had to vary and control for, or they had to decide this for themselves. Results showed significant learning gains in all conditions. No difference was found between conditions for students of different ages. Level of education was found to be important for the guidance students should receive; students of the middle level of education benefited significantly from the direct instructions compared to the other condition, whereas students with a high level of education performed equally well in both conditions.

13.1.4 Experimental Error Tool

In one study the Experimental Error tool was evaluated with 16 year old students. Students were divide d in a group that used this tool and a group that didn't use it. The experimental error tool instructed students on different aspect of experimental error. This tool differs in this sense from the other tools that it teaches students content instead of that it merely supports an inquiry task. In this study no effect of the experimental error tool on students' inquiry skills could be found, which could possibly be attributed to the fact that the tool was not offered in the local language and contained too much information to be learned in the short time available.

13.1.5 Conclusion Tool

In a study with the conclusion tool in which students aged 16 participated, the conclusion tool was evaluated in terms of conceptual knowledge gains and increase of inquiry skills. In one condition students worked with the conclusion tool and in the other condition they did not work with the tool. Results showed a significant increase in conceptual knowledge

in both conditions, as well as a significant increase in the inquiry skills of identifying variables and hypotheses. Also, a significant difference was found regarding conceptual knowledge in favour of the conclusion tool. No significant differences were found for inquiry skills.

13.1.6 Data viewer

The data viewer was evaluated in a study with 16 year olds. Their conceptual knowledge gain as well as their inquiry skills were measured. Two versions of the tool were compared; in one version students' input from previous phases was automatically transferred into the data viewer, and in the other version students had to decide for themselves what they wanted to do with the previously collected data and how they wanted to organise and visualise these data in the data viewer. Results showed significant conceptual knowledge gains only for students who had to load the data themselves, and not for students whose data loaded automatically. The opposite was found for inquiry skills; students whose data loaded automatically showed significant increases in inquiry skills, whereas students who had to do this themselves did not.

13.1.7 Reflection tool

The reflection tool shows students their time spent in a phase of the inquiry cycle in relation to a suggested time. The students are asked, in an open text filed, to reflect upon these data. In this experiment an ILS containing this tool was compared with an ILS in which students were also asked to reflect using an open text field but without the information on their time spending in the phases. Students'' reflections were assessed on their content and level of sophistication. There was no difference in content of the reflection between both conditions. Concerning the level of the reflections, students scored overall in the higher levels of reflection and students who used the open input outperformed the students who used the reflection tool. However, internet connects were specifically hindering in the reflection tool condition, which may have influenced these students' performance.

13.2 Overall conclusion

The studies described in this part of the deliverable covered various tools from the Go-Lab set of tools, different configurations of ILSs, different age groups, different labs and domains, and different assessment methods. All studies were conducted "in vivo" meaning that we have always conducted the studies in real classes under realistic circumstances. This sometimes gave specific challenges, such as hampering internet connections, and made that research conditions, in terms of time allowed for the study or number of subjects, were not always optimal. We should also consider that in all of the studies these were most probably the students' (and teachers') first encounters with inquiry learning. Despite this some general conclusions can be drawn.

First, in all of the studies in which knowledge was measured (and where no internet issues appeared) we have seen a significant increase in scores on knowledge tests. There have been no comparisons with other, more traditional, approaches (this will be done in Y4 of the project) but in any case offering online labs makes that students learn about the domain. In the case of inquiry skills such an increase was not always measured, which can be explained from the fact that for those skills to develop properly, we need more time and a prolonged training of the skills.

For a specific set of tools we could find direct effects on students' acquisition of knowledge and inquiry skills. The conclusion tool, the hypothesis scratchpad, and the experiment design tool all showed specific effects in some of the studies. In some case a comparison

was made to a condition in which the tool was not offered, in other cases a comparison between a fully specified tool and a tool that was rather “empty” (such as a hypothesis scratchpad without pre-defined terms) was made. Studies on the hypothesis scratchpad showed that most probably offering pre-defined concepts in the tool was beneficial for learning compared to letting students configure these terms themselves. In the study on the concept map, no differences between including and not offering the concept map were found but in this case the concept map did not have any predefined terms in the pull down menu. If we extend the results of the studies with the hypothesis scratchpad to the concept map, we might expect better effects when these terms will be offered in the concept map’s pull down menu. So, a second conclusion n might be that tools often support students but that they might need to be filled with domain terms in order to create an effect.

A third conclusion might be that there is no “one size fits all” solution. Several of our studies (especially the ones on the experiment design tool (EDT)) show that tools are specifically effective (differentially for knowledge and inquiry skills) for students with lower prior knowledge or for younger students. The studies with the EDT also show that in these cases there is an interaction with the difficulty of the domain involved, effectiveness of a tool for the students who need this might be more distinct when the domain gets more difficult.

Finally, also the specific configuration of a tool might matter. The studies with the data viewer, for example showed, that when the data viewer automatically incorporated data an improvement of inquiry skills was reached, whereas students who imported data themselves gained better conceptual knowledge.

When deciding which tools to include in an ILS and how these tools should be configured, it is important to consider students’ age, level of education, and prior knowledge, as well as the difficulty of the domain, and whether the goal is for students to gain conceptual knowledge, to acquire inquiry skills or both.

Part 2: Teacher evaluation

14 Teacher evaluation

14.1 Introduction

As it was described in D8.1 (Teacher evaluation part), two questionnaires, one pre- and one post- have been developed in order to collect evaluation data throughout the Go-Lab large-scale pilot activities. Teachers, under the guidance of the NCs, have been instructed to complete the pre-questionnaire once, before they started their Go-Lab implementations, in order to capture their unspoiled views. The post-questionnaire on the other hand, had to be filled in after each Go-Lab implementation so if for example a teacher had used one laboratory and one ILS, then she had to fill it in twice. The collection of the pre-and post-questionnaire data has been concluded at the end of August 2015 and the analysis of the collected data can be found below.

Pre-questionnaires have been available in the majority of languages in online form i.e. https://www.surveymonkey.com/r/go_lab_pp2_teachers_pre_en same as the post-questionnaires i.e. https://www.surveymonkey.com/r/go_lab_pp2_teachers_post_en .

14.2 Data sample

This report covers the questionnaires' responses up to the end of July 2015. In all cases the questionnaires have been filled in online by the participating teachers.

About 600 Pilot teachers[2] were expected to fill in the pre and post questionnaires during Pilot Phase B. Finally, a total of 294 unique pre and 138 post questionnaires have been completed. Research (Holtom, 2008) shows that a questionnaire response rate of 50% of participants constitutes a success, with these percentage diminishing as time from the actual activities passes (and highly influenced by whether the questionnaire is filled in right before/after the activity or at a later stage).

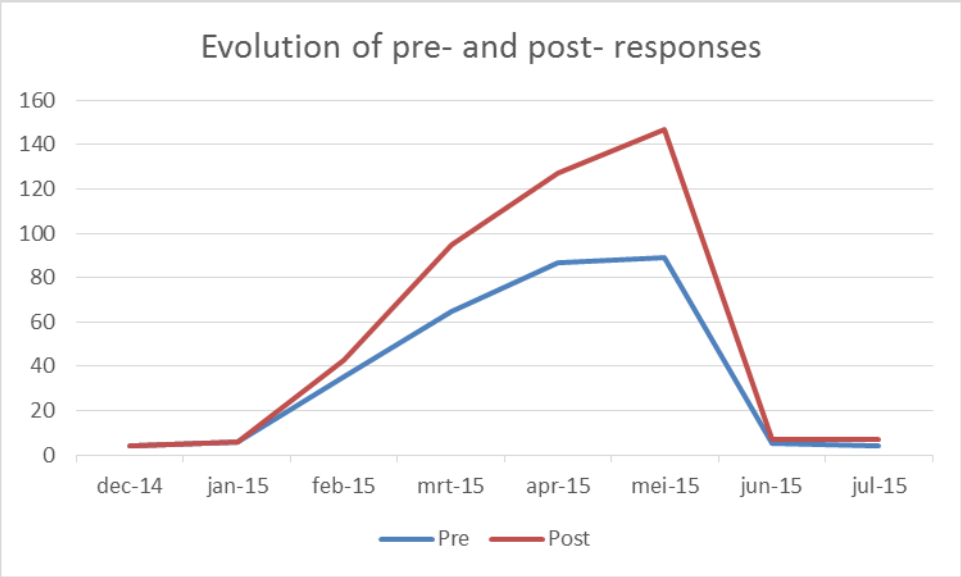


Figure 14.1. Evolution of pre- and post- responses.

The evolution of questionnaire responses along Pilot Phase B is shown in Figure 14.1. We show in this figure the responses to the questionnaires received at different moments after the launch of Pilot Phase B. The majority of pre- responses have been received between

[2] WP6 records a number of approximately 1000 teachers using the platform.

February and March 2015, thus at the beginning of the Pilot, while the majority of post-questionnaires have been received between June and July 2015 which coincides with the end of the school year in most countries.

The analysis of the pre data, Section 14.3, includes all 293 collected pre questionnaires. In order to facilitate the analysis and secure the reliability of the drawn conclusions though and during the analysis of the post data, Section 14.4, only the pre & post data of teachers who have filled in both questionnaires, has been taken into account. In this way we can ensure that drawn results are reliable and provide us with a real insight on teachers' attitudes. Consequently, in Section 14.3 we are analysing all received pre data while in section 11.3 post data and comparison to pre data is taking place only for the feedback of the 130 teachers have responded to both pre and post questionnaires.

14.3 Pre-questionnaire evaluation data

A total of 343 questionnaires has been collected between December 2014 and July 2015. After cleaning the data and removing duplications[3] the number of questionnaires came down to 293. This number corresponds to unique teachers and their breakdown per country can be seen below in Table 14.1

Table 14.1 Go-Lab Pilot phase B pre-questionnaire

Country	No of completed pre-questionnaires
Austria	20
Belgium	3
Bulgaria	1
Croatia	13
Cyprus	31
Estonia	19
France	4
Germany	16
Greece	32
Hungarian	1
Ireland	2
Italy	29
Poland	2
Portugal	10
Romania	2
Spain	93

[3] In the case of duplications, the oldest dated and most complete questionnaire has been kept while the most recent one has been removed and not taken into account for this analysis.

UK	15

It is worth mentioning at this point that in Bulgaria and Romania, due to curriculum changes, the implementation activities have started in April 2015 so more data and a more precise analysis will become available in October 2016. No questionnaires have been filled in in Switzerland and the Netherlands due to lack of registered Pilot teachers in those specific countries but the necessary actions have been taken in order to resolve this issue for the upcoming Pilot phase C and collect data from these groups as well.

After matching the pre data with the post received data, a total of 130 responses have been taken into account for the analysis and comparison with the post data in section 14.4.

While overall a significant amount of data was collected, the amount per country is limited, and therefore no country analysis will be done. In this way, we can avoid wrong assumptions that can lead to inaccurate generalisations and conclusions.

14.3.1 How do teachers intend to use Go-Lab?

In Figure 14.2 we can see teachers' replies regarding how they intend to use Go-Lab. As we can see, the majority or teachers (48%) entered Go-Lab with the intention to discover and use online laboratories. A smaller percentage (28%) was already willing to use complete ILSs that they could adapt and use with their classes while the smallest percentage (24%) was committed to implementing their own ILSs. The pattern is similar to the 1% rule in Internet, which states that 90% of the participants of a community only view content, 9% of the participants edit content, and 1% of the participants actively create new content[4]. In the Go-Lab case, though, the teachers participating appear to be more active than usual. It would be interesting to compare these results with the next cycle as teachers become more familiar with the platform, tools and ILSs.

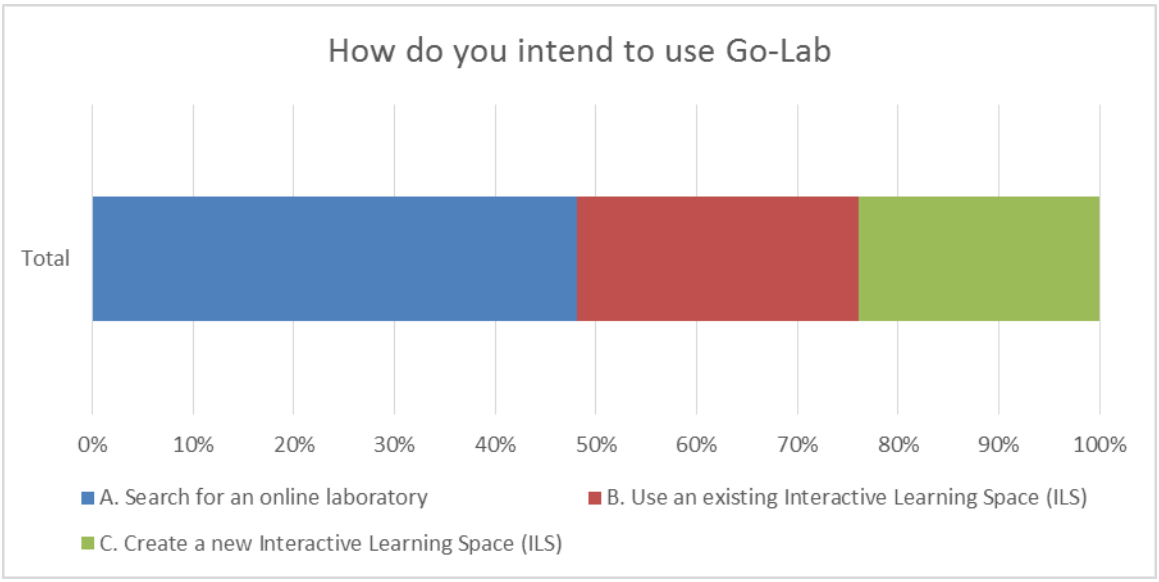


Figure 14.2. How do teachers intend to use Go-Lab?

14.3.2 How experienced are teachers regarding their science teaching

Figure 14.3 provides us information on Go-Lab teachers' experience regarding their science teaching. Teachers were asked to state whether they agreed or not with a variety of statements targeting the adequacy of their science knowledge in relation to the subjects

[4] See http://en.wikipedia.org/wiki/1%25_rule_%28Internet_culture%29

they teach, their experience in designing lesson plans that include technology, their ICT knowledge, their capacity to choose and adapt the technologies that fit better to their pedagogical needs, their student assessment methods and finally the use of different teaching methods.

In Figure 14.3 we can see that the majority of Go-Lab teachers, 92%, were confident that they have sufficient knowledge about science (i.e. Biology, Physics etc.) which allows them to teach their science classes (Figure 14.3, 4.2). This understanding teachers have is usually a working understanding of the issues involved but are rarely, as we can see in the interviews below, explicit (Samarapungavan, 1992).Their ability to assess students' performance along with their capacity to choose the appropriate technology to enhance their science classes score also quite high on their level of experience (Figure 14.3, 4.3 & 4.8).

When it comes to how Go-Lab teachers felt regarding the training they have received to use different technologies for learning science, 66% believe that it is adequate (Figure 14.3, 4.6). Consequently, a 34% feels that more training is needed in order to get them up to speed with the newest technologies. A similar percentage of Go-Lab teachers feel that they are not experienced enough when it comes to the creation of science lessons so it is safe to assume that the satisfaction of these two needs is strongly related to teachers' interest in Go-Lab (Figure 14.3, 4.9).

Overall and after taking into account teachers' responses regarding their experience in using ICT during science teaching, their capacity to create lesson plans and their ability to choose appropriate teaching approaches, we can assume that the average Go-Lab teacher is pedagogically experienced but with room of improvement in relation to technical skills and the knowledge of new technologies.

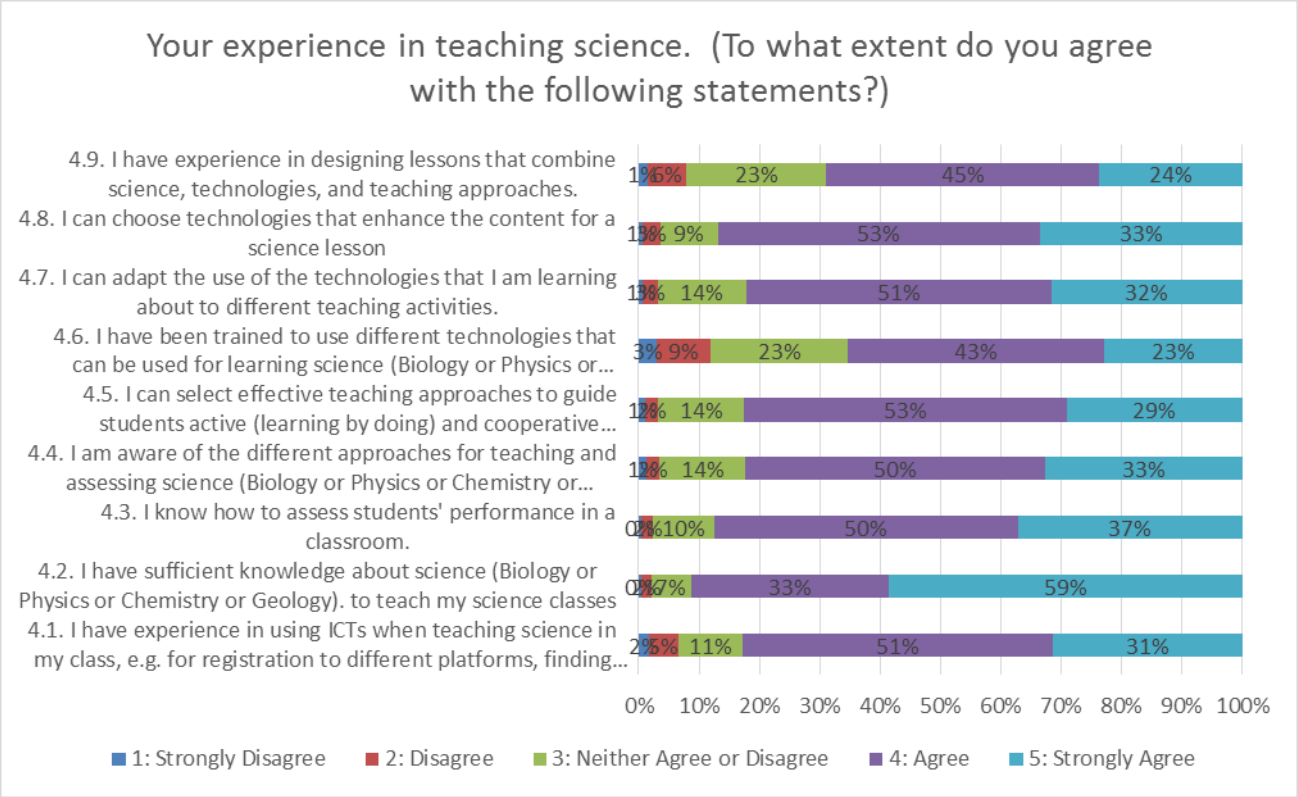


Figure 14.3. Teachers' science teaching experience.

14.3.3 How do teachers rate their technological skills

In Figure 14.4 we can see how teachers have rated their knowledge and familiarity regarding the selected technological skills.

77% of teachers were very familiar with the use of online laboratories and simulations while they were also confident in passing this knowledge to others. This result matches findings from 2013 according to which 70% of teachers around Europe have at least 5 years of ICT experience (Wastiau, Blamire, Kearney, Quittre, Van de Gaer, & Monseur 2013). A slightly smaller but still large percentage of 74% have also being using online repositories and educational material widely. Lastly, only 29% of the teachers knew how to use authoring tools i.e. Adobe Dreamweaver, Coffeecup etc.

Combining these results with the Go-Lab use indications we have seen in Figure 14.2 we can come to a number of conclusions:

- Go-Lab teachers seemed to be particularly familiar with the use of online repositories and educational resources which made them automatically much more responsive when it comes to the use of Go-Lab repository as a source of discovering online laboratories and related educational material.
- Only a small group of teachers (29%) was familiar with the use of authoring tools. Consequently, only 24% of teachers appeared to be interested in using the Go-Lab authoring tool and develop their own ILSs.

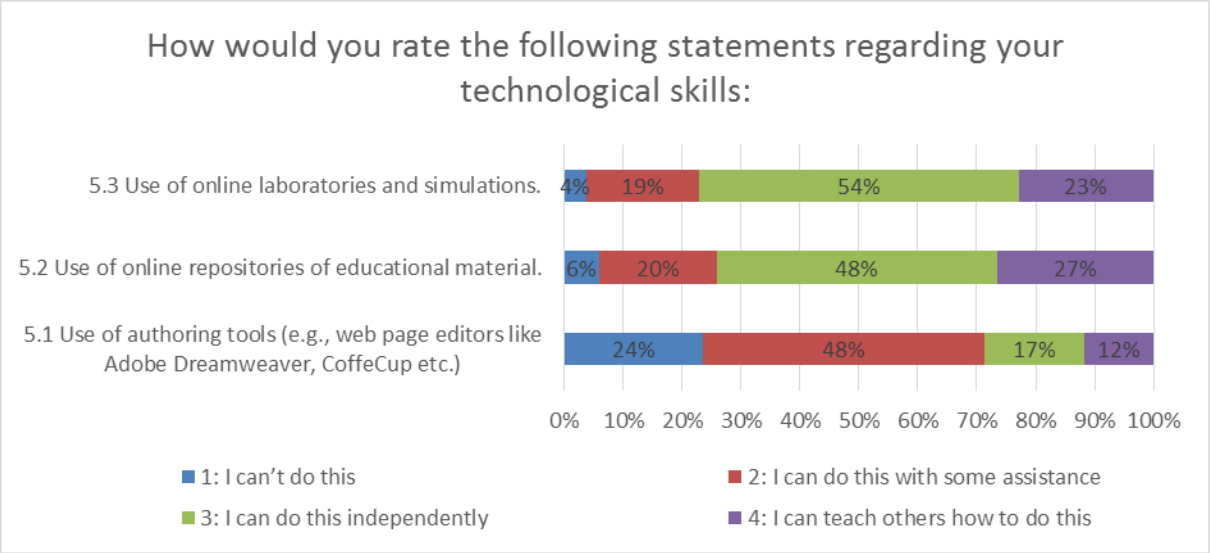


Figure 14.4. Teachers' technological skills.

14.3.4 How do teachers understand IBSE scenarios

Figure 14.5 presents how Go-Lab teachers have understood and rated a selection of teaching scenarios in relation to how they promote and develop students' IBSE skills.

The scenario that includes the analysis of laboratory collected data, the generation of conclusions and the final presentation of these conclusions to the targeted audience, has been selected by 89% of the teachers as the most representative in relation to IBSE. A few explanatory quotes that reveal teachers' views and understanding can be found below:

- "Student has to go through the experimentation process by him/herself or in a group and conduct the results from those results. He/she also need to collaborate

with his/her colleagues, understand different points of view and discuss the results and the conclusions", Croatian teacher

- "Presentation and interpretation of data is an important scientific skill. As long the activity that gave them the data is open and involves inquiry then the presentation forms part of inquiry based learning", Spanish teacher
- "Presentation of a subject to any audience requires interpretation and the adaptation of the acquired knowledge to the understanding abilities of a specific audience. Therefore, it is itself inquiry", Greek teacher

The scenarios of "Having students use graphics on the Internet to explain how gas molecules move" and "Having students to complete a lab activity or experimentation" are also perceived as good opportunities for developing IBSE skills.

A small but significant group of teachers (36%) do not seem to be convinced regarding the IBSE value of a class discussion about the arrangement of the periodic table. According to a Spanish teacher *"In this situation students are following a procedure prepared by the teacher. In my opinion true inquiry based learning should be student led with them deciding the procedure and what they want to investigate."* What is interesting in this scenario is that although it is not an obvious IBSE scenario, 64% of the teachers declare that with the right guidance and input from the teacher, the class discussion can be converted into an IBSE activity.

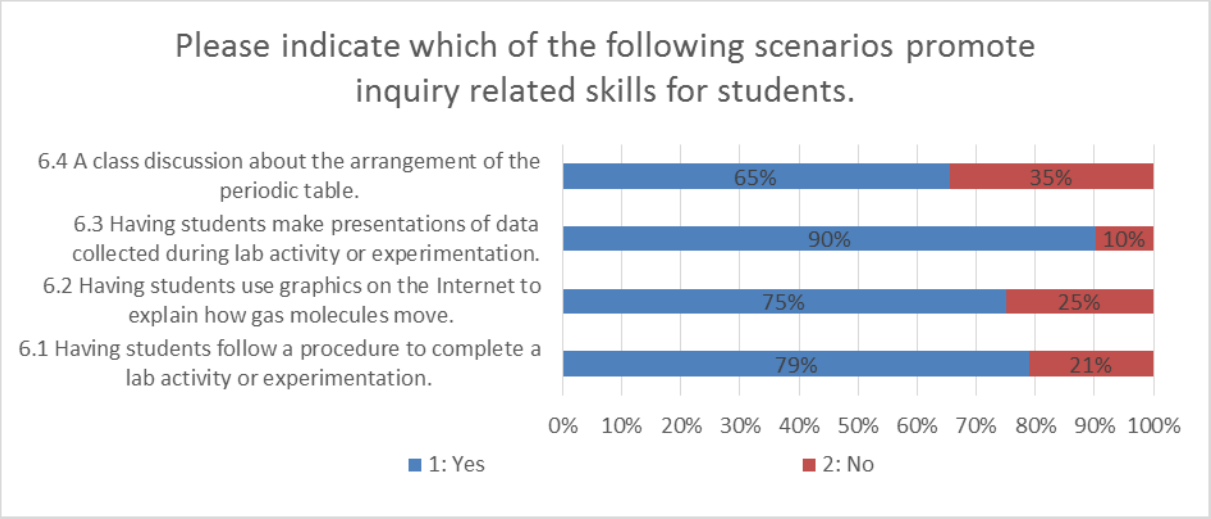


Figure 14.5. Teachers understanding of IBSE scenarios.

Figure 14.6 shows Go-Lab teachers' agreement/disagreement to a selection of statements related to the implementation of inquiry based science.

In Figure 14.6 we can see clearly that 84% of the teachers are committed to continually find better ways to teach inquiry based science which in combination to their interest in using online laboratories and finding educational material as we have seen earlier, form the basis of their involvement and interest in the project. 65% of the teachers believe that they know how to explain students to conduct inquiry based science while only 5% feel that despite their efforts they will never be able to conduct inquiry based activities as good as other approaches.

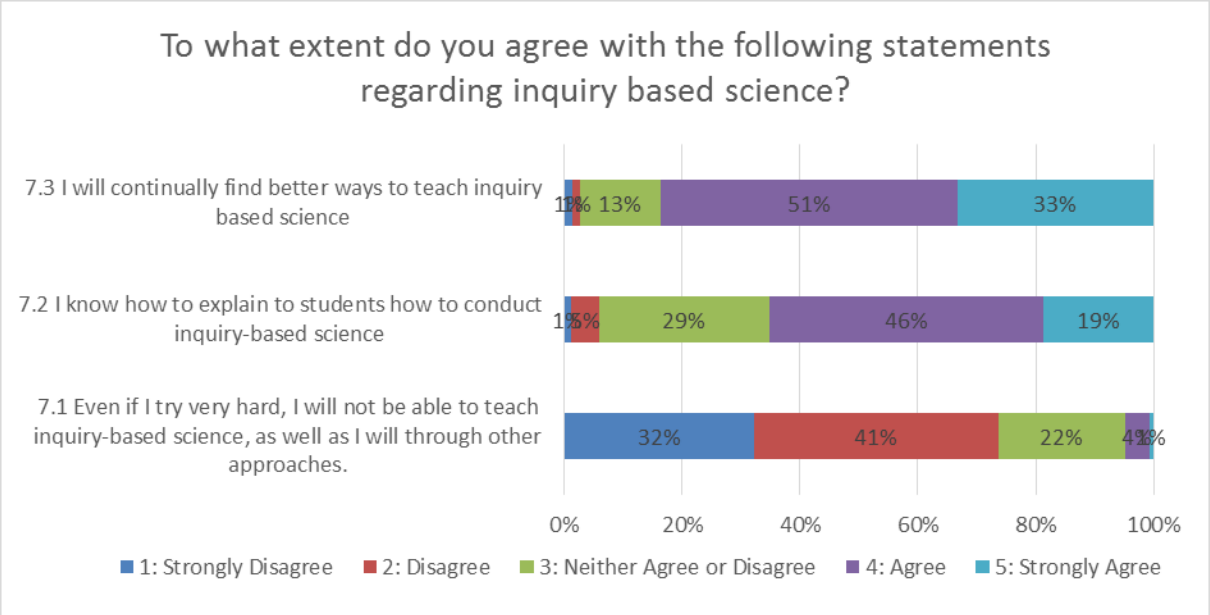


Figure 14.6. Teachers' views on inquiry based science.

14.3.5 How often do teachers intend to use specific parts of Go-Lab

Figure 14.7 provides us with information regarding how frequently Go-Lab teachers intended to use the basic three components of Go-Lab (authoring facility, repository, ILSs).

More than 70% of the Go-Lab teachers expressed their intention to use ILSs on a weekly/monthly basis, with 64% planning to do the same with the Go-Lab repository. The use of the authoring tools came third in teachers' preference with 58% of the teachers planning to use it on a monthly/less than monthly basis.

Despite teachers' tendency to use existing ILSs/repository and their reservations when it came to the use of authoring tools, there was still significant interest in their use.

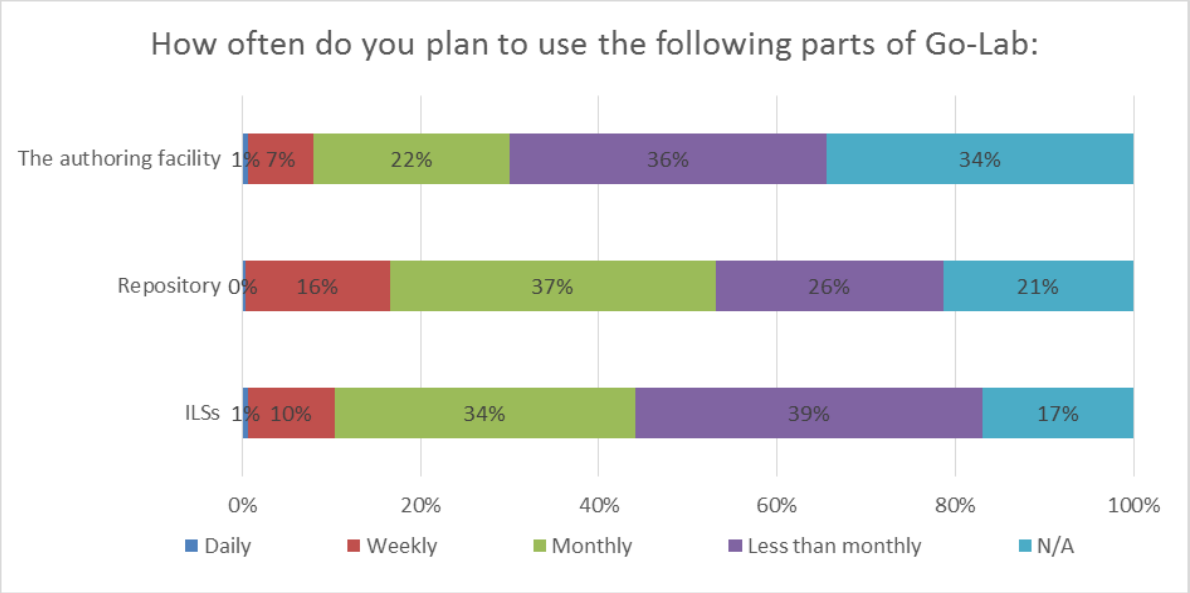


Figure 14.7. Frequency of use for Go-Lab parts.

14.4 Post-questionnaire evaluation data

A total of 160 post questionnaires has been collected between December 2014 and September 2015. After cleaning the data and removing duplications and incomplete entries, the number of questionnaires came down to 138. The breakdown of submitted questionnaires per country can be seen below in Table 3.2.

Table 14.2. Go-Lab Pilot phase B post-questionnaire

Country	No of completed pre-questionnaires
Austria	4
Belgium	2
Bulgaria	1
Croatia	12
Cyprus	12
Estonia	1
France	3
Germany	13
Greece	8
Hungarian	1
Ireland	2
Italy	12
Netherlands	1
Poland	1
Portugal	20
Romania	2
Spain	44
UK	15

This number correspond to unique teachers but as explained at the beginning, in section 11.3 we are analysing and comparing pre & post data which has been received by the same teachers. As a result and after matching teachers' codes, we arrived at a total of 130 teachers who have responded to both pre and post questionnaires.

Moreover, as one can see and as it was explained in section 14.2, the number of post questionnaires is significantly lower that the pre-questionnaires. Although this kind of behaviour is quite common in evaluation, it also indicates that a different approach needs to be taken in order to motivate teachers to fully participate in the validation process. Incentives, rewards, connection to certification are just some of the suggestions and possible solutions that will be considered, discussed extensively with the National coordinators and adopted in Pilot Phase C.

14.4.1 How have teachers used Go-Lab?

Figure 14.8 shows that 51% of the teachers have created their own ILS while 25% of the teachers have used an existing ILS. A small percentage of 25% have used Golabz only for finding an online laboratory.

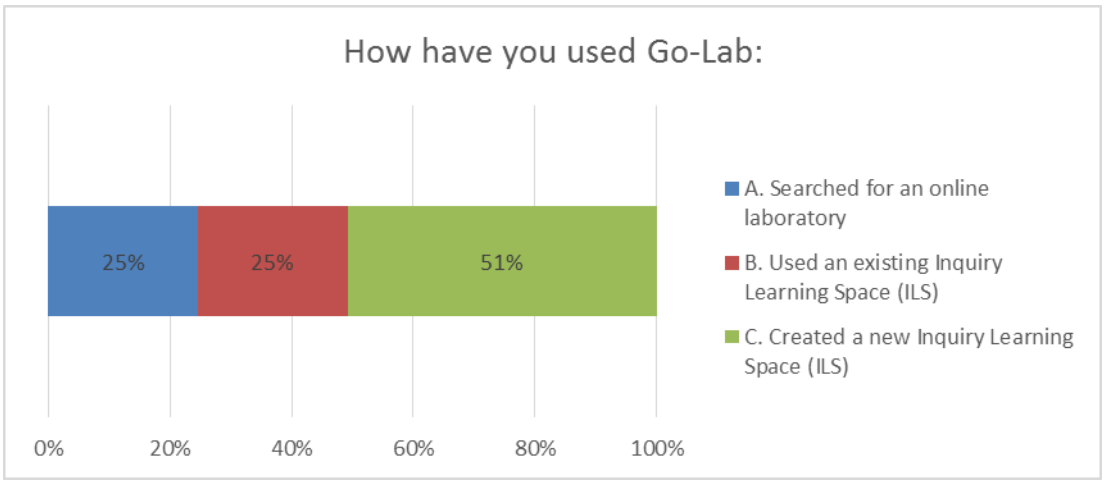


Figure 14.8. How have teachers used Go-Lab?

A comparison between Figure 14.8 and Figure 14.2 reveals that when it comes to the use of Go-Lab, teachers have exceeded their own expectations. In Figure 14.2 48% of the teachers were planning to use Golabz for finding online laboratories and for using existing laboratories with a small number of teachers declaring their intention to create their own ILS. Figure 14.8 shows a totally reversed use of Go-Lab with the majority of teachers, 51%, moving to a more active use of Go-Lab and creating their own ILS. As a result we now have less teachers, only a 25%, that are interested in just finding an online laboratory.

14.4.2 How experienced are teachers in teaching science?

Figure 14.9 provides us with an insight on teachers' experience in teaching science topics. After using Graasp, more than 78% of the teachers feel experienced enough to design lesson plans combining different elements and choosing technologies that enhance their lessons. Knowledge of various teaching approaches, adequate knowledge of science topics and experience on the use of ICT are also significant factors related to teachers' experience.

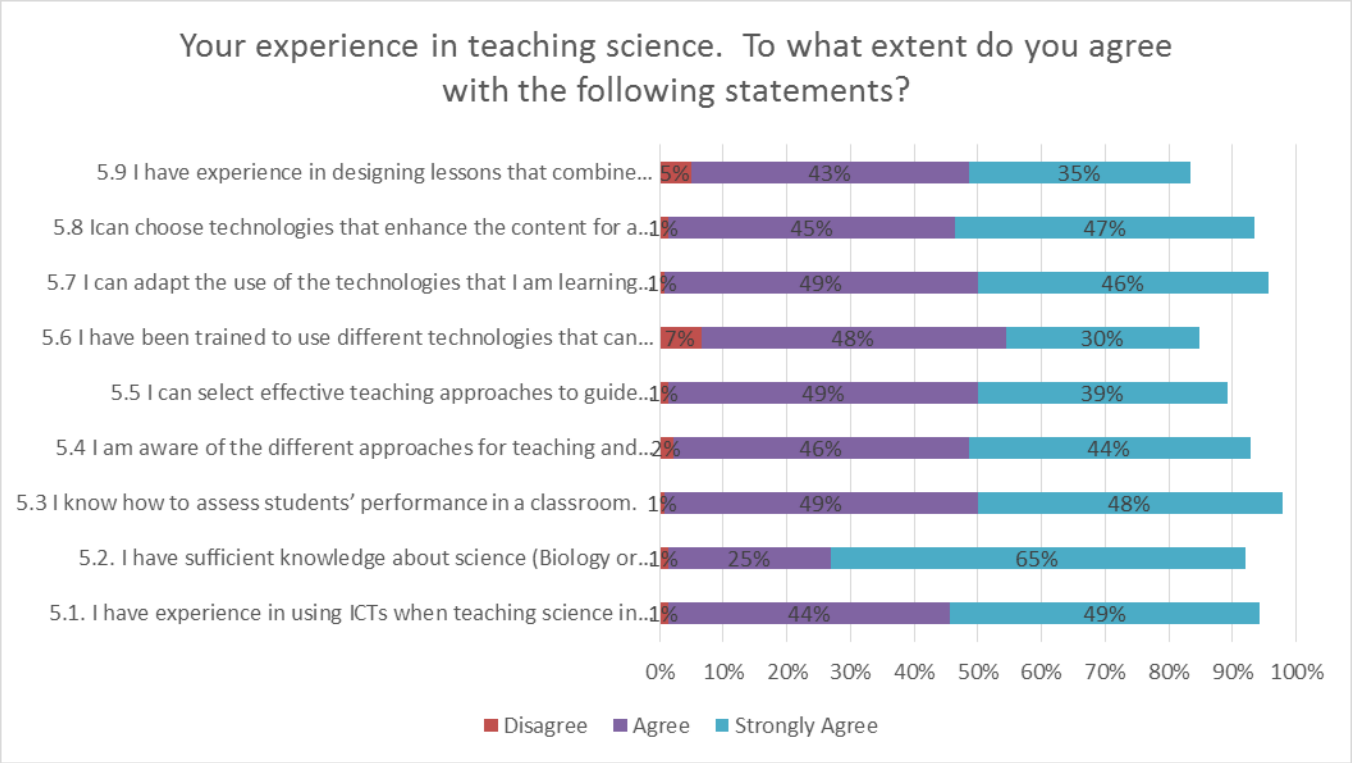


Figure 14.9. Teachers' experience in teaching science.

Comparing the above figures with Figure 14.3 we can see that the use of Go-Lab has strengthened teachers' skills when it comes to science teaching. 10-15% more teachers compared to the pre answers state that they are in a position to design lesson plans using online laboratories while a significant increase in the belief that they can choose and adapt technologies according to their needs is also noticed.

14.4.3 How do teachers rate their technological skills?

Teachers' answers regarding their experience in the use of different Go-Lab tools can be seen in Figure 14.10. According to it, more than 85% of the responders feel confident when it comes to the use of online laboratories and educational repositories. Teachers feel also strongly that they are in a position to not only use these tools independently but to also teach them to others. The use of authoring tools seems to be the exception though with 32% of the teachers stating that they need assistance to use them and 58% feeling confident to fully take advantage of them.

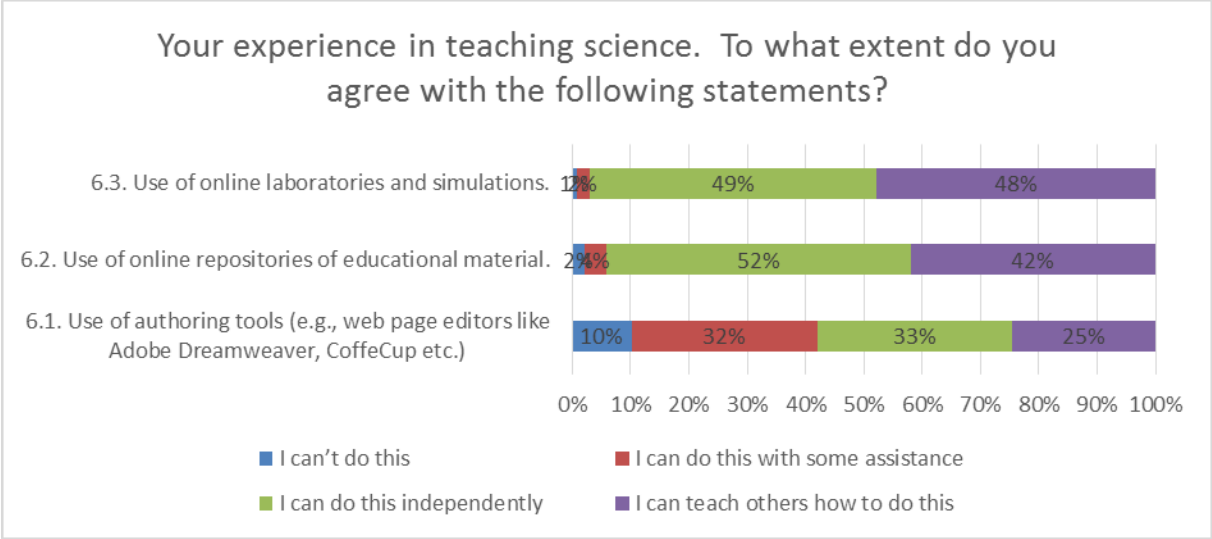


Figure 14.10. Teachers' experience in the use of online laboratories.

Looking back to teachers' pre answers, Figure 14.4, and after their exposure to the Go-Lab tools, we can see an increase in the percentage of teachers that feel capable of using online laboratories and the repository independently. The use of authoring tools still requires support but the percentage of teachers that need immediate assistance has decreased from 48% to 32%. This significant decrease proves that the use of Graasp and teachers' exposure to it, provided them with valuable skills that they are now capable of using within various authoring tools.

14.4.4 How do teachers understand IBSE scenarios?

In this question, teachers have been presented with 4 hypothetical class scenarios that they were invited to decide whether they promote the development of students inquiry skills or not. As one can see from teachers answers, all suggested scenarios can be used as opportunities to teach inquiry skills to students. The class discussion and having students following a procedure to complete a lab activity or experimentation are the ones collecting most of the criticism though with more than 30% of teachers rejecting their contribution to the acquisition of inquiry related skills.

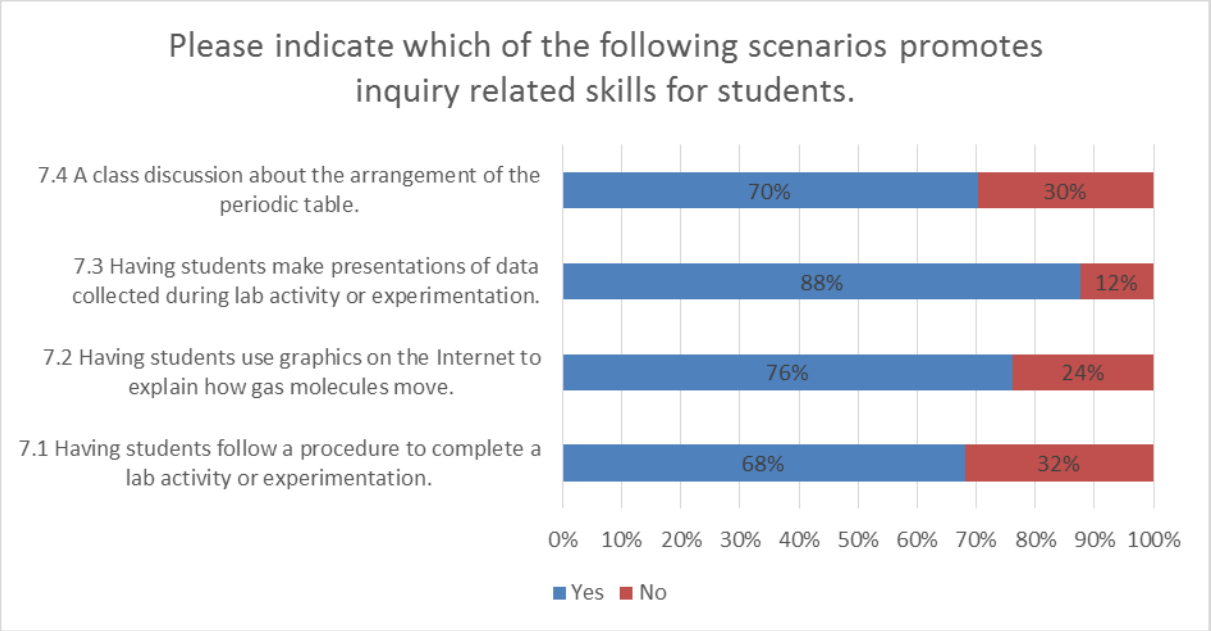


Figure 14.11. Inquiry based skills for students.

A comparison between Figure 14.11 and Figure 14.5, which describes teachers understanding of IBSE scenarios before teachers used Go-Lab, shows that for the first 3 scenarios, teachers' views remained quite stable before and after the use of Go-Lab. When it comes to the last scenario though we can see a change of opinion for more than 10% of the responders. After using Go-Lab, 10% of the teachers changed their mind and decided that having students following a procedure to complete a lab activity or experimentation does not promote their inquiry skills. More explanations about this can be found it teachers' quotes below:

- "Students should explore and research on their own solutions to the problems not always with given exact procedure", Croatian teacher
- "This isn't an inquiry related skill, since students just perform a recipe. They don't need to understand what they are doing", German teacher
- "When students follow a given procedure they do not develop inquiry related skills. First they have to try different options in order to find out and learn by themselves.", Spanish teacher

14.4.5 Teachers' views on inquiry based science

Figure 14.12 shows how teachers feel regarding their own inquiry based skills. As we can see, 84% state that after using Go-Lab they are committed to continue finding better ways to teach inquiry based science with 81% feeling confident enough to explain to their students how to conduct inquiry based science. As a result, 89% of the responders believe that they are in a position to teach inquiry-based science, as well as they can through other educational approaches.

Comparing these results with the ones on Figure 14.6, we can see that the use of Go-Lab had a positive impact on teachers' views regarding IBSE. 84% of the teachers feel encouraged to continually look for better ways to teach inquiry based science which is 20% higher than the 64% of teachers that replied positively in the pre questionnaire. Moreover 81% of teachers are now in a position to explain IBSE

to their students which again gives us an increase of approximately 20% compared to 65% that we had in the pre responses.

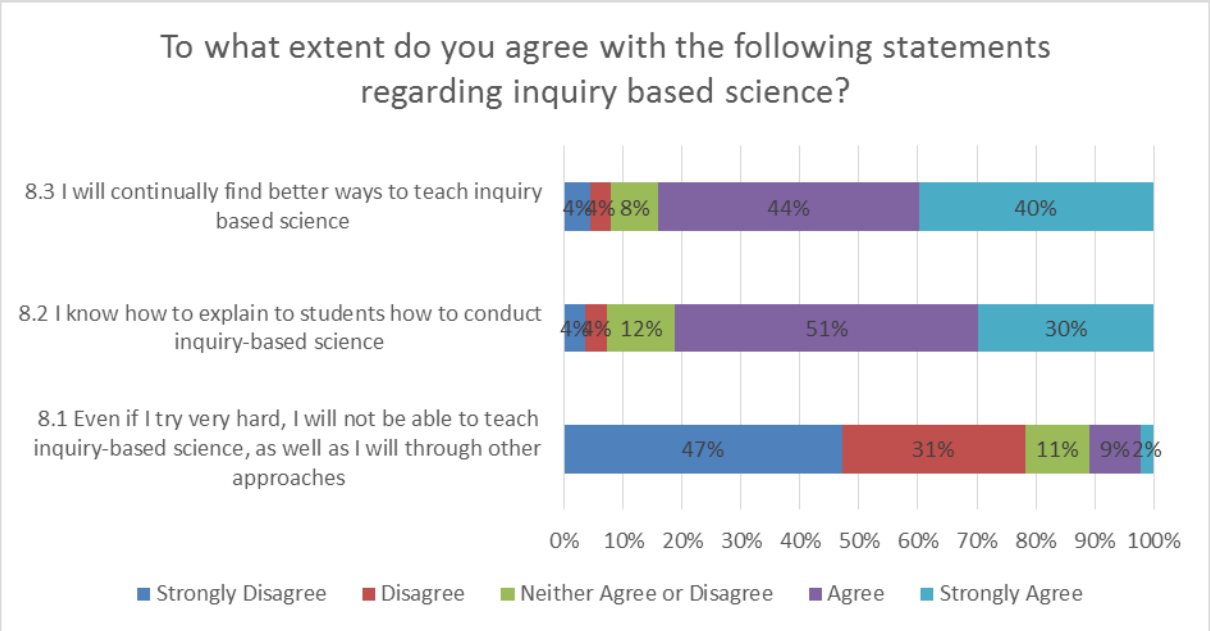


Figure 14.12. Teachers' views of inquiry based science.

14.4.6 How often do teachers use specific parts of Go-Lab

In Figure 14.13 we can see how often teachers use the different parts of Go-Lab. The online labs repository seems to be the most popular with 30% of teachers using it on a monthly basis. Existing ILSs are also popular with 29% of the responders using them monthly. When it comes to the authoring facility 67% state that they tend to use it less than monthly while 54% of the teachers recorded that same frequency for the ILSs and the repository. A small, but still significant, percentage of 15% of teachers are using all parts on a weekly basis.

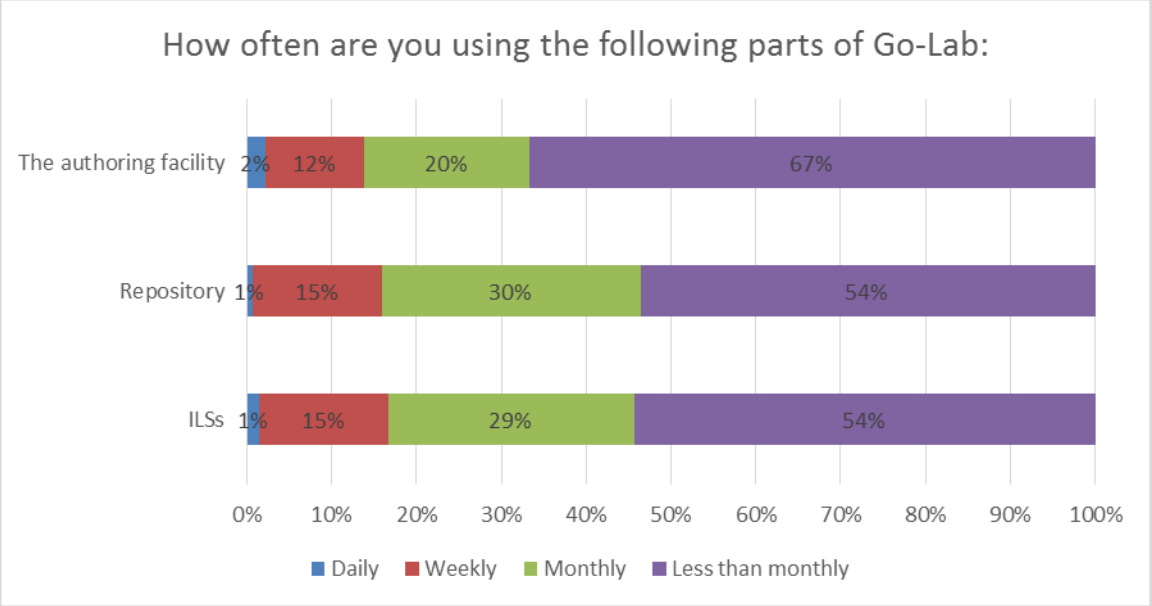


Figure 14.13. Teachers' use of Go-Lab.

Comparing the above recorded frequencies with teachers' intentions, as they can be seen in Figure 14.7, behaviours appear to be quite consistent. 10-16% of teachers intended to use the various Go-Lab parts on weekly basis, which is the actual case according to Figure 14.13. The majority of teachers, 30-40%, were planning to use the tools on a less than a month basis which has also been validated with percentages reaching up to 67% in the case of the authoring facility.

14.4.7 Go-Lab use and usability

When it comes to the usability of the different Go-Lab parts 77% of the teachers enjoyed using the Go-Lab portal while 65% found the navigation clear and understandable. 70% of the responders stated that they found the adaptation of existing ILSs easy while 62% could easily create their own ILS. Despite the variations, 84% of the teachers are willing to continue using Go-Lab in the future which shows a strong interest and commitment to further pursue the use of the portal. Figure 14.14 summarises the data.

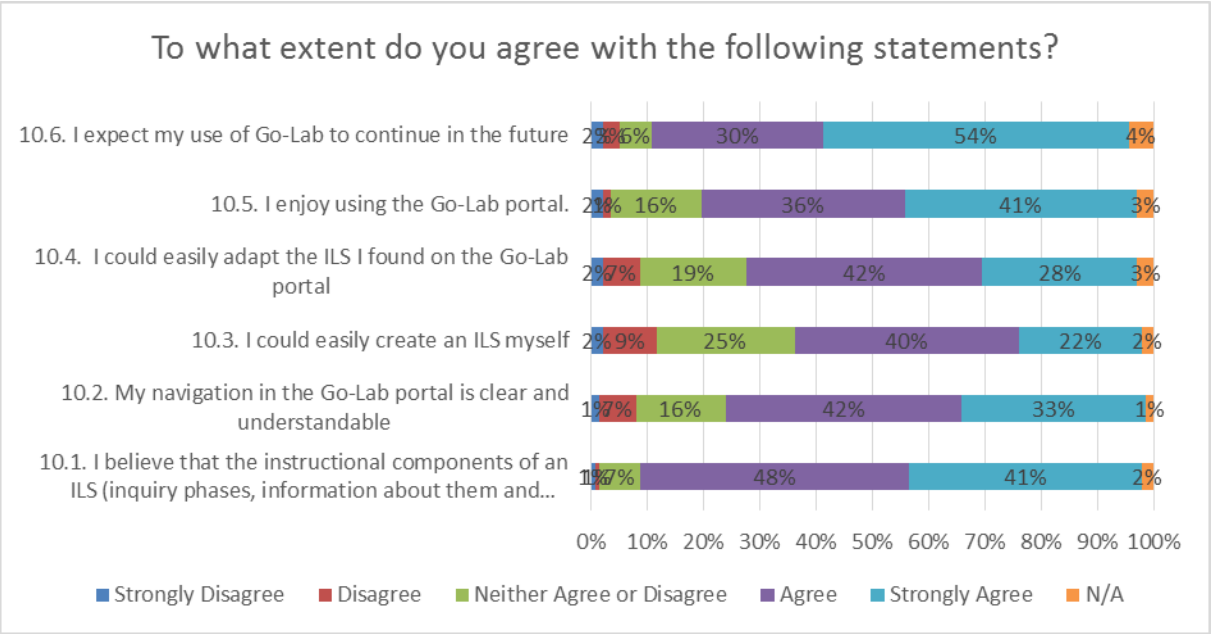


Figure 14.14. Go-Lab usability.

14.4.8 Which is the most useful component of the Go-Lab portal for science teachers and why?

In this open ended question, teachers had the opportunity to provide us with information regarding the most useful, based on their experience, component of Go-Lab. Most of the teachers avoided to provide a clear answer and expressed their satisfaction with all components. From the ones that they clearly expressed their preference the most useful components are:

- Repository of ILS
- Authoring tool

Some interesting comments and reasoning can be seen below:

- "Online labs. They are the more versatile component: they can be used and adapted in millions of ways and can fit to any class.", Portuguese teacher

- "Graasp offers the possibility of building content for all learning stages. The different steps to provide students who approach them with insights to inquiry concepts and teach them ask questions, a much needed and neglected skill.", Spanish teacher

14.4.9 Which is the least useful component of the Go-Lab portal for science teachers and why?

In this question the majority of teachers agreed that there is no Go-Lab component that it is not useful. Nevertheless, many teachers found in this question the opportunity to share a number of issues that they seem to prevent them from taking full advantage of the Go-Lab tools. The most common of these issues are listed below:

- Technical problems i.e. broken links, compatibility issues
- Lack of localised content i.e. need for ILSs in all languages
- Coverage of all subjects i.e. ILSs on mathematics are missing

Another interesting remark that is worth mentioning is that according to some teachers the existence of many Go-Lab sites (project site, lab repository, tutoring platform) is confusing. A better and more obvious connection among the different tools is essential in order to improve the flow of visitors and ensure a smooth user experience.

14.5 Conclusions

The main goal of the evaluation exercise was to monitor the impact of the use of Go-Lab tools in teachers' technical skills, IBSE knowledge, use and understanding of online laboratories. The conclusions of this section are spread over the previous sections and provide us with some first indications regarding the impact the use of Go-Lab tools has on teachers around Europe.

More specifically:

Teachers' profile: A close look at teachers teaching and technical skills reveals that a large percentage of the teachers that are interested in the use of online laboratories have quite developed pedagogical and technological skills. Thanks to the diversity of options that the Go-Lab tools offer, teachers with less experience have the possibility to start discovering the tools by using the repository and identifying labs, apps and existing ILSs that fit their needs.

Knowledge of IBSE: Most of the Go-Lab teachers have some knowledge of IBSE. The majority of teachers seem confident in teaching IBSE to their students and to design related activities. What is interesting though is that there is still a significant number of teachers that do not feel confident using IBSE and consider that they still lack skills in order to successfully apply it. Go-Lab is contributing to teachers understanding of IBSE but continuous support, good practices and training are needed in order support teachers interested in IBSE and help them fully develop their IBSE skills.

Teachers' technical skills: When it comes to their technical skills, teachers are quite confident to use online laboratories and repositories. The use of authoring tools though, is a big challenge for most teachers which also affects their intentions and ways the use the Go-Lab tools. At the end of Pilot phase B we can see a change in teachers' technical skills with a significant rise in the numbers of teachers that are stepping up and dare to use the authoring tool. The development of the tutoring platform, the various supportive materials that were made available in the course of the previous year and the training sessions that took place all around Europe, have definitely played their role and contributed to this

change. Despite the above and as stated in the ICT in education survey, support is still needed in order to encourage teachers with less advanced skills to grow and develop. Despite having access and positive attitudes towards implementing ICT in teaching and learning, teachers often find this difficult and require support – not only technical but also pedagogical (i.e., IBSE). Increasing the training provided by school staff and others to teachers of all disciplines should therefore be encouraged, including subject specific training on learning applications. (Wastiau, et al. 2013).

Use of authoring tools: The use of Go-Lab helped teachers to gain familiarity with the basic principles of authoring tools that they can use in producing their own ILS. As a result we can see a great shift regarding the use of Go-Lab. While in the pre questionnaire teachers were appearing reluctant to produce their own ILSs, in the post questionnaire we can see a large number of teachers producing ILSs and not only looking for online laboratories.

Need of training and support: The introduction of new tools in combination with IBSE, require training and support in order to provide teachers with the necessary background and skills.

Finally, regarding the response rate, additional actions need to be taken it order to motivate teachers to fully participate in the validation process. Incentives, rewards, connection to certification are just some of the suggestions and possible solutions that will be considered.

Part 3: Evaluation of the organisation

15 Organization evaluation

15.1 Introduction

Measuring the impact of Go-Lab on organizations is not based on controlled experiments but in the collection of qualitative data where we measure organizations' knowledge, attitudes etc. after they have been exposed to Go-Lab. For acquiring more in-depth qualitative information, interviews with teachers and other actors (i.e. heads of schools, parents) have been carried out while case studies will be carried out in collaboration with WP7 and during Pilot Phase C. The next sections provide more information in relation to the respective instruments.

15.2 Interviews

Interviews with teachers and other stakeholders i.e. headmasters have taken place between June and September 2015. The main aim of these interviews was to get an insight on how other school actors perceive Go-Lab and the use of online laboratories within the classroom.

15.2.1 Teachers' interviews

In collaboration with WP6 and WP7, a set of 12 teacher interviews took place. The sample teachers were all winners of the Go-Lab competition 2015 so their interviews took place online with the help of the tutoring platform. The main questions teachers were asked to answer can be seen below in Table 15.1.

Table 15.1. Go-Lab teachers' interviews: questions

Interview questions for Teachers		Answers
1.	Starting from the selection of the activity you carried out, based on what criteria did you select the specific activity? Have your students' preferences influenced your choice?	
2.	What was the highlight of your implementation activity? Was there a special moment, quote, incident that you and your students will remember?	
3.	How many activities like this would you like to do in the course of a year?	
4.	Have your students become more interested in Science as a result of this activity?	
5.	What were students' reactions during the implementation?	
6.	Have students enjoyed the activity in relation to Science?	
7.	Which part of the activity did they enjoy most/least?	
8.	Did the activity had an impact on students' collaboration or social skills?	
9.	What challenges did you face during the implementation of the specific activity?	

10.	How much time (real time) did the implementation of the activity take, including preparation?

The scripts and recordings of the interviews can be found below:

15.2.1.1 Daniela Leon, Italy

1.	**Name of teacher**	**Daniela Leone**
2.	Name of the school, Country	Istituto Comprensivo 9, Italy
3.	Type of school	-
4.	Students' age	9-10 years old
5.	ILS / online laboratory used	Meteoriti http://www.golabz.eu/spaces/meteoriti
1.	Starting from the selection of the activity you carried out, based on what criteria did you select the specific activity? Have your students' preferences influenced your choice?	My choice was following a visit to the planetarium near Bologna because we collaborated with this planetarium, taking classes and doing activities. In the planetarium there is a great collection of meteorites which students like very much and this visit aroused their curiosity about meteorites.
2.	What was the highlight of your implementation activity? Was there a special moment, quote, incident that you and your students will remember?	All the phases were interesting. Students were motivated and they were very good group and so it was very nice to work with them but anyway, I think the challenge for students was trying to approach as much as possible in real data, because in the simulation they could find as well as real data and science fiction facts, distinguish them and finding the likely events and not the unlikely events. It was a challenge for them.
3.	How many activities like this would you like to do in the course of a year?	
4.	Have your students become more interested in Science as a result of this activity?	My students are used to using digital resources for learning because they always use netbooks, PCs and interactive whiteboards and with the help of this ILS they could learn a more scientific method. The framework of the ILS forces them to be more systematic and to follow a procedure. They learned to work in a more scientific way.
5.	What were students' reactions during the implementation?	With my students there was one technical difficulty and one conceptual difficulty. The first one was that they were already used to Google Apps or Moodle, or other digital instruments to collect data and create apps. So they didn't like to use Go-Lab apps too much because it was new for them and they didn't comfortable with it. And in the end, most of them prefer to upload their files with other

		applications. Anyway the data was found and the results were found. Regarding the conceptual difficulty: my students, in the first phase of investigation about meteorites with Inquiry simulator, the first times they had fun with simulating science fiction effects, not realistic ones then they had to be more realistic and they had to try as much as possible a real impact choosing variables in order to obtain likely events. So this was a change for them and their interest was renewed by this.
6.	Have students enjoyed the activity in relation to Science?	
7.	Which part of the activity did they enjoy most/least?	
8.	Did the activity had an impact on students' collaboration or social skills?	I noticed that engagement in students went longer. They were already all very active students and interested in science but their willingness to complete all the activities in the whole phases made their engagement more long-lasting. So they wanted really to complete everything, to have everything done and especially, compare data from different groups which investigated different variables and compared the data found, discussing and approving hypothesis or validations. They got more engaged and for a longer time.
9.	What challenges did you face during the implementation of the specific activity?	Well, I'm not very experienced about meteorites but once the idea was chosen, I asked the help of an astronomer. The doctor, Marco Catalano who works at the Planetarium, he provided me with additional resources, websites with data about craters, if they really existed on Earth, which was very useful to compare the data with the simulations with the real data.
10.	How much time (real time) did the implementation of the activity take, including preparation?	We took about 3 months because the choice for the activity was made at the end of January. Then, in two months I prepared the ILS using the tutorials, the help of the lab community and then in April the students tested it and did their work. So at the end of April I collected all files of students and sent them to the conference. The longest part was creating the ILS. Students took 4-6 hours to use it in class, at school. We worked just as school, in small groups (2-3 pers.).

15.2.1.2 Filep Otilia, Hungary

1.	**Name of teacher**	**Filep Otilia**
2.	Name of the school, Country	Mészáros Lőrinc Körzeti Általános Iskola, Hungary
3.	Type of school	-
4.	Students' age	14
5.	ILS / online laboratory used	"Coloured chemistry" http://www.golabz.eu/spaces/sz%C3%ADnes-k%C3%A9mia
1.	Starting from the selection of the activity you carried out, based on what criteria did you select the specific activity? Have your students' preferences influenced your choice?	I chose this method to motivate my students because our school is in an area of Hungary where there are very disadvantaged families and children and they need to be motivated all the time. We tried first with the online labs, after that we learned about apps and then we started with ILS and they liked it a lot. The activity was also part of the curriculum.
2.	What was the highlight of your implementation activity? Was there a special moment, quote, incident that you and your students will remember?	It was when we used these tools that it go very more hours – one of the students asked me, "But teacher, next hour we will learn or we will just play?". They learned a lot of things but they take it like playing.
3.	How many activities like this would you like to do in the course of a year?	
4.	Have your students become more interested in Science as a result of this activity?	My students like a lot to work and learn with digital tools. We're used to this because I'm also a Science Computer teacher and I taught them a lot of things, tools to use and they like to learn with ILS a lot.
5.	What were students' reactions during the implementation?	The most difficult and it took a long time was to learn to use the application, the concept map, the mind map and after that when we started the ILS, they must do these kind of maps and we started using the words from the word cloud and they tried it, most of them, realised very good and correct sentences about it and after the conceptualisation and investigating phase they must go in the beginning and verify their concepts, what they created before they knew more about the subject and they corrected also, in a very good way, the concept map, they used sentences, they talked and worked with each other or in groups.
6.	Have students enjoyed the activity in relation to Science?	
7.	Which part of the activity did they enjoy most/least?	
8.	Did the activity had an impact on students' collaboration or social skills?	I think that besides the scientific and the digital competencies they get a lot of experience in social skills and competencies: they cooperate a lot, they communicate, they look through each other's work and also they learn to be more independent and work systematically and I think this is very important. To be

		able to use information, to prove their ideas, to correct themselves and learn about it.
9.	What challenges did you face during the implementation of the specific activity?	The biggest challenge was creating the ILS. To imagine in which way the students will react when they work on it and to choose the simplest things for that topic so that it is easier for them and they like to use the ILS and have a good experience in the beginning. Of course, we can do more extra things and more difficult in time, but first I wanted it to be a very good experience and for them to have a good feeling about it. Also, after that the implementation was very nice.
10.	How much time (real time) did the implementation of the activity take, including preparation?	We started to use the Online Labs since October on different lessons, activities, they especially like the PH emission. After that we plan to use the apps, because before I use the ILS they need to learn this and after that I started with an easy ILS and after that, for the content to be easily applied because I think it's important that students do it independently and not be very busy, that it's attractive and they do it in a short way. Not very busy, because in that way maybe they don't try to do until the end. I put attractive videos, with interesting exercises. They worked in groups, and also independent and some of them try it also at home. In February-March we do this ILS and some days I spent to do it in a good way and after that we tried to use together in the classroom and they liked it a lot.

15.2.1.3 Lidia Ristea, Romania

1.	**Name of teacher**	**Lidia Ristea**
2.	Name of the school, Country	Technological High School Teodora Diamant, Romania
3.	Type of school	Technological
4.	Students' age	14-15 years old (9th grade)
5.	ILS / online laboratory used	"Biodiversity of plants, birds and animals" http://www.golabz.eu/spaces/biodiversity-plantsbirds-and-animals
1.	Starting from the selection of the activity you carried out, based on what criteria did you select the specific activity? Have your students' preferences influenced your choice?	"For this activity, I worked with the students but it was the student's choice. The school in which I teach is centred on Natural Science and so it was a great pleasure for them to work on this project.
2.	What was the highlight of your implementation activity? Was there a special moment, quote, incident that you and your students will remember?	All activities are interesting for students and myself but there are students that ask me, "Teacher, can I play this moment?" or "After this activity, can I play?" and I answered "Yes you can, but with educational games, you can be a gamer, you can realize a game with you and for other students" but I recommended some applications, not all. The coding game and he said it's very difficult for him and I told him "No problem, you have to learn and for next year you can put into practice

		all the knowledge from this year, you can realize a game with the app that I recommended" and after this, he said "It's very difficult, I can't." But it is special application for realising because I participated with the students at Our Code. It was an international activity, "How can I promote programming in school?", and for this I presented some apps using programming activities: Code Game, Scratch and how I can create an app with this, Games, some online platforms and for next year I will see if some students will realize some apps that I use in their apps or for the next projects.
3.	How many activities like this would you like to do in the course of a year?	
4.	Have your students become more interested in Science as a result of this activity?	I presented my students more applications of Go-Lab, especially how they can create some apps on the Padlet and how they can include it in the project, how they can post on the platform wiki and of course, for their participation in other competitions because with this application realised by them they can participate at some competitions (local and national) especially the national one because I am coordinator of the project in school and I involve the students in many activities, especially extracurricular ones. For the students it was a good opportunity, it was special for them because they put into practice all their theoretical and practical knowledge and it's a very good application to use in the classroom. All students have an account on Google Docs, YouTube and for other activities they can post all other materials made in the school or at home.
5.	What were students' reactions during the implementation?	My students worked in a team. I selected the students that know how to select all the materials and information about this project. The title was about Bio Diversity plants and animals. And not many students can select the information and other students don't know how they can realise this using digital tools and the best students worked with team leaders because some students can't select all information, other students can't realise all materials using digital tools. After this some students had to process images because not all images on the internet are in jpg format and they had to process using some application. After this, they had to include all pictures in the application for a good visualising and after this for all activities the students could realise and finish some materials that I posted in my project and of course in the next years we have to work at the second part for this project because the students from the 9th classroom have to continue this project and 12th class, has another module for curriculum about bio diversity. Because in their curriculum there are different units, other lessons and for this they came to learn more about digital tools. In each year they have to learn more about other tools how they can use digital tools for creating, editing documentary movies, multimedia presentations.

		How they can use software to process pictures, images. For me it was a good opportunity, the first time that I use Go-Lab online laboratories and some applications and I saw more apps that I can use in the classroom. For students it was a big surprise that here saw some application that the students don't hear for this moment.
6.	Have students enjoyed the activity in relation to Science?	
7.	Which part of the activity did they enjoy most/least?	
8.	Did the activity had an impact on students' collaboration or social skills?	For implementing the project, especially using online platform, more digital tools is more time to implement but not problem for some students because they work, not too much of course, at home. For me it's a very long time to implement more projects; I want to because it's very important to involve the students in more activities for a good collaboration, communication to develop critical thinking, but it's very difficult because we don't have the time. The students have just two hours/week of ICT and just an hour of informatics, programming. I recommend to the students more applications but don't have performing computers which is very difficult for them and for me of course to realise and finish all materials. But Go-Lab was and is a good opportunity to learn more and of course, for me, I have to learn how to create more ILS because I posted on Graasp more materials but I didn't finish the one about Bio Diversity ILS and I have to finish other ILS because I think that more materials are very important to use in classroom for students, for teachers and for my colleagues of course. I recommended this platform and here I saw an app, Gear Sketch which is very important but I can't use it because I don't have a plug-in. I recommended to the students, for next year, to work with this application because it's very special. It's different with which they can work every day.
9.	What challenges did you face during the implementation of the specific activity?	For me it wasn't very difficult to realise an ILS but I have to work after finishing the classes. Because I don't have time at school, I have to realise many activities in school and extracurricular ones. I am in other activities, in other school and for this I have to work with the students on email, I have to call them. I sent one mail and I posted on the Google Sites some materials for realising at home for some students and for others realise it at school. For this, not all students finished their materials. Especially, the platform, Pladlet and documentary movie but difficult for them and for me, of course, were the technical problems: with the connection and we don't have upgraded computers and they work very, very bad but I hope for next year we'll receive other computers and for students will be easy to realise all activities. Some students, the 6th-8th grades, don't learn

		about computers or ICT and in the 9th classroom it is very difficult for them to be involved in more activities. But it is very good for me because much more students want to learn and for them it's very important to develop collaboration, communication in teams. Because I applied more projects in school and extracurricular and for them it's very good to be involved in more activities. Some students know how to create multimedia presentations, other know how to select materials and include them in the project.
10.	How much time (real time) did the implementation of the activity take, including preparation?	For this activity I worked with the students at school because not many of them have personal computers, so for the other it was very difficult to realise the activities. They liked to work at school and they had more activities to do and for this I applied in my project how I can create some apps with online platform on tablet. We have a wiki space where the students post their applications and some activities, realize some documentary movie with all activities and in the project. And the students like to be involved in many activities, especially extracurricular ones because it's a new experience for them, not, for example, just writing and reading. It is an easy activity for them because it develops collaboration, they can communicate and they can realize after their ideas. We worked about a month with two hours work/week.

15.2.1.4 Panagiota Argyri, Greece

1.	**Name of teacher**	**Panagiotia Argyri**
2.	Name of the school, Country	Πρότυπου Πειραματικού Λυκείου Ευαγγελικής σχολής Σμυρνης (Pilot Secondary school of Smirni), Greece
3.	Type of school	Secondary school
4.	Students' age	16
5.	ILS / online laboratory used	"Μαθηματικά μοντέλα πρόβλεψης του μεγέθους ενός πληθυσμού μετά t χρόνια" (Mathematical models predict the size of a population after t years")
1.	Starting from the selection of the activity you carried out, based on what criteria did you select the specific activity? Have your students' preferences influenced your choice?	I created an ILS about math model called "Logistic population growth". According to this we could predict the growth rate in population. In algebra, which I teach, students learn about math models (exponential or logarithmic models), so this activity is combined with the math curriculum but moreover, why I selected this activity: from my experience, I know that students have difficulty understanding functions and the way that two or most sizes vary – students couldn't afford to understand and combine the different representations of math models so an ILS of Go-Lab is the best selection because it enables dynamic change of parameters. So according to this, I thought that this was a good way for students to understand math models.
2.	What was the highlight of your implementation activity? Was there a special moment, quote, incident that you and your students will remember?	It was very exciting and innovative for me, the environment of Graasp. It gave me many choices: I can use observation tool, questions scratch pad, input for recording my students' responses during the inquiry. Moreover they are many useful tools for assessment of the learning process. We can use the students spent time, reflection tool, actions statistics. I think that this is a very innovative way in learning science in classroom.
3.	How many activities like this would you like to do in the course of a year?	
4.	Have your students become more interested in Science as a result of this activity?	I was working on Go-Lab with a group of 15 students, every week, for about two hours. Students enjoy this ILS and the most important thing is that they were activated in Go-Lab and worked in groups in existing ILS that present the experiments to other students. They were like investigated different scientific subjects outside of the curriculum. This activity I implemented in two groups: the first one was the one where we're usually working with Go-lab, the second group is a group that implement Go-Lab first time. So I have the idea to measure the results according to questionnaires and I have to mention that it's an alternative way of teaching science through this ILS. It

		offered many representations of scientific concepts and the lessons are not boring according to this. The students from the first group, the experienced one, take a role as mentor in this ILS to other students so in the second group, I just observed the progress of teaching because students teach other students and learn from each other. It was an exciting activity and experience for me.
5.	What were students' reactions during the implementation?	My students had difficulty in using some apps like observation tool or hypothesis tool and in this phase of activity they asked for more directions and details. They had problems with saving data and they had problems with running these apps and so in this phase we spent more time than I think. They don't have a problem in using Padlet and inductions to writing their ideas but in some apps like I told you, they have problems.
6.	Have students enjoyed the activity in relation to Science?	
7.	Which part of the activity did they enjoy most/least?	
8.	Did the activity had an impact on students' collaboration or social skills?	My students usually collaborate in classroom to complete worksheets, to carry out extra exercises so I don't think that this activity add extra positive points in social skills. The most important is that students enjoyed this activity for collaboration.
9.	What challenges did you face during the implementation of the specific activity?	I have to mention that I am a mathematician so a big challenge for me is, according to Go-Lab, to investigate experiments about physics or chemistry or biology. It was very interesting to learn with my students. We collaborate to learn more and more things. Also, I don't usually use inquiry based learning in math. The problems in teaching are different from this problems so I was interested in creating in Graasp. I mentioned that my students could learn more and more things. The most impossible was that they gave the motivation for science. According to activity some students discuss why were are performing experiments through virtual labs, why we don't test this on actual labs. So, they discussed that this is the best way for most innovative and creative results.
10.	How much time (real time) did the implementation of the activity take, including preparation?	For implementing this activity I needed two hours in one group which was using Go-Lab more and for two other groups, about three hours. I gave them more exercises about math models for homework. But I was preparing this activity for one month with writing in Graasp and finding good activities as introductions or exercises or matrix. I think I worked about 10 hours to create this ILS.

15.2.1.5 Chrystalia Lymparidou, Cyprus

1.	**Name of teacher**	**Chrystalia Lymparidou**
2.	Name of the school, Country	St. Stilianos Primary School Cyprus
3.	Type of school	Primary
4.	Students' age	11-12 years old
5.	ILS / online laboratory used	"Moon phases" http://www.golabz.eu/spaces/lunar-phases
1.	Starting from the selection of the activity you carried out, based on what criteria did you select the specific activity? Have your students' preferences influenced your choice?	I wanted to have an environment that technology could be an add-on for. For example, I didn't want to use technology for solubles for things that children could do with their hands. So I chose the lunar phases as the technology was an extreme add-on for this lesson because using telescopic charts or using telescopes or lunar calendars was much beneficial for the lesson. On the other hand, I wanted a lesson that was open enough, so many alternative explanations would come up and so children would have the social interaction in order to come up with the right explanation. And also, I wanted a topic where students hold many alternative conceptions about it, so my students did not know a lot about the moon and the lunar phases. It was somehow a new project for them and this was an extra criteria for me.
2.	What was the highlight of your implementation activity? Was there a special moment, quote, incident that you and your students will remember?	
3.	How many activities like this would you like to do in the course of a year?	
4.	Have your students become more interested in Science as a result of this activity?	Especially in this topic, I gave them an activity to perform the unit had started. I gave them a lunar calendar and a paper and asked them to observe the moon every night and to write down their observations. And unfortunately I was very disappointed that they didn't follow the activity. And because I am not a main teacher of their class and I see the children only once a week, only two periods/week. Every week I was trying to remind them and give them motivation to do it again and again. And only two or three children in each class were responding to this activity and I was very disappointed. Because in science you need first observations to make hypothesis and come to explanations, I couldn't move on with the lesson but I was very happy because after I had finished the lesson, then introduced them the lunar calendars, the telescope pictures and after they had engaged with the technology in the issue, many students were finding me after the classes and were asking me and talking issues about the moon, they have told me that after the unit they were watching the moon every night

		and trying to find explanations, they started being interested in astronomy, especially girls who were not traditionally interested in astronomy, and other planets and how to watch the sky and this was really interesting for me. I was very happy about that. I was disappointed because the moon is very interesting for young children and I was expecting them to react in the first task but they didn't. But they reacted after they were engaged with Go-Lab, they were much more motivated to observe the moon.
5.	What were students' reactions during the implementation?	They have difficulties in the environment and this was very good for us and as Claudia said, the kids were very free to use the technological tools, they need only some minutes to get adjusted to the environment and this was very good and also it was very good the fact that they were working in groups and by definition, because we rarely use 1 by on 1 computers, we had students in groups, it helped collaboration a lot. I think the difficulties are not connected with Go-lab but with other general difficulties students face. For example, they are not able to argue at a high-level. So when they had alternative explanations and I led them to make an argumentation activity in the Padlet, their argumentation was very poor but this was not because of the environment but because of their abilities. However if I had a specialized argumentation tool within Go-Lab, I might have scaffolded them and have a better result. So I find the argumentation phase is the most difficult for them, it was the main challenge. But when this phase went from phase to phase and they argue about their models, not in the computer program, but when working at groups then they overcome the difficult and they come out with selection models because they could argue face to face and their argumentation was real and more vivid. I observe those two incompatibilities – the students could not argue in the computer but they could do it face to face.
6.	Have students enjoyed the activity in relation to Science?	
7.	Which part of the activity did they enjoy most/least?	I think they enjoyed most the social media part, the Padlet. They enjoyed it very much, it was a good surprise for them when they realized what the other students was writing about the moon; they could see their classmates. This was very interesting for them and they were excited. It was the first time we were using such a tool in my classroom and I have to say that some of them, very few of them, could not use it properly and they had started cyber-bullying issues for example writing things about their classmates but it was only because it was the first time being used. Afterwards, after we explained things, things were very nice and then when they started to use other applications like the question tool or the hypothesis

		tool they were asking me “Are other classmates able to see what I’m writing now?”. They were very anxious and very excited about sharing issues and they didn’t like the fact that the others could not see their observations for example. Because we started from Padlet and the first thing they enjoyed was the social exchange. The second thing is that they really enjoyed the lunar calendars because they started finding out what was the moon when they were born, for example. They could go years back or years forward to see what would be in 2020 the moon for example and I was very happy because this has to do with the main focus of the lesson. And they enjoyed a lot the tools – they were going to different countries, to see what the lunar phase was there and they really enjoyed that.
8.	Did the activity had an impact on students’ collaboration or social skills?	The fact that they had to use the computer screen unifies children because they have to collaborate because there is only one place to write down your observations. So they were finding routines or using routines that we already had for writing down observations, for example. For writing down hypothesis or to argue about something. I have to tell that the technological tools helped them because that is the possibility that you can write the observation now and then I can write my observation later. There is room for everybody and this was very good. However, groups that were not good at those routines before I’m telling the classroom or children who were not mature enough in group work because they were being egocentric or they wanted everything to be their own, they were finding difficulties all the time. But they were very few and this happens in all lessons. But other children that usually take a background role in the group were more engaged with the technology because they had a way to separate work. “Now I can write, now you can turn on and write down.” And so I saw them engaging more in the group. If you want my observation in the way of a result I see that children with high social skills and group working skills are very happy with these environments. Children with lower abilities can enter the group and gain from these processes. But children that are very poor in communication skills and group work skills need extra work to be able to participate in such environments and as a picture, I have to describe them exactly what to do in order to survive in this environment; otherwise they were lost, they couldn’t perform good work or an accepted work for me.
9.	What challenges did you face during the implementation of the specific activity?	There were two points. The first one was the technical one. I had to learn the Graasp environment and understand how things were working over there and finding a way to learn about the tools and to implement them right because a lot of them and sometimes you get confused which one is best for the purpose you

		want. In the same time, when you first meet the application and the extras that the environment provide – it's like being in the supermarket and have one thousand pounds and you can buy everything and then you say "Ok, I will use this and this and this and then you realize that you create a very complex environment and then you have to be simple. So you have to find a simple task and all the many available processes you have to use. This was one challenge. The other challenge was to find a way to manipulate the many alternative explanations from the students and have in the same way an environment I could use but I don't know why I had in my mind the need to have a final product that someone else could use also. So, because I have worked other times in the Ministry of Education doing this job, if it wasn't open enough then only me could be the final user and I didn't like it because then it was not for use for other people so I had to make it open enough for my students but also closed enough so somebody else could use it. These were the main challenges for me.
10.	How much time (real time) did the implementation of the activity take, including preparation?	It took me about 3 months the whole project. But I didn't perceive it and finish it as a learning environment and then implemented with the students. I actually taught three 6 grade classes, Science, so I somehow used the first classroom as a pilot so I went with initial ideas, opening up and letting students react on the environment and on the tasks being proposed and afterwards I had a very complex picture of how do students react to this environment and I was modifying it for the second classroom and then again modifying and somehow finalizing for the third classroom. So it was an on-going design process and the final "result" is what has been produced by all three classes. I am going to use it as a final product but with other classrooms and not be opening it because I realize that it's very good to modify things all the time but it is not productive in some ways. So I would like to examine the effects and the results of using the environment in a stable way.

15.2.1.6 Claudia Maria Mazanti, Italy

1.	**Name of teacher**	**Claudia Maria Mazzanti**
2.	Name of the school, Country	Public secondary school "Iqbal Masih", Italy
3.	Type of school	-
4.	Students' age	-
5.	ILS / online laboratory used	"Is it good to be beautiful - understanding evolution through natural and sexual selection" - http://www.golabz.eu/spaces/it-good-be-beautiful-understanding-evolution-through-natural-and-sexual-selection and https://www.youtube.com/watch?v=jL0lzX_R-us
1.	Starting from the selection of the activity you carried out, based on what criteria did you select the specific activity? Have your students' preferences influenced your choice?	So I chose the subject because last year we made a long project on the aquarium and on guppies. So I found the activity and I thought it could be interesting for my students since they have worked already on the subject. Then I proposed the activity to them and they realized their work together but not with strong intervention from myself, they just worked by themselves. This was very interesting for me.
2.	What was the highlight of your implementation activity? Was there a special moment, quote, incident that you and your students will remember?	
3.	How many activities like this would you like to do in the course of a year?	
4.	Have your students become more interested in Science as a result of this activity?	Teacher: They shot the video during the afternoon when there were also other students that were working on other projects like theatre or other projects on Science because we had a greenhouse so we just worked all yearlong on this part plan's development. So they just work together and they chose the scenography, as you see there is a part of Science for Girls that were was sent by Chiara Tulepi. So that was their idea, they used that scenography by themselves so this was very good for me also because they understood how important this part of our work is.
5.	What were students' reactions during the implementation?	Student 1: "For me, working at this project, I haven't got any difficulties but for some of my friends it was difficult to understand English in the work projects and I helped them to translate in Italian and after they understood and worked together."
6.	Have students enjoyed the activity in relation to Science?	
7.	Which part of the activity did they enjoy most/least?	Student 2: "The part that we enjoyed most is creating and acting the video."
8.	Did the activity had an impact on students' collaboration or social skills?	Student 4: "At the beginning we were two groups. We had a contest." Teacher: So my students were divided into two groups. They were just having a contest one with each other and they were

		working separately. The original idea I had was to make them work not together, but in two groups. And the groups were mixed between two classes. But then the two groups got together. So when we arrived at the end they had to produce something and they just thought that it was better just to get together their strength and their work and they worked together. But this is something that really they decided to do, I didn't tell them go and work together. They choose to work together. So this shows that the activity gave them the opportunity to increase their collaboration skills because at the end when they had to produce something they said "It's better if we work together, if we give each other our material." Because one had the conceptual map and the other the translation. So they just worked together and I was really happy because the two classes didn't have so often this opportunity to work together on a subject. They had opportunity like theatre or trips but this was another thing. Working together and realising, as they did, it was really something new for them.
9.	What challenges did you face during the implementation of the specific activity?	The main challenge was to study the platform, to see what each activity was offered so for me was really difficult to understand what I had to do. So after the 3rd webinar, I began to understand Graasp and all the activities that were offered. For me, it was easy to choose the activities because, as I told you before, we had this project last year so I just chose the activity on the sexual selection because in that part of the year we were just studying Genetic so it was the right subject in that moment but I have to say is that I have to thank my students because they were the ones who convinced me of what this platform could be, could become for Italian students. Because at the beginning, I'm really sincere, I couldn't expect that they could work in that way so I was surprised of their ability, of the way they worked together so I think it's the project because they really were beautiful in this work and the project gave me the opportunity to see how they were special in working together in such a difficult situation. Because it was the first time for them. And the challenge I told you, is that of the technical problems but that was overcome – I had a colleague that just helped me with Java and all that technical part.
10.	How much time (real time) did the implementation of the activity take, including preparation?	There were two parts of the work. The first one was the part for me. I had to understand how Go-Lab works, I had to overcome some technical problems because we don't have really a modern connection here in my school and I had problems just to build the platform for my students. That took almost a month, I think. Then, when I was ready I proposed the activity to the students and they just met for two months, every Wednesday afternoon for two hours. So they worked on the project almost two months they told me now, for, I think, almost 20 hours of work. I choose the best students in English and Science in the two classes I had and then I mixed them. So they were Somali, also not all Italian students but

		Russian, Chinese. So they worked together just to, also work on the English part of the project. Work for the language just to practice and study science in another language. This was the first experiment for me in this kind of activity and I just was a pilot phase for me and I think I will repeat it next year with other students. But what my students told me now is that it's an activity that is good for the last year of the middle school in Italy. So for students of 13 years old, not less, because they had to manage the use of the platform, they realize this video, they studied the subject and this was really a great effort for them and they think they are mature if they are at in the last year of middle school, not before. As the other students that I have, are younger.

15.2.1.7 Ivana Gucic, Croatia

1.	**Name of teacher**	**Ivana Gucic**
2.	Name of the school, Country	Gornje Vrapce, Zagreb, Croatia
3.	Type of school	Primary school
4.	Students' age	12 (6th grade)
5.	ILS / online laboratory used	"Valovi" (waves) http://www.golabz.eu/spaces/valovi-waves
1.	Starting from the selection of the activity you carried out, based on what criteria did you select the specific activity? Have your students' preferences influenced your choice?	Last year, a group of my younger students that I'm teaching math that didn't have Physics subject at school yet, they participated in Festival of Science here in Croatia where the main topic was waves and sound and when they were preparing for the festival they were enjoying it a lot so they had so much fun during all the experiments, the researching, learning something in advance about some part of physics and when they told to the rest of the class how fun it was, the class asked me if we could do those experiments this year so I decided to take a challenge and repeat those activities. I had to add some new ones to make it easier and try to do inquiry based learning with class that doesn't know almost anything about physics so that was kind of a – yes, they had an influence on my choice.
2.	What was the highlight of your implementation activity? Was there a special moment, quote, incident that you and your students will remember?	
3.	How many activities like this would you like to do in the course of a year?	
4.	Have your students become more interested in Science as a result of this activity?	Yes, I did. They had so much fun and they are looking forward to having physics class next year. I don't know if next year when they actually start to learn physics terminology, they will have so much but I hope they will. Yeah, they are looking forward to having science subjects because they still don't have those right now in this grade.

5.	What were students' reactions during the implementation?	Everything on computers is interesting to them. They like to work on computers, they like to do hands on things because in other subjects, or subjects in Croatia, they don't have the opportunity to do that so the Science objects are something that allows them to be more creative, to discuss more, not only hearing, writing and copying something from textbooks so they had a lot of fun. I'm trying to make it fun in all of my classes, math and physics to be more hands on, to be more something like inquiry based learning.
6.	Have students enjoyed the activity in relation to Science?	
7.	Which part of the activity did they enjoy most/least?	
8.	Did the activity had an impact on students' collaboration or social skills?	I did notice that. They were more bonding because almost every activity like the team work and the group work, so there have to collaborate more, have to communicate more, compromise, cooperate, must listen actively to each other, articulate ideas, distribute their responsibilities evenly. Yeah, so it had an impact on their social skills as well.
9.	What challenges did you face during the implementation of the specific activity?	There were a lot of challenges that I had to face in the implementation of the activity because like I said the class doesn't know a lot about physics, they are not familiar with a lot of physics terminology, they have heard about terms like energy, particles, energy transfer but they didn't know how to explain it so they didn't learn about that yet so they will learn it in the next year, during the physics class, so that was challenge to make it as simple as I can. And I had to rule out all the mathematical components because the math was the problem. They didn't have that much math knowledge. So that was challenging. And of course translating to Croatian language from every English material that I had found and make it as simple as I can. So that the most challenging for me.
10.	How much time (real time) did the implementation of the activity take, including preparation?	It was around 4-5 hours. I had to make it as simple as I can because those kids don't know about physics concepts and some physics terms so I had to explain and they had to research and every single term that we used in activities before they came to class. I used the flipped classroom method.

15.2.1.8 Casper de Jong, Netherlands

1.	**Name of teacher**	**Casper de Jong**
2.	Name of the school, Country	-
3.	Type of school	-
4.	Students' age	-
5.	ILS / online laboratory used	"Gears"- http://graasp.eu/ils/54f038f38cd7e5edb8e860a2?lang=en and

		https://www.youtube.com/watch?v=nxpIM5dA7G8&feature=youtu.be
1.	Starting from the selection of the activity you carried out, based on what criteria did you select the specific activity? Have your students' preferences influenced your choice?	We use the Gear Sketch environment for two reasons: I had to use in primary school and we had to use it for the secondary school and the first two classes and we were not completely sure about the level of the students we were working with. Because of the fact that our group we were four students and we created this ILS, we came from all kind of different backgrounds so the gears were the easiest ones to understand and we were not focusing on the scientific relevance of the lab but more on the value of the educational process of the lab and we found out that with the Gear Sketch we could manage to change the variables the most and we could see what was happening with the learning process for children and what worked best. It was for us just a choice about what's the easiest one to work with and what do we understand.
2.	What was the highlight of your implementation activity? Was there a special moment, quote, incident that you and your students will remember?	
3.	How many activities like this would you like to do in the course of a year?	
4.	Have your students become more interested in Science as a result of this activity?	I thought that working with the ILS it made it more tangible. It was like they were working with something really and especially for the kids at primary school, they found it really fun, instead of working on paper, they thought that they were really doing something and they enjoyed the fact that what they did was working and there was actually happening at the screen and they were not just filling out paper. It was something that was really nice for them and I think that working with the ILS and with these online laboratories can help with classes and can help create a more interactive and more interesting classes for your students.
5.	What were students' reactions during the implementation?	
6.	Have students enjoyed the activity in relation to Science?	
7.	Which part of the activity did they enjoy most/least?	We have a luxury situation here, I believe, because our school all the students have their own laptops so we can say just "Open your laptops and start the ILS" and I also tested it a primary school with one of group members and there we used the computer room and there were 5-6 students sitting there behind their computers screens and I must say, I preferred our own school because then you have an entire class with their own devices and that works best. The age of the primary school kids was 8-12 year olds and there is a big difference between the primary school and the first class I used. It's only one year but the entire interaction

		with the ILS and how they worked with it, it's really a big difference. It's a different type of school and you can really see it.
8.	Did the activity had an impact on students' collaboration or social skills?	
9.	What challenges did you face during the implementation of the specific activity?	We encountered some technical difficulties. We were building the ILS in Graasp and during the building we found that, with other programs there were technical difficulties, bugs we found and all kinds of things like that but we all gave them through to our teacher and for the implementation in the lessons, I found it hard for the children to understand how to work with the laboratory so for that we made some instructional videos on how to work with the laboratory and the students really found it useful and they understood better how to place the gears in Gear Sketch and how to make the change in connecting the gears, which was a little bit odd to do in the first version of the Gear laboratory and at this moment I believe there was an update and it's easier to do that but when we introduce it, it was hard and that was something we found difficult to do. The Hypothesis Scratchpads, we used them as well, I used them as well in other courses not connected to an online laboratory but more to some curriculum based economics parts for my own lesson and that was very useful. I found that the Hypothesis scratchpad can be used in every kind of domain in education, not only for online laboratories.
10.	How much time (real time) did the implementation of the activity take, including preparation?	We spent a lot of time on it but it's also because of the fact that we made some instructional videos for the ILS and we used guidelines provided by a professor of mine at the UT. And I also wanted to include some parts of that in the instructional videos. So creating the instructional videos took me a lot more time than just creating the videos. And so it was for the entire ILS, we thought about stuff and we changed it I think that we spent about 30-40 hours, but not only working on the ILS it's also about reading articles and finding out theories, etc.

15.2.1.9 Karolin Mae, Estonia

1.	**Name of teacher**	**Karolin Mae**
2.	Name of the school, Country	Audentes School, Estonia
3.	Type of school	-
4.	Students' age	11th grade (16-17 years old)
5.	ILS / online laboratory used	"Is it good to be beautiful - understanding evolution through natural and sexual selection" http://www.golabz.eu/spaces/it-good-be-beautiful-understanding-evolution-through-natural-and-sexual-selection
1.	Starting from the selection of the activity you carried out, based on what	I am teaching IB Biology curriculum and we had Evolution topics the time I saw the Go-Lab call for teachers so was

	criteria did you select the specific activity? Have your students' preferences influenced your choice?	looking for an activity that would correspond to the curriculum and topics I teach in my class. Luckily I did not have to search long as I created last year a lab for Go-Lab competition that was a perfect fit at the moment. Also was interested to try it out in my classroom, how students respond to that and if they find it engaging. It was a lab that we worked on with several teachers during last year summer school and Tartu University researchers made some finishing touches. Student`s preferences of course influence a lot and as my curriculum allows me a lot of time for experiments then we can do many labs per year. Usually they prefer hands on activities.
2.	What was the highlight of your implementation activity? Was there a special moment, quote, incident that you and your students will remember?	Well my lab was about sexual vs natural selection and during the orientation phase we were discussing what influences our selection of partners and when you ask such thing from teenagers then many highlights are guaranteed :) Now we all know the specific preferences of all of our classmates as they were eager to share.
3.	How many activities like this would you like to do in the course of a year?	
4.	Have your students become more interested in Science as a result of this activity?	I can say that in general my Students are more engaged in Science lessons due to the inquiry approach.
5.	What were students' reactions during the implementation?	At the beginning it was all fun and enjoyment, but when they actually started to work on the simulation then there was also lots of confusion as modern day learners have used to quick answers and do not want to read long texts that are sometimes necessary for implementation.
6.	Have students enjoyed the activity in relation to Science?	Yes they have and hopefully thanks to that activity they will always remember the selective evolutionary pressures and can apply this knowledge in order to explain the theory of evolution
7.	Which part of the activity did they enjoy most/least?	Orientation as it was full of interesting and intriguing questions
8.	Did the activity had an impact on students' collaboration or social skills?	Yes they had to work in pairs.
9.	What challenges did you face during the implementation of the specific activity?	IT challenges-many School computers did not run java applet normally and there were some issues with that.
10.	How much time (real time) did the implementation of the activity take, including preparation?	It took 2 lessons, but they also had to write up a research paper as a home assignment that I guess took then an hour.

15.2.1.10 Nikolaos Nerantzis, Greece

1.	**Name of teacher**	**Nikolaos Nerantzis**
2.	Name of the school, Country	Special Vocational School of Serres (TEE & EPAL Eidikis Agogis), Greece
3.	Type of school	-

4.	Students' age	7-8 years old, 9-10 years old (2nd & 4th grade);
5.	ILS / online laboratory used	"Falling objects" http://graasp.eu/ils/552cb3a4680bfb937acf42be?lang=en
1.	Starting from the selection of the activity you carried out, based on what criteria did you select the specific activity? Have your students' preferences influenced your choice?	It's in the curriculum. I made an ILS with two bodies falling, I took a video of them and then we had a conversation about what is happening, why one body will fall faster than the other body, something like that. I took two classes, the 2nd and 4th grade of high school of a technical, special education high school and we make an open inquiry based project about it.
2.	What was the highlight of your implementation activity? Was there a special moment, quote, incident that you and your students will remember?	
3.	How many activities like this would you like to do in the course of a year?	
4.	Have your students become more interested in Science as a result of this activity?	Yes, through the joy of knowledge, let's say, all these activities, my students want to do science, want to learn about scientists of the past. For example, for this ILS, in one part I spoke about Galileo, how he rejected Aristotle's opinion that the lighter bodies stay up and the heavier go down. So with the whole history of science, make it more interesting. And they want to know anecdotes about science and that help me to do core science, the basic ideas of science.
5.	What were students' reactions during the implementation?	
6.	Have students enjoyed the activity in relation to Science?	

7.	Which part of the activity did they enjoy most/least?	They enjoyed it very much. I said that one of my objectives was to lead them to ask questions but then I give them nothing. They have to collaborate, they have to search, and they have to do some homework. The better part was when they make the prediction that two bodies, perhaps falling together, two same balls, and then they look at the video screenshot and it was like that. So, what is happening? When they ask me, I said "Don't ask me, you did the experiment, please find it!". It was all this interaction about them and with them. It was really joyful. They are familiar with this. I've been working with them for two years and we have sorted out what is difficult for them, I know my students and I took out the parts that are difficult for them. We are targeting the core concepts, the basic physics. Just to understand what are the basic rules and the core ideas of physics. We did the work on the whiteboard, by interacting with each other.
8.	Did the activity had an impact on students' collaboration or social skills?	
9.	What challenges did you face during the implementation of the specific activity?	The main problem was to change the program of the school. I needed to cooperate with my colleagues and to have two continuous didactical hours. That was not the problem, also my headmistress and my colleagues are very helpful so, nothing major at all. The students are familiar with these techniques, in special education we have the ability to reform to bring our new staff in the classroom, and we have this freedom. And it was very joyful, they were very happy about it. In the class we have an interactive whiteboard with a good internet connection so it was very nice.
10.	How much time (real time) did the implementation of the activity take, including preparation?	For the preparation I was able to finalise it 3-4 days. I have the prior experience from last year summer school so it was a little bit easy for me. Nevertheless, I worked in the old Graasp and now I have to take everything to the new Graasp and in the school it took us about a month, 3-4 didactical hours. Then I take the evaluation with a colleague of mine and we change some small things and we are now preparing another ILS. In this one, we want to take out the air parameter but the new one will take it in. We can make an activity to maximise the resistance when two bodies are falling. The future work.

15.2.1.11 Julieta Jimenez, Spain

1.	**Name of teacher**	**Julieta Jimenez de Llano Garcia**
2.	Name of the school, Country	Colegio Esclavas. A Coruña, Spain
3.	Type of school	-

4.	Students' age	-
5.	ILS / online laboratory used	Reglas de solubilidad, http://www.golabz.eu/spaces/reglas-de-solubilidad and https://www.youtube.com/watch?v=SuMjUvEqO9Q&feature=youtu.be
1.	Starting from the selection of the activity you carried out, based on what criteria did you select the specific activity? Have your students' preferences influenced your choice?	Since it was my first activity with the project - I also teach mathematics and physics but I studied chemistry at University --- I selected one particular lab and after that I chose a specific part of the subject that I think that could be quite difficult for the students to learn in a traditional way. It's about solubility rules so it's not very popular, so I think, but I had my students in mind, not their preferences but their best interest.
2.	What was the highlight of your implementation activity? Was there a special moment, quote, incident that you and your students will remember?	I think that our highlight was no doubt the discussion part where we shared everything. As I was telling you before they made very creative videos using TV and advertising Spanish songs and traditional children's stories. Maybe some parts that you may not understand if you are not living here but they were very creative and so we were laughing and well, I think I will remember that class for a very long. In the video where I put everything together, there are a couple of things that they had to share.
3.	How many activities like this would you like to do in the course of a year?	
4.	Have your students become more interested in Science as a result of this activity?	I think this activity definitely grew their interest in Science since they had the opportunity to do something really fun with a subject that in another situation would be maybe not very interesting to them.
5.	What were students' reactions during the implementation?	It was like a kind of game for them: it was new and interactive so it was like playing, especially in the implementation, they had a lot of fun. And, after the discussion they needed to record some videos and I think they were very creative, they did a really good job and so when they really got to understand the way they needed to do everything, they enjoyed it a lot.
6.	Have students enjoyed the activity in relation to Science?	
7.	Which part of the activity did they enjoy most/least?	I think what they enjoyed the most was in the beginning the part of the playing with the virtual lab; they had fun with this but the highlight of the whole activity was the discussion part. They needed to make up their own mnemonic device for the solubility rules and I asked them to record a short video for sharing in class so when we shared all the videos it was like the best part, also for me and for them. They worked very, very well and it was fun. And I think they truly learned the subject matter. They will be able to recall it in the future, as that was the point. That hardest part was at the beginning, because I didn't want to give them too many instructions. I wanted them to be able to read and to understand what they needed to do. It was part of the activity so they were feeling a bit lost, asking each other

		what they should do and so it was hard. But after that, when they started they were all enjoying it.
8.	Did the activity had an impact on students' collaboration or social skills?	I really think the activity contributed to improving the way they collaborate with each other and they needed to do it in the beginning. In fact, they were working in pairs for the whole activity and even when they needed to work at home, they needed to work in pairs so I think it helped in this way. Also, for the communication skills, the part that they had to record the video. In fact, they needed to record a video with their mnemonic device, the final activity, but I had asked them to record a short video with their feedback. It was short but they had to do it and this kind of things I also believe it helped.
9.	What challenges did you face during the implementation of the specific activity?	The first part was checking that all the technology required was available: computers, Wi-Fi and regarding the students I think that the most difficult part was when they realized that they needed to work on their own, to take responsibility on their learning and they are not so used to this and well, I was there to guide them but they needed to work by themselves. So it was a bit difficult for them to understand this at the beginning. And well, I also needed some communication through the activity and I wanted to include one of those apps to monitor what they are doing and sometimes maybe I wanted to say to one student: "Ok, you have been in this point for half an hour, do you have a problem, do you need help or maybe you have a question?' And I couldn't find this possibility through this ILS. I had to write traditional email and maybe they didn't read the email in real time so it was not so useful. So, I miss this kind of communication.
10.	How much time (real time) did the implementation of the activity take, including preparation?	I think that the longest part was to think about it all and the way they put it all together. I was following an online course with our Spanish coordinator and that helped a lot because we were working all together at the same pace and it was nice. I think that I was thinking about the activity maybe for 2-3 weeks and after that it was like 8 hours to design the whole ILS and for implementation it took me 2, 50 minute classes and one more for the discussion and sharing the results. I think like two extra hours at home to grade all the students' work.

15.2.1.12 Tsetsa Hristova, Bulgaria

1.	**Name of teacher**	**Tsetsa Hristova**
2.	Name of the school, Country	Bulgaria
3.	Type of school	-
4.	Students' age	14-18
5.	ILS / online laboratory used	Explorations of solar activity and sunspots – http://www.golabz.eu/spaces/sun4all
1.	Starting from the selection of the activity you carried out, based on what criteria did you select the specific activity?	I have my students very active in Astronomy so I decided to test astronomical resources. One of them was solar activity and so we made different activities about this idea and later students worked using virtual labs and other apps.

	Have your students' preferences influenced your choice?	
2.	What was the highlight of your implementation activity? Was there a special moment, quote, incident that you and your students will remember?	I think that for me it was very nice the time that students started to talk to me a lot because many of them come long distance from home and maybe one or two times/ month they go to Moscow. They were closer to me and it's great that I can to do that for my students. And they started to really enjoy our classes, always smiling, they really understand that I like to help them to learn Science.
3.	How many activities like this would you like to do in the course of a year?	
4.	Have your students become more interested in Science as a result of this activity?	Yes, after this activity. My students are very interested in Astronomy already so this activity definitely helped.
5.	What were students' reactions during the implementation?	It also fun for my students, they enjoyed it a lot. When we discussed different things, they said "Oh, it's so interesting" so I said "If you are clever, you must look at this differently, you can use it after graduation, when you finish school" and they said, "Sure, we will use the ideas". So they were eager to test it.
6.	Have students enjoyed the activity in relation to Science?	
7.	Which part of the activity did they enjoy most/least?	They liked most when they make the test but at the beginning they must prepare it a little bit. They were not so happy to make the preparations but I think that it was very important to understand how to do it.
8.	Did the activity had an impact on students' collaboration or social skills?	Yes, it surely impacted on ICT knowledge. They used different ways to work together. So they needed to study how to make this video and they liked it. I think this is the way to inspire students to study Science. To use virtual labs. Also to have a chance to have a real one to but sometimes it's not possible.
9.	What challenges did you face during the implementation of the specific activity?	Also technical problems because I have to use the internet so I go to work with my computer because it's a problem to open the computer there, but I didn't have a struggle with communication because I had email, I had Skype, I had Facebook. If one didn't work, I went to another so I had the chance to communicate every moment with my students. And also by telephone.
10.	How much time (real time) did the implementation of the activity take, including preparation?	At the beginning it took me a few weeks to prepare everything but then, we worked on this for several lessons so the time that we spend on this was quite ok.

15.2.2 Conclusions

In the above interviews a variety of countries are represented. Despite teachers' different backgrounds, teaching styles and national curricula, some common characteristic appear in the way they implemented and experienced their Go-Lab activities. More specifically:

- **Online laboratories:** Despite the fact that most teachers were aware of the existence of online laboratories, it is within Go-Lab that they make their first attempt in using them within their classroom.
- **Duration of implementation:** In all cases, the implementation of the planned activities took teachers longer than expected. Teachers had to spend additional time working on the selected ILSs on their own but also to train students on how to use specific apps. In some cases, background information and research on a specific topic were also required.
- **Students' input**: In upper secondary students had an input on the choice of ILS that they implemented within their classroom. In some cases, students had even developed ILs themselves.
- **Use of ILSs:** In most cases, teachers have used ILSs or specific laboratories as parts of larger lesson plans. In the case of ILSs, it is quite common for teachers to use only some of the inquiry cycle steps and combine them with other resources while building their lessons. In many cases, we can also see ILSs to be used in the frame of other projects and become part of larger and multilevel activities.
- **Impact on students:** The use of ILSs and online laboratories increased students' engagement. At the same time improvement on students' critical, collaboration and social skills has also been reported. It is also very interesting to see that even in the context of special education, Go-Lab has a lot of applications that can contribute to raising the interest and knowledge of kids.
- **Challenges:** The main issues that teachers had to deal with in order to achieve their Go-Lab implementations were either technical or organisational. By technical, we are talking mainly about problematic or slow internet connections while organisational problems usually include curriculum difficulties, lack of adequate preparation time, need of pedagogical support and internal school's arrangements that had to be made well in advance.

15.2.2.1 Other stakeholders' interviews

Apart from teachers, interviews were also carried out with Head of schools and policy makers. To facilitate the collection of information, interviewees were provided with the option to either attend a phone/Skype call or provide us with written answers. The template of questions the Heads of schools were asked to answer can be seen in Table 15.2.

Table 15.2. Go-Lab Heads of Schools interviews: questions

Interview questions for Stakeholders		Answers
1.	Name	
2.	Name of your country/region	
3.	Are you aware of the existence of remote laboratories and their use within the schools in your country/region?	
4.	Are you aware of any Go-Lab implementations that took place in your country/region? If yes, what were the outcomes for teachers and students?	
5.	Did Go-Lab implementations had any impact on your views regarding the importance of STEM subjects?	

6.	What are, in your opinion, the advantages of using remote laboratories within the classroom? What about the disadvantages?	
7.	What are, in your opinion, the possible problems/obstacles regarding the use of online laboratories in your country/region?	
8.	Will you encourage teachers within schools in your country/region to investigate the use of online laboratories?	
9.	Is there anything else you would like to add or comment on?	

The template of questions the policy makers were asked to answer can be seen in Table 15.3.

Table 15.3. Go-Lab policy makers' interviews: questions

Interview questions for Stakeholders		Answers
10.	Name	
11.	Name of your country/region	
12.	Are you aware of the existence of remote laboratories and their use within the schools in your country/region?	
13.	Are you aware of any Go-Lab implementations that took place in your country/region? If yes, what were the outcomes for teachers and students?	
14.	Did Go-Lab implementations had any impact on your views regarding the importance of STEM subjects?	
15.	What are according to you the advantages of using remote laboratories within the classroom? What about the disadvantages?	
16.	What are, in your opinion, the possible problems/obstacles regarding the use of online laboratories in your country/region?	
17.	Will you encourage teachers within schools in your country/region to investigate the use of online laboratories?	
18.	Is there anything else you would like to add or comment on?	

The collection of Headmasters interviews can be found below:

15.2.2.2 Sandra Fornai, Italy

1.	**Name of Head of school**	**Sandra Fornai Doctor in Medieval History**
2.	Name of the school, Country	"Iqbal Masih" of Bientina and Buti, Pisa, Italy

3.	Type of school	Public Lower Secondary School Pupils from 3 to 14 years old
4.	Are you aware of the existence of remote laboratories and their use within the classrooms?	Yes, I'm aware of them because of Go-Lab project
5.	Are you aware of any Go-Lab implementations that took place in your school? If yes, what were the outcomes for teachers and students?	Go-Lab project took place in my school for the first time last year. The activity was implemented with students of different nationalities, aged 14. The teacher involved had followed the training offered by the project.
6.	Did Go-Lab implementations had any impact on your views regarding the importance of STEM subjects?	Go-Lab project activities had an integrated approach between different subjects and developed also student's literacy skills. Science and language teachers have worked together in order to prepare and support students in this activity.
7.	What are according to you the advantages of using remote laboratories within the classroom? What about the disadvantages?	In my school Go-Lab project gave to the students the opportunity to enhance their awareness and sense of collaboration, and gave them a context to practice their creativity and skills. They realized a video on sexual selection that won the Go-Lab contest. The disadvantages we encountered were the lack of facilities and electronic devices. Having adequate equipment would work to whole classes.
8.	Will you encourage other teachers within the school to investigate the use of online laboratories?	In my school, in accordance with school reform which is now underway, I will encourage initiatives of innovation in teaching regarding using of on-line laboratories.
9.	Is there anything else you would like to add or comment on?	In this new year and based on the good experience we had with Go-Lab, our aim is to find training opportunities in the use of online laboratories for more teachers, and work on on-line laboratories with whole classes: either on activities already available, but also creating new inquiry learning space from our experiences.

15.2.2.3 Barbara Stevanin, Italy

1.	**Name of Head of school**	**Barbara Stevanin**
2.	Name of the school, Country	IC "Comuni della Sculdascia". Casale di Scodosia (Pd) Italy.
3.	Type of school	Public school
4.	Are you aware of the existence of remote laboratories and their use within the classrooms?	Yes, thanks to Go-Lab.
5.	Are you aware of any Go-Lab implementations that took place in your school? If yes, what were the outcomes for teachers and students?	We used an ILS about Crater impact. Students were very interested and curious to use this new approach.

6.	Did Go-Lab implementations had any impact on your views regarding the importance of STEM subjects?	I was aware of the importance of STEM subjects before but seeing students' motivation and engagement on the activity, reinforced my views.
7.	What are according to you the advantages of using remote laboratories within the classroom? What about the disadvantages?	The possibility to make investigations and experimentations easily and free of charge. Every classroom can become an amazing lab!
8.	Will you encourage other teachers within the school to investigate the use of online laboratories?	Yes I will do! We already have plans of training more teachers in using Go-Lab and carry out more implementation activities.
9.	Is there anything else you would like to add or comment on?	The use of Go-Lab is not easy or intuitive for every teachers and student. Practice and training are needed!

15.2.2.4 M. Mègie, France

1.	**Name of Head of school**	**M. Mègie**
2.	Name of the school, Country	Collège le Marie, France
3.	Type of school	Secondary School (11 to 15 years old)
4.	Are you aware of the existence of remote: laboratories and their use within the classrooms?	One of my teachers, M. Oubella, is involved in the project and has informed of its content and his plans of using online laboratories in his classes.
5.	Are you aware of any Go-Lab implementations that took place in your school? If yes, what were the outcomes for teachers and students?	Go-Lab has been used in ICT, Technology and Biology classes. Outcomes were very positive for both teachers and students. This year we will try to use them more, create our own ILSs and involve more teachers and students.
6	Did Go-Lab implementations had any impact on your views regarding the importance of STEM subjects?	Yes! It was a proof that it is possible to inspire young people in STEM subjects through activities organised across the school and promote links between teachers within and between subject areas. We are also aware that Go-Lab implementations have been used in Comenius and eTwinning projects dealing with STEM. This year we participate to an eTwinning project called "Sound and Light pollution "and surely Go-Lab will be used.
7.	What are according to you the advantages of using remote laboratories within the classroom? What about the disadvantages?	Remote labs allow students to conduct experiments with real equipment online. It is great for schools with low resources, like ours. Non stable Internet connection can be a problem of course but there are always alternatives and workarounds.
8.	Will you encourage other teachers within the school to investigate the use of online laboratories?	Absolutely! STEM teachers have many opportunities to better their teaching but teachers from other discipline can also use inquiry to create activities for their topics.
9.	Is there anything else you would like to add or comment on?	These projects provide us with great opportunities to better ourselves and our school. Thank you!

15.2.2.5 Vincze Szabolcs, Hungary

1.	**Name of Head of school:**	**Vincze Szabolcs**
2.	Name of the school, country:	Megyaszói Mészáros Lőrinc Körzeti Általános Iskola, Hungary
3.	Type of school	Primary school
4.	Are you aware of the existence of remote laboratories and their use within the classrooms?	Yes. I knew their existence but it was through Go-Lab that I learnt more about them. I am pleased that teachers from our school use that kind of labs. We have all conditions to use the innovative methods and motivate the children with interesting experiments in learning science.
5.	Are you aware of any Go-Lab implementations that took place in your school? If yes, what were the outcomes for teachers and students?	Yes, I know that last year the chemistry teacher used the labs at chemistry and physics hours. I talked with her and the children and they were very enthusiasts about it. I am proud that our teacher, Filep Otília participated to the summer school from Greece like one of the winners of the Go-Lab contest. I think it is a good outcome and will bring more of Go-Lab to our school.
6.	Were you aware of the IBSE teaching approach? Have you used it before?	Yes I had, but I always found it difficult to apply it. Go-Lab and its ILSs made me realise how important it is to use this approach to build students inquiry skills. I still think it is challenging to use but with a group of 2 teachers interested on the topic we will make the effort to bring more inquiry to our classes.
7.	Did Go-Lab implementations had any impact on your views regarding the importance of STEM subjects?	I have heard many times that science education is very important. But now I have seen some of its results to our students and I learnt more about a lot of innovative methods and implementations that can bring more science in the classroom. My personal aim is to follow up on this and encourages this kind of activities in our school.
8.	What are according to you the advantages of using remote laboratories within the classroom? What about the disadvantages?	The motivation of the children is growing; we can do virtually the experiments for what we have no materials to do it with really materials. No disadvantages! It is a great alternative or supplement in some ca if we do also some nice, really experiments.
9.	Will you encourage other teachers within the school to investigate the use of online laboratories?	Yes, I already do it.
10.	Is there anything else you would like to add or comment on?	I am happy that our school is a pilot school in Go-Lab project.

The responses of 2 members of the Ministries of Education in Portugal and Spain are provided below:

15.2.2.6 Fernando Franco, PT MoE, DGE of Education, Portugal

1.	**Name**	**Fernando Franco**
2.	Name of your country/region	Portugal/Lisbon
3.	Are you aware of the existence of remote laboratories and their use within the schools in your country/region?	There are some schools in Portugal that I know of. About 3-4 schools but I don't know if there are a lot of them.
4.	Are you aware of any Go-Lab implementations that took place in your country/region? If yes, what were the outcomes for teachers and students?	No, I am not aware but there is a school I know which is very interested to join Go-Lab and I will invite them to the Go-Lab website to sign up. I have also realised that we have a large numbers of schools in Portugal already using Go-Lab. Would love to find synergies and ways to support them even more.
5.	Did Go-Lab implementations had any impact on your views regarding the importance of STEM subjects?	-
6.	What are, in your opinion, the advantages of using remote laboratories within the classroom? What about the disadvantages?	There are a lot of advantages: we have a lot of schools in Portugal with no infrastructure or means of implementing a real physical laboratory. There are more possibilities like this because we have good internet also in Portugal. These remote laboratories are universal, anyone can use them (students, teachers, etc.)
7.	What are, in your opinion, the possible problems/obstacles regarding the use of online laboratories in your country/region?	In Lisbon, I am very optimistic. I don't think there are disadvantages. Think the biggest issue are the teachers because they don't feel good with that kind of technology, they prefer the hands-on experience, manual use. But I am optimistic because more teachers are coming in, with fresh, open minds.
8.	Will you encourage teachers within schools in your country/region to investigate the use of online laboratories?	Without a doubt because I belong to the MoE and I am part of the Directorate General of Education (DGE). We deal with pedagogy, and my section deals with ICT. So I encourage all teachers to implement that kind of technology. Not only laboratories but even research or other instruments. A lot of time we visit schools and talk about the use of tablets n STEM education. We have a project in which we talk about this, about using tablets in STEM education so of course, I would encourage teachers.
9.	Is there anything else you would like to add or comment on?	n/a

15.2.2.7 Juan Granados Loureda, Educational Adminstration of Galicia, Spain

1.	Name	**Juan Granados Loureda**

2.	Name of your country/region	Galicia- Spain
3.	Are you aware of the existence of remote laboratories and their use within the schools in your country/region?	We know one of the schools I am responsible is working on it from 2014. Its name is IES David Buján.
4.	Are you aware of any Go-Lab implementations that took place in your country/region? If yes, what were the outcomes for teachers and students?	I know a formative course given in a Science Museum in our city. Some of the teachers involved in this course are nowadays working on Go-Lab project.
5.	Did Go-Lab implementations had any impact on your views regarding the importance of STEM subjects?	Personally my view about STEM subjects have not changed with Go-Lab, however I think students are getting closer to these subjects because of the approach of this project
6.	What are, in your opinion, the advantages of using remote laboratories within the classroom? What about the disadvantages?	Advantages. Remote labs can introduce other experiences from different places of the world. The impossible (dangerous) things can be done. Students can repeat an experiment at home. Disadvantages. In any case, it's not available Internet connection in each home.
7.	What are, in your opinion, the possible problems/obstacles regarding the use of online laboratories in your country/region?	I think internet connection is quite difficult in our region by low speed available in rural areas
8.	Will you encourage teachers within schools in your country/region to investigate the use of online laboratories?	No doubt, after the experience with IES David Buján, we can consider to transfer the experience to other schools.

9.	Is there anything else you would like to add or comment on?	The educational administration of Galicia, through the General Department of educational innovation on charge of Mr. Manuel Vila, is a strong supporter of projects, implementing new educational use in schools. Therefore, this project would have an excellent reception in its work plan.

15.2.3 Conclusions

In the above interviews a variety of countries are represented. Despite headmasters different backgrounds, socio economic conditions and national curricula, some common characteristics and observations appear in the way they have perceived Go-Lab implementation within their schools. More specifically:

- **Knowledge of online laboratories:** Most of the Headmasters had none or limited pre-knowledge related to the existence and use of online laboratories. Many Headmasters stated that Go-Lab offered them a better understanding of what online laboratories are and how they can be used and integrated in the classrooms. They admit that using Go-Lab helps convert every classroom into a laboratory and they understand the possibilities they open to learning.
- **Use of IBL:** Inquiry Based Learning remains a difficult topic for schools around Europe. Go-Lab with its ILSs offers teachers the possibility to try, in an easy and structured way, the use of IBL with their students. Once this is done and after reflecting on the impact these activities have on students' engagement and motivation, teachers and their headmasters seem to be more open to IBL and are willing to receive more training and try more activities based on it.
- **Implementation problems:** The most common implementation problems that schools faced, and have interfered with the use of Go-Lab, were the lack of time and the instability of the available internet connection. These are common problems that have been reported extensively from teachers as well.
- **Headmaster/School intentions:** The exposure to Go-Lab has a positive impact on Headmasters perception and attitude when it comes to the use of IBL and online laboratories. All Headmasters are willing to follow up on the use of Go-Lab by training more teachers to its use, encourage multidisciplinary implementations and encourage the use of Graasp and ILSs in combination with other projects and activities (i.e. eTwinning).

When it comes to policy makers, we can make the following observations:

- **Go-Lab awareness**: Both Ministries' representatives are aware of Go-Lab and its connection to online laboratories. Due to their professional connection with specific schools that also happen to be Go-Lab Pilot schools, they have experienced in action project's impact on teachers and students.

- **Interest on online laboratories:** Both Ministries' representatives share an interest on online laboratories and their use within schools. They both recognise that the project and the access to online laboratories has a lot of advantages both teachers and students so supporting the project and its IBL approach to its implementation in schools, complements their priorities within the Ministries.
- **Possible problems:** The lack of internet connection or its instability are the only problems, in these specific countries, that can hinder the Go-Lab implementations. In Portugal for example, high speed internet connection is widely available but smaller, remote villages still face connection problems.
- **Future support:** Because of their position and mandate both Ministries are interested in supporting the project on country level and they were both asked to be kept informed on any future developments.

15.3 Case studies

In the context of Go-Lab, the purpose of the case-studies are to collect information and evidence about the use of the Go-Lab activities in schools across Europe. WP8 is keen to understand how this experience was for teachers and students but also what in their opinion, were the benefits and drawbacks of using the activities. The outcomes and impacts the activities had on students and on teachers' teaching methods plus the impact on the attitudes and perceptions of other parties (like parents, non-STEM teachers, career counsellors) will also be investigated.

Case studies will be launched in November 2015 and will take place in collaboration with WP7 and WP6 in a joined effort to obtain information. National Coordinators will be provided with a specific template and instructions that will be asked to share with their teachers. Volunteers will fill in the questionnaire and provide, if available, additional evidence of the implemented activities i.e. photos, videos etc.

The full case study template can be found in Appendix 1 - Go-Lab Case study protocol template.

Case studies will be collected on country level and the aim is to collect at least one case study per country that will provide us with an insight on how the Go-Lab implementation is organised within schools.

15.4 References

Samarapungavan, A. (1992). Scientists' conceptions of science: A study of epistemic beliefs. *Paper presented at the Annual Meeting of the American Educational Research Association, San Francisco*, April 1992.

Wastiau, P., Blamire, R., Kearney, C., Quittre, V., Van de Gaer, E., & Monseur, C. (2013). The use of ICT in education: A survey of schools in Europe. *European Journal of Education*, *48*(1), 11-27, 122.

Appendix 1 - Go-Lab Case study protocol template

Go-Lab

Global Online Science Labs for Inquiry Learning at School

Collaborative Project in European Union's Seventh Framework Programme

Grant Agreement no. 317601

WP8 Evaluation

Case study template

© 2015, Go-Lab consortium

Legal Notices

The information in this document is subject to change without notice.

The Members of the Go-Lab Consortium make no warranty of any kind with regard to this document, including, but not limited to, the implied warranties of merchantability and fitness for a particular purpose. The Members of the Go-Lab Consortium shall not be held liable for errors contained herein or direct, indirect, special, incidental or consequential damages in connection with the furnishing, performance, or use of this material.

The information and views set out in this deliverable are those of the author(s) and do not necessarily reflect the official opinion of the European Union. Neither the European Union institutions and bodies nor any person acting on their behalf may be held responsible for the use which may be made of the information contained therein.

Introduction

Thank you for agreeing to participate in the Go-Lab case study.

The purpose of the case-study is to learn about the classroom use of the Go-Lab elements in schools across Europe. We are keen to understand how this experience was for you and for your students and what, in your opinion, were the benefits and drawbacks of using those elements. We are also interested in learning the outcomes and impacts this experience had on students and on your teaching work. For these purposes, we ask you to collect multimedia records, texts and other types of evidence related to the implementation of Go-Lab in your school.

General Guidelines: Taking into consideration the variety of elements that are being used in the framework of Go-Lab, different data collecting facilities available to participating schools and the limited timeline, we keep the design of case studies flexible and open for ad hoc adjustments.

There are a number of key questions/themes that a case study needs to look into, however this information may be collected in different ways. Below you will find a list of these themes as well as a brief description of what kinds of data could be gathered. This is followed by a table that lists the research themes/questions together with the suggestions of how you might capture the evidence for each of them. We leave it to you to decide what methods to use and how much of evidence to collect on your own.

Please use this table to record what evidence you have assembled and send it back to us along with the collected data.

When we receive the data that you have kindly collected, we might wish to call you to discuss your experience in detail and to fill in the missing bits of information, if any are identified.

Research themes/questions:

1. Background information about you, your school and students that participated in Go-Lab
2. Why did you choose this laboratory/ILS?
3. Did you have to adapt the ILS in any way? If yes, what did you do?
4. How did the implementation of the ILS go within your classroom?
5. How did the students behave during the whole process and what did they learn?
6. What was good about the ILS and what were the drawbacks?
7. Would you do it again and would you recommend it to your colleagues?

Evidence: The evidence that we would like you to collect may come in a number of formats:

- multimedia (video or audio recordings, photos)

Please ensure that you have collected parental consent forms for all of the students whose faces are visible on the photo or video records that you produce

- text (written narratives and quotations of students' views)
- other evidence (samples of outcomes for students, i.e., copies of students' work, posters, pictures of other tangible outcomes)

Please continue with the table below.

Themes and evidence: Please, collect data on each theme/question in the way that best suits you. We do want to make sure that EACH THEME/QUESTION on the list will have ***some*** **form of evidence.**

15.5 Case study Template

Your Name:	**School:**	**ILS:**	**Date:**
			Answers (text):
Research Themes/Questions		Types of Evidence One type of evidence per each theme is enough (Tick the appropriate box when reporting back)	
1. Background details about **1a. your teaching background, your school and your students** **1b. other staff members in your school involved in the Go-Lab activity** **1c. any contact with representatives of organisations who created this laboratory**		LI Multimedia recording of you talking about the background details LI Text (written account of the background) LI Will discuss in a follow-up telephone call	
2. Why did you choose this laboratory/ILS? **2a. Did you have to adapt the ILS in any way?** **2b If yes, what did you do and how?**		LI Video or audio recorded testimony LI Text LI Will discuss in a follow-up telephone call	

3. How did the implementation of the activity go? **3a. the actual process of using the Go-Lab element in your classroom, 3b. students' behaviour and response to the practice**	LI Video recording of using the activity in your class LI Photos capturing the key moments of the practice LI Written description of the process	
4. Outcomes for students **4a. Students talking about their experience and if they have enjoyed the activity** **4b. Students describing what they have learned**	LI Video/audio recording of using the activity in your class LI Written description of the process LI Sample (copies) of students' work, posters or other outcomes of the practice	
5. What was good about the ILS/laboratory you have used and what were the drawbacks? **6. Would you do it again and would you recommend it to your colleagues?**	LI Video/audio recording of using the activity in your class LI Written description of the process LI Sample (copies) of students' work, posters or other outcomes of the practice	