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 D8.1

Validation and evaluation plan and evaluation matrix

Editors	Ton de Jong (UT), Evita Tasiopoulou (EUN), Zacharias Zacharia (UCY)
Date	22-07-2014
Dissemination Level	Public
Status	Final

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 Glamorgan	UoG	United Kingdom
18	Institute of Accelerating Systems and Applications	IASA	Greece
19	Núcleo Interactivo de Astronomia	NUCLIO	Portugal

Contributors

Name	Institution
Evita Tasiopoulou	EUN
Gina Mihai	EUN
Ton de Jong	UT
Siswa van Riesen	UT
Ellen Kamp	UT
Bas Kollöffel	UT
Henny Leemkuil	UT
Effie Law	ULEIC
Zacharias Zacharia	UCY
Costas Manoli	UCY
Nikoletta Xenofontos	UCY
Anna Fiakkou	UCY
Margus Pedaste	UTE
Leo Siiman	UTE
Fani Stylianidou	EA
Diana Dikke	IMC (peer review)
Pablo Orduña	Deusto (peer review)

Legal Notices

Executive Summary

This deliverable presents the structure for the evaluation activities in Go-Lab. It identifies the interventions, stakeholders, and outcomes that make up the Go-Lab evaluation matrix. The core of this deliverable is a discussion and presentation of the evaluation instruments that will be used in the evaluation to measure the impact (outcomes) of the Go-Lab interventions for the different stakeholders. For students these measures concern their knowledge, inquiry skills, attitude, motivation, and their understanding of the nature of science. Teachers are evaluated on their technological, pedagogical and content knowledge, technological skills, skills, understanding of inquiry and belief in the efficacy of inquiry and teachers' attitudes and intentions towards Go-Lab. Organisations are measured on knowledge and skills. A number of measuring instruments were developed in the project, in which case several rounds of design have been applied. Others were selected from existing measuring instruments. In the latter case this deliverable presents the selection process. For all measures the underlying rationale and the literature that was consulted is presented. The deliverable ends with a structured overview of the evaluations that will be carried out in the Go-Lab project.

Table of Contents

1 Introduction

The main objective of WP8 is to assess the impact of the major Go-Lab interventions on the participating school communities (organisations), teachers, and students. Go-Lab "interventions" refers to the use of different products of Go-Lab by our stakeholders (e.g., ILSs for students, Golabz for teachers). In order to assess the impact of these interventions, a set of evaluation criteria, constructs, and instruments needs to be developed. Through this deliverable WP8 aims to demonstrate the methodology and the background discussions that will lead to the development of the required constructs and instruments. The deliverable also provides a timeline for the foreseen activities.

At this phase of the project, Phase B, the evaluation that WP8 will carry out is quite different from the one that WP3 is implementing. WP8 evaluates the impact that the use of the Go-Lab interventions has on teachers, students, and their schools whereas WP3 evaluates the formative usability and user experience of these interventions.

The instruments constructed for both evaluations in both WPs will be closely monitored by both WP3 and WP8 in order to minimize overlaps and ensure that research questions for both WPs will be answered. However, WP8 maintains a distance from the development of Go-Lab activities, recruitment of schools and implementation of activities in schools in order to ensure the independence of the evaluation, particularly in the analysis stages. The evaluation criteria and the principles of effective evaluation have formed a basis for the subsequent development of evaluation instruments.

Work in WP8 concerns the evaluation of the use of "final" products of Go-Lab and focuses on the "outcomes" of these Go-Lab interventions. These outcomes can be measured at different levels (cognitions, motivation, attitudes) and for different stakeholders (organisations, teachers, students). This approach suggests a three-dimensional model in which all combinations can be made. These dimensions are:

a) stakeholder (students, teachers, organisations (schools/policy makers))
b) outcomes (cognitions (knowledge/inquiry skills/understanding of the Nature of Science), motivation, attitudes), and
c) interventions (inquiry learning spaces, guidance (with a focus on scaffolds), scenarios/lesson plans, and Golabz (portal)).

In practice some of the cells will be empty and of the cells that are filled we will focus on a number of specific cells. In the first following section we identify the "interventions" that will be evaluated. For each of these interventions we indicate which stakeholder is related to the intervention and what aspects are being measured. These combinations make up a series of general research questions that will be presented. In the following sections we present the theoretical considerations concerning the construct that is measured per stakeholder which leads to the actual test that we have chosen or created to measure the construct. The actual tests are presented in a series of appendices. We will end this deliverable with the methodology used and a timeline for the evaluation studies.

2 Go-Lab interventions

In Go-Lab we can distinguish a few main elements that the Go-Lab stakeholders will encounter and that are supposed to influence them. One element one may expect here and that is pivotal to Go-Lab is the online lab. In our evaluations, however, we will not focus on the online lab as such. It is known from the literature that pure labs/simulations are not effective and guidance is needed to gain results (e.g., de Jong, 2006a). Assessing the design and workability of an online lab is more a usability issue. Therefore, in our evaluation we focus on labs embedded in *guidance (complete ILSs) and lesson plans or on specific elements of the guidance*. These two Go-Lab interventions (guidance and ILSs) mainly affect the learner. The two main interventions that affect the teacher are Golabz, the portal (repository) where teachers can search for labs, apps (scaffolds), and ILSs, and the Go-Lab authoring facility that enables teachers to adapt existing ILSs or create completely new ILSs. Organisations are affected by the use of ILSs and lesson plans by their teachers and by the use of the Go-Lab portal with all its facilities (authoring, sharing etc.).

2.1 Guidance

Guidance is key to successful inquiry environments including online labs (see e.g,., d'Angelo et al., 2014). In Go-Lab we have chosen to have guidance of several types following de Jong and Lazonder (2014). The types we used are (see de Jong, 2013, pp. 15-16):

- **Process constraints:** Process constraints aim to reduce the complexity of the inquiry learning process by restricting the number of options students need to consider (e.g., offering simplified equipment).
- **Dashboard:** A dashboard provides the student with a (graphical) overview of inquiry actions (e.g., number and type of variables manipulated) or product aspects (e.g., quality of a concept map).
- **Prompts:** Prompts are reminders or instructions to carry out a certain action or learning process.
- **Heuristics:** Heuristics give students general suggestions on how to perform a certain action or learning process.
- **Assignments:** Assignments are exercises that explain students what actions to perform.
- **Scaffolds:** Scaffolds are tools that help students perform a learning process by supporting the dynamics of the activities involved. An example is a scaffold is a tool that helps the student to create an experimental design.
- **Direct presentation of information:** Offering of information that should have been the result of the inquiry process (but was not found by the student).

An overview of examples of guidance can be found in (Zacharia et al., submitted). In Go-Lab we have chosen a specific form of an overall process constraint in the form of a specific inquiry cycle (see Pedaste et al., submitted). Though we know that guidance overall is needed for an effective learning process (in terms of conceptual knowledge) much knowledge still needs to be gathered on the contribution of individual forms of guidance on knowledge acquisition and on the circumstances under which specific scaffolds are successful. Evaluation in Go-Lab will therefore focus on types of guidance found in Go-Lab with a focus on Go-Lab specific scaffolds.

- The main stakeholder here is the *student* and the main outcomes are *knowledge* and *inquiry skills*.

2.2 Inquiry Learning Spaces (ILS) and lesson plans

An Inquiry Learning Space (ILS) is the learning environment that offers students a combination of an online lab, guidance, resources, general tools, and communication facilities (see de Jong, 2013 for an overview of the design of ILSs and their components). In fact, an ILS can present the student a complete learning experience. However, next to the online material also offline material can be presented. The off line materials are added in what we have called "lesson plans". We may therefore evaluate the effect of an ILS or of the lesson plan that includes the ILS.

The literature shows that remote and virtual labs, if embedded in guidance, are just as, or even more effective for learning than hands-on laboratories for the gain of conceptual knowledge as well as for the acquisition of inquiry skills (see for example, de Jong, Linn, & Zacharia, 2013). Still we intend to evaluate the effects of complete ILS (or lesson plans) for students in comparison to other (more traditional) ways of instruction because we now have the opportunity to do the same study at different places. We will call this process *concurrent replication*. Another opportunity might be to do more *longitudinal research* and follow students who have been working with more than one ILS/lesson plan. A final reason to evaluate the effects of ILS/lesson plans is that most existing studies focus on knowledge acquisition, whereas here we will also measure inquiry skills, attitudes, motivation, and knowledge of NoS (Nature of Science).

Teachers will be using ILSs and lesson plans in their lessons. We assume this may change the teachers' technological, pedagogical, and content knowledge, their understanding and efficacy beliefs of inquiry, and their beliefs, attitudes, and intentions towards Go-Lab.

The use of ILSs and the availability of Go-Lab may also affect the organisation. By using ILSs organisations/policy makers may change their view on the mode of teaching and might become more favourable towards inquiry/online labs.

- The first main stakeholder here is the *student* and the main outcomes are *knowledge, inquiry skills, attitudes, motivation, and NoS.*
- The second main stakeholder here is the *teacher* and the main outcomes are *technological, pedagogical, and content knowledge, understanding and efficacy beliefs of inquiry, and beliefs, attitudes, and intentions towards Go-Lab.*
- The third main stakeholder here is the *organisation* (schools/policy makers) and the main outcomes are *attitudes and motivation.*

2.3 Golabz and the Go-Lab portal

Golabz (http://www.golabz.eu/) is the repository that displays all labs, apps, and ILSs (and associated lesson plans). The portal is also used to display the idea of a federation of labs following the organisation in a number of so-called "big ideas" (see, Zervas, 2013). At Golabz teachers search for labs and ILS and if they adapt or create ILSs they may search for specific scaffolds (apps). The Go-Lab portal, including the repository, will also offer all kinds of social facilities (liking etc.) and will offer the possibility of sharing ILSs. Therefore, the Go-Lab portal is mainly visited by teachers, but having Golabz available in a school may also affect the organization.

- The first main stakeholder here is the *teacher* and the main outcomes are *knowledge (TPACK and inquiry) inquiry efficacy, and beliefs, attitudes*, *and intentions towards Go-Lab.*
- The second main stakeholder here is the *organisation* (policy makers) and the main outcomes are *attitudes and motivation*

2.4 Go-Lab authoring facility

Go-Lab offers teachers the opportunity to build their own ILSs and lesson plans. In doing so they are supported by scenarios and default lesson plans (de Jong, 2014). A Go-Lab scenario describes, in a *domain independent way*, all activities, materials, and interactions for teachers and students that comprise a complete (online and offline) Go-Lab inquiry learning experience. As is described in de Jong (2014) “scenarios differ in activities included and in the combination of a) offline and online activities b) individual or collaborative actions c) distribution of activities over teachers and system, and c) sequencing of activities”. A scenario can be accompanied by a default lesson plan that has predefined texts in the different phases of the inquiry cycle included. The authoring process is further supported by an authoring facility that is integrated in GRAASP (Govaerts, 2013). The authoring facilities are part of the more comprehensive Go-Lab portal but since this is a specific part used by a subset of teachers we will regards this as a separate intervention.

- The only stakeholder in this intervention is the teacher. Working with the Go-Lab authoring facility may affect teachers’ *TPACK, understanding and efficacy beliefs of inquiry, and beliefs, attitudes, and intentions towards Go-Lab.*

The further improvement of the usability process of authoring on the basis of user feedback is the subject of Go-Lab WP3.

3 Research questions

Evaluation in Go-Lab is guided by a set of basic research questions. In the previous sections these questions have been stated implicitly and sometimes more explicitly. In this section we summarise these questions now from the perspective of the different stakeholders.

Questions aimed at students

- how do different types of Go-Lab guidance (e.g., scaffolds or different forms of the same scaffold) affect the gain of (different types of) and/or knowledge,inquiry skills?
- how do inquiry learning spaces compared to other forms of instruction affect the acquisition of knowledge, inquiry skills, attitudes, motivation, and understanding of the Nature of Science (NoS)?

Questions aimed at teachers

- do teachers have the required knowledge (technological, pedagogical, and content, TPACK) and technical skills to work with Go-Lab elements in an effective way?
- how does the use of ILSs and associated lesson plans affect teachers' understanding of inquiry learning and their belief in its efficacy, and their beliefs, attitudes and intentions towards Go-Lab?
- how does the use of the Go-Lab authoring facility affect teachers' TPACK and their knowledge of and attitude towards inquiry and Go-Lab?
- how does consulting the Go-Lab portal (www.golabz.eu) affect teachers' knowledge (e.g., the big ideas) and their attitudes and motivation towards inquiry learning and Go-Lab?

Questions aimed at organisations

- how does the use of ILSs and associated lesson plans in a school (parents/career counsellors/non-MST teachers) affect the school's' attitudes (awareness) and motivation towards inquiry learning in general and online labs in particular?
- how does the availability of the Go-Lab portal (www.golabz.eu) in a school affect the schools' attitudes (awareness) and motivation towards inquiry learning and online labs?
- how does the availability of Go-Lab affect daily practice in a school and how does it affect the attitude of the school towards STEM?

A few of the questions have to do with measuring the conditions for a successful use of Go-Lab (measuring teachers' capabilities necessary for the use of Go-Lab) but most of them are meant to measure the effects of the use different elements of Go-Lab (the Go-Lab "interventions").

These questions are the overall leading questions in Go-Lab evaluation and will be specified for each specific study or set of studies. For example, for a specific study the Go-Lab guidance will be determined, the ILS(s) used, and the specific outcomes aspects as measured.

4 Student evaluation

Students who learn with Go-Lab material, more specifically Go-Lab ILSs, will in the first place intend to gain knowledge of the domain that is involved. In close relation to that we expect students also to develop their understanding of inquiry as this is reflected in their inquiry skills. Next to these knowledge and skills issues we expect that exposure to Go-Lab ILSs also affects students attitude and motivation towards science and their understanding of how science works, in other words their understanding of the Nature of Science (NoS). These concepts and the ways we will measure them in the Go-Lab evaluations will be discussed in the next sections.

4.1 Knowledge

One of the aims of Go-Lab is to let students learn about science topics and to gain domain knowledge. Within Go-Lab we have defined educational objectives for each lab and ILS which are based on Bloom's revised taxonomy (Anderson & Krathwohl, 2001). This revised taxonomy has been adjusted for Go-Lab (see the Go-Lab Deliverables 2.1 and 1.3: de Jong, 2014; Zervas, 2013) and defines different types of knowledge and cognitive processes. Students reach these educational objectives by participating in the inquiry tasks that are offered in ILSs, guided thoroughly by different forms of guidance.

Many studies have investigated the effect of inquiry learning on knowledge gain. Scientific evidence supports the assumption that inquiry learning leads to better acquisition of conceptual knowledge (Alfieri, Brooks, Aldrich, & Tenenbaum, 2011; Deslauriers & Wieman, 2011; Eysink et al., 2009; Kollöffel & de Jong, 2013; Prince & Felder, 2006; Zacharia, 2007). A recent meta-analysis of 138 studies showed that inquiry-based instruction is associated with more conceptual understanding than other instructional practices (Minner, Levy, & Century, 2010). Eysink, et al. (2009) compared inquiry learning, self-explanation-based learning, hypermedia learning, and observational learning and showed that inquiry learning led to a higher knowledge gain compared to hypermedia learning and observational learning and to a similar amount compared to self-explanation-based learning. Eysink, et al. (2009) not only took conceptual knowledge into account, but also focused on procedural (including near and far transfer), intuitive, situational knowledge and found similar results for all four of these type of knowledge. In a meta-analysis study by Carolan, Hutchins, Wickens, and Cumming (2014) it was found that inquiry learning and similar approaches are associated with superior scores on transfer compared to other forms of active learning.

However, inquiry learning only proves to be effective when students receive adequate guidance (de Jong, 2006b). Go-Lab offers several forms of guidance as described in Section 2.1. This guidance and a few carefully selected anchor ILSs will be evaluated to determine their effect on knowledge gain. The evaluations are aimed at different types of knowledge and will take different age groups into account.

There are no standard knowledge tests that can be used to evaluate knowledge gain, since the knowledge investigated is domain dependent. Knowledge tests are developed specifically for the ILSs that are selected for evaluation and carefully validated before they are implemented. The guidance that is evaluated also cannot stand on its own, and is evaluated in a specific domain. The selected ILSs not only differ in domain, but also in targeted age group and assumed prior knowledge. These aspects are taken into account for the development of our instruments.

The meta-data of an ILS is used to determine which type of knowledge and which processes are selected for evaluation of a specific ILS. This meta-data contains information about its educational objectives for each type of knowledge and specifies which cognitive processes are targeted.

In Bloom's revised taxonomy, four different types of knowledge are distinguished (see Table 1).

Type of knowledge	Description
Factual	Knowledge of basic elements, e.g., terminology, symbols, specific details, etc.
Conceptual	Knowledge of interrelationships among the basic elements within a larger structure, e.g., classifications, principles, theories, etc.
Procedural	Knowledge on how-to-do, methods, techniques, subject-specific skills and algorithms, etc.
Meta-cognitive	Knowledge and awareness of cognition, e.g., of learning strategies, cognitive tasks, one's own strengths, weaknesses and knowledge level, etc.

Table 1. Cognitive Objectives: Types of Knowledge (Anderson & Krathwohl, 2001)

In our tests we will not focus on factual knowledge or meta-cognitive knowledge but instead focus on conceptual and procedural knowledge.

In the literature, *conceptual knowledge* is often defined in terms of knowledge about interrelations between concepts (see e.g., Hiebert & Lefevre, 1986; Krathwohl, 2002; Rittle-Johnson, Siegler, & Alibali, 2001). However, we consider knowledge about the interrelations between concepts, principles, and so on, as a *quality* (cf. de Jong & Ferguson-Hessler, 1996) of conceptual knowledge, reflecting a deeper, more advanced level of conceptual knowledge. Learning starts with only a very limited amount of conceptual knowledge, or even no conceptual domain knowledge at all, and from there it can evolve and develop into a more elaborate, compiled, and automated knowledge base, with knowledge about the (qualitative and/or quantitative) interrelations between concepts. The process of building and developing a conceptual knowledge base is assumed to be a gradual, step-by-step process. It often begins with memorization and rote learning, but with sustained learning and practice, more elaborate, integrated, compiled, and automated knowledge can be developed. Snow (1989) described the desired end states of learning in terms of "articulated, deep understanding of a domain, including the ability to reason and explain in causal terms, and to adopt multiple viewpoints about a problem or phenomenon" (p. 9). Deep-level knowledge is associated with comprehension and abstraction, with critical judgment and evaluation, and with flexibility.

Some theorists argue that development of conceptual knowledge and other forms of knowledge and skills can mutually support and stimulate each other. Such processes are called bootstrapping (Schauble, 1996) or iterative knowledge development (Rittle-Johnson, et al., 2001) that is the idea that an increase in one type of knowledge facilitates an increase in the other type of knowledge, which facilitates an increase in the first, and so on. In particular it is assumed that there must exist interrelations between the development of procedural and conceptual knowledge. *Procedural knowledge* contains actions or manipulations that are valid within a domain. This type of knowledge enables problem solvers to make transitions from one problem state to another (de Jong & Ferguson-Hessler, 1996). It is knowledge about how to do something, about domain-specific skills and algorithms, about methods of inquiry, and about criteria for determining when to use appropriate procedures (Krathwohl, 2002). The development of conceptual and procedural knowledge might stimulate or catalyze each other. For example, conceptual knowledge helps learners to recognize and identify key concepts when studying or diagnosing a problem. As a result, a better conceptual understanding of the problem will increase the likelihood that the learner will select the appropriate problem solving procedure (enhancing procedural skills). In turn, reflecting on or self-explaining the conceptual basis of procedures can help learners to become aware of which concepts play a key role in a problem (Rittle-Johnson, et al., 2001). Some evidence for the existence of bootstrapping processes is found in the domain of mathematics and recently, in a study by Kollöffel and de Jong (2013) in the domain of electricity theory. Their results showed strong effects of simulation-based inquiry learning in the domain of electricity theory on the acquisition of both conceptual and procedural knowledge, which was remarkable because the instruction focused on the acquisition of conceptual knowledge but not on the acquisition of procedural knowledge. Analyses of errors suggested that inquiry learning with simulations enhanced the students' ability to analyse circuit problems and this in turn lead them to be more adequate in selecting appropriate problem solving procedures.

As mentioned in this section we will distinguish, apart from the type of knowledge, qualities of knowledge that we like to measure in Go-Lab. The main two qualities are the structure of the knowledge and its level, in other words how "deep" the knowledge is. These two qualities are further elaborated in the next two sections.

4.1.1 Knowledge structures: schemas and Big Ideas

The structure of knowledge is an important quality of knowledge (de Jong & Ferguson-Hessler, 1996). This has partly to do with how the knowledge base is organized and integrated. The integratedness of domain knowledge is best described as structural knowledge (Jonassen, Beissner, & Yacci, 1993). Structural knowledge is the knowledge of how concepts within a domain are interrelated (Jonassen, 2000). The structure can be described for example in terms of its organization in schemata or in terms of its hierarchic organization. To start with the latter, a schema contains the different types of knowledge (conceptual and procedural) required for task performance. Domain knowledge includes several schemas. Knowledge can also be described in terms of its hierarchical organization, that is lower-level, domain-specific knowledge connected to higher-level, broader, more general knowledge. The higher-level knowledge can still be related to the domain, but it may also include knowledge and concepts that transcend the domain of instruction, such as what is called the Big Ideas of Science. Big ideas are overarching concepts that connect multiple concepts, procedures, or problems within or even across domains or topics (Baroody, Cibulskis, Lai, & Li, 2004). According to Baroody, Feil, and Johnson (2007), big ideas are integral to achieving a deep understanding of both concepts and procedures. Baroody and his colleagues argue that big ideas invite students to view knowledge "*as cohesive or structured rather than as a series of isolated procedures, definitions, and so forth. In particular Big Ideas invite students to look beyond surface features of procedures and concepts and see diverse aspects of knowledge as having the same underlying structure*" (p. 126). There are various ways of assessing structural knowledge, such as constructing graphical representations including concept maps and graphs, but also verbal assessments such as essay questions. Perhaps the most common one is the use of concept mapping. Students can use a concept map to construct an external representation of the key concepts and variables and use arrows to indicate the interrelations between them. This may apply to interrelations between low-level, domain-specific concepts, but also to the interrelations between domain-specific concepts (small ideas) and big ideas. Constructing a graph depicting the relationships between concepts and variables can also be effective for assessing structural knowledge (Jonassen, 2000), although this may be limited to domain-specific knowledge rather than connecting small and big ideas. An alternative form of assessing structural knowledge can be achieved by using essay questions, for example by asking students to describe factors that influence a certain phenomenon by identifying important concepts and describing the relations between those concepts (Gijlers & de Jong, 2013). Assessing the students' structural knowledge of big ideas in science is beyond the aims and scope of the Go-Lab project, but is definitely worthwhile pursuing in the future. The focus in Go-Lab will be on assessing domain understanding including structural knowledge and conceptual and procedural knowledge. But knowing if students acquire these types of knowledge is not enough, we also need to assess how deep their knowledge is. This will be discussed in the following section.

4.1.2 Depth of knowledge

In the introduction of this section on knowledge we mentioned deep-level knowledge being associated with comprehension and abstraction, with critical judgment and evaluation, and with flexibility. Here, we will use a revised version of Bloom's taxonomy to assess the depth of knowledge (Anderson (Ed.) et al., 2001; Krathwohl, 2002). The original taxonomy was published in 1956 and aimed at the cognitive domain (Bloom, Engelhart, Furst, Hill, & Krathwohl, 1956). Bloom and colleagues also developed taxonomies for the affective domain and for the psychomotor domain. The focus here will be on the cognitive domain. In the revised taxonomy six knowledge levels are distinguished: Remember ("Knowledge" in the original taxonomy), Understand (previously: Comprehension), Apply, Analyse, Evaluate, and Create (previously: Synthesis). The processes are thought to reflect a hierarchy based on difference in complexity, but the strict hierarchy from the original taxonomy was abandoned, leaving room for categories to overlap each other (Krathwohl, 2002).

4.1.2.1 Levels of knowledge

The revised taxonomy has been adjusted for Go-Lab (see the Go-Lab deliverables 2.1 and 1.3: de Jong, 2014; Zervas, 2013). Four categories are distinguished instead of six (see Table 2).

Process	Description
To remember	To help the student recognize or recall information
To understand	To help the student organize and arrange information mentally
To apply	To help the student apply information to reach an answer
To think critically and creatively	To help the student think on causes, predict, make judgments, create new ideas

Table 2. Cognitive Objectives: Processes

The most basic level is the *Remember*-level. At this level the student is able to recall previously learned material, relevant terminology, specific facts, or procedures related to information and/or course topics. The next level is the *Understand*-level. At this level, a student is able to grasp the meaning of information (facts, definitions, concepts, etc.) that has been presented. Then follows the *Apply*-level. At this level, a student is able to use previously acquired knowledge and/or skills in new or different situations or in problem solving. The highest level is about *Thinking critically and creatively*. In fact, this level merges the three highest levels of Bloom's revised taxonomy, the Analyze-, Evaluate-, and Create-level. At the (merged) level, a student is able to: break information down into its constituent parts and detecting how the parts relate to one another and to the overall structure or purpose, judge the value of information and/or sources of information based on criteria and standards, and creatively or uniquely apply prior knowledge and/or skills to produce new and original thoughts, ideas, processes, etc.

4.1.2.2 Applying the levels of knowledge to inquiry learning

Papadouris and Constantinou (2009) argued that the accumulation of experiences with natural phenomena through active exploration, investigation, and interpretation provides a basis for developing conceptual knowledge. In inquiry learning, students learn through exploration and application of scientific reasoning. They are actively engaged in finding out what is happening instead of just witnessing something being presented. They need to make predictions, design experiments, analyze and interpret the collected data, and formulate answers to their research questions (see e.g., Chi, Slotta, & de Leeuw, 1994; Jaakkola & Nurmi, 2008; Kollöffel & de Jong, 2013; Muller, Bewes, Sharma, & Reimann, 2008; Strike & Posner, 1985; Tao & Gunstone, 1999; Trundle & Bell, 2010; Zacharia, 2007). These processes and activities involved in inquiry learning require at least some knowledge at the Remember and the Understand level. At the Remember-level, the student can remember concepts, but may not really understand them. At the Understand-level, the student can, for example, explain or restate concepts. These qualities of knowledge are required to be able to start thinking about the concepts in a domain. This is the point where the student can enter the third level, the Apply-level. At this level, the student can for example recognize and identify concepts in problems and situations. The problems and situations can be identical or different from those used in the training situation. When students have to apply their knowledge to problems or situations that are identical to those in the training situation, this is called *identical transfer*. If the task is different, but similar (analogical) it is called *near transfer*, and if the task is new, structurally different or more complex, it is called *far transfer* (Carolan, et al., 2014). The Apply-level is a pre-requisite for entering the fourth level. At this level, the student can examine and assess concepts in order to draw conclusions in terms of cause and effect, make inferences, or find evidence to support statements/arguments, justify and defend his/her ideas about a concept, present arguments in a convincing way, find errors, evaluate and judge information from others about concepts, create their own thoughts and ideas about concepts and apply them in new ways.

4.1.2.3 Integrating inquiry learning and the knowledge levels

Ideally, there are strong connections between the learning objectives, learning activities, and assessment in a course. This helps to assure the consistency and coherence of the instruction and assessment. The knowledge levels framework can be used to align the objectives, activities, and assessments. There are several lists available with active verbs or action words that are (more or less) typical for a specific knowledge level. These active verbs can be used to design, select, and/or shape learning objectives, activities, and assessments. This paper contains a list of suggested active verbs for each knowledge level (see Appendix 1).

4.1.3 The structure of knowledge assessment

In order to facilitate comparisons between the knowledge levels of students but also between and across domains, assessment instruments need to include items at all knowledge levels. Moreover, for each level there need to be sufficient numbers of items in order to assure valid and reliable measurements. In Table 3 a basic structure for assessment is presented.

Item	Sub-item	Knowledge type	Knowledge level	Example
Item 1 (Difficulty level: Fair)	a.	Conceptual	Remember	• Define concept A
	b.	Conceptual	Understand	• Give examples of A
	c.	Conceptual	Apply: Identical transfer	• Identify A in problem P
	d.	Procedural	Apply	• Calculate A
	e.	Conceptual	Think critically and creatively	• Predict what happens with A if R1 • Recommend using A or Z in situation S1. Argue. • Propose a new way to use A.
	f.	Conceptual	Apply: Near transfer	• Identify A in problem N1
	g.	Conceptual	Apply: Far transfer	• Identify A in problem F1
Item 2 (Difficulty level: Higher than Item 1)	a.	Conceptual	Remember	• Define concepts B and C
	b.	Conceptual	Understand	• Compare/Give examples of B and C
	c.	Conceptual	Apply: Similar transfer	• Identify B and C in problem P2
	d.	Procedural	Apply	• Calculate B and C
	e.	Conceptual	Think critically and creatively	• Predict what happens with B and C if R2. • Recommend using B or C in situation S2 Argue. • Propose a new way to use B and C.
	f.	Conceptual	Apply: Near transfer	• Identify B and C in problem N2
	g.	Conceptual	Apply: Far transfer	• Identify B and C in problem F2
Item 3 (Difficulty level: Higher than Item 2)	a.	Conceptual	Remember	• Define concepts D, E, and G
	b.	Conceptual	Understand	• Compare/Give examples of D, E and G
	c.	Conceptual	Apply: Similar transfer	• Identify D, E and G in problem P3
	d.	Procedural	Apply	• Calculate D, E and G

Item	Sub-item	Knowledge type	Knowledge level	Example
	e.	Conceptual	Think critically and creatively	• Predict what happens with D, E and G if R3. • Recommend using D, E, and G in situation S3 Argue. • Propose a new way to use D, E and G.
	f.	Conceptual	Apply: Near transfer	• Identify D, E and G in problem N3
	g.	Conceptual	Apply: Far transfer	• Identify D, E and G in problem F3

Table 3. Basic structure for assessment

The items presented in Table 3, are ordered by increasing complexity. Items 1a-g are fairly difficult, items 2a-g are more complex, and item 3 a-g contains the most complex items. If necessary, more items can be added. Moreover, it is not strictly necessary to maintain the order of the items, but keeping this order may be helpful for research purposes, such as analysing and diagnosing errors made by students. For example, the causes of errors/failures at the "Think critically and creatively"-level might be traced back to incomplete or erroneous knowledge at the Remember, Understand, or Apply-levels. In order to detect such patterns, the sub-items (a-i) need to build upon each other. The examples are mere examples, and they can be modified using the active verbs presented in the appendix. The most important thing is that each sub-item should have a clear connection with a specific level of knowledge. This structure might be used in different domains and the connections of domain-specific sub-items with the (generic) knowledge levels can facilitate comparisons across domains.

Repeated measures within domains can be used to detect developing knowledge. By using pre- and post-tests, the progress of students from basic levels on conceptual understanding towards more advanced levels can be detected. Perhaps, such models in combinations with formative assessments can provide data that can be used to adapt the learning activities and feedback provided during the learning process. By aggregating data from tests to the more general levels of conceptual knowledge, may allow comparison of progress across different domains.

Appendix 1 gives examples of knowledge items at different levels and for different knowledge types. Sets of "verbs" to construct these items are given to support the construction of new items for a specific domain.

4.2 Inquiry skills

Within Go-Lab students participate in inquiry learning. Inquiry learning is an educational approach in which students actively construct their own knowledge by participating in processes of inquiry similar to those of professional scientists (Keselman, 2003). Inquiry learning can be very effective for gaining inquiry skills (Alfieri, et al., 2011; Furtak, Seidel, Iverson, & Briggs, 2012; Minner, et al., 2010) and it prepares students for practicing science in their follow-up studies and careers. Within Go-Lab students can follow the Go-Lab inquiry cycle, or parts of it, by orienting themselves on the topic of inquiry, formulating research questions and/or hypotheses, planning and conducting an experiment, drawing conclusions, reflecting upon their inquiry process, and communicating their findings. However, research reveals that students find the processes involved in inquiry learning very difficult (de Jong et al., 1998; Kuhn, Black, Keselman, & Kaplan, 2000). In order for it to be an effective approach it is important to guide students (de Jong, 2006a). Go-Lab offers guidance, as explained in Section 2.1, for each inquiry phase to help students conduct sound experiments and acquire inquiry skills within an Inquiry Learning Space (ILS). Both the guidance that is offered, and a selection of typical ILSs are evaluated in terms of students' acquirement of inquiry skills. The evaluations are aimed at all phases of inquiry. Students of all Go-Lab age levels, 10-18 years, participate in the evaluations. The methodology, as explained in Section 5, to evaluate the increase in inquiry skills is carefully selected to fit the target group.

Inquiry skills are measured in different ways. In the conceptual model that is used in Go-Lab, the inquiry process is viewed as a strategic competency, entailing both epistemic and procedural aspects (P. M. Kind, 2013; Kuhn & Pease, 2008). In other words, students not only need to know how to perform inquiry operations, but also need the ability to evaluate why and when inquiry skills are performed appropriately. For example, a well-known strategy used by researchers to draw valid, causal inferences, is to Vary One Thing At a Time, also referred to as the Control of Variables Strategy (CVS). This skill involves both applying epistemic knowledge about dependent, independent and control variables and criteria for defining their cause–effect relationships, and it also involves understanding procedural knowledge about how to do "fair testing" (P. M. Kind, 2013).

This section provides an overview of the instruments that are used to evaluate the acquirement of inquiry skills of students working with ILSs and/or receiving Go-Lab guidance. Lists of existing, validated instruments and their characteristics were made. These were used to select the instruments that would meet the following Go-Lab requirements. First, as mentioned above, the instruments should be appropriate for target groups of different ages (10-18 years). Second, instruments were needed that allow relatively quick measurements of inquiry skills in large groups. These instruments should be complemented by instruments that allow more extensive, more detailed, in-depth measurements, perhaps at the cost of being applicable in large groups of students. Third, Go-Lab covers many science, technology, engineering and mathematics domains, and (prior) domain knowledge can bias the measurement of inquiry skills, so the assessments needed to be general rather than domain-specific. Fourth, although many measurement instruments can assess whether students possess inquiry skills, only very few of them can also assess the level of these skills. At least one instrument should enable assessing the depth of inquiry skills. Finally, the instruments should be designed in such a way that parallel tests could be constructed. The following inquiry skills tests were reviewed:

- Control of Variables Strategy test (CVS-test) (Chen & Klahr, 1999)
- Earthquake Forecaster (Kuhn & Pease, 2008)
- How science works (P. M. Kind, 2013)
- Lawson's Classroom Test of Scientific Reasoning (LCTSR) (Lawson, 2000)
- Paper towels and Bugs test (Alonzo & Aschbacher, 2004; Baxter, Shavelson, Goldman, & Pine, 1992; Shavelson, Baxter, & Pine, 1992)
- Performance of Process Skills (POPS) (Mattheis & Nakayama, 1988)
- SAVE Science (Timms et al., 2012)
- Science ASSISTments (Timms, et al., 2012)
- Science Process Skills Test (SPST) (Molitor & George, 1976)

- Science Reasoning Test (SRT) (Rifkin & Georgakakos, 1996)
- Scientific Inquiry Literacy Test (ScIngLiT) (Wenning, 2007)
- SimScientists Calipers II (Timms, et al., 2012)
- Test of Basic Process Skills in Science (BAPS) (Marshall, 1991)
- Test Of Graphing in Science (TOGS) (McKenzie & Padilla, 1984)
- Test of Integrated Process Skills (TIPS) (Dillashaw & Okey, 1980)
- Test of Integrated Process Skills II (TIPSII) (Burns, Okey, & Wise, 1985)
- Test of Integrated Science Process Skills (TISPS) (Beaumont-Walters & Soyibo, 2001; Maltheis et al., 1992)
- Test of Integrated Science Processes (TISP) (Tobin & Capie, 1982)
- Test of Science Processes (TSP) (Butzow & Sewell, 1972; Tannenbaum, 1971)
- Test of Scientific Literacy Skills (TOSLS) (Gormally, Brickman, & Lutz, 2012)
- The South Eastern Regional Vision for Education (SERVE) Science Process Skills Test (Turpin & Cage, 2004)

After reviewing the tests listed above, only few turned out to meet Go-Lab criteria: the paper towels and Bugs test (Alonzo & Aschbacher, 2004; Baxter, et al., 1992; Shavelson, et al., 1992) for primary education, the Test of Integrated Process Skills (TIPS) (Dillashaw & Okey, 1980) and Test of Integrated Process Skills II (TIPSII) (Burns, et al., 1985; Maltheis, et al., 1992) as short inquiry skills test for secondary education, and How science works (P. M. Kind, 2013) as an extensive inquiry skills tests for secondary education. The inquiry skills tests that have been selected on the basis of these criteria are described below.

4.2.1 Inquiry skills tests for primary education (age 10-12 years)

First, we will discuss the tests that aim at the youngest Go-Lab population, kids at higher levels of primary education. These children mostly have hardly if any understanding of jargon like “hypotheses”, “variables”, “controlled experiments”, and so on, but still, this population does demonstrate understanding of investigating the world around them. Some inquiry skills tests focus in particular on this target group. The inquiry skills of the Go-Lab population between 10-12 years old will be assessed by the "Paper Towels" and “Bugs”-tests (Alonzo & Aschbacher, 2004; Shavelson, et al., 1992). In the, "Paper Towels"-investigation, students determine which of three kinds of paper towels soaks up the most and least water. Following the Paper Towels investigation, students complete a "scientific notebook" in which they describe the steps taken in their investigation and the variables controlled. In the "Bugs"-investigation, students conduct experiments with bugs (in a computer simulation) to determine the bugs' preferences for various environments. Two experiments manipulate a single independent variable: dampness (damp or dry) and darkness (light or dark). The third experiment is a 2 x 2 factorial combining the dampness and darkness variables. For each of the Bugs experiments, students (a) draw a picture of their experimental setup, (b) describe the steps, and (c) explain the basis for arriving at conclusions. More similar investigations are available (Alonzo & Aschbacher, 2004).

More details about both tests for inquiry skills for young children are displayed in Appendix 2A. For this age level the test needs to be translated into the children’s native language.

4.2.2 Short inquiry skills tests for secondary education (age 12-18 years)

The first, quick, large-scale tests for age 12-18 are the Test of the Integrated Science Process Skills (TIPS) developed by Dillashaw and Okey (1980) and the TIPS II (Burns, et al., 1985), which is an extended version of the original TIPS, and will be used as a parallel test. These tests measure the following inquiry skills:

- Identifying dependent, independent, and controlled variables
- Operationally defining variables
- Identifying testable hypotheses
- Data and graph interpretation
- Experimental design

The test contains 36 multiple choice items, each with four alternatives. It is recommended for grades 7-12, and the time to complete the test is 25-50 minutes. Test reliability, Cronbach's Alpha, is 0.86. A sample test item from the TIPS-II is displayed in Figure 1. In this item, students have to select a testable hypothesis, given a description of variables involved in an investigation.

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

A. The more carbon dioxide a bean plant gets the more light it needs.
B. The more starch a bean plant produces the more light it needs.
c. The more water a bean plant gets the more carbon dioxide it needs.
D. The more light a bean plant receives the more carbon dioxide it will produce.

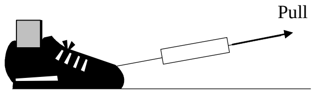
Figure 1. Sample test item TIPS-II (Burns, et al., 1985)

More sample items are displayed in Appendix 2B.

4.2.3 Extensive inquiry skills tests for secondary education (age 12-18 years)

The second, more extensive test that will be used, is the "How Science Works"-test developed by (P. M. Kind, 2013). This test contains combinations of multiple-choice (mc), open, and ordered mc questions. The test 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. For example, with regard to 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. In Figure 2, an example is presented of an item that is used to assess the students' reasoning using knowledge about uncertainty in measurement.

School experiment on friction
Daniel, Philip and Tom investigated how a trainer slips on different surfaces. They put a mass in the trainer and measured the pull (with a Newton-meter) needed to drag the trainer along. They tested each surface four times.



Here are their results:

Type of surface	Pull / force (Newtons)			
	First time	**Second time**	**Third time**	**Fourth time**
Playground	11	12	13	21
Grass	14	13	13	14
Classroom carpet	8	9	8	9
Polished floor	5	7	7	7

A) They thought they had done everything the same. Why didn't they get the same results each time they measured the same surface? (Tick one box each line)

	Agree	**Disagree**
a) They were not as accurate as they thought. Being more accurate the measurements would have been the same	☐	☐
b) Measurements never will be exactly the same, however hard you try to get it accurate	☐	☐
c) The surfaces must have got slippier each time they did their test	☐	☐
d) There must have been something wrong with the Newton- meter as they repeated the measurement	☐	☐

B) How should they decide which results they should use? (Choose one)
a) Add up measurements from all trials and divide by 4 to get the average
b) Take away irregular (odd) measurements, then get average among the rest
c) Choose measurements that are the same (occur several times)
d) Choose the lowest value, because this is the least force that was needed

C) The boys disagreed about the conclusion on which surface needed most force. Who do you agree with? Choose one and explain why.
☐ Daniel said the playground,
☐ Philip said you couldn't tell,
☐ Tom said the grass

This is because…..…………………………………………………………………………………

Figure 2. Example item of "How science works" test (P. M. Kind, 2013)

Item A in Figure 2 asked why measurements are not the same over four repetitions. Item B asked for a strategy to handle uncertainty in measurements. Here, two response alternatives give credit: students scored 1 point for alternative a, to average all measurements, and 2 points for alternative b, which suggested removing the irregular measurement before averaging the other measurements. Item C asks students to select the conclusion they most agreed with. This is thought to impose a higher cognitive demand on the students. It requires them to relate tabulated data and to compare results between fictitious students. The reliability scores, Cronbach's alpha, are between 0.78-0.83. More test items are displayed in Appendix 2C.

4.2.4 Connecting assessment data with other measures

The selected instruments all allow the construction of parallel tests. By using one version of the test at the beginning of the inquiry learning activities and the other at the end of it, possible progression of inquiry skills can be assessed. The data from the assessments can also be used for analyses of the learning outcomes and knowledge gain, Furthermore, they can be useful for explaining the students' inquiry actions as observed in the ILS's.

The brief test can be used for larger scale evaluations, the longer test can be used in more experimental on site-studies.

4.3 Attitude

Attitudes of students are becoming a more important aspect of learning in the literature (see e.g., Felder & Brent, 2005). Recently, more and more focus is set on attitudinal aspects under the hypothesis that learning mainly occurs if the student has willingness to learn (positive attitude towards learning a particular topic). Osborne, Simon, and Collins (2003) argue that students' attitudes and motivation towards science need more research. A positive attitude is on the one hand a pre-requisite for successful learning in Go-Lab, but on the other hand students' positive attitude towards science can be seen as a general expected outcome of learning in Go-Lab – one of the goals to use Go-Lab Inquiry Learning Spaces is to increase attractiveness of science for students. Therefore, students' attitudes are of interest in evaluating the use of Go-Lab. The leading research question in this context is "*How do students' attitudes change by using Go-Lab Inquiry Learning Spaces?"*

The importance of attitudes on learning science, technology, engineering and mathematics (STEM) has been revealed in several studies. Springer, Stanne, and Donovan (1999) presented a meta-analysis that demonstrated how various forms of small-group learning are effective in promoting greater academic achievement, more favourable attitudes toward learning, and increased persistence through STEM courses and programs (SMET in their sequencing of the subjects).

In characterizing and finding changes in students' attitudes towards science education and inquiry, the model of dual attitudes can be considered (Wilson, Lindsey, & Schooler, 2000). This model makes a distinction between intrinsic and extrinsic attitudes and asks researchers to consider if achieved attitudinal changes are only explicit or implicit as well. In the context of Go-Lab, it is important while our aim is to achieve a balance between these two and it is an interesting research question, if one or other type of attitude can be correlated with students' or teachers' knowledge or inquiry skills and should be considered in providing adapted personalised feedback or support. However, the science attitude test items finally selected for our studies have been not characterized based on the measures of explicit and implicit attitudes and it needs to be analysed based on our data.

In selecting the instrument for evaluating science attitude we followed a review made by Blalock et al. (2008). These authors made an extensive analysis of instruments used for evaluating science-related attitude and categorized these into five groups based on what has been evaluated, specifically: (i) attitudes toward science, (ii) scientific attitudes, (iii) nature of science, (iv) scientific career interests, and (v) other. In addition, they analysed the quality of psychometric properties of these instruments. They created a rubric to analyse different instruments. The following aspects were evaluated: theoretical background for instrument development (up to 3 points), reliability of the test (internal consistency, test-re-test, standard error of measurement, up to 9 points), validity (content, discriminant, congruent, contrasting groups, factor analyses, up to 9 points), dimensionality (up to 6 points), development and usage (if instrument development and evaluation is described, up to 1 point). The test by Germann (1988) test got the highest score (22 out of 28) but it was indeed not very highly evaluated.

In the following search of measures we found an instrument developed in the context of physics education (that is related to the Go-Lab) by P. Kind, Jones, and Barmby (2007) and this one was finally selected to be used in the project evaluations. The reason for selecting this was that they presented an extensive theoretical background for instrument development, reliability of the test (internal reliability of all measures was higher than 0.7), they validated it by two studies were participated 932 and 668 students in particular, they applies factor analysis to increase its validity, and they described the development and evaluation of this instrument. These are the measures described by Blalock, et al. (2008) as the ones needed to ensure a high quality tests for measuring pupils' attitudes towards science. Using a factor analysis they distinguished eight different measures of attitudes towards science: Learning science in school (6 items), Self-concept in science (7 items), Practical work in science (8 items), Science outside of school (6 items), Future participation in science (5 items), Importance of science (5 items), General attitude towards school (8 items), and Combined interest in science (17 items). The last scale consisted of a selection of items from the other scales and, therefore, this factor was not included in the Go-Lab version of the instrument. The factor "general attitude towards school" is not of interest of Go-Lab project and has been left out in the adaptation process. Thus, in conclusion an instrument consisting of six factors describing science attitude will be applied in Go-Lab. These factors will be evaluated through asking student agreement with 34 statements that are measured in a five point Likert scale 'Strongly agree', 'Agree', 'Neither agree nor disagree', 'Disagree', 'Strongly disagree'.

Although the statements of the Go-Lab instrument have been divided into groups according to the factor where they belong to, in evaluations these should be sequenced randomly. The grouping into factors is only important in analysing the data in order to find in which aspects Go-Lab Inquiry Learning Spaces have an effect.

Students' attitude towards science could be correlated to their motivation and some of the attitudinal and motivational aspect could overlap with each other. However, in Go-Lab studies these two have been evaluated separately and later it should be analysed if these two can be combined and a new instruments can be designed so that both attitude and motivation can be evaluated.

The attitude questionnaire can be found in Appendix 3.

4.4 Motivation

Motivation can be seen as one important predictor of success when learning within Go-Lab. In our studies we understand motivation as a process that instigates and sustains a goal-directed activity (Schunk, Pintrich, & Meece, 2008). If the positive attitude toward science can be mostly seen as a general outcome of inquiry process in Go-Lab, then motivation is a more specific pre-requisite for active change of behaviour.

In evaluating students' and teachers' motivation we make a distinction between intrinsic and extrinsic aspects as it has been widely acknowledged in research in psychology (see for example Amabile, Hill, Hennessey, & Tighe, 1994). Intrinsic motivation can be operationalised through the following characteristics: a person is interested, curious, and usually focused on the task. In case of extrinsic motivation the outcomes of learning are prevalent (grades, prizes, etc.) over the task itself. The same aspects will be used in selecting motivation scales for Go-Lab evaluations. However, in addition, students' self-efficacy is evaluated while motivational aspects and self-efficacy tend to be strongly related.

In evaluating students' self-efficacy both efficacy expectations and outcome expectations (Bandura, 1977) of students will be described.

In Go-Lab we choose to apply an instrument that allows both motivational and self-efficacy aspects, namely the questionnaire developed by Glynn (2011) which has been tested for construct validity with nonscience majors (Glynn, Brickman, Armstrong, & Taasoobshirazi, 2011)[1]. This instrument has been designed taking into account extensive analysis of other instruments used for evaluating students' motivation in science. Therefore, it is appropriate for using in Go-Lab. One more reason to use it is its availability in several other languages in addition to English (German, Spanish, Greek, Russian, etc.). This instrument has been widely applied and, therefore, it is validated in repeated studies and the findings from Go-Lab might be compared with these.

The science motivation instrument designed by Glynn (2011) consists of 25 items that have been divided by factor analysis into five scales. In adapting the instrument to be focused on the aims of Go-Lab we have left out the scale for grade motivation, while of our interest are the other four scales: students' intrinsic motivation (and personal relevance), self-efficacy (and assessment anxiety), self-determination, and career motivation.

Similarly to the science attitude instrument (previous section) these factors are evaluated through asking students' agreement with statements that are measured on a five point Likert scale. There is, however, a difference. For science attitude only the agreement with the statements is asked for. In the case of measuring a more personal action-related aspect questions about the students' feeling during their learning process must be presented. According to Glynn (2011) the respondents should answer the following question: "In order to better understand what you think and how you feel about your science courses, please respond to each of the following statements from the perspective of "When I am in a science course…". If adapted to the context of Go-Lab the question to be asked is the following: "In order to better understand what you think and how you feel about your science learning, please respond to each of the following statements from the perspective of "When I am learning science then I feel that …" Each of the statements is evaluated in a five point Likert scale where the options are: 'never', 'rarely', 'sometimes', 'often', 'always'. The statements included in the Go-Lab motivation instrument are divided in four scales and are the following: Intrinsic motivation, self-efficacy, self-determination, and career motivation. Because the scale of grade motivation has been left of as an out of context scale in Go-Lab project the final instrument to be used in Go-Lab evaluations has 21 items.

Similarly to the instrument for evaluating students' attitude towards science the items in the motivation instrument should be sequenced randomly in evaluations. Grouping into factors is again important in analysing the data.

The motivation questionnaire can be found in Appendix 4.

4.5 Understanding of the Nature of Science

One more aspect that is expected to be affected in the use of Go-Lab is students' understanding of the Nature of Science (NoS). NoS has been studied extensively for about 60 years. In early 90's Lederman (1992) made a review of understanding the NoS by students and teachers but concluded that there is no singularity in several aspects of NoS. It was recommended to study specific aspects of NoS in their variety. In context of Go-Lab we follow his understanding about what NoS is. Lederman states that NoS refers to the idea

[1] This 25-item questionnaire can be found at http://www.coe.uga.edu/smq/files/2011/10/SMQII-Glynn.pdf

that science is a way of knowing, or the values and beliefs inherent to scientific knowledge and its development. In their critical analysis of instruments for measuring NoS Lederman, Abd-El-Khalick, Bell, and Schwartz (2002) explain that most of these have several issues why simple NoS-tests are not applicable. For example, often it is expected that the respondents understand the statements similarly to the developers. In contrast they argue for using open-ended questions and interviews to understand students' NoS (see for example Driver, Leach, Millar, & Scott, 1996). They finally found that a 10-item questionnaire (Abd-El-Khalick, Bell, & Lederman, 1998) could be applied in evaluating students' NoS. In a study of Lederman, et al. (2002) the same questionnaire has been provided with descriptions how the collected data should be analysed. Thus, it is also applicable in the Go-Lab project in order to find an answer to the following research question of our interest: *how does students' understanding of NoS change after they have been studying with Go-Lab ILSs?*

Lederman's test was administered to college undergraduates and graduates, and pre-service secondary science teachers. It has been developed through long period (adapted from previous versions developed based on literature review and tested in practice). Lederman, et al. (2002) applied it in written format but in addition they interviewed a reasonable sample of students. Typical time for the test is 45–60 minutes. However, according to Lederman, et al. (2002) the time limits shouldn't be set. Each question should be given separately. It should be mentioned to the respondents that there are no right or wrong answers.

After data has been collected the answers should be qualitatively analysed in order to describe students' NoS through the following aspects: 1) scientific knowledge is tentative; 2) empirical; 3) theory-laden; 4) partly the product of human inference, imagination, and creativity; 5) and socially and culturally embedded; 6) the distinction between observation and inference; 7) the lack of a universal recipe like method for doing science; 8) and the functions of and relationships between scientific theories and laws. These NoS aspects have been emphasized in several science education reform documents (American Association for the Advancement of Science, 1989, 1993; Millar & Osborne, 1998; National Research Council, 1996). The findings from the Go-Lab project should show in what extent and in which aspects Go-Lab learning environments are suitable to contribute in achieving students' understanding of NoS. Evidences regarding these aspects can be found from the answers to these 10 questions. In some cases (e.g., contradictory answers) it is important to validate some data by interviews (which allows us to interpret some specific ideas that could be otherwise misunderstood in written format).

The NoS questionnaire is displayed in Appendix 5.

5 Teacher evaluation

Another goal of the Go-Lab evaluation is to examine the impact that Go-Lab has on teachers' knowledge, skills, and affective domain constructs (i.e., beliefs, attitudes and intentions). Right below we take each one of these variables/aspects separately and break it down to the constructs we are aiming to measure.

5.1 Technical, pedagogical, and content knowledge

The idea behind measuring teachers' knowledge is to investigate whether teachers have the technical, pedagogical, and content knowledge required to teach with computer technology environments, such as Go-Lab. Further, we would like to know if these aspects of teacher knowledge develop after use of Go-Lab facilities such as the federations of labs and the Go-Lab authoring facility. For the development of our measuring instruments we decided to use the Technological Pedagogical Content Knowledge (TPACK) framework Figure 3). According to its developers, the TPACK framework extends Shulman's idea of Pedagogical Content Knowledge (PCK) (Shulman, 1986, 1987). TPACK signifies the interplay of content, pedagogy, and technology, which provides a new unique lens for capturing/perceiving how particular aspects of content could be organized, adapted, and represented for teaching in a technology oriented context.

We selected the TPACK framework because it is one of the very few frameworks that blends PCK with technology knowledge (Schmidt et al., 2009). In fact, it blends three basic forms of knowledge, namely the Content Knowledge (CK), the Pedagogy Knowledge (PK), and the Technology Knowledge (TK). The TPACK framework also clarifies what kinds of knowledge emerge at the intersections between CK, PK and TK, namely the Pedagogical Content Knowledge (PCK), the Technological Content Knowledge (TCK), the Technological Pedagogical Knowledge (TPK), and the Technological Pedagogical Content Knowledge (TPACK) (see Figure 3). However, given the science orientation of Go-Lab, we further complemented TPACK with the Technological Pedagogical Science Knowledge (TPASK) framework (Jimoyiannis, 2010), which also originated from TPACK, but has a more clear science orientation. In particular, we altered the Content Knowledge, Pedagogical Content Knowledge, the Technological Content Knowledge and the Technological Pedagogical Content Knowledge parts of the instrument to make these four constructs specific to science. All alterations were based on principles from the Jimoyiannis (2010) work.

Overall, for measuring teachers' knowledge in the context of Go-Lab, we focus on the following constructs:

- Content (science) Knowledge (CK)
- Pedagogy Knowledge (PK)
- Technology Knowledge (TK)
- Pedagogical Content Knowledge (PCK)
- Technological Content Knowledge (TCK)
- Technological Pedagogical Knowledge (TPK)
- Technological Pedagogical Content Knowledge (TPACK)

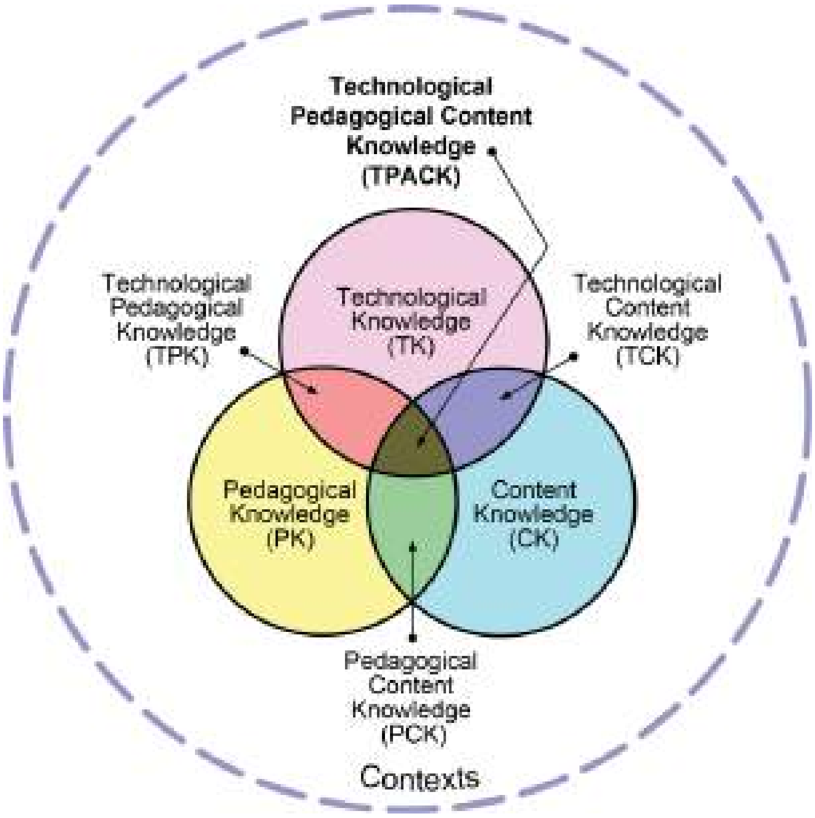


Figure 3. The TPACK framework. Reproduced by permission of the publisher, © 2012 by tpack.org.[2]

For measuring knowledge in the context of Go-Lab, we used the TPACK survey (Schmidt, et al., 2009; see Table 2). Specifically, the TPACK survey consists of seven constructs:

- Content Knowledge (CK)
- Pedagogy Knowledge (PK)
- Technology Knowledge (TK)
- Pedagogical Content Knowledge (PCK)
- Technological Content Knowledge (TCK)
- Technological Pedagogical Knowledge (TPK)
- Technological Pedagogical Content Knowledge (TPACK)

Each one of these constructs has been checked for validity (face, content, and construct) and reliability (Cronbach's alpha) by Schmidt and his colleagues (Schmidt, et al., 2009). Table 4 below provides information on the number of items included in each construct, whether validity was checked and the reliability measured for each construct (see Table 3; for more details see Schmidt, et al., 2009). It should be noted that certain constructs were altered to address the science orientation of Go-Lab. To do so we complemented TPACK with the Technological Pedagogical Science Knowledge (TPASK) framework (Jimoyiannis, 2010). In particular, we altered the Content Knowledge, Pedagogical Content Knowledge, the Technological Content Knowledge and the Technological Pedagogical Content Knowledge parts of the instrument to make these four constructs specific to science. All alterations were based on principles from the Jimoyiannis (2010) work. In particular,

[2] From: http://www.matt-koehler.com/tpack/tpack-explained/

Jimoyiannis has taken the original TPACK, identified its constitutional components and turned it into a framework that made explicit the connections among science (content), pedagogy, and technology (for details see Jimoyiannis, 2010, pp. 1262-1264). In our case, we have taken TPACK removed all items referring to other subject domains besides science (e.g., mathematics, literature) and altered all subject domain-related items to refer only to science [according to the Jimoyiannis (2010) principles. For instance, we proceeded and removed the notion of black-box science and made specific that the term science in the items represents particular sciences (i.e., physics, or chemistry, or biology, or geology). For example, for the science content knowledge construct we altered the item “I have sufficient knowledge about science” to “I have sufficient knowledge about science (Biology or Physics or Chemistry or Geology).” The addition of the clarification in brackets, right after the item, removes the notion of back-box science and clarifies to the teacher that we are referring to the specific science s/he studied and not all sciences in general (note that this clarification has a meaning only if the teachers specifies in the instrument his/her expertise [i.e., physicist, chemist, biologist, geologist]). In Appendix 6 we present the final version of the questionnaire (after all alterations).

Construct	Number of items	Reliability measures (Alphas)	Validity checks
CK*	4	.78**	YES***
PK	7	.87	YES
TK	6	.86	YES
PCK*	4	.87**	YES***
TCK*	4	.86**	YES***
TPK	9	.93	YES
TPACK*	4	.89**	YES***

Table 4 Information concerning the TPACK survey

*Altered constructs.

** This number concerns the initial construct of TPACK. In our version this construct was altered to focus only on science. New Cronbach alphas will be calculated when new data are gathered.

*** Only face and content validity were checked. Construct validity will be calculated after Go-Lab related data are gathered.

5.2 Teachers’ technology skills questionnaire

Despite the fact that our instrument for measuring knowledge (see the previous Section 5.1) focused on technology and content (science), it did not address explicitly teachers’ computer technology skills and understanding of inquiry. Both of these aspects were identified in the literature as factors that affect the integration of technology in inquiry oriented context (e.g., van Joolingen & Zacharia, 2009; Williams, Coles, Wilson, Richardson, & Tuson, 2000). Moreover, the fact that Go-Lab involves both aspects, a computer supported and an inquiry oriented learning environment at the same time, commands for capturing/measuring the status of teachers’ computer technology skills and inquiry understanding. The latter would also allow us to understand if the integration of Go-Lab in a teacher’s instruction is affected by these factors, as stated in the literature.

In the case of measuring teachers’ computer technology skills, we selected the Technology Skills, Beliefs, and Barriers scale developed by Brush, Glazewski, and Hew (2008). Given that we were interested in measuring only teachers’ computer technology skills, from the aforementioned scale, we isolated only the Technology Skills part. Overall, the development of the whole instrument was developed based on data from a total of 176 teachers. All of its scales demonstrated valid and reliable measures. For the Technology Skills part the Cronbach-Alpha score was .95. Face and content validity were checked through the use of an expert panel. The construct validity was also checked and found to be strong (Brush, et al., 2008).

In terms of the constructs involved in the Technology Skills part, it consists of six computer oriented constructs: Basic Operation, Productive Software, Communication, Electronic References, World Wide Web and Multimedia. The total number of items is 32. No alterations were made to this scale. The instrument was selected on the premise that it covers a wide range of basic (everyday life) computer skills, which are also involved in the use of the Go-Lab platform.

The teachers' technology skills questionnaire is presented in Appendix 7.

5.3 Understanding of inquiry and efficacy beliefs

For measuring teachers' understanding of inquiry (e.g., what phases and skills are involved) we used the work of Kang, Orgill, and Crippen (2008) In the case of measuring what is inquiry and what are its phases/steps we developed two questions ourselves (WP8 Go-Lab researchers). The first one is open-ended and asks teachers to define inquiry, whereas the second one is close-ended and asks teachers to select the option that better describes all the phases/steps involved in inquiry (see questions 1 and 2 of the instrument in Appendix 8). Both questions were examined for face and content validity by a panel of experts on inquiry and deemed to be appropriate.

In the case of measuring teachers understanding of the skills involved when enacting inquiry, we adopted six open-ended items from the work of Kang, et al. (2008). Even though these six items were originally designed and included in an interview protocol which was the data collection method used by Kang, et al. (2008), we used the same wording in our questionnaire. Issues of face and content validity were also considered for these adopted items. Reliability would be calculated after data are collected from Go-Lab implementations.

For measuring teachers' efficacy beliefs about inquiry-based science teaching, we used the Science Teaching Efficacy Beliefs Inventory (STEBI) instrument (Enochs & Riggs, 1990), which we altered to measure teachers' efficacy beliefs in using inquiry in their science teaching and which we thus labelled ISTEBI). The original instrument has 23 items divided on in two scales, Personal Science Teaching Efficacy Belief (13 items) and Teaching Outcome Expectancy Scale (10 items). For the purposes of Go-Lab we selected only the Personal Science Teaching Efficacy Belief (13 items) scale, which we altered to measure efficacy beliefs about inquiry-based science teaching (see Appendix 9). This is a well-known worldwide used instrument with high levels of validity and reliability. The reliability analysis of the Personal Science Teaching Efficacy Scale produced an alpha coefficient of .90. All kinds of validity were checked, including the construct validity through a factor analysis.

5.4 Go-Lab beliefs, attitudes, and intentions

In addition to measuring teachers' teachers' efficacy beliefs in using inquiry in their science teaching, we decided to further measure other affective domain related constructs, such as teachers' beliefs, attitudes and intentions concerning the use of Go-Lab. The idea behind this addition was to examine what the teachers' believe about the use of Go-Lab, what their attitudes towards the use of Go-Lab are, and whether they intent to use Go-Lab in the future. A close review of the relevant literature denotes that in order to understand if an individual has accepted a technology as part of its routine (e.g., use it during teaching), you need to have measures on all of these three constructs (i.e., beliefs, attitudes and intentions). For us (Go-Lab consortium) such measures are of great essence, because they relate to the sustainability of Go-Lab (especially after the project is completed). In doing so, we adopted the Technology Acceptance Model (TAM - Davis, 1989; Davis, Bagozzi, & Warshaw, 1989). Specifically, we used an adaptation of TAM by Gardner and Amoroso (2004). TAM is grounded on the Attitude-Behaviour Theory of Reasoned Action (TRA), which portrays a model of how beliefs, attitudes, and intentions relate (Ajzen & Fishbein, 1980; Fishbein, 1979; Fishbein & Ajzen, 1975).

As in TRA, TAM links the beliefs about Perceived Ease of Use (PEU) and Perceived Usefulness (PU) with attitude (AT) towards using computers, behavioural intention (BI) and actual use (computer use). TAM suggests that when users are presented with a computer technology, PEU and PU influence their decisions about how and when they will use it (Davis, et al., 1989).

We have selected TAM because, over the years, researchers have successfully used the TAM framework to examine users' acceptance toward several computer-based technology constructs, such as Graphic User Interface, World Wide Web, e-learning, WebCT (Agarwal & Prasad, 1999; Moon & Kim, 2001; Ngai, Poon, & Chan, 2007; Yuen & Ma, 2008).

Overall, for measuring teachers' beliefs, attitudes and intentions in the context of Go-Lab, we focused on the following constructs:

- Perceived Usefulness of Go-Lab
- Perceived Ease of Use of Go-Lab
- Attitude Toward Using Go-Lab
- Behavioural Intention to Use Go-Lab

In order to measure teachers' beliefs, attitudes, and intentions concerning the use of Go-Lab, we used the Gardner and Amoroso (2004) instrument, which was based on the Technology Acceptance Model (TAM). Based on the existing literature, TAM is considered a valid and reliable model for predicting user acceptance in relation to information technologies (Chau, 1996; Davis, et al., 1989) and is one of the most influential research models used today (Gardner & Amoroso, 2004).

For the purpose of the Go-Lab project, we had to modify the Gardner and Amoroso (2004) instrument items to refer to Go-Lab. Moreover, our revised instrument was shortened to reflect only the constructs that interested us, namely the *Perceived Usefulness of Go-Lab* (6 items)*, the Perceived Ease of Use of Go-Lab* (6 items)*, the Attitude Towards the Use of Go-Lab* (4 items), and *the Intention of Using Go-Lab* (5 items) (see Appendix 10).

6 Evaluation at the organisational level

An organisation is a social unit of people that is structured and managed to meet a need or to pursue collective goals. In education, organizations (schools) have a management structure that determines the activities and relationships between these activities of the teaching staff, and subdivides and assigns roles, responsibilities, and authority to carry out different tasks.

The structure of organisations/schools with their various levels, actors, subjects, topics and internal structures makes the evaluation of any kind of impact a quite complicated and challenging process.

Within Go-Lab, we will aim to address and measure the impact the use of Go-Lab will have on both knowledge the attitude of organisations although we realise that such measurements will be relative and representative only for the part of the organisation that will be exposed to the use of the Go-Lab instruments. Evaluation in this case will mainly address school principals/directors, members of the extended school team and parents.

When discussing Go-Lab, attitude indicates a predisposition or a tendency to respond positively or negatively towards the main Go-Lab ideas, activities and proposed tools. Attitudes, in this case, influence an organisation's choice of action and its responses to certain opportunities, solutions, incentives, and rewards that Go-Lab offers.

Three major components of attitude that are relevant and worth investigating in relation to Go-Lab are:

1. **Affective:** These include the emotions or feelings of organisations' members towards the use of online laboratories and Go-Lab activities
2. **Cognitive:** These processes are involved in the acquisition and understanding of Go-Lab related knowledge as long as the formation of beliefs and attitudes towards the use of online laboratories.
3. **Evaluative:** These attitudes determine the organisation's positive or negative response to Go-Lab.

As we can see above, one of the first issues that we would like to investigate is the organisation's feelings when it comes to the use of online laboratories. *Has the school been exposed to the use of online laboratories before? How do teachers feel about the use of online laboratories? Are teachers/parents/career counsellors/non-MST teachers interested in inquiry learning in general and online labs in particular?* Answers to these questions will help us get an idea on the affective attitudes of the organisation which can later on can be used further in order to analyse the relation between these attitudes and the final evaluative attitudes of the organisation as a whole.

Focus will also be given on investigating *how the availability of the Go-Lab portal (www.golabz.eu) in a school affects the schools' attitudes (awareness) and motivation towards inquiry learning and online labs.* In this case, we will investigate how the availability of the Go-Lab portal within an organisation affects the daily practices followed within the school, expanding its influence to the formation of lesson plans, teaching styles, classroom evaluation techniques etc.

Information in this case will be collected via interviews with members of the teaching staff and case studies.

Information regarding attitudes will be collected via interviews and targeted case studies.

A first set of interview and case studies will take place around the end of Phase B in order to test our instruments and approach, while the main load will take place during Phase C.

6.1 Interviews

Interviews will be held in close collaboration with National Coordinators. A limited sample of organisations/schools will be selected taking into account organisation's level of experience when it comes to the use of online laboratories, size of the school and the type of the school (primary/secondary). The lead teacher and other school representatives (i.e., career counsellors, other MST and non-MST teachers, parents) will be interviewed following an interview script provided by WP8. The script will include questions concerning: the purpose of the activity; why the specific activity was chosen; how the activity was planned and carried out; who was involved in the implementation and what was his/her role; what the outcomes of using the activity were; the strengths and weaknesses of the activity; any adaptations if the activities were used again. The interview schedule may be face-to-face, by telephone, online (i.e. Skype) or by email. Outcomes of the interview will be conveyed in English to the evaluation team based on notes of the interviewer or email transcript. The interview notes / transcript will be confidential to the national coordinator and the evaluation team.

A first draft of the interview script can be found on Appendix 12.

6.2 Case studies

The purpose of these case-studies is to learn about the classroom use of the Go-Lab elements in schools across Europe and its impact on the organisation's attitude. We are keen to understand how this experience was for not only for the teachers but for the entire school and what, in their opinion, were the benefits and drawbacks of using Go-Lab. We are particularly interested in learning the outcomes and impacts this experience had on teachers' and school's attitudes. For these purposes, we will be asking teachers to collect multimedia records, texts and other types of evidence related to the implementation of Go-Lab in their school. The exact nature of the work and collected evidence will of course depend on the type of activity, the language in which it is conducted and of course the overall level of school involvement. Anonymity will be preserved apart from the country in which the activity took place.

A draft template for the Go-Lab case studies can be found in Appendix 13.

7 Methodology

7.1 Evaluation in STEM education

Evaluation theory and practice in the field of STEM education is limited. Existing practices and guidelines tend to promote general principles of good evaluation design and practice, and focus on identifying appropriate questions and criteria for specific initiatives in STEM education (Cullen, Sullivan, Junge, & Britain, 2007, p. 35).

Evaluation in the field of science education has a number of distinctive characteristics, which affect the choice of evaluation criteria, strategies, and methodologies. First, scientific and technological knowledge in the contemporary world is rather turbulent, its social value and cultural meanings are contested and this is reflected in a multiplicity of perspectives on how science and the social world should engage with each other. Thus, cultural differences of various scales (national, regional, or school level) often intervene in the delivery of STEM projects and result in a range of project outcomes and impacts. Hence effective evaluation is impossible without a proper understanding of the exact 'delivery chain' that links initiatives with their end-users and reveals important contextual factors (see Figure 4).

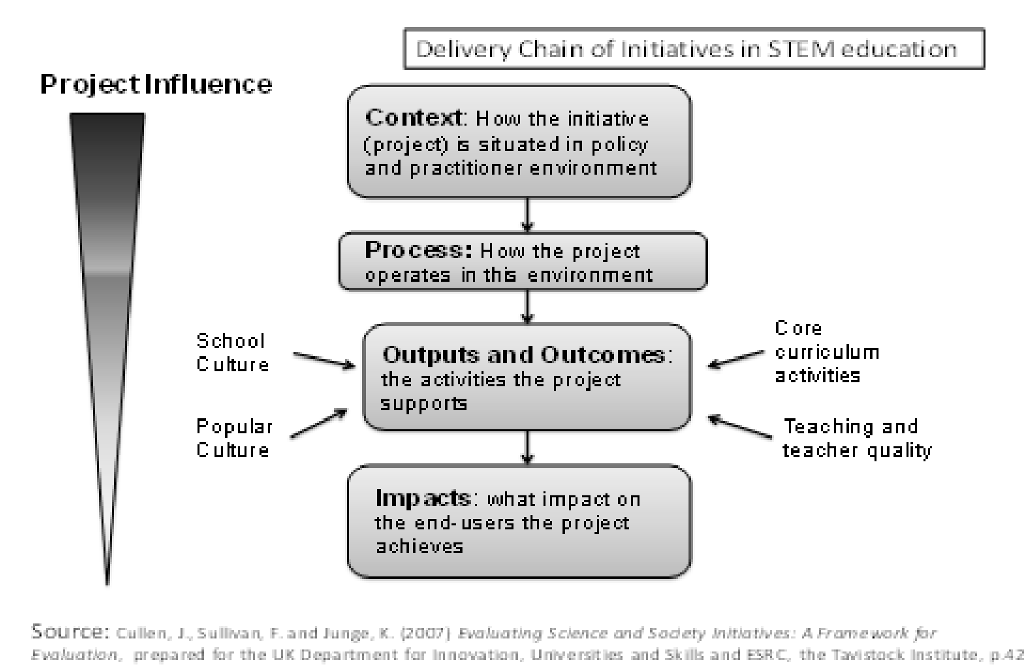


Figure 4. Delivery chain of initiatives in STEM education

Second, most of the general initiatives in the area of STEM education aim at promoting some kind of change in attitudes, awareness, skills or behaviours of students, educators or their wider environment. Attitudes express our feelings about an object, person, issue, or behaviour (Fishbein & Ajzen, 1975). Attitudes towards STEM subjects, careers and industries are a complex mixture of what is known about them objectively, what is felt about them, favourably or unfavourably, and what one's intention or behaviour towards involvement in STEM is. Knowledge of an issue is potentially a precursor to the development of, or change in, attitude. It is thus to be expected that the knowledge that students gain through industry-education activities could influence attitudes towards STEM industries and thence future behaviour, although the relationship is not straightforward.

Early studies on attitudes and beliefs were based on the assumption that attitudes could be used to predict behaviour, and that changes in attitudes should lead to changes in practice (Jones & Carter, 2007). However, many studies have now cast doubt on the simplicity of this linear relationship (Ajzen & Fishbein, 1980; Hungerford & Volk, 1990; Jones & Carter, 2007; Zint, 2002). Attitudes are more likely to change if individuals are offered successive opportunities to engage with the object or issue, especially if this is direct experience (Fazio & Zanna, 1981). Thus, when the examination of impact on future behaviour of students and teachers is not feasible, the major focus of evaluation should be on what was gained from activities, attitudes towards STEM industries and desire for further involvement.

Finally, many initiatives in STEM education are in essence 'complex intervention processes', each including a set of very different activities and practices with diverse aims, target groups, expected outcomes, 'scenarios' of implementation, timelines, etc. Additionally, evaluation is often administered on initiatives which are "work in progress" and they can undergo a substantial change in the further process of implementation. This requires the evaluation plan to remain an open and dynamic system and to be flexible enough to accommodate ad hoc changes and revisions. To allow flexibility and open-endedness evaluation needs to employ different methods and collect a wide range of data, which means that methodological triangulation (Denzin, 2006) should be a default choice for the evaluation design. Hence, most evaluations use a combination of pre-test/post-test questionnaire surveys, interviews and focus groups.

Our plan for evaluation of the Go-Lab project thus recognizes these three factors:

- the complexity of the initiative;
- a focus on the immediate impact of Go-Lab instruments for students and teachers;
- flexibility in coping with any changes in the project

7.2 Investigation approaches in Go-Lab

In Go-Lab, taking into account the diversity of our research questions and the variety of target groups, we will adapt different types of approaches.

7.2.1 Students

7.2.1.1 Approach

For students, we will use classical (quasi-) experimental designs where we will compare groups of students using different (versions of) Go-Lab interventions.

These designs will address students under two different dimensions: before and after they use a very specific Go-Lab intervention. Normally, these are pre-test post-test set-ups of controlled experiments where students' knowledge, inquiry skills etc. are measured both before and after the chosen activity.

7.2.1.2 Design

Students will receive tests and questionnaires if possible in a digital form allowing students to fill them in online. During specific experiments which require the presence of teachers and/or Go-Lab partners or National Coordinators, students will be asked to complete the questionnaires right before and after the experimentation. In this case, the organisers of these sessions will need to foresee adequate time before and after the experiment/intervention in order to allow students to complete the questionnaires. The inquiry test for young children (see Section 4.2.1 and Appendix 2A) requires the use of physical material. This test will only be applied if circumstances allow this. Inquiry tests for young children often rely on physical experiences.

7.2.1.3 Timeline

Phase B

In phase B of the project (October 2014 – June 2015) focus will be given on the effects of individual scaffolds (e.g., the experimental design tool, the conclusion tool etc.).

Phase C

During phase C of the project (October 2015 – June 2016) teachers' familiarity with Go-Lab and ILSs is expected to be at its peak. In phase C student evaluation will shift to the effect of ILSs on students' knowledge, inquiry skills, attitudes, motivation, and understanding of NoS. For knowledge an immediate effect is expected, for inquiry skills an immediate effect may be present but this is expected to become more prominent after using different ILSs, the effects on attitude, motivation, and understanding of NoS are expected to require a prolonged experience with ILSs. The possibility to evaluate these therefore depends on schools using a series of ILSs.

7.2.2 Teachers

7.2.2.1 Approach

For measuring the impact of Go-Lab on teachers and organizations we will not follow controlled experimental designs but we will measure teachers' knowledge, attitudes etc. before and after they have been exposed to Go-Lab interventions.

7.2.2.2 Timeline

Phase B

The impact of the Go-lab project on teachers' knowledge, skills, and affective domain constructs (i.e., beliefs, attitudes, and intentions) will be evaluated during the large scale pilot activities that will take place between October 2014 and June 2016 (Phase B and C). In phase B focus will also be given on acquiring more in-depth qualitative information by means of case studies and interviews (see also the organisational level evaluation). If time allows and during the in-depth experimentation, we will also measure teacher characteristics and use the teacher questionnaires as presented in Section 5.

Phase C

Phase C is similar with Phase B in terms of organisation and evaluation aims. Additionally and since teachers' interactions will be more mature, more extensive and more focused with the system, emphasis will be given in collecting data about their experience and the impact of Go-Lab interventions. For this we will use the questionnaire as displayed in Appendix 11. This questionnaire contains elements for the other questionnaires presented in Chapter 5.

7.2.3 Organisations

7.2.3.1 Approach

For measuring the impact of Go-Lab on organisations/schools and policy makers a set of interviews and case studies will be used.

7.2.3.2 Timeline

Phase B

Some initial interviews will be organised at the end of Phase B. The aim of these draft instruments will be to test our approach and make the necessary adaptations before applying them on a larger scale.

Phase C

In Phase C the adapted instruments will be used in order to collect data related to the impact of Go-Lab interventions within schools. Special focus will be given to policy-makers' impressions and feedback that can provide the project with a good basis on the type of policy recommendations which can contribute to the wider use and adaptation of the online laboratories.

7.3 Evaluation timeline

The development of our evaluation timeline is based on the principles of effective evaluation (Rossi, Lipsey, & Freeman, 2004) and follow the guidelines for building a successful evaluation design (Cullen, et al., 2007). The evaluation timeline defines a sequence of research methods that provide a vigorous and efficient way of measuring the extent to which project aims have been achieved. The stages of evaluation need to be coordinated with the timeline of the project while the choice of methods and instruments is based on the assessment of availability and value of data.

The evaluation timeline of the Go-Lab has been drafted taking into account the work and development of the other work packages and partners and more specifically those of the Technical cluster. Experience has also been drawn upon previous European projects run by European Schoolnet (e.g., InSpire, Stella, PENCIL, inGenious), which had many similarities with the Go-Lab project and thus provided the most valuable information about data, logistics and procedures required for the evaluation.

Pre-Pilot phase (phase A)

Due to the status of the Go-Lab portal development, the months between February and June 2014 have been used as the Pre-Pilot phase during which WP8 partners concentrated on testing, validating, and adjusting the chosen instruments.

Phase B (October 2014 – June 2015)

In Phase B though, students' and teachers' evaluation instruments are available (see this deliverable) and will form an integral part of teachers' Go-Lab related activities providing WP8 and the rest of the consortium with a better insight on the Go-Lab impact and its variations from actor to actor.

The timeline of Phase B evaluation activities and the instruments to be used can be seen in Figure 5. As it can be seen, between October 2014 and June 2015, the WP8 instruments will be used during both the experimentation and extensive Pilot activities. The instruments from Chapter 5 will mainly be used during the experimentation activities while the questionnaire from Appendix 11 will be used by teachers participating in the large scale pilot activities.

To avoid confusion, it is worth mentioning that the label/names used below are plainly for internal purposes related to the creation and management of the questionnaires and the collection of data and do not, in any way, contradict the more extensive names of the instruments we have been using throughout this document. Moreover, the student questionnaires mentioned during the large scale pilots refer mainly to the attitude and motivation instruments provided earlier on.

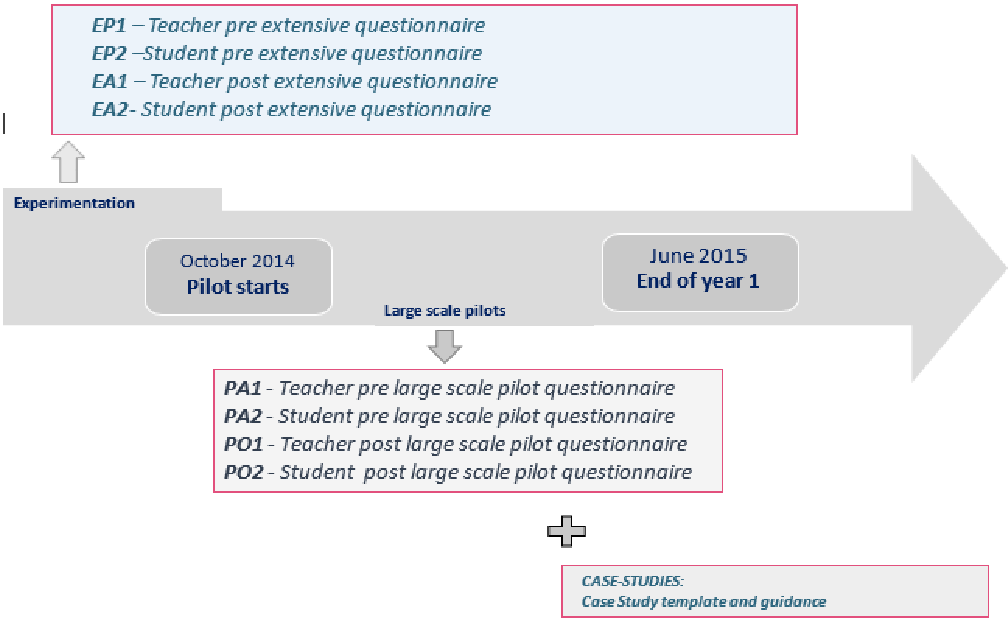


Figure 5. Pilot Evaluation Timeline – Phase B

In parallel, a number of case studies with teachers participating in the pilot activities will also be collected, providing insights on the pilot and on the impact of Go-Lab on teachers' skills.

Phase C (October 2015 – June 2016)

Similar to Phase B, the timeline of Phase C evaluation activities and the instruments to be used can be seen in Figure 6. As it can be seen, between October 2015 and June 2016, the WP8 instruments will be used during both the experimentation and extensive Pilot activities. Focus will now be given to the Pilot activities and the impact of Co-Lab interventions of both teachers and students. The collection of case studies will also be continued.

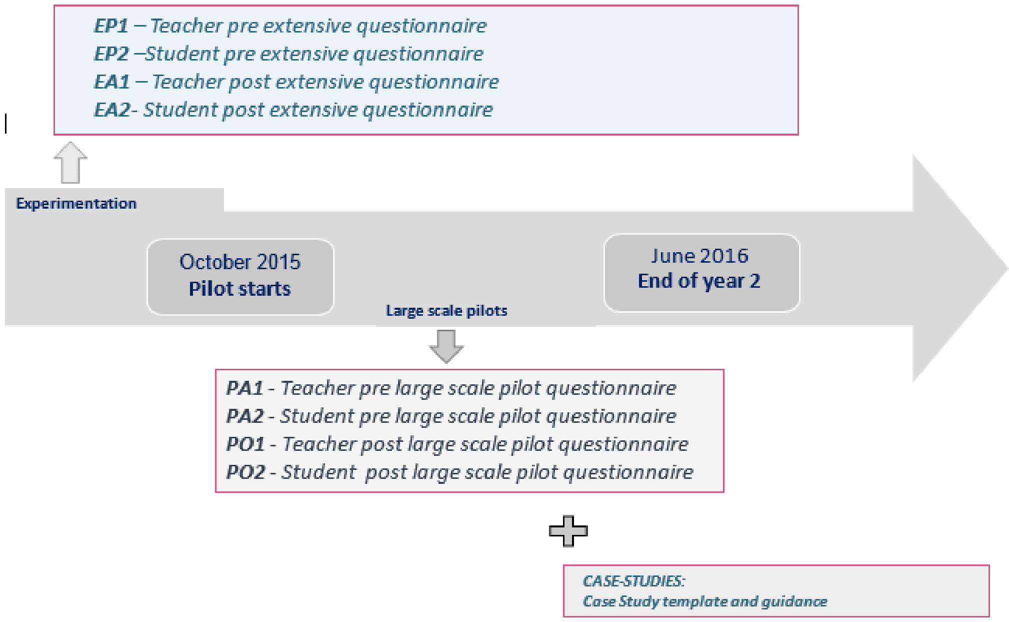


Figure 6. Pilot Evaluation timeline - Phase C

Validation of evaluation instruments

As it has been discussed earlier, the questionnaires that have been composed for use in the Go-Lab evaluation are based on well validated existing questionnaires, though often adaptations had to be made or new elements had to be created in order to fully correspond to project's evaluation aims and needs. As a result, it was decided that an extra validation round is needed before using these questionnaires on a large scale. As can be seen in Table 5, a pre-testing validation of the questionnaires (Phase A) will be done in July, with teachers participating in the Go-Lab Summer School, in the context of the project. The main objectives of the Summer School will be to introduce to teachers the use of online virtual experimentations and remote laboratories as well as inquiry-based science teaching techniques in order to help them develop, improve and enhance their teaching skills and practices.

The main evaluation will then be organized in two phases (phases B and C) during the project that will take place after the completion of the development and testing of the validation tools, according to the timeline specified in Table 5:

Phases	Number of teachers/schools involved	Timeline
Phase A: Composition of instruments and validation	50 teachers	July 2014
Phase B	500 schools	September 2014-August 2015
Phase C	1000 schools	September 2015-August 2016

Table 5. Timeline of evaluation and validation activities

The data collected through the Go-Lab Summer School will enable us to get an insight on teachers' views before they receive any Go-Lab related training and measure the impact that the training will have on their knowledge, attitudes and perceptions. In this way and based on the received feedback, final recommendations/adaptations will be made to the questionnaires in order to validate them throughout the pilot activities. Therefore, the data obtained during this phase will be used to improve the instruments described in this document, if necessary; this will be done before the start of pilot/evaluation activities in order to avoid inconsistencies in collected data.

7.4 Conclusion

In this deliverable we have outlined the most important Go-Lab interventions their expected outcomes and the ways to measure these outcomes. In defining effects and interactions, in choosing the instruments, and in setting up the methodology, choices have been made. Making these choices was necessary and in complex and multifaceted realities like the ones we encounter when implementing Go-Lab in real educational settings new choices need to be made when the actual investigations start. New research questions will then pop-up and practical constraints will determine what can actually be done. For example, we have selected and designed quite a few questionnaires that most probably cannot be all administered. Depending on the actual situation and possibilities researchers will need to make new choices and adaptations so that the emerging constraints are satisfied. The information presented in this deliverable will then function as a resource for designing experiments and it will ensure a level of consistency over the different investigations.

8 References

Abd-El-Khalick, F., Bell, R. L., & Lederman, N. G. (1998). The nature of science and instructional practice: Making the unnatural natural. *Science Education, 82*, 417-436.

Agarwal, R., & Prasad, J. (1999). Are individual differences germane to the acceptance of new information technologies? *Decision Sciences, 30*, 361-391.

Ajzen, I., & Fishbein, M. (1980). Understanding attitudes and predicting social behaviour.

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.

Alonzo, A. C., & Aschbacher, P. R. (2004). *Value-added? Long assessment of students' scientific inquiry skills.* Paper presented at the Proceedings of assessment for reform-based science teaching and learning, a symposium at the annual meeting of the AERA. San Diego.

Amabile, T. M., Hill, K. G., Hennessey, B. A., & Tighe, E. M. (1994). The work preference inventory - assessing intrinsic and extrinsic motivational orientations. *Journal of Personality and Social Psychology, 66*, 950-967.

American Association for the Advancement of Science. (1989). *Science for all Americans*. New York: Oxford University Press.

American Association for the Advancement of Science. (1993). *Benchmarks for science literacy*. New York: Oxford University Press.

Anderson (Ed.), L. W., Krathwohl (Ed.), D. R., Airasian, P. W., Cruikshank, K. A., Mayer, R. E., Pintrich, P. R., et al. (2001). *A taxonomy for learning, teaching, and assessing: A revision of Bloom's taxonomy of educational objectives* (Complete edition ed.). New York: Longman.

Anderson, L. W., & Krathwohl, D. R. (2001). *A taxonomy for learning, teaching, and assessing: A revision of Bloom's taxonomy of educational objectives*. New York: Addisson Wesley Longman.

Bandura, A. (1977). Self-efficacy: Toward a unifying theory of behavioral change. *Psychological review, 84*, 191-215.

Baroody, A. J., Cibulskis, M., Lai, M. L., & Li, X. (2004). Comments on the use of learning trajectories in curriculum development and research. *Mathematical Thinking and Learning, 6*, 227-260.

Baroody, A. J., Feil, Y., & Johnson, A. R. (2007). An alternative reconceptualization of procedural and conceptual knowledge. *Journal for Research in Mathematics Education, 38*, 115-131.

Baxter, G. P., Shavelson, R. J., Goldman, S. R., & Pine, J. (1992). Evaluation of procedure-based scoring for hands-on science assessment. *Journal of Educational Measurement, 29*, 1-17.

Beaumont-Walters, Y., & Soyibo, K. (2001). An analysis of high school students' performance on five integrated science process skills. *Research in Science & Technological Education, 19*, 133-145.

Bixler, B. (n.d.). The ABCDs of writing instructional objectives. Retrieved from http://www.personal.psu.edu/bxb11/Objectives/ActionVerbsforObjectives.pdf

Blalock, C. L., Lichtenstein, M. J., Owen, S., Pruski, L., Marshall, C., & Toepperwein, M. (2008). In pursuit of validity: A comprehensive review of science attitude instruments 1935-2005. *International Journal of Science Education, 30*, 961-977.

Bloom, B. S., Engelhart, M. D., Furst, E. J., Hill, W. H., & Krathwohl, D. R. (Eds.). (1956). *Taxonomy of educational objectives: The classification of educational goals. Handbook 1: Cognitive domain*. New York: David McKay.

Brush, T., Glazewski, K. D., & Hew, K. F. (2008). Development of an instrument to measure preservice teachers' technology skills, technology beliefs, and technology barriers. *Computers in the Schools, 25*, 112-125.

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.

Butzow, J. W., & Sewell, L. E. (1972). An investigation of introductory physical science using the test of science process. *Journal of Research in Science Teaching, 9*, 267-270.

Carolan, T. F., Hutchins, S. D., Wickens, C. D., & Cumming, J. M. (2014). Costs and benefits of more learner freedom: Meta-analyses of exploratory and learner control training methods. *Human Factors: The Journal of the Human Factors and Ergonomics Society*.

Chau, P. Y. K. (1996). An empirical assessment of a modified technology acceptance model. *Journal of management information systems, 13*, 185-204.

Chen, Z., & Klahr, D. (1999). All other things being equal: Acquisition and transfer of the control of variables strategy. *Child Development, 70*, 1098-1120.

Chi, M. T. H., Slotta, J. D., & de Leeuw, N. (1994). From things to processes: A theory of conceptual change for learning science concepts. *Learning and Instruction, 4*, 27-43.

Cornell University. (n.d.). Action words for Bloom's taxonomy. Retrieved from http://www.cte.cornell.edu/documents/Assessment%20-%20Blooms%20Taxonomy%20Action%20Verbs.pdf

Cullen, J., Sullivan, F., Junge, K., & Britain, G. (2007). *Evaluating science and society initiatives: A framework for evaluation*: Tavistock Institute.

d'Angelo, C., Rutstein, D., Harris, C., Bernard, R., Borokhovski, E., & Haertel, G. (2014). *Simulations for STEM learning: Systematic review and meta-analysis*. Menlo Park, CA: SRI International.

Davis, F. D. (1989). Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS quarterly*, 319-340.

Davis, F. D., Bagozzi, R. P., & Warshaw, P. R. (1989). User acceptance of computer technology: A comparison of two theoretical models. *Management Science, 35*, 982-1003.

de Jong, T. (2006a). Computer simulations - technological advances in inquiry learning. *Science, 312*, 532-533.

de Jong, T. (2006b). 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.

de Jong, T. (Ed.). (2013). *D1.1: Go-Lab learning spaces specification*: Go-Lab project.

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

de Jong, T., & Ferguson-Hessler, M. G. M. (1996). Types and qualities of knowledge. *Educational Psychologist, 31*, 105-113.

de Jong, T., & Lazonder, A. W. (2014). The guided discovery principle in multimedia learning. In R. E. Mayer, J. J. G. van Merriënboer, W. Schnotz & J. Elen (Eds.), *The Cambridge handbook of multimedia learning* (second ed., pp. 371-390). Cambridge: Cambridge University Press.

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

de Jong, T., van Joolingen, W. R., Swaak, J., Veermans, K. H., Limbach, R., King, S., et al. (1998). Self-directed learning in simulation-based discovery environments. *Journal of Computer Assisted Learning, 14*, 235-246.

Denzin, N. K. (2006). *Sociological methods: A sourcebook, aldine transaction (5th edition)*.

Deslauriers, L., & Wieman, C. E. (2011). Learning and retention of quantum concepts with different teaching methods. *Physical Review Special Topics - Physics Education Research, 7*, 010101.

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

Driver, R., Leach, J., Millar, R., & Scott, P. (1996). *Young people's images of science*. Buckingham, UK: Open University Press.

Enochs, L. G., & Riggs, I. M. (1990). Further development of an elementary science teaching efficacy belief instrument: A preservice elementary scale. *School Science and Mathematics, 90*, 694-706.

Eysink, T. H. S., de Jong, T., Berthold, K., Kolloffel, B., Opfermann, M., & Wouters, P. (2009). Learner performance in multimedia learning arrangements: An analysis across instructional approaches. *American Educational Research Journal, 46*, 1107-1149.

Fazio, R. H., & Zanna, M. P. (1981). Direct experience and attitude-behavior consistency. *Advances in experimental social psychology, 14*, 161-202.

Felder, R. M., & Brent, R. (2005). Understanding student differences. [Review]. *Journal of Engineering Education, 94*, 57-72.

Fishbein, M. (1979). A theory of reasoned action: Some applications and implications. *Nebraska Symposium on Motivation, 27*, 65-116.

Fishbein, M., & Ajzen, I. (1975). *Belief, attitude, intention and behavior: An introduction to theory and research*. Reading, MA: Addison-Wesley.

Furtak, E. M., Seidel, T., Iverson, H., & Briggs, D. C. (2012). Experimental and quasi-experimental studies of inquiry-based science teaching. *Review of Educational Research, 82*, 300-329.

Gardner, C., & Amoroso, D. L. (2004, 5-8 Jan. 2004). *Development of an instrument to measure the acceptance of internet technology by consumers.* Paper presented at the System Sciences, 2004. Proceedings of the 37th Annual Hawaii International Conference on.

Germann, P. J. (1988). Development of the attitude toward science in school assessment and its use to investigate the relationship between science achievement and attitude toward science in school. *Journal of Research in Science Teaching, 25*, 689-703.

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

Glynn, S. M. (2011). Science motivation questionnaire II. from http://www.coe.uga.edu/smq

Glynn, S. M., Brickman, P., Armstrong, N., & Taasoobshirazi, G. (2011). Science motivation questionnaire II: Validation with science majors and nonscience majors. *Journal of Research in Science Teaching, 48*, 1159-1176.

Gormally, C., Brickman, P., & Lutz, M. (2012). Developing a test of scientific literacy skills (tosls): Measuring undergraduates' evaluation of scientific information and arguments. *CBE-Life Sciences Education, 11*, 364-377.

Govaerts, S. (Ed.). (2013). *D5.2: Specifications of the Go-Lab portal and app composer*: Go-Lab project.

Hiebert, J., & Lefevre, P. (1986). Conceptual and procedural knowledge in mathematics: An introductory analysis. In J. Hiebert (Ed.), *Conceptual and procedural knowledge: The case of mathematics* (pp. 1-27). Hillsdale, NJ: Lawrence Erlbaum Associates, Publishers.

Hungerford, H. R., & Volk, T. L. (1990). Changing learner behavior through environmental education. *The journal of environmental education, 21*, 8-21.

Jaakkola, T., & Nurmi, S. (2008). Fostering elementary school students' understanding of simple electricity by combining simulation and laboratory activities. *Journal of Computer Assisted Learning, 24*, 271-283.

Jimoyiannis, A. (2010). Designing and implementing an integrated technological pedagogical science knowledge framework for science teachers professional development. *Computers & Education, 55*, 1259-1269.

Jonassen, D. H. (2000). Toward a design theory of problem solving. *Educational Technology Research and Development, 48*, 63-85.

Jonassen, D. H., Beissner, K., & Yacci, M. (1993). *Structural knowledge. Techniques for representing, conveying, and acquiring structural knowledge*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Jones, G. M., & Carter, G. (2007). Science teacher attitudes and beliefs. In S. Abell & N. Lederman (Eds.), *Handbook of research on science education* (pp. 1067-1104). Mahwah, NJ: Lawrence Erlbaum Associates.

Kang, N.-H., Orgill, M., & Crippen, K. (2008). Understanding teachers' conceptions of classroom inquiry with a teaching scenario survey instrument. *Journal of Science Teacher Education, 19*, 337-354.

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

Kind, P., Jones, K., & Barmby, P. (2007). Developing attitudes towards science measures. *International Journal of Science Education, 29*, 871-893.

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

Kollöffel, B., & de Jong, T. (2013). Conceptual understanding of electrical circuits in secondary vocational engineering education: Combining traditional instruction with inquiry learning in a virtual lab. *Journal of Engineering Education, 102*, 375-393.

Krathwohl, D. R. (2002). A revision of Bloom's taxonomy: An overview. *Theory into Practice, 41*, 212-218.

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

Kuhn, D., & Pease, M. (2008). What needs to develop in the development of inquiry skills? *Cognition and Instruction, 26*, 512-559.

Lawson, A. E. (2000). Classroom test of scientific reasoning. *Arizona State University*.

Lederman, N. G. (1992). Students' and teachers' conceptions of the nature of science: A review of the research. *Journal of Research in Science Teaching, 29*, 331-359.

Lederman, N. G., Abd-El-Khalick, F., Bell, R. L., & Schwartz, R. S. (2002). Views of nature of science questionnaire: Toward valid and meaningful assessment of learners' conceptions of nature of science. *Journal of Research in Science Teaching, 39*, 497-521.

Maltheis, F. E., Spooner, W. E., Coble, C. R., Takemura, S., Matsumoto, S., Matsumoto, K., et al. (1992). A study of the logical thinking skills and integrated process skills of junior high school students in north carolina and japan. *Science Education, 76*, 211-222.

Marshall, J. E. (1991). Construct validity of multiple-choice and performance-based assessments of basic science process skills: A multitrait-multimethod analysis.

Mattheis, F. E., & Nakayama, G. (1988). Development of the performance of process skills (pops) test for middle grades students.

McKenzie, D., & Padilla, M. (1984). *Effect of laboratory activities and written simulations on the acquisition of graphing skills by eighth grade students*.Unpublished manuscript, New Orleans, LA.

Millar, R., & Osborne, J. (Eds.). (1998). *Beyond 2000: Science education for the future*. London: King's College.

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*, 474-496.

Molitor, L. L., & George, K. D. (1976). Development of a test of science process skills. *Journal of Research in Science Teaching, 13*, 405-412.

Moon, J.-W., & Kim, Y.-G. (2001). Extending the tam for a world-wide-web context. *Information & Management, 38*, 217-230.

Muller, D. A., Bewes, J., Sharma, M. D., & Reimann, P. (2008). Saying the wrong thing: Improving learning with multimedia by including misconceptions. *Journal of Computer Assisted Learning, 24*, 144-155.

National Research Council. (1996). *National science education standards*.

Ngai, E. W. T., Poon, J. K. L., & Chan, Y. H. C. (2007). Empirical examination of the adoption of webct using tam. *Computers & Education, 48*, 250-267.

Osborne, J., Simon, S., & Collins, S. (2003). Attitudes towards science: A review of the literature and its implications. *International Journal of Science Education, 25*, 1049-1079.

Papadouris, N., & Constantinou, C. P. (2009). A methodology for integrating computer-based learning tools in science curricula. *Journal of Curriculum Studies, 41*, 521 - 538.

Pedaste, M., Mäeots, M., Siiman, L. A., de Jong, T., van Riesen, S. A. N., Kamp, E. T., et al. (submitted). Phases of inquiry-based learning: Definitions and inquiry cycle.

Prince, M. J., & Felder, R. M. (2006). Inductive teaching and learning methods: Definitions, comparisons, and research bases. *Journal of Engineering Education, 95*, 123-138.

Rifkin, T., & Georgakakos, J. H. (1996). Science reasoning ability of community college students. *ERIC Digest, ED393505*.

Rittle-Johnson, B., Siegler, R. S., & Alibali, M. W. (2001). Developing conceptual understanding and procedural skill in mathematics: An iterative process. *Journal of Educational Psychology, 93*, 346-362.

Rossi, P. H., Lipsey, M. W., & Freeman, H. E. (2004). *Evaluation: A systematic approach (7th edition)*. London, UK: Sage.

Schauble, L. (1996). The development of scientific reasoning in knowledge-rich contexts. *Developmental Psychology, 32*, 102-119.

Schmidt, D. A., Baran, E., Thompson, A. D., Mishra, P., Koehler, M. J., & Shin, T. S. (2009). Technological pedagogical content knowledge (tpack): The development and validation of an assessment instrument for preservice teachers. *Journal of Research on Technology in Education, 42*, 123.

Schunk, D. H., Pintrich, P. R., & Meece, M. L. (2008). *Motivation in education: Theory, research, and applications (3rd ed.)*. Upper Saddle River, NJ: Pearson.

Shavelson, R. J., Baxter, G. P., & Pine, J. (1992). Performance assessments: Political rhetoric and measurement reality. *Educational Researcher, 21*, 22-27.

Shulman, L. S. (1986). Those who understand: Knowledge growth in teaching. *Educational Researcher, 15*, 4-14.

Shulman, L. S. (1987). Knowledge and teaching: Foundations of the new reform. *Harvard Educational Review, 57*, 1-23.

Snow, R. E. (1989). Toward assessment of cognitive and conative structures in learning. *Educational Researcher, 18*, 8-14.

Springer, L., Stanne, M. E., & Donovan, S. S. (1999). Effects of small-group learning on undergraduates in science, mathematics, engineering, and technology: A meta-analysis. *Review of Educational Research, 69*, 21-51.

Strike, K. A., & Posner, G. J. (1985). A conceptual change view of learning and understanding. In L. West & L. Pines (Eds.), *Cognitive structure and conceptual change* (pp. 211-231). Orlando, FL: Academic Press.

Tannenbaum, R. S. (1971). The development of the test of science processes. *Journal of Research in Science Teaching, 8*, 123-136.

Tao, P. K., & Gunstone, R. F. (1999). The process of conceptual change in force and motion during computer-supported physics instruction. *Journal of Research in Science Teaching, 36*, 859-882.

Texas Tech University. (n.d.). Writing and assessing course-level student learning outcomes. Retrieved from http://www.depts.ttu.edu/opa/resources/docs/writing_learning_outcomes_handbook3.pdf

Timms, M., Clements, D. H., Gobert, J., Ketelhut, D. J., Lester, J., Reese, D. D., et al. (2012). New measurement paradigms.

Tobin, K. G., & Capie, W. (1982). Development and validation of a group test of integrated science processes. [Article]. *Journal of Research in Science Teaching, 19*, 133-141.

Trundle, K. C., & Bell, R. L. (2010). The use of a computer simulation to promote conceptual change: A quasi-experimental study. *Computers & Education, 54*, 1078-1088.

Turpin, T., & Cage, B. N. (2004). The effects of an integrated, activity-based science curriculum on student achievement, science process skills, and science attitudes. *Electronic Journal of Literacy through Science, 3*, 1-16.

van Joolingen, W. R., & Zacharia, Z. C. (2009). Developments in inquiry learning. In N. Balacheff, S. Ludvigsen, T. de Jong, A. Lazonder & S. Barnes (Eds.), *Technology-enhanced learning* (pp. 21-37): Springer Netherlands.

Wenning, C. J. (2007). Assessing inquiry skills as a component of scientific literacy. *Journal of Physics Teacher Education Online, 4*, 21-24.

Williams, D., Coles, L., Wilson, K., Richardson, A., & Tuson, J. (2000). Teachers and ICT: Current use and future needs. *British Journal of Educational Technology, 31*, 307-320.

Wilson, T. D., Lindsey, S., & Schooler, T. Y. (2000). A model of dual attitudes. *Psychological Review, 107*, 101-126.

Yuen, A. H. K., & Ma, W. W. K. (2008). Exploring teacher acceptance of eβ€•learning technology. *Asia-Pacific Journal of Teacher Education, 36*, 229-243.

Zacharia, Z. C. (2007). Comparing and combining real and virtual experimentation: An effort to enhance students' conceptual understanding of electric circuits. *Journal of Computer Assisted Learning, 23*, 120-132.

Zacharia, Z. C., Manoli, C., Xenofontos, N., de Jong, T., Pedaste, M., van Riesen, S. A. N., et al. (submitted). Identifying potential types of guidance for supporting student inquiry in using virtual and remote labs: A literature review.

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

Zint, M. (2002). Comparing three attitude-behavior theories for predicting science teachers' intentions. *Journal of Research in Science Teaching, 39*, 819-844.

Appendix 1. Example student knowledge items

In this appendix we present examples of verbs and how they can be used in constructing test items used for assessing the levels of knowledge students have achieved. The active verbs are a compilation from various lists available on the internet (Bixler; Cornell University, n.d.; Texas Tech University, n.d.). To further support the process example items are also presented.

Remember

In Figure 7, active verbs are suggested that can be used for formulating learning objectives focusing on recalling previously learned material, relevant terminology, specific facts, or procedures related to information and/or course topics. The verbs can also be used as a basis for designing learning activities.

Suggested verbs:
Define; Describe; Draw; Label; List; Name; Recall; Recite; Recognize; Reproduce
Other suggestions:
Copy; Count; Discover; Duplicate; Enumerate; Examine; Identify; Locate; Match; Memorize; Observe; Omit; Outline; Quote; Read; Record; Relate; Repeat; Retell; Select; State; Tabulate; Tell; Underline; Visualize; Write

Figure 7. Active verbs for formulating objectives, activities and/or items for assessing knowledge at the Remember-level

The verbs and the learning objectives and activities based upon these verbs, can be used as input for formulating items for the assessment instruments. In the example test item presented in Figure 8, students can give the names of the quantities (I, V, and R) without really understanding what they mean. This is an example of labelling.

Given is Ohm's law: $I = V/R$ What do the I, V, and R stand for?

Figure 8. Example of test item assessing the Remember-level

Understand

In Figure 9, active verbs are suggested that can be used for formulating learning objectives, activities, and test items focusing on grasping the meaning of information (facts, definitions, concepts, etc.).

Suggested verbs:
Associate; Compare; Compute; Describe; Determine; Estimate; Explain; Extrapolate; Generalize; Give examples; Predict; Recognize; Relate; Report; Summarize; Tell
Other suggestions:
Ask; Choose; Cite; Cite examples of; Classify; Contrast; Convert; Defend; Demonstrate; Differentiate; Discover; Discriminate; Discuss; Distinguish; Express; Extend; Give in own words; Group; Identify; Illustrate; Indicate; Infer; Interpret; Judge; Locate; Observe; Order; Paraphrase; Pick; Practice; Represent; Research; Respond; Restate; Review; Rewrite; Select; Show; Simulates; Trace; Transform; Translate

Figure 9. Active verbs for formulating objectives, activities and/or items for assessing knowledge at the Understand-level

Test items such as the one presented in Figure 10, are already a bit more advanced and complicated as compared to the item in Figure 8. Here it is no longer reproduce the name of a concept, but the student needs to be able to give some description or explanation of a concept(s).

What is an electric current? Explain.

Figure 10. Example of test item assessing the Understand-level

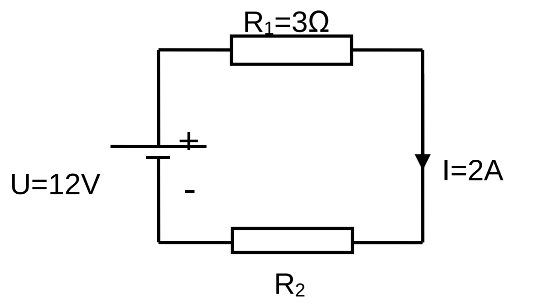
Apply

In Figure 11, active verbs are shown that can be used for formulating learning objectives, activities, and test items focusing on using previously acquired knowledge and/or skills in new or different situations or in problem solving.

Suggested verbs:
Calculate; Construct; Discover; Examine; Experiment; Explain; Generalize; Graph; Interpolate; Operationalize; Predict; Solve
Other suggestions:
Act; Add; Administer; Apply; Articulate; Change; Chart; Choose; Classify; Collect; Complete; Compute; Demonstrate; Determine; Develop; Divide; Employ; Establish; Illustrate; Initiate; Interpret; Interview; Judge; List; Manipulate; Modify; Operate; Practice; Prepare; Produce; Record; Relate; Report; Schedule; Show; Simulate; Sketch; Subtract; Teach; Transfer; Translate; Use; Utilize; Write

Figure 11. Active verbs for formulating objectives, activities and/or items for assessing knowledge at the Apply-level

An example of a test item at the “Apply” level is presented in Figure 12.



Given the circuit displayed above. Calculate the resistance of R_2 (in Ω).

Figure 12. Example of test item assessing the Apply-level

Think critically and creatively

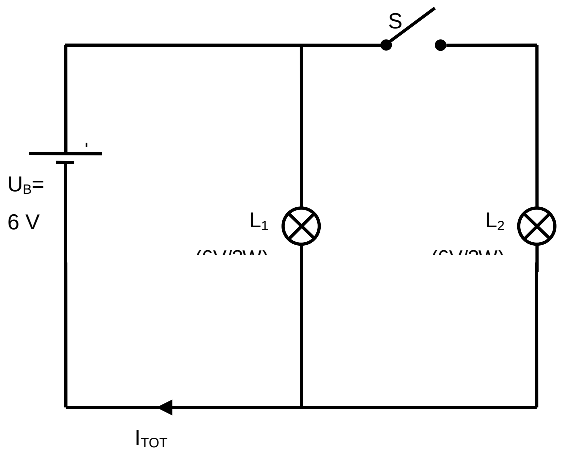
In Figure 13, active verbs are suggested that can be used for formulating learning objectives, activities, and test items focusing on breaking information down into its constituent parts and detecting how the parts relate to one another and to an overall structure or purpose. Students examine and assess concepts, draw conclusions in terms of cause and effect, make inferences, find evidence to support statements/arguments, justify and defend ideas about concepts, present arguments in a convincing way, find errors, evaluate and judge information based on criteria and standards, and create or uniquely apply knowledge and/or skills to produce new and original thoughts, ideas, processes, etc.

Suggested verbs:
Analyze; Argue; Assess; Break down; Combine; Compare; Compile; Conclude; Connect; Consider; Construct; Contrast; Convince; Correlate; Create; Criticize; Deduce; Defend; Diagnose; Diagram; Differentiate; Discriminate; Dissect; Distinguish; Explain; Find errors; Generalize; Generate; Infer; Integrate; Interpret; Judge; Justify; Predict; Prescribe; Question; Recommend; Reframe; Relate; Revise; Rewrite; Select; Speculate; Structure; Support; Synthesize; Systematize; Test; Weigh;

Other suggestions:
Adapt; Analyze; Anticipate; Appraise; Argue; Arrange; Assemble; Assess; Break down; Calculate; Categorize; Choose; Classify; Collaborate; Collect; Combine; Combine; Compare; Compile; Compose; Conclude; Connect; Consider; Construct; Contrast; Convince; Correlate; Create; Criticize; Criticize; Debate; Decide; Deduce; Defend; Design; Detect; Determine; Develop; Devise; Diagnose; Diagram; Differentiate; Discriminate; Dissect; Distinguish; Divide; Drive; Editorialize; Estimate; Evaluate; Examine; Experiment; Explain; Explain; Express; Facilitate; Find errors; Focus; Formulate; Generalize; Generate; Grade; Group; Hypothesize; Identify; Illustrate; Imagine; Infer; Infer; Inspect; Integrate; Interpret; Intervene; Invent; Inventory; Judge; Justify; Justify; Make; Manage; Measure; Modify; Negotiate; Order; Organize; Originate; Outline; Persuade; Plan; Point out; Predict; Predict; Prepare; Prescribe; Prioritize; Produce; Propose; Question; Rank; Rate; Rearrange; Recommend; Reconstruct; Reframe; Relate; Related; Reorganize; Report; Revise; Rewrite; Role-play; Schematize; Score; Select; Select; Separate; Set-up; Simulate; Solve; Specify; Speculate; Structure; Subdivide; Substitute; Summarize; Support; Support; Survey; Synthesize; Systematize; Test; Test; Transform; Utilize; Validate; Value; Weigh; Write

Figure 13. Active verbs for formulating objectives, activities and/or items for assessing knowledge at the "Think critically and creatively"-level

An example of a test item at the "Think critically and creatively" level is presented in Figure 14.



Given the circuit displayed above. Light bulb L_1 is shining. Peter is measuring the current at I_{TOT}. When switch S is turned on, Peter notices that the current remains unchanged. Argue if that is normal or not.

Figure 14. Example of test item assessing the Evaluate-level

Appendix 2A. Test for inquiry skills for young children

(Shavelson, et al., 1992) (used with permission of the author)

PAPER TOWELS

TASK: Students are asked to find out which paper towel can hold, soak up or absorb the most and least amount of water.

TIME: Allow 50 minutes.

MATERIALS: available at: **http://web.stanford.edu/dept/SUSE/SEAL/**, including:

- Administration Instructions
- Student Notebook (see Figure 15 for an excerpt)
- Scoring form

B. Here are some questions about your experiment. Answer each of the questions "yes" or "no".

Were all the paper towels the same size? ____________________

Were all the paper towels completely wet? ____________________

Did you use the same amount of water to

get each paper towel wet? ____________________

Did you let each towel soak in the water

or the same amount of time? ____________________

C. How did you know from the experiment which paper towel holds, soaks up or absorbs the most water and which paper towel holds, soaks up or absorbs the least water?

Most: ____

Least: ____

D. Francisco thinks all of the paper towels must be completely wet before you can decide which paper towel holds the most water and which holds the least. Sally does not think the paper towels have to be completely wet. What do you think?

____

____

Figure 15. Excerpt from Student Notebook Paper Towels task

BUGS

TASK: Students are asked to conduct a series of scientific experiments to find out the preferences of bugs for light or dark environments, damp or dry environments, and combinations of light/dark and damp/dry environments. See Figure 16 for an example of instructions for students.

TIME: Allow 50 minutes.

MATERIALS: available at: **http://web.stanford.edu/dept/SUSE/SEAL/**, including:

- Administration Instructions
- Student Notebook
- Computer simulation notebook
- Computer simulation instructions
- Scoring forms

Experiment #3

If macbugs are given a choice between light and damp, light and dry, dark and damp, or dark and dry, which one would they choose?

NOTES:

Draw a picture of your experiment below, with X's to show where the bugs were at the end. If you made more than one trial draw the ones that gave you your answer.

RESULT: Mark an X next to your result.

Light and damp _____ **Light and dry** _________

Dark and damp _____ **Dark and dry** _________

How did you know from your experiment which one macbugs choose?

____

____

Figure 16. Excerpt from Instructions in Bugs task

Appendix 2B. Test of Integrated Process Skills

(Burns, et al., 1985; Dillashaw & Okey, 1980). (Used with permission of the authors.) This appendix shows a number of sample test items.

<table>
<tr><th>Objectives</th><th>Sample Test Item</th></tr>
<tr><td>I. Given a description of an investigation, identify suitable operational definitions for the variables.</td><td>A study of auto efficiency is done. The hypothesis tested is that a gasoline additive will increase auto efficiency. Five identical cars each receive the same amount of gasoline but with different amounts of Additive A. They travel the same track until they run out of gasoline. The research team records the number of miles each car travels. How is auto efficiency measured in this study?
A. The time each car runs out of gasoline.
B. The distance each car travels.
C. The amount of gasoline used.
D. The amount of Additive A added.</td></tr>
<tr><td>2. Given a description of an investigation, identify the manipulated, responding and controlled variables.</td><td>Marie wondered if the earth and oceans are heated equally by sunlight. She decided to conduct an investigation. She filled a bucket with dirt and another bucket of the same size with water. She placed them so each bucket receive the same amount of sunlight. The temperature in each bucket was measured every hour from 8:00 a.m, to 6:00 p.m. What is the dependent or responding variable?
A. Kind of water placed in the bucket.
B. Temperature of the water and soil.
C. Type of material placed in the buckets.
D. Amount of time each bucket is in the sun.</td></tr>
<tr><td>3. Given a description of variables involved in an investigation, select a testable hypothesis.</td><td>Susan is studying food production in bean plants. She measures food production by the amount of starch produced. She notes that she can change the amount of light, the amount of carbon dioxide, and the amount of water the plants receive. What is a testable hypothesis that she could study in this investigation?
A. The more carbon dioxide a bean plant gets the more light it needs.
B. The more starch a bean plant produces the more light it needs.
C. The more water a bean plant gets the more carbon dioxide it needs.
D. The more light a bean plant receives the more carbon dioxide it will produce.</td></tr>
<tr><td>4. Given a description of an investigation and obtained data, identify a graph that represents the data and describe the relationship between the variables.</td><td>A researcher is testing a new fertilizer. Five small fields of the same size are used. Each field receives a different amount of fertilizer. One month later the average height of the grass in each plot is measured. Measurements are shown in the table below.
<table>
<tr><td>Amount of Fertilizer (kg)</td><td>Average Height of Grass (cm)</td></tr>
<tr><td>10</td><td>7</td></tr>
<tr><td>30</td><td>10</td></tr>
<tr><td>50</td><td>12</td></tr>
</table></td></tr>
</table>

<table>
<tr>
<td></td>
<td>
<table>
<tr><td>80</td><td>14</td></tr>
<tr><td>100</td><td>12</td></tr>
</table>

Which graph represents the data in the table?

</td>
</tr>
<tr>
<td>5. Given a hypothesis, select a suitable design for an investigation to test it.</td>
<td>A greenhouse manager wants to speed up the production of tomato plants to meet the demands of anxious gardeners. She plants tomato seeds in several trays. Her hypothesis is that the more moisture seeds receive the faster they sprout. How can she test this hypothesis?

A. Count the number of days it takes seeds receiving different amounts of water to sprout.
B. Measure the height of the tomato plants a day after each watering.
C. Measure the amount of water used by plants in different trays.
D. Count the number of tomato seeds placed in each of the trays.</td>
</tr>
</table>

Appendix 2C. How Science Works

(P. M. Kind, 2013) (used with permission of the author). This appendix shows a number of sample test items.

Sample item 1: School experiment on dissolving sugar

Jasmine was asked to do an experiment to find how long it takes some sugar to dissolve in water. What advice should you give Jasmine to tell her how many repeated measurements she should make? (Choose one)

- Two or three measurements are always enough
- She should always make 5 measurements
- If she is accurate she only needs to measure once
- She should go on taking measurements until she knows how much they vary
- She should go on making measurements until she gets two or more the same

Sample item 2: Student experiments on battery powered buggy

Alice is investigating the speed of a battery powered buggy.

- She can make a buggy with small wheels or large wheels
- She can make a light buggy or a heavy buggy (with a 500g load)
- She can use ordinary batteries or long-life batteries.

She wants to find out if these make any difference to the speed of the buggy. She makes many measurements and these are her means.

	Wheel size	Load	Type of batteries	Time (in sec) for 5 m
Experiment 1	Small	Heavy	Ordinary	8.6
Experiment 2	Large	Light	Ordinary	7.5
Experiment 3	Large	Heavy	Long-life	8.3
Experiment 4	Small	Light	Ordinary	7.5

a) What do these results tell you about the effect of wheel size on the time for 5m? (Choose one)

A. Large wheels make the buggy use less time for 5m
B. Large wheels make the buggy use more time for 5m
C. Wheel size makes no difference to the time for 5m

b) Which two experiments are needed to work this out?

Experiments ____

c) What do these results tell you about the effect of weight on the time for 5m? (Choose one)

A. A heavy load makes the buggy use less time for 5m
B. A heavy load makes the buggy use more time for 5m
C. A load makes no difference to the time for 5m

d) Which two experiments are needed to work this out?

Experiments ____

Sample item 3: Can we use a scientific method?

Two of the questions below can not be answered by using a scientific method. Identify these and explain why.

A. How many birds are living in the UK?
B. Do birds prefer food from people's feeding trays?
C. What species of bird is most beautiful?
D. Why do birds sing?
E. Should birds be kept in captivity?

Because ………………………………………………………………………………………

……………………………………………………………………………………………

……………………………………………………………………………………………

Appendix 3. Students' attitude questionnaire

The following test is adapted from (P. Kind, et al., 2007) (The test is used with permission of the authors).

1. Learning science in school (6 items)

a) We learn interesting things in science lessons.
b) I look forward to my science lessons.
c) Science lessons are exciting.
d) I would like to do more science at school.
e) I like Science better than most other subjects at school.
f) Science is boring.

2. Self-concept in science (7 items)

a) I find science difficult.
b) I am just not good at Science.
c) I get good marks in Science.
d) I learn Science quickly.
e) Science is one of my best subjects.
f) I feel helpless when doing Science.
g) In my Science class, I understand everything.

3. Practical work in science (8 items)

a) Practical work in science is exciting.
b) I like science practical work because you don't know what will happen.
c) Practical work in science is good because I can work with my friends.
d) I like practical work in science because I can decide what to do myself.
e) I would like more practical work in my science lessons.
f) We learn science better when we do practical work.
g) I look forward to doing science practicals.
h) Practical work in science is boring.

4. Science outside of school (6 items)

a) I would like to join a science club (if there would be available such clubs).
b) I like watching science programmes on TV.
c) I like to visit science museums.
d) I would like to do more science activities outside school.
e) I like reading science magazines and books.
f) It is exciting to learn about new things happening in science.

5. Future participation in science (3 items)

a) I would like to study more science in the future.
b) I would like to study science at university.
c) I would like to have a job working with science.
d) I would like to become a scientist.

6. Importance of science (3 items)

a) Science and technology is important for society.
b) Science and technology makes our lives easier and more comfortable.
c) The benefits of science are greater than the harmful effects.

Appendix 4. Students' motivation questionnaire

The following test is adapted from the Science Motivation Questionnaire II © 2011 Shawn M. Glynn (Glynn, 2011). Used with permission.

1. Intrinsic Motivation

a) The science I learn is relevant to my life.
b) Learning science is interesting.
c) Learning science makes my life more meaningful.
d) I am curious about discoveries in science.
e) I enjoy learning science.

2. Self-Efficacy

a) I am confident I will do well on science tests.
b) I am confident I will do well on science labs and projects.
c) I believe I can master science knowledge and skills.
d) I believe I can earn a grade of "A" in science.
e) I am sure I can understand science.

3. Self-Determination

a) I put enough effort into learning science.
b) I use strategies to learn science well.
c) I spend a lot of time learning science.
d) I prepare well for science tests and labs.
e) I study hard to learn science.
f) Scoring high on science tests and labs matters to me.

4. Career Motivation

a) Learning science will help me get a good job.
b) Knowing science will give me a career advantage.
c) Understanding science will benefit me in my career.
d) My career will involve science.
e) I will use science problem-solving skills in my career.

Appendix 5. Students' NoS test

This test is taken from Abd-El-Khalick, et al. (1998) and reprinted with permission.

1) What, in your view, is science? What makes science (or a scientific discipline such as physics, biology, etc.) different from other disciplines of inquiry (e.g., religion, philosophy)?
2) What is an experiment?
3) Does the development of scientific knowledge require experiments?
   a) If yes, explain why. Give an example to defend your position.
   b) If no, explain why. Give an example to defend your position.
4) After scientists have developed a scientific theory (e.g., atomic theory, evolution theory), does the theory ever change?
   a) If you believe that scientific theories do not change, explain why. Defend your answer with examples.
   b) If you believe that scientific theories do change: (a) Explain why theories change= (b) Explain why we bother to learn scientific theories? Defend your answer with examples.
5) Is there a difference between a scientific theory and a scientific law? Illustrate your answer with an example.
6) Science textbooks often represent the atom as a central nucleus composed of protons (positively charged particles) and neutrons (neutral particles) with electrons (negatively charged particles) orbiting that nucleus. How certain are scientists about the structure of the atom? What specific evidence do you think scientists used to determine what an atom looks like?
7) Science textbooks often define a species as a group of organisms that share similar characteristics and can interbreed with one another to produce fertile offspring. How certain are scientists about their characterization of what a species is? What specific evidence do you think scientists used to determine what a species is?
8) It is believed that about 65 million years ago the dinosaurs became extinct. Of the hypotheses formulated by scientists to explain the extinction, two enjoy wide support. The first, formulated by one group of scientists, suggests that a huge meteorite hit the earth 65 million years ago and led to a series of events that cause the extinction. The second hypothesis, formulated by another group of scientists, suggests that massive and violent volcanic eruptions were responsible for the extinction. How are these different conclusions possible if scientists in both groups have access to and use the same set of data to derive their conclusions?
9) Some claim that science is infused with social and cultural values. That is, science reflects the social and political values, philosophical assumptions, and intellectual norms of the culture in which it is practices. Others claim that science is universal. That is, science transcends national and cultural boundaries and is not affected be social, political, and philosophical values, and intellectual norms of the culture in which it is practices.
   a) If you believe that science reflects social and cultural values, explain why. Defend your answer with examples.
   b) If you believe that science is universal, explain why. Defend your answer with examples.

10) Scientists perform experiments/investigations when trying to find answers to the questions they put forth. Do scientists use their creativity and imagination during their investigations?
    a) If yes, then at which stages of the investigations you believe scientists use their imaginations and creativity: planning and design, data collection, after data collection? Please explain why scientists use imagination and creativity. Provide examples if appropriate.
    b) If you believe that scientists do not use imagination and creativity, please explain why. Provide examples if appropriate.

Appendix 6. TPACK Questionnaire as modified for Go-Lab

TK (Technology Knowledge)	Strongly Disagree	Disagree	Neither Agree nor Disagree	Agree	Strongly Agree
1. I know how to solve my own technical problems	1	2	3	4	5
2. I can learn technology easily.	1	2	3	4	5
3. I keep up with important new technologies	1	2	3	4	5
4. I frequently play around the technology	1	2	3	4	5
5. I know about a lot of different technologies	1	2	3	4	5
6. I have the technical skills I need to use technology	1	2	3	4	5
SCK (Science Content Knowledge)	**Strongly Disagree**	**Disagree**	**Neither Agree nor Disagree**	**Agree**	**Strongly Agree**
7. I have sufficient knowledge about science (Biology or Physics or Chemistry or Geology).	1	2	3	4	5
8. I have the knowledge required to teach science (Biology or Physics or Chemistry or Geology).	1	2	3	4	5
9. I have a very good understanding of science (Biology or Physics or Chemistry or Geology).	1	2	3	4	5
10. I have mastered science content (Biology or Physics or Chemistry or Geology)	1	2	3	4	5
PK (Pedagogical Knowledge)	**Strongly Disagree**	**Disagree**	**Neither Agree nor Disagree**	**Agree**	**Strongly Agree**
11. I know how to assess student performance in a classroom.	1	2	3	4	5
12. I can adapt my teaching based-upon what students currently understand or do not understand	1	2	3	4	5
13. I can adapt my teaching style to different students	1	2	3	4	5
14. I can assess student learning in multiple ways	1	2	3	4	5
15. I can use a wide range of teaching approaches in a classroom setting	1	2	3	4	5
16. I am familiar with common student understandings and misconceptions	1	2	3	4	5
17. I know how to organize and maintain classroom management	1	2	3	4	5
PCK (Pedagogical Content Knowledge)	**Strongly Disagree**	**Disagree**	**Neither Agree nor Disagree**	**Agree**	**Strongly Agree**
18. I can select effective teaching approaches to guide students thinking and learning in science	1	2	3	4	5

(Biology or Physics or Chemistry or Geology).					
19. I am aware of the different approaches for teaching science (Biology or Physics or Chemistry or Geology).	1	2	3	4	5
20. I know pedagogical theories/models that apply to teaching science (Biology or Physics or Chemistry or Geology)	1	2	3	4	5
21. I know teaching strategies that could be used for improving teaching science (Biology or Physics or Chemistry or Geology)	1	2	3	4	5
TCK (Technological Content Knowledge)	**Strongly Disagree**	**Disagree**	**Neither Agree nor Disagree**	**Agree**	**Strongly Agree**
22. I can select effective technologies for understanding and doing science (Biology or Physics or Chemistry or Geology)	1	2	3	4	5
23. I am aware of the different technologies that can be used for understanding and doing science (Biology or Physics or Chemistry or Geology)	1	2	3	4	5
24. I have been trained to use different technologies that can be used for learning science (Biology or Physics or Chemistry or Geology)	1	2	3	4	5
25. Several technologies exist for understanding and doing science	1	2	3	4	5
TPK (Technological Pedagogical Knowledge)	**Strongly Disagree**	**Disagree**	**Neither Agree nor Disagree**	**Agree**	**Strongly Agree**
26. I can choose technologies that enhance the teaching approaches for a lesson	1	2	3	4	5
27. I can choose technologies that enhance students' learning for a lesson	1	2	3	4	5
28. My teacher education program has caused me to think more deeply about how technology could influence the teaching approaches I use in my classroom	1	2	3	4	5
29. I am thinking critically about how to use technology in my classroom	1	2	3	4	5
30. I can adapt the use of the technologies that I am learning about to different teaching activities	1	2	3	4	5
31. I can select technologies to use in my classroom that enhance what I teach, how I teach and what students learn	1	2	3	4	5

32. I can use strategies that combine content, technologies and teaching approaches that I learned about in my coursework in my classroom	1	2	3	4	5
33. I can provide leadership in helping others to coordinate the use of content, technologies and teaching approaches at my school and/or district	1	2	3	4	5
34. I can choose technologies that enhance the content for a lesson	1	2	3	4	5
TPACK (Technological Pedagogy and Content Knowledge)	**Strongly Disagree**	**Disagree**	**Neither Agree nor Disagree**	**Agree**	**Strongly Agree**
35. I can design lessons that appropriate combine science, technologies and teaching approaches	1	2	3	4	5
36. I can teach lessons that appropriate combine science, technologies and teaching approaches	1	2	3	4	5
37. I know how to blend science, technologies and teaching approaches for teaching purposes	1	2	3	4	5
38. I use science, technologies and teaching (all together) in my teaching	1	2	3	4	5

Appendix 7. Teachers' Technology Skills Questionnaire

Basic Operation	I can't do this	I can do this with some assistance	I can do this independently	I can teach others how to do this
1. Create, save, copy & delete files; move or copy files onto hard disks or floppy disks; find files on a hard disk or a floppy disk; create folders and move files between folders	1	2	3	4
2. Print an entire document, selected pages, and / or the current page within a document	1	2	3	4
3. Cut, paste, and copy information within and between documents	1	2	3	4
4. Troubleshooting: When my computer freezes or an error message comes up, I can usually fix the problem	1	2	3	4
5. Troubleshooting: I know the things to check if my computer doesn't turn on	1	2	3	4
6. Viruses: I can use anti-virus software to check my computer for viruses	1	2	3	4
Productivity Software	I can't do this	I can do this with some assistance	I can do this independently	I can teach others how to do this
7. Word Processors: Use the functions of a word processor to format text (font colours and styles), check spelling / grammar	1	2	3	4
8. Word Processors: Use advanced features of a word processor such as headers / footers, tables, insert pictures	1	2	3	4
9. Spreadsheets: Use the basic functions of a spreadsheet to create column headings and enter data.	1	2	3	4
10. Spreadsheets: Use advanced features of a spreadsheet (e.g., using formulas, sorting data, and creating charts / graphs)	1	2	3	4
11. Presentation: Create a presentation using predefined templates	1	2	3	4
12. Presentation: Create a presentation with graphics, transitions, animation, and hyperlinks	1	2	3	4
13. Classroom Management: Use an electronic/ computer grade book	1	2	3	4
Communication	I can't do this	I can do this with some assistance	I can do this independently	I can teach others how to do this
14. Email: Send, receive, open, and read email	1	2	3	4
15. Email: Use advanced features of email (e.g., attachments, folders, address books, distribution lists)	1	2	3	4
16. Listservs: Subscribe to and unsubscribe from a listserv	1	2	3	4

Electronic References	I can't do this	I can do this with some assistance	I can do this independently	I can teach others how to do this
17. Searching: Use a search tool to perform a keyword / subject search in an electronic database (e.g., CD-ROM, library catalogue)	1	2	3	4
18. Use advanced features to search for information (e.g., subject search, search strings with Boolean operators, combining searches)	1	2	3	4
World Wide Web	I can't do this	I can do this with some assistance	I can do this independently	I can teach others how to do this
19. Navigate the WWW using a web browser (e.g., Netscape Navigator, Internet Explorer, AOL)	1	2	3	4
20. Use more advanced features of a web browser (e.g., creating, organizing, and using bookmarks; opening multiple windows; using reload / refresh and stop buttons)	1	2	3	4
21. Use advanced features of a web browser (e.g., install plug-ins, download files and programs, download images)	1	2	3	4
22. Use a search engine (e.g., Yahoo, Lycos, Google) to search for information on the Web	1	2	3	4
23. Use a web authoring tool (e.g., Netscape Composer or FrontPage) to create basic web pages with text and images	1	2	3	4
24. Format web pages using tables, backgrounds, internal and external links	1	2	3	4
25. Upload web page files to a server	1	2	3	4
Multimedia	I can't do this	I can do this with some assistance	I can do this independently	I can teach others how to do this
26. Drawing / Painting: Create simple shapes such as lines, circles, rectangles, and squares with a drawing program (e.g., Appleworks, Kidpix)	1	2	3	4
27. Drawing / Painting: Use advanced features of a drawing program (e.g., layering, grouping objects, changing fill and outline colours)	1	2	3	4
28. Authoring: Create and modify a simple multimedia product using an authoring tool such as Hyperstudio	1	2	3	4
29. Digital Images: Import a digital image (e.g., clip art, photograph) into a document	1	2	3	4
30. Digital Images: Use various tools (e.g., digital camera, scanner) to capture a digital image	1	2	3	4
31. Use a photo editing tool (e.g., Photoshop, PhotoDeluxe) to manipulate a digital image	1	2	3	4
32. Desktop Publishing: Use desktop publishing software (e.g., Publisher, PageMaker) to create a newsletter, pamphlet, or award certificate	1	2	3	4

Appendix 8. Teachers' Understanding of Inquiry Questionnaire

Measuring Teachers' Understanding of Inquiry

1. What is inquiry?

____

2. Which of the following combinations better describe all the phases/steps involved in inquiry?

   a. Investigation – Conclusion – Discussion
   b. Orientation – Conceptualization – Investigation – Conclusion – Discussion
   c. Orientation – Investigation – Conclusion
   d. Orientation – Conceptualization – Investigation – Conclusion
   e. Orientation – Investigation – Conclusion – Discussion

3. Please indicate which of the following scenarios[3] promotes inquiry related skills.

   a. Giving students a white powder and asking them to determine what the powder is. YES NO

   Please explain why Yes or No.

   ____

   b. Having students follow a procedure to complete a lab. YES NO

   Please explain why Yes or No.

   ____

   c. Having students classify substances based upon their observable properties. YES NO

   Please explain why Yes or No.

[3] Derived from Kang, et al. (2008).

d. Having students use graphics on the Internet to explain how gas molecules move. YES NO

Please explain why Yes or No.

e. Having students make presentations of data collected during a lab. YES NO

Please explain why Yes or No.

f. A class discussion about the arrangement of the periodic table. YES NO

Please explain why Yes or No.

Appendix 9. Inquiry Science Teaching Efficacy Belief Instrument (ISTEBI)

	Strongly Disagree	Disagree	Neither Agree nor Disagree	Agree	Strongly Agree
1. I will continually find better ways to teach inquiry-based science	1	2	3	4	5
2. Even if I try very hard, I will not teach inquiry-based science as well as I will through other approaches	1	2	3	4	5
3. I know the steps necessary to teach science concepts through inquiry effectively	1	2	3	4	5
4. I will not be very effective in monitoring inquiry-based science experiments	1	2	3	4	5
5. When a student has difficulty understanding an inquiry process, I know how to help the student to understand it better	1	2	3	4	5
6. I understand inquiry well enough to be effective in teaching science through inquiry	1	2	3	4	5
7. I know how explain to students to conduct inquiry-based science	1	2	3	4	5
8. I will typically be able to answer students' questions about inquiry	1	2	3	4	5

Appendix 10. The TAM instrument as it was adapted for Go-Lab

Perceived Usefulness of Go-Lab	Strongly Disagree	Disagree	Neither Agree or Disagree	Agree	Strongly Agree
1. Using Go-Lab can enable to accomplish tasks more quickly	1	2	3	4	5
2. Using Go-Lab can improve my performance	1	2	3	4	5
3. Using Go-Lab can make it easier to do my tasks	1	2	3	4	5
4. Using Go-Lab in my job/school can increase my productivity	1	2	3	4	5
5. Using Go-Lab can enhance my effectiveness	1	2	3	4	5
6. I find Go-Lab useful in my job/school	1	2	3	4	5
Perceived Ease of Use	**Strongly Disagree**	**Disagree**	**Neither Agree or Disagree**	**Agree**	**Strongly Agree**
7. Learning to use Go-Lab is easy for me	1	2	3	4	5
8. I find it easy to get what I need from Go-Lab	1	2	3	4	5
9. My interaction with Go-Lab is clear and understandable	1	2	3	4	5
10. I find Go-Lab to be flexible to interact with	1	2	3	4	5
11. It is easy for me to become skilful at using Go-Lab	1	2	3	4	5
12. I find Go-Lab easy to use	1	2	3	4	5
Attitude Toward Using Go-Lab	**Strongly Disagree**	**Disagree**	**Neither Agree or Disagree**	**Agree**	**Strongly Agree**
13. I have fun interacting with Go-Lab	1	2	3	4	5
14. Using Go-Lab provides me with a lot of enjoyment	1	2	3	4	5
15. I enjoy using Go-Lab	1	2	3	4	5
16. Using Go-Lab bores me	1	2	3	4	5
Behavioural Intention to Use Go-Lab	**Strongly Disagree**	**Disagree**	**Neither Agree or Disagree**	**Agree**	**Strongly Agree**
17. I always try to use Go-Lab to do a task whenever it has a feature to help me perform it	1	2	3	4	5
18. I always try to use Go-Lab in as many cases/occasions as possible	1	2	3	4	5
19. I plan to use Go-Lab in the future	1	2	3	4	5
20. I intend to continue using Go-Lab in the future	1	2	3	4	5
21. I expect my use of Go-Lab to continue in the future	1	2	3	4	5

Appendix 11 – Large scale Pilot Teacher questionnaire

The selected questions below (the most representative) are related to the research questions; therefore a mapping between the research questions and the constructs that are going to be measured is shown.

Identify the underlying concepts that we like to know in the general part of the questionnaire: This form is designed to provide valuable background information, focusing on various aspects of teaching and learning in your school: knowledge of teaching science, knowledge of instructional approaches and technologies i.e. your pedagogical experience related to computer supported inquiry learning, general technical skills.

Explain why we like to know the underlying concepts: We would like to investigate whether teachers have the content, pedagogy and content knowledge required to teach with computer technology environments, such as Go-Lab.

Knowledge	Type of question	Related research question	What are we measuring?
1. I have experience in solving computer problems when teaching science in my class like registration to different platforms, finding resources/learning materials, perform online scientific experiments with students, etc.	1-Strongly Disagree, 2-Disagree, 3-Neither Agree or Disagree, 4-Agree, 5-Strongly Agree	Does previous technical experience has an impact on the way teacher's approach and perceive Go-Lab?	Measuring teachers' technical background when it comes to the use of web tools/platforms. We can also investigate how teachers' technical experience is connected to the adaptation of Go-Lab and teachers' intentions, beliefs (in combination with question 22 and onward)
2. I have sufficient technical skills to understand and use new technologies as Learning Management Systems (LMS), interactive whiteboards, simulations, and online laboratories	Multiple choices (multiple answers)		Measuring teachers' confidence in using and understanding new technologies
3. I believe I have the necessary pedagogical background to teach my science classes	1-Strongly Disagree, 2-Disagree, 3-Neither Agree or Disagree, 4-Agree, 5-Strongly Agree	What is the impact of teachers' prior pedagogical knowledge towards the stimulation of students' interest/motivation in learning STEM subjects with Go-Lab	Measuring teachers' confidence regarding the adequacy of their pedagogical background. We can also investigate how teachers' pedagogical background is connected to the adaptation and teachers' intentions regarding the use of Go-Lab (in combination with question 22 and

Knowledge	Type of question	Related research question	What are we measuring?
			onward)
4. I have received training on how to assess students' science performance in a classroom.			Measuring how many teachers have received training on assessing students' science performance.
5. I am aware of a wide range of teaching science approaches in a classroom setting			Measure teachers' knowledge of science teaching approaches. Interesting to investigate if there is a relation between this knowledge and the adoption of Go-Lab
6. I am aware of the different approaches that can be used for teaching science as active (learning by doing) and cooperative learning			Interesting to investigate whether there is a relation between this knowledge and the adoption of Go-Lab
7. I have received training on the use of technologies like a whiteboard, LMSs, simulations, online labs etc. that can be used for teaching science	Multiple choices (multiple answers)	What is the effect of teachers' prior technological knowledge on adopting Go-Lab and using it with students?	Do teachers receive training on how to use certain tools and technologies?
8. I can adapt the use of the technologies that I am learning about to different teaching activities	1-Strongly Disagree, 2-Disagree, 3-Neither Agree or Disagree, 4-Agree, 5-Strongly Agree	What is the effect of teachers' prior technological knowledge on adopting Go-Lab and using it with students?	??
9. I can choose technologies that enhance the content for a science lesson	1-Strongly Disagree, 2-Disagree, 3-Neither Agree or Disagree, 4-Agree, 5-Strongly Agree		Do teachers feel that they have the freedom to choose to use technologies that can enhance their science teaching?

Knowledge	Type of question	Related research question	What are we measuring?
10. I have experience in designing lessons that combine science, technologies, and teaching approaches appropriately	Rating scale 1-Strongly Disagree, 2-Disagree, 3-Neither Agree or Disagree, 4-Agree, 5-Strongly Agree	Is there a connection between teacher knowledge of science teaching approaches and the way students perceive science?	??

Skills: Technology Skills	Type of question	Related research question	What are we measuring?
11. I am familiar with the use of authoring tools (e.g., web page editors like Adobe Dreamweaver, CoffeCup etc.)	1-I can't do this, 2-I can do this with some assistance, 3-I can do this independently, 4- I can teach others how to do this	How does the use of authoring tools help teachers in using the Go-Lab platform (i.e. adapt and existing or create a new ILS)?	Do teachers have any previous knowledge of using web page editors? Do teachers have any previous knowledge of using photo editing tools? (Is this needed while using some of our online labs?)
12. I am familiar with the use of online repositories of educational material.			
13. I am familiar with the use of online laboratories and simulations.			
14. I enjoy experimenting with new tools and technologies, a whiteboard, an LMS (like e.g. Blackboard), simulations, online labs			Do teachers enjoy learning to use new technologies and tools? (their attitude towards "new" will give us an indication on how willing they will be to discover how to create their own ILS)

Inquiry	Type of question	Related Research question	What are we measuring?
15. What is the inquiry approach in science education?	**Open-ended**	How does the use of the Go-Lab authoring facility affect teachers' knowledge of inquiry and technical skills (TPACK) and their motivation and attitude towards inquiry and online labs?	Previous knowledge and understanding of IBSE
16. Please indicate which of the following scenarios promotes inquiry related skills for students. a. Having students follow a procedure to complete a lab activity or experimentation. b. Having students use graphics on the Internet to explain how gas molecules move. c. Having students make presentations of data collected during a lab activity. d. A class discussion about the arrangement of the periodic table.	**Rating Scale:** 1-Yes 2-No With one comment field per row (Please explain why Yes or No)	How does consulting the Go-Lab portal (www.golabz.eu) affect teachers' knowledge (e.g., the big ideas) and their attitudes and motivation towards inquiry learning and online labs?	Practical understanding of IBSE. Investigate whether teachers can recognise IBSE within real examples of IBSE educational practices.
17. Even if I try very hard, I will not be able to teach inquiry-based science, as well as other approaches	**Rating scale:** 1-Strongly Disagree, 2-Disagree, 3-Neither Agree or Disagree, 4-Agree, 5-Strongly Agree	How does the use of ILSs and associated lesson plans affect teachers' attitudes and motivation towards inquiry learning in general and online labs in particular?	Do teachers feel confident in using IBSE in their teaching?
18. I know how to explain to students how to conduct inquiry-based science		How does consulting the Go-Lab portal (www.golabz.eu) affect teachers' knowledge (e.g.,	

19. I will continually find better ways to teach inquiry based science		the big ideas) and their attitudes and motivation towards inquiry learning and online labs?	

Beliefs, Attitudes, and Intentions	Type of question	Related research question	
20. How often are you using the following parts of Go-Lab: ILSs, lesson plans, repository, the authoring facility?	Rating scale: Daily Weekly Monthly Less than monthly	Do the existing Go-Lab features fulfill the teachers' needs?	Measure how often the various tools are being used.
21. I believe that the instructional components of an ILS (inquiry phases and their information and the student scaffolds) support the performance of my students	**Rating scale** 1-Strongly Disagree, 2-Disagree, 3-Neither Agree or Disagree, 4-Agree, 5-Strongly Agree 6- N/A	How does the use of the Go-Lab authoring facility affect teachers' knowledge of inquiry and technical skills (TPACK) and their motivation and attitude towards inquiry and online labs? How does the use of ILSs and associated lesson plans affect teachers' attitudes and motivation towards inquiry learning in general and online labs in particular?	Are teachers convinced that ILSs can have an impact on students' performance?
22. My navigation in the Go-Lab portal is clear and understandable			Can teachers find what they are looking for, in go-Lab portal, easily?
23. I could easily create an ILS myself			
24. I could easily find and use an existing ILS			
25. I could easily find, adapt and use an ILS			
26. I enjoy using the Go-Lab portal.			How many teachers enjoy using the portal?
27. I expect my use of Go-Lab to continue in the future			How many teachers will use Go-Lab in the future?
28. Which Go-Lab features you dislike/like (Authoring tool, repository of online laboratories, laboratories & simulations	**Rating scale** 1-dislike 2-like least 3-neither like or dislike 4-like 5-like most With one comment field per row (Please explain why you have	Do teachers find the current interface/features of Go-Lab user-friendly and practical?	Which are the features that teachers like/dislike most and why? Answers will allow us to draw conclusions on what works and

	chosen the selected option)		what not.
29. In your opinion, which are the strongest or weakest features of the Go-Lab portal?	Open-ended	How does consulting the Go-Lab portal affect teachers' knowledge (e.g., the big ideas) and their attitudes and motivation towards inquiry learning and online labs?	

Appendix 12– Go-Lab Interview on Organisation attitudes (draft)

1. What is your role within the organisation (or relationship with the organisation)
2. What is the first thing that comes to your mind when you hear about Science and Science teaching? (for parents)
3. Have you used/heard of online laboratories before your Go-Lab experience?
4. How would you characterize that first experience, if any? Positive/negative, useful/not useful and why?
5. Was it easy for you to use www.golabz.eu? Could you understand its structure and find the information you were looking for?
6. Did www.golabz.eu motivate you in any way to use and discover online laboratories?
7. What were the main difficulties you have encountered while using the Go-Lab portal, if any?
8. Were you familiar with the concepts of inquiry based learning before using Go-Lab portal?
9. Did www.golabz.eu had any impact on your understanding of inquiry based learning?
10. Did the Go-Lab portal and your encounter with it had any impact on your teaching style and lesson planning? If yes, please provide us with an example.
11. Did the Go-Lab portal and your encounter with it changed your views on Science and Science teaching? How? (for parents)

Appendix 13 – Go-Lab Case study protocol template

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 up 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.**

Your Name:	School:	ILS:	Date:

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	o Multimedia recording of you talking about the background details o Text (written account of the background) o 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?	o Video or audio recorded testimony o Text o Will discuss in a follow-up telephone call
5. How did the implementation of the activity go? 5a. the actual process of using the Go-Lab element in your classroom, 5b. students' behaviour and response to the practice	o Video recording of using the activity in your class o Photos capturing the key moments of the practice o Written description of the process
6. Outcomes for your class and school 6f. Teachers talking about their experience and its impact to their class and school	o Video or audio records of teacher's comments about the activity and its outcome o Text (written quotes of what your students were saying during or after the activity) o Samples/copies of outcomes related to the activity
7. What was good about the ILS/laboratory and what were the drawbacks? **8. Would you do it again and would you recommend it to your colleagues?**	o Video or audio recording of you talking about the issues you faced o Text written account o Will discuss in a follow up telephone call

Please, send this form together with the collected data.

Thank you!