Formative usability report on the early prototype of the Go-Lab portal

Effie Law

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 D3.2

Formative usability report on the early prototype of the Go-Lab portal

Editor:	Effie Law
Date:	30th October 2014
Dissemination Level:	PU
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	École 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
Effie Law, Rob Edlin-White, Matthias Heintz, Samaneh Soleimani	ULEIC
Urmas Heinaste, Margus Pedaste	UTE
Constantinos Manoli, Nikoletta Xenofontos, Zacharias Zacharia	UCY
Ton de Jong, Henny Leemkuil, Siswa van Riesen, Bas Kollöffel, Ellen Wassinkamp	UT
Evita Tasiopoulou, Gina Mihai	EUN
Eleftheria Tsourlidaki, Sofoklis Sotiriou	EA
Javier Garcia-Zubia, Olga Dziabenko	UDEUSTO
Angelos Alexopoulos	CERN
Denis Gillet (internal reviewer), Adrian Holzer, Sten Govaerts	EPFL
Michael Auer (internal reviewer)	CUAS

Legal Notices

Executive Summary

The key objective of WP3 is to identify, update and integrate, on an ongoing basis, requirements for developing the Go-Lab Portal. Methodologically, WP3 adopts a participatory design (PD) approach grounded in the established User-centred Design (UCD) frameworks, which also underpin Usability and User Experience (UX) methodologies. In Year 2 (Phase A), WP3 is responsible to collect data on usability and user experience of the Go-Lab design artefacts, ranging from individual scaffolds to an entire online lab to inform the refinement of the pedagogical specifications (WP1) for such artefacts and the improvement on their design and development (WP4, WP5).

Specifically, we provide teachers and students with access to prototypes of parts of or the whole system in increasing levels of fidelity, some iteratively in different levels of refinement, in order to elicit their feedback in usefulness, enjoyability, learnability, memorability, effectiveness, and other pragmatic as well as hedonic qualities. Several complementary HCI approaches have been adopted and adapted for PD data collection and analysis, including:

1) Face-to-face PD workshops with a range of engaging activities that encourage participants to provide feedback on the artefacts under scrutiny. Different instruments were deployed, including traditional paper-and-pen and its software-supported counterpart PDot for capturing feedback, surveys, audience response system, interviews, and observations;
2) Online Core Group Teachers (CGT) involving 17 participants from eight countries in Europe, who are highly motivated in following the development of Go-Lab and are willing to provide prompt feedback to specific questions posed to them on a regular basis;
3) Heuristic Evaluation applied to evaluate some scaffolds and ILs by usability researchers without involving end-users; the results have directly been fed to the development team for improving the prototype before testing it with users.

From November 2013 to October 2014, 21 face-to-face PD events of various scales involving end-users have been conducted in nine countries whereas 6 PD activities have been conducted online with the Core Group Teachers. Altogether these PD activities involved 158 teachers and 221 students from 9 countries. The PD events took place at schools, research/training centres, and other venues such as the Go-Lab Summer School. The data are predominantly subjective self-reports of teachers and students on their experiences and opinions after using specific Go-Lab artefacts for specified tasks.

Overall, the participating teachers and students have highly been persuaded of the potential benefits of Go-Lab. However, they have found some of the components in their current form and some aspects of the integration of the components not intuitive to use. Encouraging is that they have provided improvement suggestions as well as ideas for new features; data thereof have been analysed, assimilated and communicated to the pedagogical and development teams.

The wide variety of PD events with end-users have provided a steady flow of information back to the project leadership and thereby influenced system development direction and priorities. Apart from a set of specific recommendations derived from the PD results, a genera redesign proposal is the use of a style guide, focusing on visual design, content, interaction design, data handling, and online help, to ensure consistent look and feel and thus positive user experience.

Table of Contents

1. Introduction

1.1 Overview of WP3 Tasks in Year 2

The overarching goal of Task 3.2, Task 3.3, and Task 3.4 in Year 2 is to evaluate as well as enhance usability and user experience of the Go-Lab design artefacts, ranging from individual scaffold apps to the entire Inquiry Learning Spaces (ILSs) and labs. Evaluation feedback is communicated via written documents as well as physical/online meetings to the Go-Lab pedagogical and technical teams, providing them with empirically and analytically grounded information on how to improve the Go-Lab artefacts from their respective perspectives. Requirements for new features (i.e., creative options) and recommendations for enhancing the quality of the existing ones are mostly originated from Go-Lab end-users, namely teachers and students, as well as from the researchers. We aim to evaluate both *pragmatic* qualities (e.g., efficient, effective, ease of use, error free, highly intuitive and learnable) and *hedonic* qualities (e.g., fun, enjoyment, engagement, pleasure, and aesthetically pleasing) of the Go-Lab artefacts. For this purpose, we have adopted and adapted a selection of usability and user experience evaluation methods (lab-based, field-based, asynchronous, and remote) while taking the contextual constraints into careful considerations, for instance, the limited timeframe and heavy workload of teachers.

Clearly, the WP3 tasks need substantial inputs from WP1, WP4 and WP5. Besides, WP3 has been collaborating closely with WP7 concerning the recruitment of teachers/schools and with WP8 to scope the focus of the respective evaluative activities.

1.2 Resources and Opportunities

Seven partners (UT, EA, EPFL, EUN, ULEIC, UCY and UTE) are formally engaged in this work package. ULEIC as the WP leader has a role of coordinating and facilitating the work, and reporting on it. All partners work to engage teachers, students and schools, conduct evaluations and report them as well as contribute to wider aspects of WP3. Different partners have been able to dedicate different amounts of time to the WP in this particular phase. Recruitment of schools is more successful in some countries and more challenging in others, due to cultural differences and the workload and level of work-related stress experienced by teachers in some countries. School term timetables in different countries mean opportunities for research are not always when the project most needs it. For these reasons we have supplemented in-school studies with other methods including recruiting a core group of teachers, who conduct evaluations remotely on a regular basis, and also the use of researcher-based heuristic evaluations.

1.3 Outline

Section 2 presents the conceptual and methodological frameworks underpinning the WP3's activities. We summarize key ideas of the four highly interrelated notions in the field of Human-computer Interaction (HCI): User-Centred Design (UCD), Participatory Design (PD), Usability and User Experience (UX). Then we outline the three major types of PD approaches employed for Year 2 work, namely the face-to-face PD approach, remote PD approach, and researcher-based analytic evaluation approach. Next we describe the strategies on managing and scoping a potentially large number of PD studies to address a range of Go-Lab design artefacts.

Section 3 focuses on the 21 events conducted with different *face-to-face PD approach*es, including paper-based and software-supported with PDot. We report the consolidated findings rather than elaborating study-by-study. Specifically, we have aggregated and analysed the empirical data, and report the results of ten scaffolds, two labs and two ILSs. For each of these design artefacts, we first present an overall evaluation, followed by some fine-grained descriptions of individual issues, which are sorted by frequency and accompanied with recommended remedies elicited by the participants and/or researchers. In cases where the findings have already been presented to the development team, we also summarise their responses so far. We dedicate a sub-section on the Go-Lab Summer School from which a relatively large amount of data was collected; a holistic view of the PD activities in this event and some complementary quantitative results are presented. Besides, given its specific methodology – video analysis, we report the most recent PD session as the last sub-section of this rather extensive part of the deliverable.

Section 4 reports on the 6 studies using the *remote PD approach* involving a core group of 17 teachers (CGT) distributed over eight countries. CGT have been given Go-Lab related task on a regular basis via email and returned their responses in the same way. They have been consulted for specific questions such as the perceived usability and usefulness of Golabz.

Section 5 presents the findings of 4 researcher-based analytic studies, known as Heuristic Evaluation, on a set of Go-Lab scaffolds, labs and ILSs. HEs were conducted to identify usability problems, which were sorted in terms of importance and communicated to the pedagogical and technical teams. This enabled some of the usability problems, especially the severe ones, to be fixed before rolling out the prototypes to be tested with real end-users.

Section 6 provides a summary of the main findings of the PD studies in Year 2. In addition, some specific and general recommendations derived from these findings for improving the Go-Lab Portal are presented.

Section 7 discusses the interplay between end-user feedback and redesign work, describing how the findings of the PD work have been communicated to the development team. This section also discusses how the teachers' perceptions of the Go-Lab Portal differ from the students', and how the outcomes of the researcher-based analytic evaluations (i.e., Heuristic Evaluation) are (in)consistent with those of the end-user empirical evaluations.

Section 8 concludes the deliverable by reflecting on the challenges and limitations in implementing the PD work in Year 2. The plan for WP3 work in Year 3 is also presented in this section.

The **Appendices** provide some examples of PD study protocol, instrument, and report of findings and detailed responses of the development team to the findings. To avoid this becoming an excessively long document, we have not included full details for every study. These are available to Go-Lab partners on Graasp.

2. Conceptual and Methodological Frameworks

2.1 Overview of the WP3 Frameworks

In the Go-Lab project, we aim to involve representative end-users of the final system – both teachers and students – to understand their requirements for such a system. To achieve this aim, it is essential to engage end-users in the design process and to encourage them to provide creative input. This is vital for the success of Go-Lab to ensure, as far as possible, that the final system will meet the needs of a wide range of teachers and students, so that they will choose to use it and they can use it in a way which is effective, enjoyable, satisfying, efficient and worthwhile. Good design by experts in pedagogical frameworks, educational technologies, and human-computer interaction paradigms can make a substantial contribution, but it is also highly valuable to have design input and feedback from user communities, especially in terms of usability and user experience.

In Year 2, WP3 focuses on formative evaluation to identify diagnostic information how to improve the system under development. This entails providing end-users with the access to prototypes of parts of or the whole system in increasing levels of fidelity in order to elicit their feedback on different pragmatic (e.g., ease of use) and hedonic qualities (e.g., fun). Thus there will be user involvement throughout the project, feeding into a cycle of continuous refinement of objectives, requirements and design.

Go-Lab aims to implement the project's goals at a large scale in Europe. Ten countries, including Austria, Belgium, Cyprus, Estonia, Germany, Greece, the Netherlands, Portugal, Spain and the UK, have been selected for the first phase of the project. Stakeholders from these countries are involved in a range of PD activities.

2.2 Synopsis of Key HCI Concepts Applied in Go-Lab

It is deemed not necessary to provide detailed descriptions of the key HCI concepts applied in Go-Lab, including User-centred Design (UCD), participatory design (PD), usability, and user experience (UX); such elaboration is not the focus of this document and would make it unduly long. Nonetheless, we present a synopsis of these concepts in order to provide a backdrop of our empirical and analytic work.

UCD is a broad philosophy of involving users in the process of designing a system from the early conceptual phase to the final deployment phase. This is to ensure users' voices are heard, and their needs as well as expectations are addressed. Consequently, the system can have high usability and thus high user acceptance as well as adoption. Since its inception in the 1980s (e.g., Gould & Lewis 1985; Norman & Draper, 1985), UCD has been regarded as a cornerstone of the field of Human-computer Interaction (HCI). Forging hybrid research and practice between HCI and software engineering has been endeavoured for decades, as shown by advocating the use of the UCD framework for enhancing software development processes.

UCD is highly relevant to many research and technical development (RTD) issues across disciplines and to underpin a wide range of projects, not only in the field of HCI but also in technology-enhanced learning (TEL), intelligent content and semantics as well as future internet. The significance of UCD can be epitomized by the birth of ISO 13407:1999 - "Human-centred design

processes for interactive systems", which has recently been replaced by ISO 9241-210:2010[1] in which the notion “User Experience” (UX) is defined as “*A person's perceptions and responses resulting from the use and/or anticipated use of a product, system or service.”* This rather crude definition cannot differentiate itself from usability (ISO 9241-11:1998[2]), which is defined as “*The extent to which a product can be used by specified users to achieve specified goals with effectiveness, efficiency and satisfaction* [the three usability metrics] *in a specified context of use*.”

UX, broadly speaking, descends from UCD, focusing on the experiential aspect of human-machine interactions. In the field of HCI, there has been a shift of emphasis along several dimensions since about 15 years ago: from cognition to emotion, from pragmatic to hedonic, from productivity to experiential quality, from quantitative to qualitative methods, and some other evolvements (e.g., Hassenzahl, 2008; Hassenzahl & Tractinsky, 2006; Bargas-Avila & Hornbaek, 2011; Law et al., 2009; Vermeeren et al., 2010).

One of the driving forces for the extension of usability to UX is the increasingly popularity of games research. With the focus on gamers’ affective responses such as fun, challenge, affect, immersion, flow, the traditional usability metrics of effectiveness and efficiency are deemed insufficient (Law & van Schaik, 2010). Nonetheless, the fuzziness of experiential qualities makes them difficult to be measured (Law et al., 2014). Furthermore, the UX evaluation methods (UXEM) are largely drawn from the traditional usability ones (UEM) (e.g., Tullis & Albert, 2008; Hartson & Pyla, 2012). Nonetheless, UXEMs are mostly qualitative with heavy use of narrative-based methods (e.g., Tuch et al., 2013). Despite attempts to demarcate usability and user experience (Roto et al., 2010), their relations remain ambiguous. Some researchers and practitioners opt to use UX as an umbrella term to subsume usability (e.g., Thüring & Mahlke, 2007) and its associated metrics, whereas some (erroneously) treat them as synonyms (e.g., Tullis & Albert, 2008).

Participatory Design (e.g., Muller & Druin, 2010; Simonsen & Robertson, 2012) is a methodological approach to design which places a strong emphasis on involving users or potential users or representatives of users in the end to end development of artefacts which are intended for their use. This begins in determining needs or establishing requirements; continues by involving them in contributing creatively to design ideas; continues further by involving users in reviewing design ideas or trying out mock-ups or early prototypes to provide confirmation of efficacy or critique or suggestions for improvement, and concludes by involving them in evaluating the final product. In a PD team, it is recognised that users are experts in their own needs; specifically in knowing what functional and usability features will be beneficial to them and what usability obstacles might be most detrimental to product acceptance, adoption and usage.

The purpose of PD is to create designs which gain widespread acceptance and come to be used effectively, efficiently, enjoyably and safely by the target user groups (not just professional designers!). PD has a strong focus on outreach to user communities, understanding their needs and creating designs informed by a high level of insight into the user’s world; their working contexts and needs and priorities. PD is therefore likely to lead to better products than would emerge from a more inward looking design team composed solely of technology enthusiasts with rather homogeneous skills, capabilities and enthusiasms. Such teams tend to design products which are effective for

[1] Ergonomics of human-system interaction -- Part 210: Human-centred design for interactive systems

[2] Ergonomic requirements for office work with visual display terminals (VDTs) -- Part 11: Guidance on usability

themselves and those like them, but less effective for people with quite different lives and capabilities and preferences and motivations.

2.3 Objectives of the PD work in Go-Lab

The objective of the PD work in Go-Lab is to support the ongoing processes of product design, development and refinement by providing a flow of feedback from end-users, particularly in areas of usability and user experience, to enable the development of a highly usable and attractive system.

Each PD study is expected to fulfil one or more of the following purposes:

- to test developed system components for usability, user experience and whether they provide valued functionality;
- to assess teachers' and students' acceptance and engagement with aspects of the system, or system concepts;
- to verify (or challenge) working assumptions;
- to provide additional design ideas or to evaluate early design ideas illustrated with mock-ups.

Earlier studies tend to focus on providing design ideas, critique and suggested improvements; later studies tend to be more evaluative. Based on results of PD activities, recommendations for modifying the design of specific components of the Go-Lab artefacts and for informing the plan and implementation of future evaluative activities in Go-Lab are identified.

2.4 Three major UCD approaches in Go-Lab

Models of user participation in PD often involve close collaboration between participants and researchers, usually but not always being based on face to face interactions. While *face-to-face PD approach*es are our primary and preferred style of engagement for many studies, we have also deployed alternative approaches, namely *remote PD approach* and *researcher-based analytical evaluation,* to address different needs and contextual constraints.

2.4.1 Face-to-Face PD Approaches

Paper-based PD – the Layered Elaboration approach:

Given the target group of teachers (adults) and students (children), we decided to apply a PD approach that has already been successfully used with children: Layered Elaboration (Walsh et al., 2010). This is a rather new PD method where feedback is given on acetate sheets that are put on top of mock-up printouts. This has several advantages compared to comments and scribbles directly done on the printout itself. First of all, the mock-up itself is not altered and stays intact, preventing the participants to feel like they are "destroying" something, thus lowering the inhibition level to provide feedback. Additionally, the acetate sheet of the next participant can be layered on top of the existing feedback, thus allowing comments on the original mock-up as well as on the redesign ideas of previous participants. An advantage of this method from the researcher perspective is that all acetates regarding one mock-up printout (e.g., from different groups providing feedback using the Layered Elaboration approach) can be superimposed to see areas of the screen where, for instance, especially much feedback has been given on.

We argue that the Layered Elaboration approach is especially well suited for a face-to-face setting, as it can be flexibly applied and adjusted to be suitable for individual and group-based work,

depending on the requirements and conditions of the study at hand. At the same time (e.g., compared to questionnaires or discussions) it allows the gathering of actual design feedback (e.g., drawings on the actual mock-up instead of textual answers on questions regarding the mock-up) of a rather large number of participants (which can be challenging in discussions, especially when working with children, where it needs a very skilled facilitator to keep the participants on track and make sure all opinions are heard).

Software-supported PD – PDot:

While the custom built PDot, which stands for Participatory Design Online Tool, enables remote engagement of participants in usability studies, it can also effectively be applied in a face-to-face PD setting (Figure 2.1).

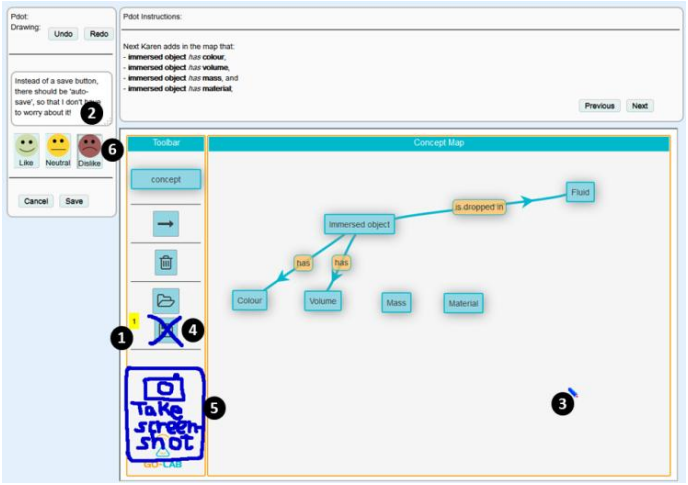


Figure 2.1: PDot used to give feedback on Concept Mapper (a learning tool). The left hand panel and upper right panel are PDot and the lower (bigger) area is the tool itself (white numbers in black circles added for referencing in the text). If the user wants to comment on a specific object on the user interface, she can click that object. A yellow Post-it icon will appear (1) and then she can provide a comment in the text box on the left panel (2). She can also use the cursor as a free-hand drawing tool (3) to cross out existing objects (4), sketch a new element (5), and so on. Besides, the user can indicate her emotional response by selecting one of the smiley icons (6).

Although a variety of paper-based methods have been used successfully in PD activities, they can become tedious or even impracticable in a distributed project setting like the one in Go-Lab. Amongst others, we highlight three reasons that motivated us to develop PDot: (i) involving distributed users and (ii) disseminating results to distributed stakeholders within the project; (iii) addressing the limitations of existing online annotation tools (Heintz et al., 2014).

With the universal approach the Go-Lab is pursuing (i.e., one portal for all the target groups with heterogeneous backgrounds), it is very important to gather inputs from a variety of prospective users. For instance, the requirements of a biology teacher in the UK might differ substantially from the design ideas of a physics teacher in Greece. But even within countries there can be diverse teacher and student requirements. Capturing requirements from these distributed users can be very costly and time consuming because of travelling costs. Digital tools can support participants and researchers in their respective tasks of sharing and analysing feedback. Another challenge for paper-based data is how to share them economically with different stakeholders in the project. For instance, the participants' scribbles and comments on the mock-ups can be useful for the HCI researchers in Leicester as well as the designers and developers in Lausanne. Also, software tools can support the project by enabling direct online access to the gathered data from anywhere.

Nonetheless, from the user perspective, there are some limitations to be considered when using a tool instead of paper-based methods to gather feedback: a computer and Internet access are required and it is less natural to write and scribble digitally on the screen as compared to providing feedback with a pen on paper.

2.4.2 Remote PD Approach: Core Group of Teachers

While face-to-face PD activities enable us to capture rich data from end-users, it normally takes time to identify a right venue and a right schedule that fits all parties involved. In certain conditions some relatively fast feedback can be very helpful. For instance, a creative idea may be appealing to a pedagogical/technical designer, but whether it is potentially useful for end-users remains unknown. Before investing further resources in elaborating the idea, it is deemed practical if end-users can be consulted. Their acceptance (or rejection) of the idea together with their reasoning will enable the designer to make an informed decision how the idea should be handled. A remote PD arrangement is a viable solution for this scenario.

Indeed, one effective means to obtain prompt feedback on some 'half-baked' idea or work-in-progress from end-users is through computer-mediated communication (CMC). With the support of WP7, we have recruited a group of teachers, known as *Core Group of Teachers*, which currently consists of 17 teachers from 8 countries across Europe. They are contacted via email, being asked to carry out some specific tasks and return us with feedback within a certain period of time. To enable CGT to integrate the remote PD activities into their routine, the tasks are given out on a biweekly basis. The flexibility that asynchronous CMC provides – no geographical or temporal barrier – encourages the teachers to get involved easily. Furthermore, such continuous user involvement aligns well with the UCD philosophy. Nonetheless, a limitation of this remote approach is the fluctuating response rate, depending on the workload of CGT in their everyday job.

2.4.3 Researcher-based Analytical Evaluation Approach

It is a commendable practice in the field of HCI to perform analytic evaluation such as Heuristic Evaluation (HE) to identify significant issues of a system by usability specialists prior to testing it with end-users (e.g., Nielsen, 1994). Essentially, HE involves walking through a prototype to identify any feature violates one or more of a set of ten usability heuristics (e.g., visibility of system status). HE, in contrast to user-based evaluation, does not involve any end-user, and relies on the expertise and experience of a usability specialist, who is knowledgeable of the heuristics and ideally is also a domain-specific expert ("double experts"). A main outcome of HE is a list of usability problems (UPs), usability principles each UP violates and the impact (i.e., severity and frequency) each has. Such evaluation feedback can somehow help the development team fix the UPs, thereby improving the overall usability of the prototype and eventually enhancing end-users' acceptance.

In following this practice, a team of three usability specialists had performed HE on a set of Go-Lab artefacts, which were subsequently evaluated with the teachers in the Summer School. Results of the HE were in the meantime communicated to both pedagogical and technical teams, and consequently the artefacts were improved.

2.5 Strategies for Managing PD Work

2.5.1 Challenges

A significant challenge for this project is to conduct PD sessions in multiple countries with different languages, using teachers in different science subjects, with varying levels of IT aptitude and enthusiasm, and with different ages of pupils and different curriculum expectations. Teachers in different countries also have different pressures and patterns of work, and different motivations for

use of online labs and for taking part in studies. Schools in different countries have different attitudes to allowing researchers into their premises and to ethical requirements. Also Go-Lab partners in different countries have different levels of access to, and rapport with, teachers.

As the number and the specificity of Go-Lab design artefacts have been increasing during Year 2, the number and the specificity PD studies would increase accordingly. It implies that either a bespoke PD protocol (PDP) for individual artefact or heavy customisation of a generic PDP (Section 3) is required. Consequently, in principle there would be a diversity of PD studies varying in the following variables: (i) the type of Go-Lab artefact to be studied; (ii) the number of teachers and/or students involved; (iii) the amount of time available for the study; (iv) the physical setting and equipment available (e.g., shared computers or one each); (v) curricular constraint (e.g., meeting specific learning/teaching objectives); (vi) any particular research questions that the pedagogical and technical team may have about the artefact under study; (vii) the amount of data already collected on the usability of the artefact being studies.

It would entail nontrivial resources to create PDPs and prepare associated materials for individual artefacts. To facilitate the participation of end-users, there is also a need for study materials to be translated. All these require a large number of researchers and participants. Practically it is very challenging, especially the number of teachers and students being able to get involved in evaluation activities is relatively modest for various reasons (e.g., tight school timetable; some of the reasons mentioned in the previous paragraph). Another compelling concern is that the empirical findings would be (too) large in breadth (provided that a significant number of planned PD studies could really be realized) but shallow in depth. Such piecemeal findings might not allow us to draw any solid conclusion. Hence, there is a need to prioritize the artefacts to be evaluated. Consequently, a *PD Study Catalogue* has been developed to manage the aforementioned challenges.

2.5.2 Structuring and sequencing the study programme

As illustrated in Figure 2.2 below, the Go-Lab Portal end-user usability tests can be conceived of in three levels; level 1 are formative foundational studies testing of individual components, ideas or mock-ups; level 2 are also formative evaluations, testing segments of the system (e.g., using ILSs or mini-ILSs); level 3 involves end-user tests of the whole system and are the summative evaluations. During the formative studies new requirements can arise leading to new components and more level 1 studies. The earlier studies towards the base of the pyramid are very diverse and tend to need individual protocols, whereas a smaller number of protocols can cover the higher level studies. The lower level studies can often be more qualitative and creative, whereas higher level studies are more quantitative.

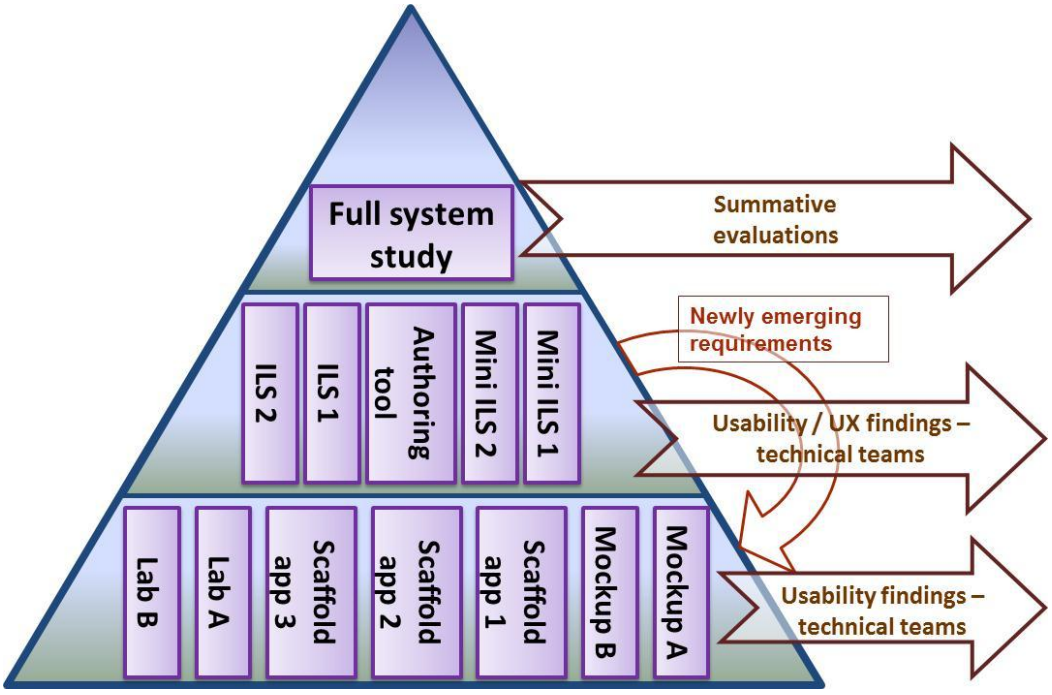


Figure 2.2. Study types and levels

2.5.3 Recruitment Strategies

There are some very clear areas of human and contextual variability which are likely to affect user experience and could be covered by targeted recruitment. For teachers these might include age, gender, nationality, language, years of teaching experience, subject taught, and experience with various other ICTs. Student recruitment might cover similar factors, with a more specific emphasis on age.

All WP3 partners have contacted schools in their country to explore opportunities to conduct PD studies. Due to cultural and political diversity, and differences in the way education is organised across Europe, recruitment methods and approaches and opportunities vary from country to country. Recruitment to studies has therefore sometimes been systematic and purposeful and sometimes involved informed opportunistic approaches, always aiming to access as wide a range of end-user perspectives as possible, while recognising the constraints of school holiday closures and teacher workloads.

As described at Section 3.1 below, WP3 as a whole has successfully engaged with a very broad and diverse range of schools, teachers and students from across Europe.

2.5.4 Workflow and Study Catalogue

To address the challenges mentioned in Section 2.5.1, the following workflow has been developed:

1. The WP3 lead partner is responsible for maintaining and regularly publishing a catalogue of end-user study work. This is currently implemented as an Excel spreadsheet, illustrated in the diagrams below, with one row for each requested study. All project partners identifying their need for end-user research studies are requested to provide a brief description of the research

objective and any specific system components affected, and optionally the information on project need, importance (high/medium/low) and urgency.

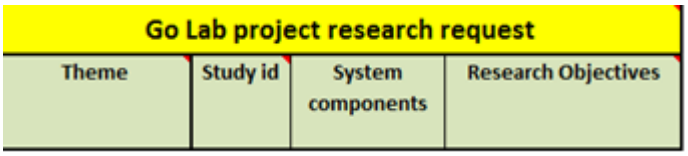


Figure 2.2a. PD Study Catalogue – Partner request records

2. The overall project management team are given the opportunity on a regular basis to revise the relative priorities and urgency of request studies.

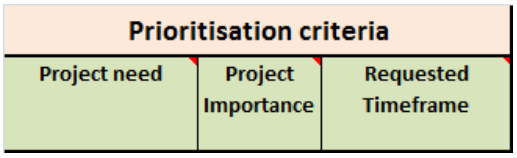


Figure 2.2b. PD Study Catalogue – Prioritisation criteria

3. The WP3 lead partner specifies a provisional Outline Research Plan for each study, consisting of research methods, data to be collected, proposed number and type of participants, estimated session duration, and the equipment, study resources and software needed. This information is then added to the catalogue.

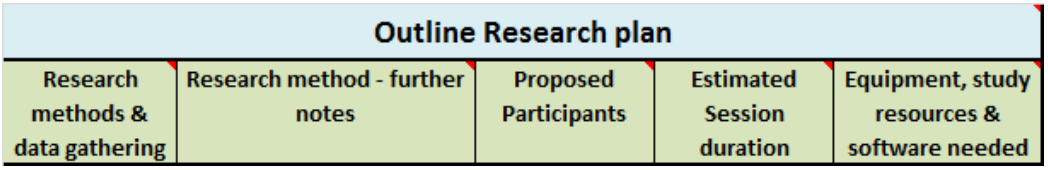


Figure 2.2c. PD Study Catalogue – Outline research plan

4. Starting with the highest priority and most urgent studies, the WP3 lead partner requests frozen copies of the relevant software components, and then conducts a Participatory Design Planning Workshop for each study. The relevant software components are then reviewed to ensure they are sufficiently stable and bug-free and usable to be used in a PD session with likelihood of an effective session and without a risk of reputational damage. The study originator is then contacted to gain a greater understanding of the study requirements. The Outline Research Plan will then be verified or modified, and resources for the study are created in English. The resources could include an overall study script, a scenario specifying tasks for the participant to attempt, any demonstration material (e.g., PowerPoint, videos), questionnaires, interview templates, observer recording sheets, etc. The location of the artefacts and work in progress are documented in the PD Study Catalogue “Preparation” columns.

Preparation	
Equipment, study resources & software available	Next actions

Figure 2.2d. PD Study Catalogue - Preparation

5. Asynchronously with this process, all WP3 partners are actively recruiting participants (teachers and students) and organizing PD sessions. As soon as an event is arranged, the partner informs the WP3 lead partner of the date, country, number and type of participants available and time available. This is recorded by the WP3 lead partner in a second table of PD sessions.

Research opportunity					
Partner	location	date	number of teachers	number of students	time available

Figure 2.2e. PD Study Catalogue: Research opportunity

6. Using up to date information on priorities and opportunities, ULEIC specify which study or studies should be conducted, and provide information on resources available.

Study instructions	
studies to conduct	Resources to use

Figure 2.2f. PD Study Catalogue: Study instructions

7. The partner conducting the study organizes for the materials to be translated into the relevant language. They should if possible conduct a pilot study with either a real participant or with someone not familiar with Go-Lab acting as participant. Based on the pilot outcome they may need to refine or modify the materials.

8. The partner should then conduct the study, and collect and label all data clearly and store it securely with appropriate backups. Within 2 working days they should report completion of the study to WP3 using the completion report template.

9. Finally the partner should upload all resources used, raw data collected (translated into English) and analytical results to the relevant Graasp folder.

10. The WP3 lead partner tracks completion using the following categories in the Participatory Design Sessions table:

Study Completion		
Session completed?	analysis completed?	reporting completed?

Figure 2.2g. PD Study Catalogue: Study completion

11. The WP3 lead partner reviews the data and analyses together with findings from any other PD sessions for this study, and provides project guidance to the relevant WPs.

12. The WP3 lead partner aggregates all findings for reporting in the relevant deliverables.

Having taken relevant factors into consideration, the WP3 lead partner decided to focus on three types of PD study (Table 1) which WP3 partners have been asked to carry out, thereby allowing sufficient amount of data to be collected for drawing more solid conclusion. WP301 and WP302 focus on Splash Lab and Sinking & Floating ILS and on Electricity Lab and Ohm's Law ILS, respectively. The main rationales are as follows: (i) Sinking & Floating ILS implements a typical five-phase Inquiry Learning Cycle whereas Ohm's Law represents an alternative, using more than five phases; (ii) Both ILSs integrate a relatively high number of scaffold apps; (iii) Both labs are owned by Go-Lab; formative evaluation could have significant impacts on their future development. WP303 is complementary to study scaffold apps that are not covered by the other two studies.

WP3 PD Study Catalogue			**Outline Research Plan**			
Study no	**Lab and ILS**	**Scaffold Apps**	**Research methods**	**End-users**	**Time**	**Equipment, resources & software**
WP301	Splash - Sinking and Floating	Concept mapper, Questioning scratchpad, Experiment Design tool, Scratchpad – personalised	Composite study with the use of ILS, PDot, Printed feedback booklets, Questionnaires, Group discussion	>30 teachers from several countries	1.5 to 2 hours	Facilitator script and notes, questionnaires annotated Powerpoints, feedback booklet
WP302	Electricity Lab - Ohm's law	Concept mapper, Hypothesis Scratchpad, Experiment Design tool, Data Viewer tool, Notepad tool, Scratchpad - personalised, Spell Checker, Calculator, Drop Files.	Composite study with the use of ILS, PDot, Printed feedback booklets, Questionnaires, Group discussion	>30 teachers from several countries	1.5 to 2 hours	Facilitator script, Questionnaire, PowerPoints, Feedback booklet
WP303	N/A	all other available scaffold apps	Composite study with the use of working scaffold apps and of Power-Point for yet to be built tools. Feedback with questionnaires, PDot, Audience response system, Group discussion	>15 teachers from at least 2 countries	1 to 1.5 hours	Facilitator script, Questionnaire, PowerPoints, Audience response system.

Table 2.1. Three main types of PD study in Year 2

3. Face-to-face Participatory Design Studies

There were altogether 21 face-to-face PD sessions conducted during Year 2. Table 3.1 shows an overview. Results thereof are analysed, consolidated and presented in the following sub-sections.

Event ID	Date	Partner	Location	No. of teacher/ student	Description
PD01	18/11/2013 22/11/2013	EPFL & UT	Bali / Indonesia	18 / 0	ILS authoring
PD02	14/01/2014	EPFL	Switzerland	1 / 13	1.5 hrs ILS usage in class
PD03	22/01/2014	UCY	Cyprus	7 / 0	1.5 hours
PD04	28/01/2014	EPFL	Switzerland	1 / 0	ILS / ILS authoring (4 hours)
PD05	06/03/2014	UCY	Cyprus	12 / 0	Feedback on scaffolds (CM, HS) with paper-and-pen by pre-service teachers
PD06	06/03/2014	ULEIC	United Kingdom	6 / 0	Feedback on scaffolds (CM, HS) with PDot; group discussion (2 hours)
PD07	07/03/2014	ULEIC	United Kingdom	0 / 30	Feedback on scaffolds (CM, HS) with paper-and-pen and PDot by university students (2 hours)
PD08	12/03/2014	ULEIC	United Kingdom	0 / 30	Feedback on scaffolds (CM, HS) with PDot by primary school students (1 hour)
PD09	24/03/2014 25/03/2014	UT	Netherlands	0 / 27	Feedback on scaffolds (HS, EDT) and Splash Lab by primary school students
PD10-A	31/03/2014	UCY	Cyprus	17 / 0	Feedback on the Electricity ILS by pre-service teachers: Sessions 1, 2 &3 (3 * 1.5 hours)
PD10-B	03/04/2014	UCY	Cyprus	17 / 0	
PD10-C	07/04/2014	UCY	Cyprus	17 / 0	
PD11	24/04/2014	EPFL	Switzerland	3 / 0	Graasp redesign and ILS (7 hours)
PD12	04/05/2014	EA	Portugal	23 / 0	Workshop on Big Ideas of Science (4 hours)
PD13	09/05/2014	EPFL	Switzerland	1 / 12	1.5 hours ILS usage in class
PD14	13/05/2014	EA	Greece	15 / 0	Workshop on Big Ideas of Science (4 hours)
PD15	13/05/2014	UDEUSTO	Spain	1 / 0	Feedback on text editor and wiki scaffold with paper-and-pen by a teacher
PD16	26/05/2014	UDEUSTO	Spain	1 / 25	Feedback on scaffolds (HS, EDT) with paper-and-pen by a teacher and secondary school students.
PD17	27/05/2014	EA	Greece	12 / 0	Workshop on Big Ideas of Science
PD18	23/06/2014 - 26/06/2014	UT	Netherlands	0 / 53	27 3rd grade students (14-15) and 26 4th grade students (15-17) using the Electricity lab and filling in questionnaires.
PD19	24/06/2014	ULEIC	Denmark	0 / 11	Feedback on Splash ILS: Sinking and Floating with PDot by university students
PD20-A	13/07/2014 - 18/07/2014	ULEIC	Greece	39 / 0	Observation notes on the teachers' interactions with Go-Lab artefacts
PD20-B	13/07/2014 - 18/07/2014	MENON	Greece	39 / 0	WP6 activities at the Go-Lab summer school resulting in some WP3-related findings on the authoring environment
PD20-C	13/07/2014 - 18/07/2014	EPFL	Greece	39 / 0	Observation notes on the teachers' interactions with Go-Lab artefacts
PD20-D	15/07/2014	ULEIC & EPFL	Greece	39 / 0	Feedback on scaffolds with audience response system
PD20-E	15/07/2014	EPFL	Greece	25 / 0	Feedback on specific features of Golabz
PD20-F	16/07/2014	ULEIC	Greece	33 / 0	Feedback on Electricity ILS with PDot (2 hours)
PD20-G	17/07/2014	ULEIC	Greece	39 / 0	Feedback on Splash ILS with PDot (1.5 hours)
PD21	07/10/2014	EA	Greece	1 / 20	Video-recorded lesson using Splash mini-ILS

Note: PD20 denotes the Go-Lab Summer School (13-18 July 2014, Marathons, Greece) where a series of PD sessions were held, involving all or a subset of 39 teacher participants; they are differentiated by suffix A-G.
HS = Hypothesis Scratchpad; CM = Concept Mapper; EDT= Experiment Design Tool

Table 3.1. Face-to-face PD studies in Go-Lab Year 2

3.1 Participant profiling

In the study programme it was recognised that it was important to elicit feedback from both students and teachers, since they approach the system with different needs, perspectives, concerns, capabilities, purposes, and priorities, and may have different (and potentially conflicting) usability and functionality requirements. Some components of the system (for instance ILS authoring) are used only by teachers. Both sets of perspectives are important.

It was also recognised that we should attempt to access as broad a range of science teachers and students as possible within the target demographic, to include perspectives from a range of ages, nationalities, and subject specialities. Given the immense diversity of teachers and students, and the nature of our usability studies, and the large number of Go-Lab components to be tested, it was judged more important to cover a breadth of Go-Lab components rather than to try for a completely representative sample of participants for each component; that is we prioritise system coverage over certainty and generalisability of findings.

Over the course of the study programme this year, we have conducted 21 face-to-face Participatory Design events, lasting from 1 hour to 7 hours (mode: 1.5 hours) in 9 countries, involving participants of 16 nationalities. The 158 science teachers have been involved; more than 75% are female. Physics has been the most common speciality, but they have also included biology, mathematics, chemistry and general science teachers. 221 students have taken part, with an age range of 9-20 (mode: 14-16) years old.

In addition, rather than assuming that the findings from our participants can be generalised to a wider population, we have often endeavoured to gain an understanding of the likely usability experience of a wider population by asking teachers questions about the usability experience in three forms: “for you”, “for other science teachers” and “for your students”. The premise is that teachers may have insights which are not obvious to the project partners about the likely generalisability of their usability assessments.

As described in later sections, the face-to-face activities have been supplemented with remote studies using a Core Teacher Group (Section 4), and with Heuristic Evaluations conducted by Human-Computer Interaction researchers (Section 5), in order to obtain an even wider range of perspectives. The effectiveness of this sampling approach and a comparison of the findings from these different approaches are discussed at Section 7.2.

3.2 Participatory Design Protocols

In Year 1, Participatory Design Protocol (PDP) was developed for evaluating early Go-Lab design artefacts which were low-fidelity mock-ups. The main use of PDP was to provide the WP3 partners with actionable guidelines (with customization wherever deemed appropriate) to carry out PD activities in their local sites. In Year 2, a much larger range of artefacts with higher fidelity have been available. It was necessary to revise PDP of Year 1 substantially.

Nonetheless, due to the wide range of studies to be conducted – involving varying numbers of teachers and/or students, sessions of various durations, different system components or research questions to be evaluated, different facilities and equipment, session facilitated by researchers or teachers, and a host of other factors, almost every PD session has a bespoke protocol, though

some materials (e.g., questionnaires or PowerPoints) may be reused. Occasionally, some sessions could only be arranged in a short notice because of the events rescheduling in the institutions concerned. In these cases, protocols have taken the form of general guidance and a toolbox of resources rather than a formal script.

As it is deemed impractical to include in this deliverable the PDPs of all individual studies, which are made accessible in a Graasp folder, for the sake of reference, we show two illustrative samples of PDPs and related resources: ***Appendix A – PDP Splash ILS*** and ***Appendix B – PDP Electricity ILS***. Typical components of a face-to-face PD workshop from a participant's perspective include:

- Attending presentations on the introduction to the objectives and procedure of the study;
- Interacting with the prototype components under evaluation to attempt the related tasks;
- Providing feedback using paper-and-pen or software-supported tool PDot, filling in questionnaires, participating in interviews, and being observed unobtrusively.

3.3 Consolidated Results and Recommendations

Given the volume of data collected in the 21 PD studies, this document would be overwhelmingly intense if all the related results were presented study-by-study. Hence, a better strategy is to consolidate the findings with respect to individual components of the system; a method of presentation also requested by the development team members. Nonetheless, as shown in Table 3.1, the event PD20 A-G "Go-Lab Summer School" contains seven independent sessions, a separate section (Section 3.3) covering the main findings of these sessions is deemed necessary. Similarly the event PD21 "Video-recorded lesson using Splash mini-ILS" is reported separately since it uses a specific methodology.

The studies conducted for WP3 have sometimes produced findings of more interest to other work packages. Some of these findings are mentioned here, but more importantly they have been passed on to the relevant WPs directly. As this is a formative evaluation report, and so most of the detailed comments recorded in this section describe areas which could be improved. There is no attempt to list all the many aspects of the system which are already perfectly functional and usable as this would serve no useful purpose.

In many cases the findings listed below have been identified and communicated prior to this report, and in some cases the development team's actions are already planned or even completed which fully or partially satisfy these findings. The responses so far provided by the development team leaders are summarised in the sections below, and a fuller set of responses is provided in ***Appendix E*** and ***Appendix F***. The responses demonstrate that the development priorities are influenced by the findings of the usability studies. In most cases they describe changes which are at the planning or design stage, or are very newly implemented, so it is not yet possible to determine whether they fully satisfy the user issues. Most studies have evaluated more than one Go-Lab prototypes (i.e., scaffolds, ILS), and often have given rise to serendipitous anecdotal findings about others. The value and significance of the findings is more recognisable when they are collated by Go-Lab component, and this analysis is also of more use to the Go-Lab pedagogical and technical teams. Nonetheless, we begin with some general findings.

In the following Usability Observation tables, the 'Frequency/Source' column reports two values: For Frequency, the following codes are used: Frequently (*F*; more than 10 times), Sometimes (*S*; 5-9

times), Occasionally (*O*; 2-4 times), or Rarely (*R*; once). For Source, as the findings of some components have been obtained from both participating students and teachers, we mark the column with "S" if the finding is from students only, "T" if it is from teachers only, and "ST" if the finding is reported by both students and teachers. Otherwise, when no explicit source is given, the item is from the teachers. In Section 7.2, we discuss the comparative effectiveness of different types of study with different participant groups.

3.3.1 General Findings

This sub-section describes various recurrent themes emerging from the body of research. These have been communicated to the Development Team and they are planning to make many appropriate software changes. Their responses to most of the actionable points are provided in ***Appendix E - Development Team Responses to General Findings***.

- ***Overall impression***

Based on the data of multiple PD studies, we tend to conclude that there are strong positives from teachers on the potential of *the Go-Lab system* (i.e., a generic term used here to refer to the set of design artefacts with which the PD participants have been exposed to and/or interacted with), especially they often report a strong intention to use. However, the system has been perceived as somewhat awkward to use in parts, and needs work before it is fully ready to be deployed. In other words, they believe it has good potential, good content, good facilities but needs refinement for usability and perhaps other reasons. This is exactly what the project might expect at this stage.

- ***Other tools reported as used***

It has become apparent that many teachers in various contexts use other online tools in teaching situations. There have been many unsolicited anecdotal comments, including suggestions that parts of Go-Lab should be a bit more like Moodle products with which the teachers are familiar. Examples named include: MicroWorlds Ex; Algodoo, Virtual Lab Electricity, Stagecast Creato, Thermolab, Kidspiration, Inspiration, Stellarium, and Pt.lab2go.net. In addition, examples of quiz tools named are: HotPotatoes, Moodle (SCORM), Google module, Quiz Faber, and Flubaroo.

- ***Sign-on security***

There are also sometimes undesirable real time interactions if two browsers are logged into the same ILS at once using the same nickname. Maybe a sign-on process could prevent this, perhaps offering a facility to close the one already logged on. Also, the system seems to remember one's log-on from session to session, with no clear way of logging off. This can cause problems for shared computers (e.g., in school computer labs).

- ***Storage and retrieval of work***

Frequent questions on where a user's work is stored and how it can be retrieved were raised. So were the questions on privacy as well as plagiarism risk. There was also a related observation that tools do not save their content if the user reloads a page. Some kind of auto-save functionality consistently applied might be beneficial. Better still when a person signed on, the system could present a list of ILSs in progress and offer a resume function.

- ***Undo and Redo, Retyping***

It would be beneficial for UNDO and REDO facilities to be provided where possible. In some ILSs, it seems necessary to enter data more than once – e.g., in EDT and lab. Ideally all components in an ILS ought to link so data is pre-populated where possible

- ***Scaffold toolbar visibility***

Several teachers could not find the pull down toolbar. Some tools in the scaffold drop-down toolbar appear too small for use in some circumstances. The facility to make them wider is not visible or not recognisable.

- ***Consistent interaction design***

Scattered comments converge to the point that it would be good if all components in an ILS had the same interaction design (e.g., drag and drop paradigm). It is confusing if one scaffold is different from the other in terms of visual aesthetics.

- ***Help features***

There were frequent requests for help information or examples, which could be implemented as links to YouTube demo videos, showing users how to deploy various labs or scaffolds. There should be a consistent design for help or guidance for all tools. Besides, it is recommended to use well-known iconography with the support of textual explanation, being presented underneath an icon or being provided as "tool tip" which pops up when one hovers over an icon.

- ***Wish for facilities to monitor students' work and support assessment***

Classroom management seemed a serious concern for most teachers, who, for instance, expressed their wish to view all concept maps being created during a lesson, and suggested that such monitoring facilities could be provided through a dashboard. Teachers were also keen to demonstrate teaching effectiveness for school assessment or inspections. There were also requests for a print function for student use (e.g., within Concept Mapper) and help teacher marking.

- ***Facilities to support group working***

Having students worked in groups is thought to be good pedagogical practice by some teachers. Two models: (i) groups of less able students are given slightly easier objectives and at the same time with more support than groups of abler students are given; (ii) each (mixed ability) group member has different complementary role assigned.

- ***Distinguishable interactive and non-interactive components***

Some users "classified" the content of an ILS tab as either being "text" or "image", not recognizing that the non-textual parts can have different degrees of interactivity (from "non" for pictures to "some" for videos and "fully" for scaffolds and labs). Also, sometimes text decorations used to make the learning content more exciting and engaging were in violation with common "Internet rules" for (hyper)text, e.g., underlined text on a Website usually being/showing/highlighting a (hyper)link. From a usability perspective the user experience could in this case be improved by:

  - Providing (visual) clues (e.g., appropriate icons next to the content) to aid the user in his or her task to semantically structure the page (i.e., this is textual learning content, this is a video containing learning content, this is a learning task to accomplish, this is the lab that should be used to conduct the experiment). Those clues should also be consistent throughout all ILSs to support the user further in recognizing the type of content independent of the current ILS.
  - Make teachers aware of the "Internet rules" for text and other content (like "Text that is no hyperlink, should not be underlined. Use bold or italic text decoration instead to emphasize text fragments."), e.g., by providing a "How to design a Go-Lab ILS" booklet.
  - The ILS design provided by the Go-Lab designers and developers should also reflect conventional Internet and design rules (e.g., links should be blue and underlined).

- ***Relevant information for ILSs in GoLabz***

When selecting an ILS in Golabz, a key criterion to be provided is how long it typically takes to work with it with students, enabling teachers to estimate whether (or when) it may fit in a school timetable.

- ***Multi-lingual components***

A frequent comment has been that tools are not currently available in the participant's own language.

- ***Usable scrolling***

All ILSs should support sideways scrolling for cases when the window is not wide enough. It is advisable to avoid scrolling within scrolling, where possible,

- ***Browser issues***

Some apparently contradictory comments seem caused by the fact that some components of the system are sometimes rendered differently, depending on which browser and which version is used. Ideally the system should appear identically in any up-to-date version of any of the three main browsers used by teachers: Microsoft Internet Explorer, Google Chrome and Mozilla FireFox (cf. the survey results reported in D3.1). Another more compelling issue is browser refresh; some links (including the user's own name at the top of an ILS) would take the user all the way back to the start of the ILS and cause the user to lose her work.

3.3.2 Scaffold Apps

In the ensuing text, we present consolidated findings for individual scaffold apps, which have been tested in one or more of the 21 PD studies listed in Table 3.1 (including the data from the seven sessions of the Summer School). The Development Team have been asked to document their responses to the findings. Such responses are categorized (Table 3.2) and summarised below each component. More detail of the responses is provided in ***Appendix F - Development Team detailed responses to some of the Specific Usability Findings.***

Response category	Description
Software already changed and verified by end users	Component already changed to remedy the usability observation, and remedy verified by subsequent user studies
Software changed – not yet retested	Component already changed to remedy usability observation; remedy not yet verified by subsequent user studies
Software change – work in progress	Component change to remedy this usability observation is in progress; not ready for testing.
In design	Change to remedy this usability observation is being designed.
Planned	Change to remedy this usability observation is planned.
Willing	Development teams agree that a change should be implemented to remedy this usability observation.
Considering	Development teams recognise the usability observation and are considering whether a change should be implemented to remedy it.
Questioning importance	Development teams are not yet decided whether the work involved in remedying this usability observation is justified.
Do not intend to address the issue	Development teams believe the usability benefits of remedying this observation do not justify the costs.

Table 3.2. Categories of Development Team responses

- ***Hypothesis Scratchpad (HS)***

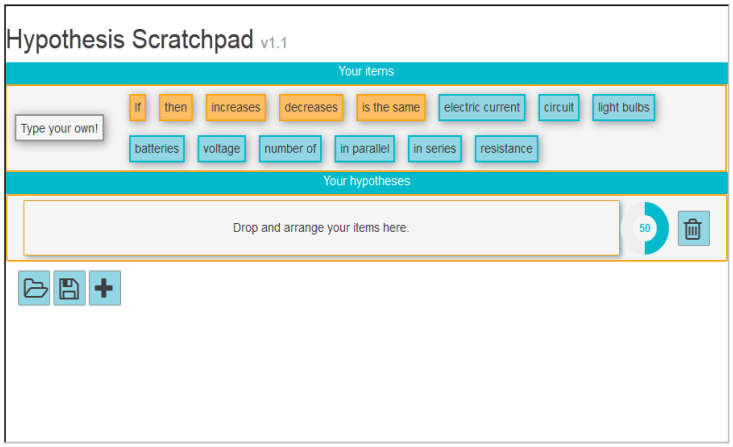


Figure 3.1. Screenshot of Hypothesis Scratchpad

The Hypothesis Scratchpad scaffold (http://www.golabz.eu/app/hypothesis-tool) allows students to formulate hypotheses to guide their following experiments and data analysis. Hypotheses are created with this app by dragging items (teacher-provided or individual words in boxes) and dropping them one after the other to form the desired sentence.

Overall evaluation of Hypothesis Scratchpad (HS):
Generally well liked, compared to other scaffold tools, and perceived as useful and fairly easy to learn and to use, but would benefit from some usability improvements and minor bug fixes.

	Usability Observation	Frequency/ Source	Recommended response
HS1	List of supplied words is very specific for a particular lesson and for older students who understand “gradient”. But it would be huge and unnavigable if all words needed for all lessons were supplied. Would be nice for teacher or ILS writer to be able to provide a list.	Sometimes/ ST	Consider an enhancement to allow custom vocabularies to be supplied
HS2	Users could not delete a single word from hypothesis – bin icon deleted the whole hypothesis	Sometimes/ S	Style Guide - interaction
HS3	When user deleted all text from “type your own” box, the box disappears and becomes unusable	Sometimes/ S	
HS4	UNDO and REDO needed	Sometimes/ ST	Style Guide
HS5	Sequence the lexicon alphabetically, or by most recently used. Also when words are typed in “type your own” box, add them to the lexicon for this user.	Sometimes/ S	Alphabetic is probably better.
HS6	Editing a hypothesis – e.g., inserting a missing word – is not easy (“difficult, frustrating and error-prone” according to one user). Drag doesn’t seem to work and some participants found “type your own” easier. It’s also difficult to rearrange word order. One user suggested when dragging an element between two others, it would be nice if they separated and a space opened up automatically.	Sometimes/ ST	Enhance facilities for editing hypotheses.
HS7	Deleting boxes requires very precise dragging over the bin, else it doesn’t happen also large boxes with lots of text are impossible to delete.	Sometimes/ ST	Tune / improve deletion sensitivity.
HS8	No confirmation on save or delete	Sometimes/	Style Guide

		ST	
HS9	No spell checker	Occasionally/ S	Possible enhancement – add spell checker
HS10	Possible to delete words from supplied list by dragging to bin, then it gets frustrating if you need that word	Occasionally/ S	Remove the ability to delete pre-supplied words in this way
HS11	When several hypotheses are written, tools at the bottom of the screen vanish	Occasionally/ S	Provide scrolling in this case
HS12	Interaction inconsistent with Experiment Design Tool	Occasionally/ ST	Style Guide - interaction
HS13	Colour coding of words is not self-explanatory / consistent	Occasionally/ S	Use a consistent approach or just one colour (or let teacher specify – see HS1)
HS14	Too easy to delete all by mistake	Occasionally/ S	Provide warning before deleting
HS15	"type your own" box should start off as blank, not contain previously typed entry.	Occasionally/ ST	Suggest it always starts with "type your own" in. See HS3.
HS16	Text in boxes can't be edited.	Occasionally/ S	Perhaps working as designed so no action except providing help text.
HS17	Add / Save / fetch icons: purpose unclear and not working	Occasionally/ ST	Make them work and provide help text
HS18	You can drag objects off the screen and into oblivion	Occasionally/ S	Don't let this happen
HS19	Should be able to change colour of boxes	Occasionally/ S	Possible enhancement
HS20	Hints useful but not noticed soon enough	Rarely/ T	Put hints nearer to the top?
HS21	"Plus" icon not recognised as enabling user to add another hypothesis	Rarely/ ST	Style Guide
HS22	The box in which you build up your hypothesis should be bigger	Rarely/ S	Consider making it bigger
HS23	Provide a larger selection of logical operators (IF, AND, OR, ELSE etc)	Rarely/ ST	Possible enhancement. See also HS1.
HS24	Target audience would prefer a brighter and more youthful aesthetic.	Rarely/ S	Consider visual design
HS25	There are no accessibility features for users with limitations – e.g., to change colours for people who are colour blind or fonts for those who are partially sighted	Rarely/ S	Consider accessibility features
HS26	Hypotheses should be numbered	Rarely/ S	Add numbers to hypotheses
HS27	Help text not very helpful, and would be more readable if broken up – e.g., with bullet marks.	Rarely/ S	Style guide

Table 3.3a. Summary of the findings, and recommendations for Hypothesis Scratchpad (HS)

Response category	Usability observations
Software changed – not yet retested	HS1 (partial), HS3, HS5 (improve), HS6, HS7, HS8 (partial), HS10, HS17, HS18
Software change – work in progress	HS1 (partial)
Willing	HS2, HS8 (partial), HS12, HS13, HS19
Considering	HS4, HS14
Questioning importance	HS9

Table 3.3b. Development Team response summary to Hypothesis Scratchpad (HS)

- ***Experiment Design Tool (EDT)***

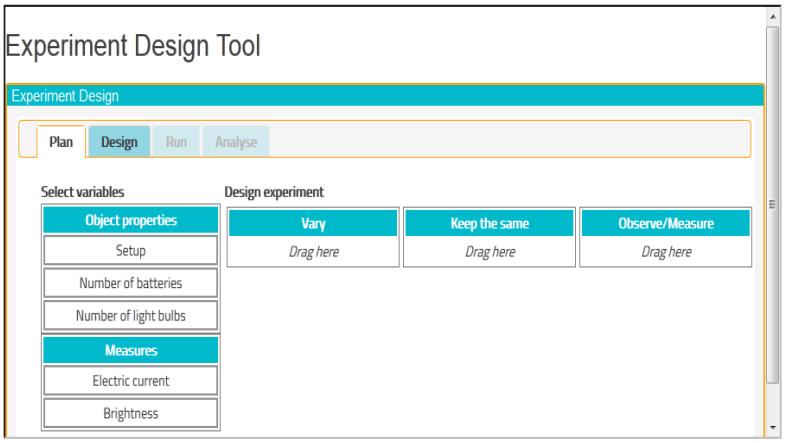


Figure 3.2. Screenshot of Experiment Design Tool

Note: not all studies through the year may have used the same version of the scaffold.

The Experiment Design scaffold (http://www.golabz.eu/apps/experiment-design-tool) creates an environment for the learners to design scientific experiments. There are pre-defined set of variables in this tool that the learners need to choose among them which ones vary, stays constant and supposed to be measured during the experiment.

Overall evaluation for Experiment Design Tool (EDT):
Generally well liked, compared to other scaffold tools, and perceived as useful and fairly easy to use. For instance, it attracted more LIKEs and fewer DISLIKES than a typical scaffold tool in PDot, and was the most frequent response to “best part of the system” (cf. p.16 of Appendix A) and also scored well in the “stickers” exercise in that study (cf. p.10 of Appendix A). The impressions were less good when it was used in the “Sinking and floating” ILS, with many negative comments in the feedback booklet and PDot. Perhaps the discrepancy was due to different versions, or perhaps participants focussed on different parts of it in different sessions. It seems that some teachers used only part of the tool, and overall it is regarded as rather complex. It would benefit from better help information and guidance, or even (as a couple of teachers suggested) a video tutorial.

	Usability Observation	Frequency/ Source	Recommended response
EDT1	Needs explanatory material and guidance or examples or video tutorial,	Frequently ST	Provide online help and/or tutorial video, explaining purpose of each function as well as how to use it.
EDT2	Should be able to continue without having to define all the variables	Frequently T	If there’s a good pedagogical or scientific reason for this restriction, then it should be communicated to the user better. If not, remove this restriction.
EDT3	Padlock not noticed, and/or facility to drag padlock not recognised	Frequently ST	Provide more on screen help or better visual clues.
EDT4	RUN tab not understood as it doesn’t run the lab	Frequently ST	Consider splitting the tool so that different functions can be presented to students at different phases of the ILS. The RUN tab is probably not needed until after the lab is used.

EDT5	Navigation from tab to tab sometimes constrained and not well understood	Sometimes T	Make the options clear. Provide explanatory messages when something not possible is attempted. Consider whether tabs are the best visual metaphor, or is a "Wizard" style interaction more appropriate.
EDT6	Plan tab has an annoyingly persistent confirmation message with a tick box to stop seeing the message which doesn't work.	Sometimes T	Make tick box work as users expect.
EDT7	Design tab asks you to specify values for mass, volume and density, but in fact they need to be consistent; one needs to be derived from the other two. It's possible to type and see an inconsistent set with no error message.	Sometimes S	If the data in this tab is supposed to be merely aspirational (what the student hopes to do) then this needs to be clear. (In practice a design process would probably involve selecting a material, then a size or weight).
EDT8	Not consistent with (other scaffold apps)	Occasionally S	Style Guide
EDT9	"data has to be retyped"	Occasionally T	Style Guide (implement Vault; pre-populate fields in scaffolds if data is available in ILS)
EDT10	To change a value you must delete the whole row!	Occasionally S	Provide an edit facility. Style Guide.
EDT12	Adjustment of range of variables it too limited, Why can't you have an object with volume of $1cm^3$?	Rarely S	Extend ranges
EDT13	"It seems very specific for a particular experiment. Is it generalisable?"	Rarely ST	Fundamental design decision to be made. Do we want multiple EDTs or a single, rather generic one?
EDT14	No facility to plan more than one experiment	Rarely T	Consider providing this
EDT15	On Analyse tab, it's not clear, obvious or intuitive that you have to drag icons to Data Collector.	Rarely T	Provide better help and/or visual clues.
EDT16	Too many choices!	Rarely T	Consider a more guided interaction rather than presenting lots of facilities at once.
EDT17	Why no facility to save results?	Rarely T	Style Guide – data strategy
EDT18	The list of "Measures" should visually indicate which ones have already been used – e.g., by changing colour or removing them from the list.	Rarely T	A possible usability enhancement
EDT19	One "measure" ("volume of displaced fluid") appears twice	Rarely T	Minor bug fix.
EDT20	Slider for "specify a value for density objet" should also allow a value to be typed.	Rarely S	Provide a text box.

Table 3.4a. Summary of the findings, and recommendations for Experiment Design Tool (EDT)

Response category	Usability observations
Planned	EDT2
Willing	EDT1, EDT3, EDT4, EDT5, EDT6, EDT7

Table 3.4b. Development Team response summary to Experiment Design Tool (EDT)

Note that the Experiment Design tool interacts with whichever lab is being used, and some of the issues reported as EDT usability issues may in fact be caused by the lab in use. Usability reports by

participants in PD activities tend to categorise issues based on visual appearance, not on an understanding of the way components interact.

- ***Concept Mapper (CM)***

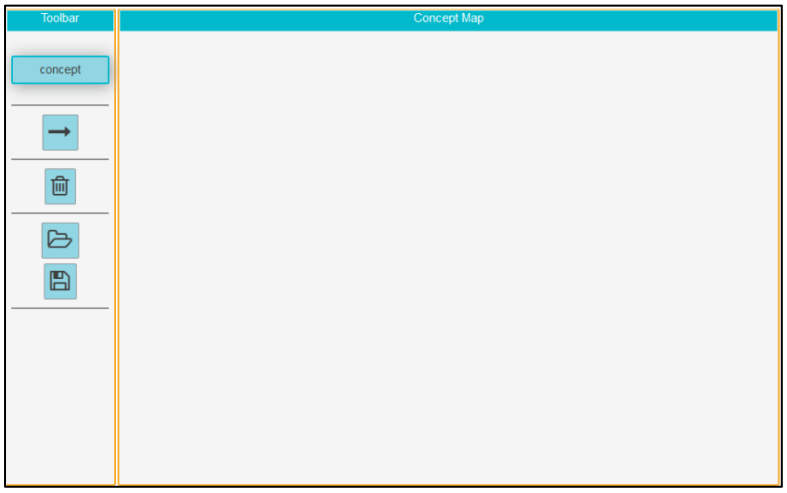


Figure 3.3. Screenshot of Concept Mapper

The Concept Mapper scaffold (http://www.golabz.eu/content/go-lab-concept-mapper) allows learner to create concept maps in order to get an understanding of the key concepts within a scientific domain and how they relate to each other. To create a concept map, user can easily drag and place concept box from the toolbar into the right side map and then connect them with use of arrows.

Overall evaluation of Concept Mapper (CM):
Despite its apparent simplicity and usefulness, this scaffold app attracted poor scores. For instance, it was most frequently nominated as the worst part of the system (p.16 of Appendix A), probably due to some usability problems listed below. Some participants used words such as “frustrating” and “annoying” in their responses. Some teachers advised using it with caution and not too early in the ILS as it can stifle creativity by guiding the student’s thoughts.

	Usability Observation	Frequency/ Source	Recommended Response
CM1	Concept boxes use drag and drop paradigm but arrows change mode and use a “click both ends” paradigm.	Frequently S	Use a consistent style within (and preferably between) apps. (Style Guide). Suggest drag and drop for both.
CM2	Mode change to arrow more not sufficiently visible – other things seem to stop working	Frequently ST	Either eliminate arrow mode (See CM1) or make it more visible. Grey out any tools or icons which are temporarily inactive.
CM3	Delete facility deletes the whole lot	Frequently ST	Style Guide
CM4	No way to delete arrows, or “Deleting arrow by re-drawing it in reverse is not intuitive”	Frequently ST	Suggest drag to bin as way of deleting arrows
CM5	Provide an UNDO facility	Sometimes T	Style Guide
CM6	Provide a facility to delete a single concept	Sometimes ST	Style Guide
CM7	Not clear why “select or type” and “concept” buttons are different.	Sometimes ST	Suggest single “concept” button which allows both types of interaction

CM8	Provide a facility for teachers to customise the list of pre-written concepts as it will vary from lesson to lesson	Occasionally T	Enhancement request
CM9	Arrow mode should switch off automatically after and arrow is created	Occasionally ST	Either eliminate arrow mode (see CM1), or make the mode more visible (see CM2). If mode not eliminated, consider whether or not switching back to normal mode automatically would be an improvement.
CM10	Needs help facility	Occasionally ST	Style Guide
CM11	Need bigger list of concepts if this is general purpose	Occasionally S	Consider custom lists of concepts – either by ILS writer or by teacher
CM12	It would be nice to be able to change the size of a box for emphasis	Rarely S	Possible enhancement, though a facility to change colour may be more useful
CM13	No facility to have double headed arrow, or two arrows between the same two concepts, or a connection between a concept and itself	Rarely S	A possible enhancement
CM14	Needs accessibility features – e,g, to customise fonts or colours for visually impaired users	Rarely S	Possible enhancement
CM15	When you want to delete something, why is there not a cross beside it? Why do you have to drag it to the bin?	Rarely S	Possible extra feature (Style Guide), but drag to bin seems widely understood paradigm.
CM16	When you double click or select a box to add text, existing text should be highlighted so that it can be deleted with a single key press.	Rarely/ S	Provide this facility
CM18	Make drop-down lists in alphabetical order	Rarely./ S	A usability enhancement
CM19	Allow users to add a title to their work	Rarely/ S	Functional enhancement request
CM20	Bin must ask for confirmation / make bin red to warn of dangers!	Rarely/ S	Style Guide - confirmation
CM21	Add a spell checker	Rarely/ S	Functional enhancement request
CM22	Transparent text box on arrow is hard to read	Rarely/ ST	Good minor usability enhancement
CM23	Allow pictures	Rarely/ S	Functional enhancement request
CM24	Provide UNDO, REDO, PRINT, IMPORT and EMAIL buttons	Rarely ST	Style guide

Table 3.5a. Summary of the findings, and recommendations for Concept Mapper (CM)

Response category	Usability observations
Software changed – not yet retested	CM7, CM8, CM11, CM14, CM18, CM20
Planned	CM10
Willing	CM1, CM2, CM3, CM4, CM6, CM9, CM12, CM13, CM15, CM16, CM22
Considering	CM5
Questioning importance	CM21
Other	CM19 is available by existing functionality which the user didn't recognise; CM23 & CM24 – development team needs more clarification / detail before they can respond.

Table 3.5b. Development Team response summary for Concept Mapper (CM)

- ***Data Viewer Tool (DVT)***

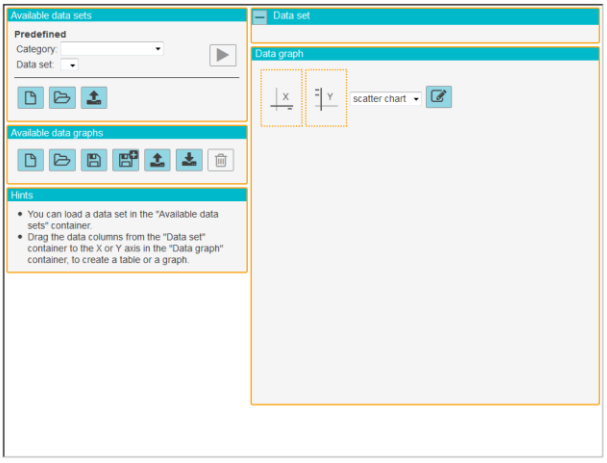


Figure 3.4. Screenshot of Data Viewer Tool

The Data Viewer scaffold (http://www.golabz.eu/apps/data-viewer) provides different functionalities and features to visualise and organise data from the experiments. Data visualisation in this tool can be performed by use of bar chart, scatter plot, table and so on. Learners can drag data columns from the “Data set” container to the “Data graph” to observe the relationship between variables in a table, chart or a graph.

Overall evaluation for Data Viewer Tool (DVT):
This tool does not appear in many ILSs and so has had comparatively little attention in PD activities. The comments below are therefore not particularly conclusive. Some of the function appears to have been incorporated in other tools (particularly the Experiment Design Tool).

Examples of qualitative responses:

- One user said “couldn’t use it”
- “can’t find own data”
- “after trying simulation, Greek letters and comments begin to appear”. (Interpretation: these were probably browser messages of some sort, which were therefore in Greek as the computer was in Greece.)
- Provide scales, units and labels on axes of all graphs.

- ***Questioning Tool / Questioning Scratchpad (QS)***

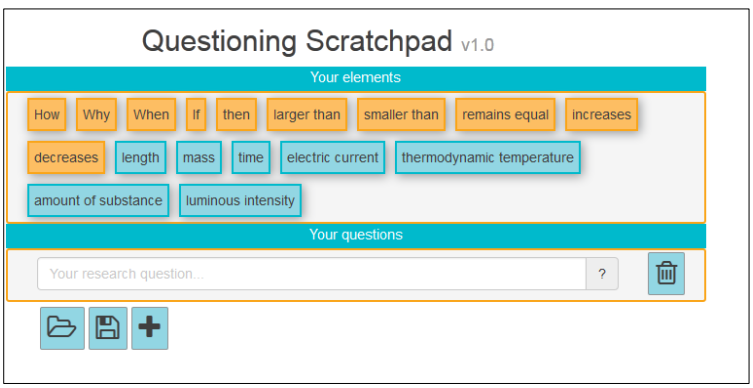


Figure 3.5. Screenshot of Questioning Scratchpad

The Questioning Scratchpad scaffold (http://www.golabz.eu/apps/questioning-scratchpad) allows learners to formulate research questions. In this tool, pre-defined domain terms have been provided (in addition to free text typing) to help students create their queries. These pre-defined words can be dragged and dropped into the question pad.

Overall evaluation for Questioning Scratchpad (QS):
Mixed and borderline evaluations, some uncertainty (some identified it as the best part of the ILS and some as the worst). One teacher likes it because it is creative. One says it is way too difficult, even for secondary school students.

	Usability Observation	Frequency	Recommended Response
QS1	Blank spaces between words disappear	Frequently	Fix usability issue
QS2	Would be improved if it wasn't limited to so few pre-defined words	Rarely	
QS3	Questionable whether this fulfils a useful purpose or is a mere subset of Hypothesis Scratchpad … but students who need more support, or younger students might benefit.	Rarely	Keep the tool and highlight its applicability in the GoLabz portal

Table 3.6a. Summary of the findings, and recommendations for Questioning Scratchpad (QS)

Response category	Usability observations
Software changed – not yet retested	QS2
Willing	QS1

(No response is needed for QS3)

Table 3.6b. Development Team response summary for Questioning Scratchpad (QS)

- ***Conclusion Tool (CT)***

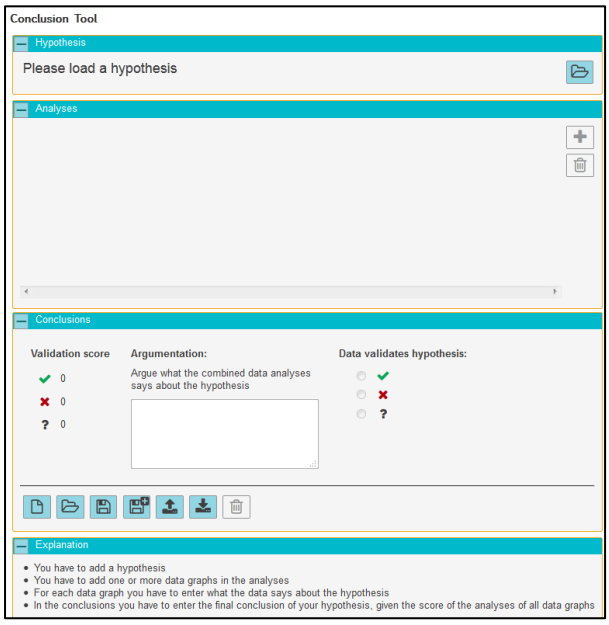


Figure 3.6. Screenshot of Conclusion Tool

Conclusion Tool allows students to retrieve their work such as hypothesis, graphs, and data and so on that they have created in previous phases and try to establish a conclusion based on them. To work with scaffold, students can upload previous works and justifies their conclusion accordingly.

Overall evaluation for Conclusion Tool (CT):
This is a relatively simple tool which provides students with an opportunity to record their conclusion. Few comments and responses were provided partly due to its simplicity and party because it was at the end of the ILS and not all participants reached the end.

Some specific qualitative responses:

- "Useful"
- "Needs online help"
- "Not understood"

- ***Quiz Master Tool***

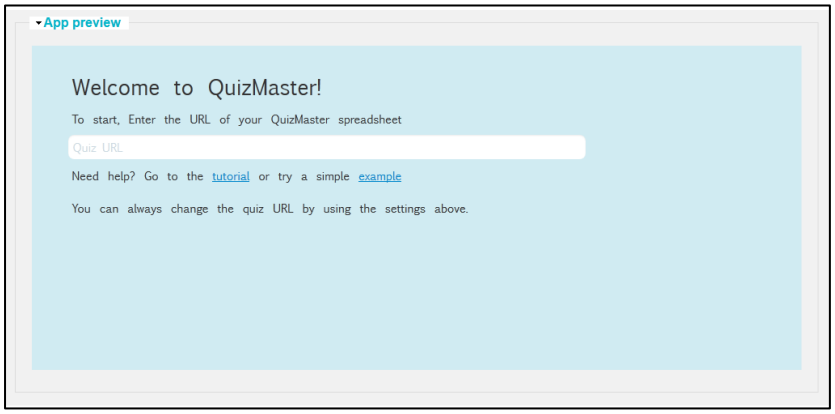


Figure 3.7. Screenshot of Quiz Master Tool

The Quiz Master tool (http://www.golabz.eu/apps/quizmaster) allows teachers to assess students by creating quizzes. Teachers can easily put quiz questions into a Google Spreadsheet along with right and wrong answers. The tutorial to create quiz with Google Spreadsheet is also provided in the above link.

Overall evaluation for Quiz Master (QM)
In response to the questions on the perceived/anticipated *usefulness* and *ease of use* of QM, the majority of the teacher participants in the studies tended to agree that QM could be useful, and that QM was/would be easy for them and their students to use, but were not sure whether their peers would find it so. Despite the estimated positive value, some teachers commented on some usability issues such as the need to install Google Drive. Nonetheless, many of the teachers were already familiar with other web-based quiz tools such as HotPotatoes, Moodle (SCORM), Google module, Quiz Faber, Flubaroo. The prior experience probably shaped their perception as well as anticipation of QM.

- ***Calculator***

Figure 3.8. Screenshot of Calculator Tool

The Calculator scaffold (http://www.golabz.eu/apps/calculator) provides simple math calculations such as addition, subtraction, division, multiplication, exponentiation and square root for the users.

Overall evaluation of Calculator:
This tool was scored well for the metric 'usability', 'intention to use', and 'no need for improvement', but it was scored less for being needed. As one teacher remarked "Some students may prefer to use their own calculator". Another teacher asked for more advanced functions (e.g., sine, cosine, reciprocal).

- ***YouTube Widget (YTW)***

Figure 3.9. Screenshot of YouTube Widget

YouTube widget allows teachers to add proper learning videos to the inquiry space.

Overall evaluation for YouTube Widget:
It was generally regarded as a simple, usable and familiar tool. It has good educational value, especially for engaging students' interest.

Some specific qualitative responses for YouTube Widget:

- Preferably always show an image of the video, not just a link, in ILS
- Avoid videos with too much advertising content – especially when it affects the image of the video in the ILS
- Provide a maximise button if possible

- ***NotePad / Notes tool (NP)***

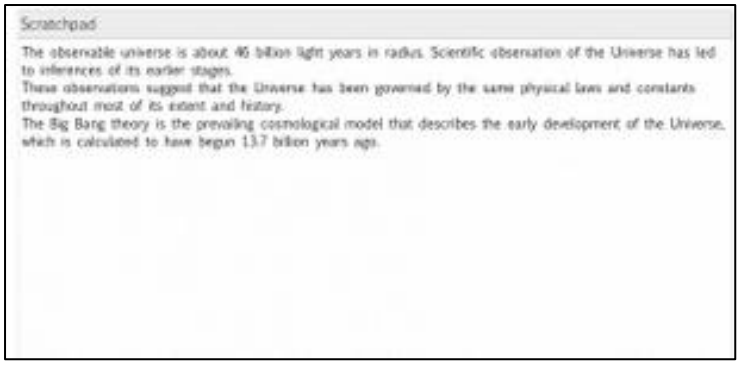


Figure 3.10. Screenshot of NotePad / Notes Tool

The NotePad scaffold (http://www.golabz.eu/apps/input-box) is a note taking tool. Students can simpely write down their notes. To save the notes for later usage, each student needs an user account in the inquiry space. (*Note*: It seems there have been various versions of note-taking tools, all with very similar and simple functionality.)

Overall evaluation of NotePad:
It is a simple and useful tool, but would be best to have a single tool for all note-taking. Some enhancements could be beneficial if it is intended to be heavily used.

	Usability Observation	Frequency	Recommended Response
NP1	Tool sometimes not found – the pull-down toolbar at the bottom of the screen isn't very noticeable	Sometimes	Style Guide + help information (in Go-Lab, not in every ILS)
NP2	The facility to enlarge toolbar apps is not intuitive and is invisible.	Occasionally	Add visible "handles" for enlarging function, or provide something more versatile.
NP3	When loading, at first a blue symbol appears:  and then a green symbol:  is flashing. This might irritate the participant, as it is not obvious why they appear and what they do.	Rarely	Remove distracting displayed emblems
NP4	If two browsers are open with the same ILS and nickname "logged in", editing the notepad content is not possible. Also probably various other unwanted interactions when two instances of the same ILS are in use by the same nickname.	Rarely	Some kind of warning when the second login occurs. Better still some kind of password protection.
NP5	If the user writes a lot (which might happen as the notes tool is used in several phases to answer multiple questions or take notes) the note pad tool becomes unscrollable and therefore unusable.	Rarely	Enable scrolling

Table 3.7a. Summary of the findings, and recommendations for NotePad (NP)

Some specific qualitative responses:

- "We can't write. Tool invades all of the screen (almost)" … the user somehow had got it covering other screen elements
- Useful

Response category	Usability observations
Software changed – not yet retested	NP2, NP3, NP5
Considering	NP4

Table 3.7b. Development Team response summary for NotePad (NP)

- ***Untested scaffolds***

Here below is the list of scaffolds (mainly newer or so far little-used ones) for which empirical data are not yet collected:

- Measurement error tool
- Data analysis tool
- Reflection tool
- Heuristics tool
- Prompts/Assignments
- Wiki app
- Input box
- File drop
- Resource view

3.3.3 Go-Lab Owned Labs

The term 'Go-Lab Owned Labs' refers to the labs, which have been created by the Go-Lab partners themselves rather than those "plugged in" from external sources. Two labs – Splash and Electricity - have been evaluated in altogether five of the PD studies (Table 3.1).

- ***Splash Lab***

Overall evaluation of Splash Lab (SL):
The Lab received mixed reviews: One teacher commented "easy to navigate and enjoyable", another remarked "many troubles here". It was observed in the related studies that almost half of the participants had to ask for help. Generally, the visual design was regarded as attractive and some features were evaluated to be good, but some features were perceived to be complex and were thus not attempted by the participants.

	Usability Observation	Frequency/ Source	Recommended Response
SL1	Problems with entering values: sliders are coarse grained	Frequently ST	Could provide a more sophisticated control, but possibly fixing SL2 may mean this no longer matters.
SL2	Problems with entering values: entering your own numbers doesn't work despite the box appearing active. (No-one correctly guesses that you have to type a semicolon after the number!)	Frequently ST	Style Guide. Make enter and/or tab key effective everywhere in Go-Lab when entering values in boxes.
SL3	Needs help and a video tutorial	Frequently S	Style Guide
SL4	Padlock symbols are not visible for some users (browser issues) and almost no-one guesses that they can be dragged.	Frequently ST	Make them bigger and make the ability to move them more obvious in some way.
SL5	Results table not understood	Frequently ST	See SL6

SL6	Tables and graphs not understood, not wanted (by some) and not used in our sessions. Makes the screen rather busy too. ("Left hand side is good")	Sometimes ST	Consider delivering these as separate tools or steps. Or have a more guided workflow or wizard interaction so they're not all presented at once.
SL7	Some users sometimes find that the lab window expands or moves upwards, obscuring the "object properties" window.	Occasionally T	It can be a genuine bug or a browser issue
SL8	Slider control for density is wrong as it's not a value you can in practice adjust gradually. Better to have a pull-down menu to choose a material.	Rarely ST	Consider this for density, and possible for other potential variables where discreet rather than continuously variable choices are needed.
SL9	It would be nice of changing the volume on the EDT changed the size of the visual representation of the ball.	Rarely T	Possible minor improvement
SL10	Results table must have units beside numbers – sets a very bad example to students	Rarely T	Should correct this
SL11	Graph tool doesn't appear to work	Rarely ST	See SL6
SL12	Relative Density tab not understood	Rarely S	Help / tutorial materials

Table 3.8a. Summary of the findings, and recommendations for Splash Lab (SL)

Response category	Usability observations
Software changed – not yet retested	SL1, SL2, SL3, SL7
Considering	SL5, SL6
Do not currently intend to address the issue	SL10
Other	SL4 & SL8 need clarification / discussion SL9 may need to be referred to Experiment Design Tool owner

Table 3.8b. Development Team response summary for Splash Lab (SL)

- ***Electricity Lab (to be renamed "Circuit Simulator")***

Overall evaluation of Electricity Lab:

On balance, most users found the lab felt easy to use, effective and enjoyable, with an attractive layout and appearance (one user said "a bit like Lego"), but more online help and a tutorial video would be useful.

	Usability Observation	**Frequency/ Source**	**Recommended Response**
EL1	When asked half of participants wanted more "how to use" information.	Frequently ST	Style Guide. More help, examples and tutorials.
EL2	Meters were thought less intuitive and less usable than other components. E.g., you need to drag a small part of the diagram, not the whole (even though dragging the whole thing appears to work visually).	Frequently S	More explanatory text or enlarge the diagram, making it easier to drag.
EL3	Users wanted more space for building circuits	Sometimes ST	Consider revising screen layout
EL4	Support enter, not just tab key, when typing in a box	Rarely S	Usability improvement. Style Guide.
EL5	Some component symbols better understood than others	Rarely S	Provide tool tips

EL6	Provide a quicker and easier way to replace burnt-out light bulbs?	Rarely S	Possible minor usability improvement
EL7	Power button is not intuitive and doesn't really belong in the meters window.	Rarely S	Consider making it possible to click on a switch on the actual circuit board
EL8	Make it possible to type in a voltage rather than using a slider	Rarely S	Possible minor usability improvement
EL9	Hints were good but should be easier to find.	Rarely S	Place them at the top rather than the bottom of the lab
EL10	Add Capacitor and coil components, then suitable for experiments involving electro-magnetism.	Rarely T	Possible worthwhile enhancement or provide a facility for teacher to specify what components will be available.
EL11	Over-complex screen with many features; not easy to know what to do first.	Rarely T	Consider split into multiple tools: (circuit design, data collecting, graphing) each of which can be put into the ILS (if wanted) at the appropriate point.

Table 3.9a. Summary of the findings, and recommendations for electricity Lab (EL)

Response category	Usability observations
Software changed – not yet retested	EL4
Software change – work in progress	EL1 (partial)
Planned	EL9
Willing	EL1 (partial), EL2, EL3, EL8
Considering	EL10
Other	EL11 – feature (simpler version of lab) available, but not chosen by author of ILS

Table 3.9b. Development Team response summary for Electricity Lab (EL)

3.3.4 Go-Lab Designed ILS

As mentioned in Section 2.5, the two ILSs have been chosen because they represent a typical and alternative combination of inquiry learning cycle phases and the range of scaffold apps covered. Note that a lot of responses to the two ILSs were hugely affected and influenced by the overall study experience including the components within the ILS. Overall both ILSs attracted some good comments. Limiting factors were that the overall designs were not fully coherent; the implementation of *Vault* may mitigate this issue.

Note: The findings of the two ILSs are more content-related rather than design-oriented, and they have already been communicated to different teams using different channels (e.g., emails, regular WP online meetings). Informal group discussions of the findings have been undertaken, and such group-based responses are not presented in this section.

- ***Splash Lab: Sinking and Floating ILS***

These findings were based on the version of this ILS at
http://graasp.epfl.ch/metawidget/1/783b7b75afb9aa5843ab2c54366220c5f728d3ea

Here below we present two sets of results: the first set (upper table) is more content-related observations from which some enhanced usability design can be derived; the second set (lower table) is more design-related issues.

	Content-related (c) Observation	Frequency	Derived Usability Finding
cSSF1	Nice to have a story as a starting point	Occasionally	The learning content should be provided in an engaging and exciting way. It could be even further improved by adding an animation, which reads out loud the text and explains it.
cSSF2	Missing guidelines and information in the Conclusion phase	Occasionally	Provide an input box below each question to allow students to note down their answers. Maybe provide a conclusion input field at this page as well.

Table 3.10a. Summary of content-related observations and derived usability findings for Splash – Sinking and Floating ILS (SSF)

	Usability Observation	Frequency	Recommended Response
SSF1	Lab presented in conceptualisation phase – caused lots of misunderstanding.	Frequently	Make it very clear that this is only to explore capabilities and facilities of lab, not to do stuff.
SSF2	Welcome-page is perceived well, although having the username being called “nickname” is confusing.	Occasionally	Reword: Replace “nickname”, e.g., by “You can login to this ILS by choosing a username you like.”
SSF3	Participants prefer numbered lists instead of bullet point lists.	Occasionally	Instead of bullet point lists use numbered lists.
SSF4	Participants liked the tips regarding the use of the Concept Mapper in this ILS.	Occasionally	Provide similar support for all scaffolds.
SSF5	Consider presenting Splash Lab before EDT so they have an idea what’s possible?	Rarely	Possibly, but make the purpose very clear. (In a later version of this ILS, the lab was presented early and it caused confusion)
SSF6	Image of scales in investigation phase has a figure for mass without units – not good science!	Rarely	Show units (grams)
SSF7	More visuals and put it into bigger frame in Discussion phase	Rarely	Instead of using pure text in discussion page, provide visual elements such as pictures.
SSF8	Going back from the link in Conceptualization phase under Questioning Scratchpad leads you back to the beginning of the ILS!	Rarely	It is a bug which needs to be fixed.
SSF9	The current visualization of the different phases in “tabs” is not clear enough for some of the users.	Rarely	Similar visualizations as in tabbed browsing should be used to visualize the phases in tabs.

Table 3.10b. Summary of usability findings and recommendations for Splash – Sinking and Floating ILS (SSF)

- ***Electricity Lab: Ohm’s Law ILS***

The findings below relate to the version of the ILS at
http://graasp.epfl.ch/metawidget/1/27e621bed3c8e57d4aca8a97f8d494a6ec2ab556

Here below we present two sets of results: the first set (upper table) is more content-related observations (cEOL) from which some enhanced usability design can be derived; the second set (lower table) is more design-related issues (EOL).

	Content-related Observation	Frequency	Derived Usability Finding
cEOL1	The icon for each phase of the ILS is very helpful for the students (for visualization). [Comment about image of reading person]	Occasionally	Each phase should be assigned an individual icon (consistently used in all ILSs), to aid the user in recognizing the current phase. This would support the usability concept of recognition rather than recall.
cEOL2	[Arrow from Sticky Note and Pen Icon to this comment] It's good to mark tasks.	Occasionally	Visual clues should be used throughout the text to aid the user in understanding the different sections on the page.
cEOL3	This video should be also embedded inside the ILS and not shown in YouTube. [Comment on Video 1 just being a link]	Occasionally	YouTube videos should be embedded in the ILS and not provided as links.
cEOL4	Good that this concept map appears directly not with a web-link	Occasionally	Not only videos, all (external) learning material should be embedded, if possible, to provide an integrated user experience without shifts of context (e.g., when navigating to another page and back).
cEOL5	Great to add overview what the phase is about [Comment on first content paragraph]	Rarely	Each phase should start with a short introduction, providing guidance for the user.
cEOL6	A visual order is good [Comment between second and third content paragraph]	Rarely	Make sure the text in the ILS is well divided and positioned to improve its readability.
cEOL7	no motivating start [Comment on the first paragraph in the Orientation phase]	Rarely	Present the learning material in an engaging way.
cEOL8	Why showing a video? [Circle around Video 3] It seems that the students are familiar with a series circuit and a parallel circuit, so that they could do and create a setup.	Rarely	Keep it simple. Present only as much information as necessary, to prevent information overload.
cEOL9	[Arrow to link to video "Electric Circuits: The Basics"] should there be a sound? (commentary) [as there is sound in the videos, it was maybe just not working for the participant]	Rarely	Provide additional (“backup”) means to get the learning content (e.g., subtitles in videos).
cEOL10	[Arrow from instruction text] Maybe there should be more guidelines on how to construct the concept map.	Rarely	Make sure to explain tasks to be conducted with scaffolds and how to use them in enough detail as guidance for the students.
cEOL11	An important aspect of inquiry cycle. Giving driving questions or not? We think that it depends on the age of students. [On the Pdot step m2] [Comment on the questions in the bullet point list]	Rarely	Provide driving questions as guidance where appropriate.
cEOL12	students should be given an example from daily life	Rarely	Speak the user’s language (not only for the interface but also in the content).
cEOL13	I like the video [about electric circuits], it would be better to have an interactive way of presentation	Rarely	Develop a scaffold explaining electric circuits in an interactive way.

Table 3.11a. Summary of content-related observations and derived usability findings for Electricity Ohm’s Law ILS (EOL)

	Usability Observation	Frequency	Recommended Response
EOL1	The concept of nicknames is not clear to the users (e.g., what they are used for, if they need to be always the same). E.g., *"What the nickname is for? What's the use for it, must it always be the same*?"	Occasionally	Explain the current system and its background (e.g., anonymity, ability to save and load data per user) in a prominent position or use a more thorough user administration, which the users are already used to from other sites (e.g., moodle).
EOL2	[Arrow from nickname input field to this comment] "*I think nickname is not very useful. If it's a first step to have users, saving tasks and accessing them, then it's OK. Otherwise, it's unnecessary*"	Rarely	Provide thorough user management facilities.
EOL3	[Phases Tabs] "*We love this set of tabs. They are not the same we use yesterday. Data Interpret was missing, discussion was after conclusion, and communication was not there*."	Rarely	Option to generate a "non-standard" ILS is well received by the teachers. The process of adding, removing, and repositioning tabs should be as easy as possible.
EOL4	[End of page] "*It's good that you have this sentence but maybe it could be organized like that...if student is not reading it through and watching video links he cannot proceed to next stage. For example teacher can put timing to each page.*"	Rarely	Add option to set a timer for each phase to the ILS authoring facilities.
EOL5	[End of the page] There could be a possibility to add a poll to the end and only then students can continue if they have filled it	Rarely	Implement scaffold which offers a quiz to be added at the end of a phase to "guard" moving to the next one.
EOL6	"*Can eduCanon tool be integrated*?" (https://www.educanon.com/) [Circle drawn around Video 3]	Rarely	Check if that is possible and let users know what external content can be included and how.
EOL7	"*There could be a programme link to making groups using random sampling*." [Comment on instructions between Concept Mapper and Hypothesis Scratchpad]	Rarely	Implement a scaffold providing this facility.
EOL8	"*All these tools have to be translated in other languages*."	Rarely	Translate all scaffolds and labs to all target languages or if this is too much effort provide an easy to use option for the teachers to do so.
EOL9	"*Why learn something so specific as making graphics for an specific lab when you can use other general programs?"*	Rarely	Allow the usage of more general programs.

Table 3.11b. Summary of usability findings and recommendations for Electricity Ohm's Law ILS (EOL)

3.3.5 ILS Authoring Toolset

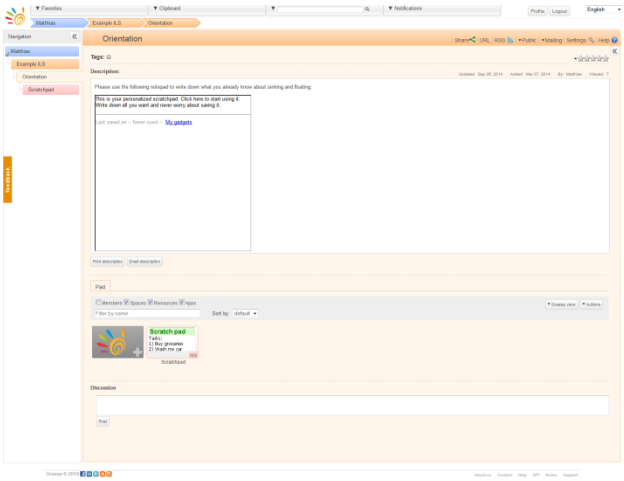


Figure 3.11. Screenshot of ILS Authoring Tool

The main authoring toolset for teachers to develop their own ILS is the Graasp platform (http://graasp.epfl.ch/#url=welcome_space). First, they log in to Graasp, search for the lab they want to use in their ILS in the Golabz repository (http://www.golabz.eu/labs) and then click on the button “Create inquiry space”. This will create the basic structure of the ILS with the five default phases in their Graasp space, including the lab they choose in the investigation phase. The Graasp editing facilities are then used to add content and scaffolds to the different phases of the ILS.

Overall evaluation of ILS Authoring Toolset:
Due to the intermediate state of the Graasp authoring facilities for ILSs, some usability issues have been discovered that are only present in the current version. For the envisioned final version of the authoring toolset some of the processes and interactions required to produce the desired result will be completely different from their current versions. To help shaping the final design of that functionality, issues found regarding those elements are still reported here. At the same time the new design ideas have been tested to gather feedback on them.

In general, the participants find the authoring facilities easy to use, but sometimes struggle when it comes to details. This can partly be attributed to the intermediate state of the toolset (ongoing development with implementing improvements and coding additional functionality) described above. For example, currently teachers need to upload pictures to a separate Graasp space in order to include them in their ILS. In its final state, this functionality of adding pictures to an ILS will be a seamlessly integrated process of choosing a file from one’s hard drive to be included. Also some usability issues have been found which should be addressed as part of the upcoming redesign of the Graasp platform. For instance, the text editor has too many icons, which represent some functionality not needed by the teachers. To avoid information overload, those should be removed, or the Graasp platform is not speaking the user’s language, because the existing general terms are not tailored to the current target group of teachers (e.g., the text area to enter the content of the ILS is labelled “Description” and not “Learning Content”). Although the ILS authoring toolset can thus be used for its intended purpose by the teachers, their user experience could benefit from applying some changes.

The findings below are derived from the data collected in the face-to-face PD sessions, observations at Summer School, and Heuristic Evaluation RH02; all of which produced very similar sets of findings.

	Usability Observation	Frequency	Recommended response
AT1	To include an app into the ILS, the teacher has to find it in GOLABZ, and then right click on “source code” and then use “copy shortcut” and then paste from the clipboard into the ILS Authoring tool.	Frequently	A user interaction more like the “Widget store” would be much preferable. However it may be best to avoid the word Widget (see below).
AT2	The Graasp authoring environment is not always speaking the users/teachers language, e.g. • To include an app into the content, it is necessary to find the “Insert widget” icon. The word “widget” (in UK English at least) is almost unknown except to computer programmers and therefore may be a bit alienating to teachers. • The word “Description” suggests that text there will describe the space, whereas teachers are expected to insert the learning content there. • The teacher notes are pasted into “Discussion” section. It is not immediately clear to the teacher that the Discussion section is invisible to students. • The button for accessing the student view is labelled “Share”, which is somewhat misleading. The button also provides a URL for student access.	Frequently	Try throughout to use wording oriented to teachers and educational work, e.g. • “Insert scaffold” • “Learning Content” • “Teacher notes” Incorporate some visual cues for teachers (e.g., background colour, labels) to indicate which parts will be visible to students and which parts will not. • Re-label button, or create a new button to access student view.
AT3	The process to include YouTube videos requires teachers to copy & paste (only) the video ID, not the whole URL, which they are more used to and which is also easier than finding the appropriate part of the URL or ID at the video.	Sometimes	Use whole URL instead of video ID to include YouTube videos.
AT4	Default values for scaffolds, labs, images, videos and other multimedia elements to be included in the learning material are sometimes too small, leading to usability issues (e.g., scrolling within scrolling) occurring in the ILS. Manually adjusting them is possible, but should not be possible in most of the cases.	Sometimes	Specify appropriate default values. Drag handles could be provided for resizing instead of having to specify the size in pixels.
AT5	The process for getting images in the learning content is convoluted (first uploading it to a Graasp space and then including it at the desired position in the editor by copy & pasting its URL).	Sometimes	Simplify the process by providing an “include image” dialog, where the image can be selected from your hard drive, not only specified with an Online source.
AT6	We believe Graasp doesn’t work (very well) on one of the most popular browsers in use, i.e., Internet Explorer.	Sometimes	All parts of the product should work on all popular browsers used by teachers which were identified in D3.1 page 13 as Chrome, Firefox and Internet Explorer.

AT7	The Graasp authoring environment provides a lot of functionality, which is not necessarily needed for the authoring task, e.g. • Editor has many icons, and no obvious organisation to them. Some icons are probably not widely recognised by teachers. • The screen has many features which in a teaching context might cause a distraction; e.g., star rating, Tags, Favourites, clipboard, notifications, "Path: p", arrow symbol to "show right column" and the right column itself, etc.	Occasionally	Simplify the interface where possible, to make it easier for the teacher to focus on the current task.
AT8	When editing the phase name (e.g., Conceptualisation, Orientation etc.) from within a space, you have to click once in the area, then type and then press a small green tick mark, which is easy to miss. The tab key doesn't accept the text. Also the red cross alternative acts as cancel rather than delete.	Rarely	The simplest solution would be to use single click or tab entry to the field to enter edit mode and auto-save when the user navigates away from the input field. This might also require some kind of Undo function.
AT9	The description field has a completely different interaction. To edit the contents, the user has to hover in the field, and then a previously invisible pencil-shaped control becomes visible in the top left hand corner of the field. This control can't be seen by the user if the screen scrolling means the top left hand corner of the field is not in view. Clicking on the pencil navigates to edit mode.	Rarely	The simplest solution would be to use single click or tab entry to the field to enter edit mode and auto-save when the user navigates away from the input field. This might also require some kind of Undo function.
AT10	When the mouse hovers in the Description field, after a moment a tool tip sometimes appears explaining that double click will enter edit mode. However the tool tip does not appear if the mouse pointer is in certain sorts of content – e.g., a YouTube video box.	Rarely	Avoid invisible controls which appear only when the pointer hovers somewhere. The simplest solution would be to use single click or tab entry to the field to enter edit mode and auto-save when the user navigates away from the input field. This might also require some kind of Undo function.
AT11	For some functions it is necessary to navigate to a multi-coloured hand icon with a "+" symbol, which then changes to reveal several controls including new app, new resource etc.	Rarely	The hand icon is rather abstract and not recognisable to teachers. The interface would be easier to understand if the controls were more visible, or there was some clue (e.g., a pull down list symbol) to suggest there are hidden controls.
AT12	When editing Description is complete, the user has to click on an inconspicuous SAVE button else data is lost. Also there is no warning about unsaved data if the user navigates away without saving.	Rarely	The simplest solution would single click to enter edit mode and auto-save when the user navigates away from the input field. This might also require some kind of Undo function. As a minimum, there should be a warning if a user actions means data may be lost.
AT13	If the user accidentally tries to edit text while in view mode, by selecting text and then	Rarely	Either eliminate view mode, or find some way to prevent this browser

	using the back-arrow key to delete, in the browser we used (Chrome) the back-arrow key was interpreted as a browser back instructions taking us into a very different screen. We believe this sort of unexpected and dramatic change could potentially be disturbing and alienating to some teachers.		interaction with the back-arrow key, which most people expect is used active only when text editing.
AT14	When editing the Description, a large selection of icons appears at the bottom of the box. Some icons are probably not widely recognised by teachers.	Rarely	Consider having editing icons at the top of the screen. Remove any which are not necessary and organise those that remain so as to speed up icon searching. Consider adding text (rather than just tool tips) to aid recognition of rarely used icons.
AT15	Vault is showing in the student view of the ILS, but is not one of the phases and should thus not be displayed as a tab.	Rarely	Don't show Vault as one of the tabs.

Table 3.12a. Summary of the findings and recommendations for ILS Authoring Toolset (AT)

Response category	Usability observations
Software changed – not yet retested by users	AT2, AT4, AT5, AT6, AT8 (possibly partial), AT9, AT10, AT12, AT13, AT14, AT15
Software change – work in progress	AT7
Planned	AT1
Considering	AT11
Other	AT16 – function exists

Table 3.12b. Development Team response summary for ILS Authoring Toolset (AT)

3.3.6 GOLABZ Portal

Studies involving the Go-Lab portal (www.golabz.eu) have been conducted and are mainly reported in deliverables D5.4 and D2.2, but also include CGT Task 2 (Section 4.2). This material is not duplicated here.

3.4 Go-Lab Summer School

As indicated in Table 3.1, seven PD activities were conducted during the Go-Lab Summer School (13-18 July 2014, Marathons, Greece) coordinated by EA. The School involved altogether 39 teachers from 9 European countries. A large amount of data was collected with different approaches and instruments, including paper/online questionnaires, audience response system, printed booklets for paper-based annotation, online annotation with PDot, flipcharts, observations, and interviews.

As examples, ***Appendix A – PDP for Splash ILS*** illustrates how PD20-G was run and what types of data were collected whereas ***Appendix C - Results of Electricity ILS*** shows the analysis and reporting of data collected in PD20-F. Most of the prototype-specific (i.e., scaffold apps, labs, and ILS) qualitative findings of PD20A-H have already been incorporated into Section 3.2.1-3.2.6. In this section, we present a holistic report on this event from the PD perspective and also provide some complementary quantitative findings.

3.4.1 Overview of key observations

- There was a high level of energy and enthusiasm among the teachers, who showed a willingness to learn and exploit the Go-Lab system and held a belief that it is good and they want to use it.
- A minority tended to believe certain features are not yet easy enough for their students to use.
- The system would be much improved with more labs to suit different ages of students and science subjects.
- The usability and learnability of Graasp needs to be further enhanced, e.g., Graasp should provide a more robust functionality to prevent the loss of data (NB: It transpired that a significant redesign of Graasp is foreseeable, including better visual and usability design as well as more stability, reliability and trustworthiness).
- Many teachers wanted extra features to facilitate group work in classes. They also wanted better password protection so students could not copy each other's work.
- The proposed Quiz tool seemed to be a very popular idea, as was the dashboard and particularly the feature to allow students to submit work for assessment.
- Teachers also requested other outputs to facilitate evaluation and assessment of student work.

3.4.2 Survey with Audience Response System

A tool called Audience Response System was deployed to capture the participating teachers' opinions on a set of specific questions. Specifically, individual teachers were given a small handheld device (Figure 3a-e), which is wirelessly connected to a laptop configured to receive responses. The tool can then instantly visualize the aggregated results in the form of bar charts. This ability to provide immediate feedback proved particularly motivating and engaging, as shown by the high response rates for all the questions presented. Here below we present the related findings:

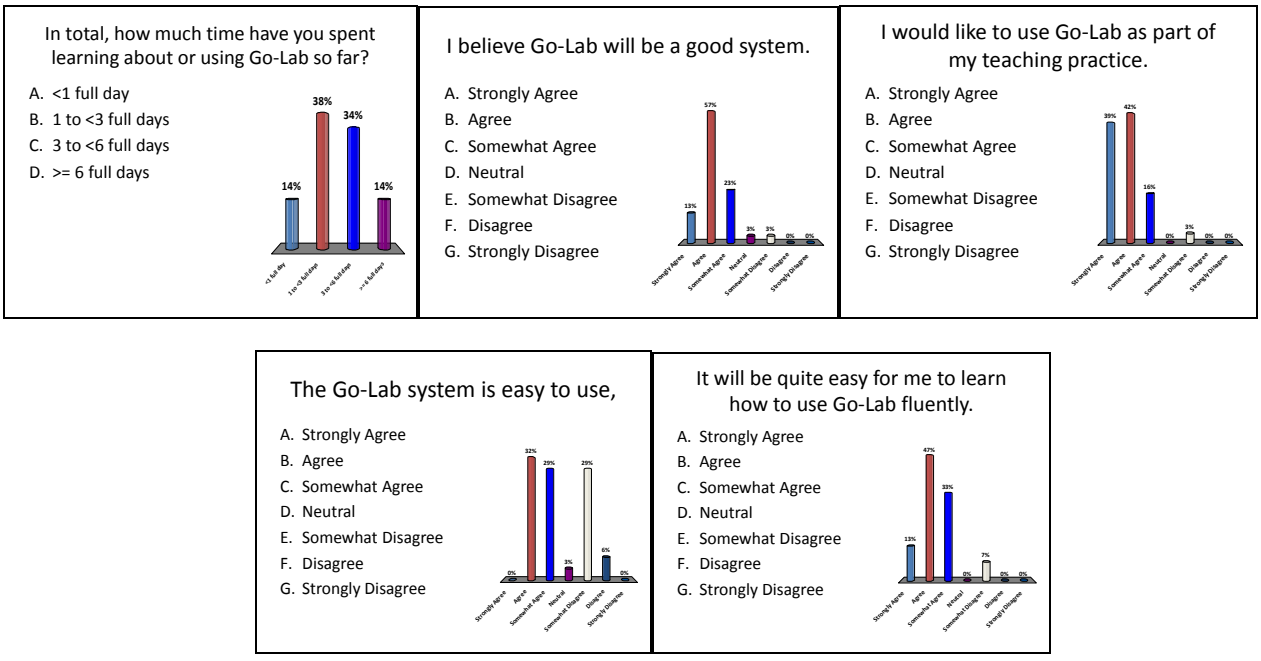


Figure 3.12a-e. Teachers' responses to the questions on the acceptance of the Go-Lab system

On average, the teachers had a significant level of expertise of Go-Lab. The teachers had a very strong belief that Go-Lab would be a good system and that they would like to use it in their teaching practice. Ease of use was less certain, though still more agree than disagree. Nearly all thought

they would be able to learn to use it fluently. These very positive results were from a highly skilled and selected sample of science teachers, but nevertheless give good grounds for optimism, as well as highlighting that usability improvements would be beneficial. This may have been influenced by the usability of the ILS editing process which the participants had been attempting.

The teachers were also polled very simply on a series of prototypes and mock-ups of scaffold tools (cf. the qualitative findings integrated into Section 3.2.2), having had demonstrations and time to experiment with the prototypes and verbal descriptions and screenshots of the mock-ups. Table 3.13 shows the results (all as percentages of the teachers who voted). The column "Score" is an arithmetic mean[3] derived as follows:

*Score = ("Strongly Agreed %" * 3 + "Agreed %" * 2 + "Somewhat Agreed %" * 1 + "Neutral %" * 0 + "Somewhat Disagreed %" * (- 1) + "Disagreed %" * (- 2) + "Strongly Disagreed %" * (- 3)) / 100*

Thus a Score of 0 would indicate participants were on average 'neutral' (N); a score of 2.00 would indicate that on average participants Agree (perhaps with an equal number on 'strongly agreed' (SA) and 'somewhat agreed' (SWA)), and a score of -1.5 would indicate that on average participants came midway between 'somewhat disagreed' (SWD) and 'disagreed' (D).

It is encouraging to note that all scores are positive. Of particular interest is that the two tools, which are especially relevant to practical teaching and classroom management tasks, namely QuizMaster and Teacher Dashboard, received the highest scores. The very highest score was that for a dashboard screen showing when pupils have submitted their work; possibly because the teacher participants thought it might help them in marking.

Scaffold	Statement	SA	A	SWA	N	SWD	D	SD	Score
Questioning tool	I would use this tool	33	33	14	17	3	0	0	1.76
Hypothesis tool	I would use this tool	23	40	9	11	9	9	0	1.31
Hypothesis tool	This tool is easy for my students to use	12	29	18	29	6	3	3	0.91
concept mapper	I would use this tool	25	36	17	11	8	3	0	1.50
concept mapper	This tool is easy for my students to use	6	35	24	18	6	6	6	0.76
Experiment Design tool	I would use this tool	6	34	6	47	6	0	0	0.86
Experiment Design tool	This tool is easy for my students to use	3	6	26	50	9	3	3	0.23
Data Viewer	I would use this tool	23	17	11	46	0	0	3	1.05
Data Viewer	This tool is easy for my students to use	6	6	23	55	10	0	0	0.43
Quiz Master	I would use this tool	41	48	4	7	0	0	0	2.23
Quiz Master	This tool is easy for my students to use	38	52	7	3	0	0	0	2.25
Quiz Master	This tool is easy for me to use	25	25	12	33	0	4	0	1.29
Conclusion tool (new version)	I would use this tool	40	30	20	10	0	0	0	2.00
Conclusion tool (new version)	This tool is easy for my students to use	4	31	31	19	15	0	0	0.90

[3] Jamieson (2004) challenges the use of arithmetic means to summarise Likert scale data, on the ground that the categories are ordinal and not necessarily interval. However, it is a common practice as she recognises, including in HCI studies (e.g. Brooke, 1996), and arguably provides a more useful summary than mode or median.

Teacher dashboard - overview screen	I would use this tool	70	23	3	0	0	3	0	2.53
Teacher dashboard - # connected students	I would use this feature	59	34	3	0	3	0	0	2.45
Teacher dashboard - activity log	I would use this feature	39	39	10	3	0	3	6	1.81
Teacher dashboard - # students by phase	I would use this feature	48	45	6	0	0	0	0	2.40
Teacher dashboard - phase for each student	I would use this feature	47	37	7	3	0	0	7	2.01
Teacher dashboard - average time per phase	I would use this feature	59	34	0	7	0	0	0	2.45
Teacher dashboard - # phase changes per student	I would use this feature	19	22	11	15	4	22	7	0.43
Teacher dashboard - pdf submission by student	I would use this feature	55	45	0	0	0	0	0	2.55

Note: SA = Strongly Agreed; A = Agreed; SWA = Somewhat Agreed; N = Neutral; SWD = Strongly Disagreed; D=Disagreed; SD = Strongly Disagreed.

Table 3.13. Summer School teacher responses to Audience Response system questions

3.5 Live lesson video analysis (PD21)

A teacher at Ellinogermaniki Agogi (EA; an international school in Greece) had conducted a lesson in English with 20 students (aged 14 years), who were asked to work through the Splash (Sinking and Floating) mini-ILS in a classroom setting under the supervision of the teacher. The whole lesson had been video recorded and was played back on a large screen to the Go-Lab partners, who were attending a meeting at EA some days later. The video playback was also attended by the teacher and 8 of the students who had taken part. After the video playback, the 8 students and teacher responded to questions about the lesson posed by the Go-Lab partners present.

As most of the observations derived from the video were related to the practical issues of classroom and learning management using Go-Lab, at least for this particular class, they are more relevant to WP1. Nonetheless, some insights into the design issues can be gained, which are briefly described subsequently. There was an evident high level of motivation and enthusiasm in both students and teacher, and a belief that Go-Lab could become a valuable and engaging learning aid. Some of the main usability matters concerned are:

- the availability and stability of Internet connections in schools and homes;
- the need for systems to work identically on all mainstream browsers;
- the way interactive components might take attention away from non-interactive components such as static text;
- the perceived need for time to learn the system before a productive lesson can be delivered;
- the learnability of the Go-Lab portal for the first-time users;

While the usability problems identified in this specific PD workshop are already covered in its preceding ones, it enabled us to reflect further on potential practical challenges facing teachers in the practical world of the classroom when deploying educational technologies such as Go-Lab. The challenges include classroom management (especially when students tackle the ILS at different speeds), need for student assessment, need for robust technical infrastructure, and need for lessons to fit curriculum.

4. Remote Participatory Design Approach: Core Group of Teachers

Seventeen teachers originating from nine different countries have been recruited as so-called "Core Group of Teachers" through the Go-Lab National Coordinators with the continuous support from WP6 and WP7 partners, especially the WP7 leader. The number of CGT will be increased in the course of the project's lifetime. The WP3 lead partner is responsible to coordinate and sustain the involvement of the CGT members.

The main role of CGT is to provide fast comments and improvement suggestions for Go-Lab work. They are contacted on a biweekly basis via emails and required to carry out specific tasks with Go-Lab design artefacts and to share their feedback within a short period of time. The strategy is to engage CGT incrementally, not so overwhelm them with demanding tasks. Hence, the first four tasks are relatively straightforward, which nevertheless provide us with useful feedback (Table 4.1). Subsequently, we will provide them with more challenging ones with the use of PDot and other online sharing tools to capture richer qualitative data. Here below we describe the design of the six completed CGT activities and results.

Event ID	Date	No. of responses	Description
OPD01	20/06/2014	12	Introduction about CGT; A questionnaire on overall impressions of the Go-Lab system and on collaboration strategies
OPD02	04/07/2014	6	Five tasks with GoLabz; self-timed and error-counted; 5-item close-ended questionnaire
OPD03	01/08/2014	9	Five-step usage of QuizMaster; self-timed and error-counted; 5-item close-ended questionnaire
OPD04	15/08/2014	7	Cross-checking the findings of Summer School
OPD05	05/09/2014	9	Introduction to the use of PDot with which to explore and annotate Concept Mapper
OPD06	21/09/2014	5	Exploring widgets http://www.wolframalpha.com/widgets/gallery/ to comment on their usability and usefulness.

Table 4.1. Overview of six tasks of Core Group of Teachers

The Development Team has been consulted before deciding which studies to conduct with the CGT. The findings from CGT studies are shared by email and in meetings with the development team. They have been responsive in understanding the usability issues and using this as one input to their development priorities. For ease of reading we do not include their detailed responses in this already large document.

4.1 CGT Task 1

An introductory message was sent to individual members of CGT, explaining their role. An email-based questionnaire "CGT Task 1" consisting of seven questions on the overall impressions of Go-Lab and on the collaborative strategy for the group was administered. The CGT members responded per email to the WP3 lead partner.

The CGT Task 1 received twelve responses. The questions and summarized responses are as follows (Table 4.2):

Q1. Have you explored the two main Go-Lab websites, www.go-lab-project.eu and www.golabz.eu . NOT AT ALL / A LITTLE / MODERATELY / QUITE A LOT / EXTENSIVELY A1: Most people had explored “moderately”. 4 said “a little” and 3 “quite a lot”.
Q2. How easy or difficult do you think you would find it to use Go-Lab? VERY EASY / QUITE EASY / SOMEWHAT DIFFICULT / QUITE DIFFICULT / VERY DIFFICULT A2: Everyone said either “quite easy” or “somewhat difficult” with a slight majority on “quite easy”. One person said it would be easier when translated into their native language.
Q3. How difficult or easy do you think you would find it to learn to use Go-Lab? VERY EASY / QUITE EASY / SOMEWHAT DIFFICULT / QUITE DIFFICULT / VERY DIFFICULT A3: Most said “Quite easy” or “Somewhat difficult” with one person responding “Quite difficult”.
Q4. *What impact do you think the use of Go-Lab could have on your students learning science? A4: Everyone who responded had positive beliefs or at least positive hopes.*
Q5. What day of the week should we send out an email to be most useful for you? A5: Friday was a very clear favourite.
Q6. Are you happy for us to use the data you provide, in summarised form, to inform the product development direction? A6: Everyone agreed
Q7. Are you happy for us to share the data you provide with other teachers in the group in a less anonymous way? A7: Everyone agreed

Table 4.2. Questions and responses of CGT Task 1

*The elaborate answers to Q4 above are presented in verbatim below:

- As Go-Lab put the students in the centre of the learning-teaching process, it can be very useful. We need to clarify what part of the activities is the main point for getting the improvement.
- I think and hope that the use of such Go-Lab platforms offer a great motivation for students, because the use of online laboratories should be new for many students and in addition, this approach allows an investigative and inquiry-discovery learning, so that students develop their own questions and explore possible solutions using the laboratories.
- It could be nice if the teacher use it at classroom, but could be great if the use of the tools is exploited by students to achieve the concepts.
- I think the impact could be huge when you can use virtual labs in classroom. Specially those labs that it's difficult to do it in class due to economic or logistics reasons.
- I would like to use it for experiments with expensive equipment. Or to let them experience the equipment before the experiment or to let them think about research questions and how to make a measurement plan
- I think there are great potentialities for students from the use of Go-Lab. During past three years I was involved as pilot teacher within inGenious Project and I had the chance to use some digital practices, such as Sensore Adventure (http://lgfl.skoool.co.uk/sensors.aspx) for example, that students liked very much
- I guess that it probably depends on the student. The impact could be very positive and cannot be negative.
- Positive, I am sure that it will be a friendly environment for them. Much more interesting than blackboard or word edited notes.
- It will be a new experience. And I hope that they are more awake to science
- Mayor impact on every teacher first, then on a student. A remote lab is a real thing that gives our students the right to be a part of the experiment, no matter where they live or how clever they are.

- The virtual lab work allows students to check the relationship between the school theory and practice. It motivates students to learn and let them construct new scientific hypotheses and check them quickly.

4.2 CGT Task 2

The description of CGT Task 2, which was distributed via email, and responses were collected in the same manner, is presented in Table 4.3 below:

Please enter www.golabz.eu into the address of a browser on your computer – preferably Google Chrome or Mozilla Firefox. Please explore the site for a few minutes, checking out various features, and then attempt the following tasks. A. Find the lab called Weblab-DEUSTO Aquarium. B. Find the name of an ILS (Inquiry Learning Space) which uses this lab. C. Find a list of labs which are suitable for teaching astronomy to 12-15 year old students. D. Find the name of the app which can be used for typing mathematical formulae. E. Look for a lab which would be useful to you in your own lessons. Then please respond to the following statements by indicating your level of agreement (STRONGLY AGREE / AGREE / NEITHER AGREE NOR DISAGREE / DISAGREE / STRONGLY DISAGREE), and where possible provide some text to say why. 1. I found the GoLabz website easy to use. Please explain why. 2. I can think of ways that the GoLabz website could be improved to make it easier to use. Please describe one or more changes which you believe would improve the site. 3. I found the selection criteria on the right side of some screens were useful to narrow my searches. Please explain. 4. I think the system would be better if there were additional selection criteria. Please explain. 5. I think some of the selection criteria would be of no use to me. Please say what criteria were not much use to you.

Table 4.3. Description of CGT Task 2

Results of CGT Task 2

Only 6 responses were received; school holidays affected the availability in some countries. Table 4.4a and Table 4.4b display the results.

Task	Description	Success		time (mins)					Error	
		yes	no	<1	1	2	3 to 5	6 to 10	0	1 to 3
A	Find the lab called Weblab-DEUSTO Aquarium	**5**		**5**					**4**	
B	Find the name of an ILS which uses this lab	**3**	**1**	**1**	**1**	**1**	**1**	**1**	**3**	**1**
C	Find a list of labs suitable for teaching astronomy to 12-15 year olds	**5**		**2**	**1**	**2**			**4**	
D	Find the name of an app which can be used for typing mathematical formulae	**5**		**3**	**1**	**1**			**4**	
E	Look for a lab which would be useful to you in your own lessons	**3**	**1**	**1**	**1**	**1**	**1**		**3**	

Table 4.4a. Results of the five subtasks on Golabz of CGT Task 2

Questions		SA	A	N	D	SD
Q1	I found the GOLABZ website easy to use	2	4			
Q2	I can think of ways that the GOLABZ website could be improved to make it easier to use.		6			
Q3	I found the selection criteria on the right side of some screens were useful to narrow my searches.	3	2	1		
Q4	I think the system would be better if there were additional selection criteria.	1		3	2	
Q5	I think some of the selection criteria would be of no use to me		1	2	1	2

Note: SA= Strongly Agree; A=Agree; N=Neutral; D=Disagree; SC=Strongly Disagree

Table 4.4b. Results of the questions posed in CGT Task 2

Here we summarise the most illuminating and frequent comments:

- Several people used mainly the search function. One did task B by paging through, and another by "surfing". One didn't notice filtering until answering Q3;
- There were many comments that the website content and the labs and apps it links to should be available in more languages;
- One teacher asked in answer to Q2: "What about evaluation and assessment";
- Help facility or guidance for less experienced users would be good;
- It would be useful to have more descriptive information on labs and how teachers have used them practically in their lessons;
- Some want to have a better process to log in and save tools
- Graasp can be somewhat hard for "common teachers" to understand;
- Teaching subject (Physics, Biology) is the most important filtering criterion; other useful selection criterion (Q4) could include the type of student activity (e.g., hypothesising, graphing, data collecting)
- One general but interesting question: "… where are the results of the apps recorded? I mean, if I use the hypothesis scratchpad and I save my hypothesis, where am I saving it? Can I recover from other place? Is it like a LMS?"

***Discussion points**:*

After reflecting on the responses given to the CGT Task 2 as well as on observations in the Summer School, the following points are noted for further consideration:

- Tasks A, C & D were quick and easy for most teachers.
- Task B (finding an ILS which uses a specified lab) was not easy for some teachers, and one was unsure whether it was correct when they had found one.
- Task E ("find a lab which would be useful to you in your own lessons") attracted understandably varied responses.
- Q1 & Q2 suggest these teachers find the site easy to use but also can think of improvements. From qualitative data in this study and also observed at Summer School, it is clear that much navigation was done by scrolling through pages or using the search function. If the number of labs becomes large, these methods will be less useful and filtering will be more necessary.
- Q4 attracted mixed responses. Some people think too many search criteria creates unnecessary complexity; others think the more options the better as one can't always predict what will be useful. However Q5 attracted very few Agrees; no-one wanted to drop any of the existing filtering criteria.

- Lab descriptions would benefit from some more practical teacher-oriented and plain English descriptions about how they could be used in classes, perhaps text provided by other teachers who have used the lab in their lessons.
- Having all material available in multiple languages would be a major improvement.

4.3 CGT Task 3

This task focuses on QuizMaster and consists of five steps (Table 4.5). The task description and responses were communicated via email.

Please time yourself on each step of this task and record:

- Task completion: Yes or No
- Time on task in minutes: <1, 1, 2, 3,…
- Number of mistakes: 0, 1-3, 4-6, 7-9, >=10

Step 1. Please enter www.golabz.eu into the address of a browser on your computer – preferably Google Chrome or Mozilla Firefox.
Use the site to find an App called Quiz Master, and find out how to use Quiz Maser by clicking on "App Preview". In the box which pops up, click on "Tutorial" and a new window should appear entitled "How to create a Quiz spreadsheet".

Step 2. Click on "Drive" in the tutorial screen, log on to Google and (if not already installed) install Google drive.

Step 3. Within Google drive, click on CREATE and create a simple spreadsheet as described in the tutorial. The spreadsheet should have 2 questions:

i. "What is the SI unit of resistance" with three possible answers of "Volt", "Ampere" and "Ohm"
ii. "How long does it take for the earth to orbit the sun" with two possible answers of "One day" and "One year".

Step 4. As described in steps 3 & 4 of the tutorial, publish the spreadsheet to the Web and copy the URL and paste it into the field in the Quiz Master tool.

Step 5. Now, taking the role of a student, try out the functionality of the Quiz using the QuizMaster tool.

Table 4.5. Description of CGT Task 3

Results of CGT Task 3

Quantitative and qualitative results for the five steps of the task with the QuizMaster tool and for the five-item questionnaire are presented in Table 4.6a, Table 4.6b and Table 4.6c, respectively.

Step	Description	Success		time (mins)						error		
		yes	no	<1	1	2	3 to 5	6 to 10	more	0	1 to 3	4 to 6
1	Find QuizMaster and its tutorial	8	1	4	3	1		1		5	4	
2	Install Google Drive	9		7	1	1				9		
3	Create Google spreadsheet	9				1	3	5		3	4	1
4	Publish spreadsheet and paste URL	9		2	2	1	1	2	1	5	4	
5	Try quiz as a student	4	5	3		2	2			4		

Table 4.6a. Five-step results of CGT Task 3 with the QuizMaster tool

Question		SA	A	N	D	SD
Q1	I believe the Quiz Master tool could fulfil a useful function	2	3	2		
Q2	I believe the Quiz Master tool is easy for me to use	2	4	1		
Q3	I believe the Quiz Master tool is easy for other teachers to use	1	1	2	2	1
Q4	I believe the Quiz Master tool is easy for my students to use	4	3			
Q5	I believe the Quiz Master tool should be improved	5	2			

Table 4.6b. Questionnaire results of CGT Task 3 with the QuizMaster tool

Step1	Needs more explanation and examples of spreadsheets for those who don't use Google Drive.
	Couldn't find Quiz Master using the search tool or the filters on the right hand side, so went into apps tab and browsed to find it.
	Would be good to be able to list the apps according to the number of downloads
Step2	Would be useful to have a tutorial for Google Drive for new users
	I entered Google drive from the other way (the link didn't work for me). I already had it installed.
Step3	Colleagues might prefer a faster tool (like google module or quiz Faber) even if more difficult to put in explanation for right or wrong answers
	Have never used a Google spreadsheet before
	There is a powerful product called HotPotatoes. You should use this or at least use it as a model.
Step4	Many people could not answer the second question as there was no visible scrolling function
	difficulty getting the correct URL
	even if publishing is not difficult, there are some missing information in the tutorial (no mention to renaming the spreadsheet) and some errors displayed while following the tutorial
Step5	It's not clear how a student will access the quiz. Do I have to share a URL with them?
Q1	If teacher is online (and homework) it means we really exist.
	Can be used to integrate assessment into the process of working with labs without need to go out to other tools.
	There is a script called Flubaroo on Google spreadsheets to analyze student responses to quizzes and give grades. The only problem on that is that it is hard work to make quiz in Google Forms for teachers. So it seems more easy with this tool.
	Can I see results of students? There should be a way to limit answers to one click.
Q2	Quizmaster is difficult to use on a tablet or iPad
Q3	Most teachers are not familiar with Google Drive
	(Strongly disagree): It depends on age and experience. A Youtube tutorial with slow instructions would have been more useful, along with a detailed text. We learn by seeing, listening, reading
	(Agree) if you are confident with spreadsheets. If not, you need some time to learn how it works.
	(Strongly Agree) but I would like to know how to grade a student's responses and I would like a summary of the responses.
	Some will not like to use Google drive
Q4	(Strongly agree) but we need an option to maintain student name for grading purposes
Q5	They are much better than us at using the web
	Test it with some really inept teachers and see where the instructions have loopholes
	I find it easy to use though it requires some time to prep[are all the columns. Maybe it would be easier and faster finding them already prepared, with a check/uncheck tool that allows teachers to select which to use.
	Will be more useful if student can log in and teacher can control who does the tests in order to evaluate them. Maybe the way to improve it can be to make it compatible with MOODLE (SCORM) and give the possibility to insert the app in a MOODLE course. (The app and maybe the other labs).
	feature to get summary of responses and grade assessment
	The quiz should be more like HotPotatoers
	Where are the answers saved? How many attempts do students have?

Table 4.6c. Qualitative comments on the CGT Task 3 with the QuizMaster tool

4.4 CGT Task 4

The findings of the PD activities in the Summer School (Section 3.3) were essentially based on a group of science teachers who were already quite well-informed about Go-Lab. We aimed to know whether the CGT teachers might have different opinions through the responses to the CGT Task 4 (Table 4.7). Quantitative and qualitative findings are presented in Table 4.8a and Table 4.8b.

Below we summarise our findings about the usability of the Go-Lab system in terms of eight statements and two questions. The findings are mainly based on input from some teachers with a high level of involvement and expertise in the system. We would like you to read through each of the eight statements and indicate to what extent you agree with it, not only for yourself but also for other science teachers you know or work with, and of science students (aged 9 to 18). Please also answer the two open-ended questions.

1. The GoLabz website provides an easy to use portal for finding GoLabz resources.
2. Go-Lab ILSs are easy enough to learn and use.
3. The ILS authoring process can encourage teachers to write their own ILSs.
4. The system should have extra features for evaluating and assessing student work.
5. The system should provide features to support students working in groups as well as individually.
6. Scaffolds apps provided by Go-Lab, in their current form, are suitable for use in a classroom setting.
7. The system should have online help on how to use different components.
8. Students using a Go-Lab ILS will be able to focus on the science learning.
9. Which of Scaffold Apps in Go-Lab you would definitely NOT use for your science teaching at school? Why?
10. Which other web-based tools which you have used (or known) for science learning in classroom and would like to have it/them included in Go-Lab?

Table 4.7. Description of CGT Task 4

Questions			SA	A	N	D	SD
Q1	The GoLabz website provides an easy to use portal for finding GoLabz resources.	For me	3	4			
		for science teachers	2	3	2		
		for students	2	4			1
Q2	Go-Lab ILSs are easy enough to learn and use.	For me	4	2	1		
		for science teachers	1	3	2	1	
		for students	2	3	1	1	
Q3	The ILS authoring process can encourage teachers to write their own ILSs.	For me	2	3	1	1	
		for science teachers		2	3	1	1
Q4	The system should have extra facilities for evaluating and assessing student work	For me	4	2	1		
		for science teachers	4	2	1		
		for students					
Q5	The system should provide features to support students working in groups as well as individually.	For me	3	3	1		
		for science teachers	3	3	1		
		for students	3	3	1		
Q6	scaffold apps provided by Go-Lab, in their current form, are suitable for use in a classroom setting.	For me		5	2		
		for science teachers		3	3	1	
		for students	1	3	3		
Q7	The system should have online help on how to use different components	For me	3	2	1	1	
		for science teachers	4	2	1		
		for students	4	2			1
Q8	Students using Go-Lab will be able to focus on science learning						
		for students	2	3	2		

Table 4.8a. Quantitative results of CGT Task 4

Q1	• Students have no problems with new technologies, but some teachers still have; • Is very affordable find what we seek; • The GoLabz portal is easy to survey; • It could be divided into subjects and a general part useful for all subjects. For example, I teach only maths and physics, so I'd look only into the general part and my two happy islands;
Q2	• Some ILSs must be improved. Have to be more intuitive. The visual part has to be more attractive. • It depends on the specific ILS, but on the whole they are easy to use.
Q3	• The platform we use to build ILSs still has many flaws. Needs to be improved. There are many items that are not immediate. Needs to be simplified. • According to my experience teachers often have only a very small amount of time to develop own extensive proposals/materials. Those teachers who are developing new materials for school are usually publishing these materials in paper form. Therefore it is necessary to get teachers acquainted with the ILS editor and authoring process.
Q4	• It is necessary for students that they have sufficient learning time, especially during an IBSE process and the development of understanding through inquiry. This part is completely distinguished from the part of assessment. Formative assessment is essential to inquiry-based learning, but attention must be given to summative assessment if learning is to be improved through formative assessment and IBSE. Both the skill and understanding dimensions of IBSE need to be assessed in combination. To enable the facilities of evaluation and assessment the system should have this feature. • There could be examples into which we would only change the words and the formulae so as to minimise the first impact with the new tool.
Q5	Good idea. Gives the possibility to learn together.
Q6	• Some of the apps are not appealing, like the periodic table. The concept mapper are interesting but has some flaws. For example if we want an arrow and Then you want to delet that same arrow, the process for doing it has no logic, is not common. • On the whole, all scaffolds offer a specific useful tool. Unfortunately not all of them worked properly in their current form. • I have not enough experience to affirm it categorically (but) I would say it would be a good improve to include video tutorials about using each tool.
Q7	• Online help is always needed. People should have the right to comment at every time when a problem occurs. • According to my experience, teachers use online help when it is necessary, while students often don’t want to confess that they need help. But the system should render the possibility for (online) help. Lots of examples (eve with films YouTube) type are very helpful.
Q8	This depends on the specific ILS if the following terms are performed: IBSE is variously defined in terms of: Practices of scientific investigation; Answering questions; Developing understanding through investigation; Behaving as a scientist
Q9	RSS gadget. There are so many websites we can use to get information. This app is not necessary. No one is useless. a. Not all of them are in their final form, so it's early to say. b. I have not explored all of them, so it's not fair to pin point one. c. Before involving students, I would certainly try and test an app thoroughly. I don't already know, but I would not remove any app. Maybe, the calculator tool is not so necessary. === Concept Mapper (students not free enough to make hypothesis; it probably takes more time to make the map than to think of the actual concepts) Mindmeister (too complicated for school purposes) Hypothesis tool (students not free enough to make hypothesis) Checkspelling (students should be responsible for their spelling) Both Scratchpads (could be faster done on a piece of paper) Experiment Design Tool (students not free enough to think out of the box) Flickr (belongs to yahoo) ===

Q10	Stellarium. There are some interested labs in pt.lab2go.net I always use Wikipedia but this is too obvious for everyone. I don't know exactly what web-based means in that case, but anyway I miss a tool for building tables. PhET and Easy Java Simulations (PhET and Compadre websites). Google Drive, Docs, Spreadsheet, Presentations, Forms, Fusion Tables, and Calendar. Diigo or Delicious link's aggregators (to organize and share links in a teaching topic). Issuu or Scribd to publish study notes for my students. Skype or Google Hangout to help them at home. Gmail or Google Groups to discuss a topic or homework. It can be used to help them at home Twitter to instant communication, group's work and homework. Blogs and Facebook to disseminate. Flickr or Picasa to photo sharing. YouTube for videos. Powtoon or toondoo for animation and comics. Soundcloud for audio notes or other student's work. wikidot to create and maintain wikis. Google Sites to collect homework. Kahoot to create quizzes. Online whiteboards to annotation, tasks.

Table 4.8b. Qualitative results of CGT Task 4

In general these findings confirm the findings derived from observations, discussions and formal studies carried out during Summer School (Section 3.3). The Portal is good, ILSs once written are easy enough to use, and students would be able to focus on their science. However, the scaffold apps would benefit from some improvements, and useful online help could be added throughout the system, perhaps using YouTube videos or examples. Some CGT participants felt ILS authoring should be made easier, mainly for the benefit of other science teachers, as most respondents were more confident with it themselves. They also felt that features to support group working and also to help with feedback and marking would be valuable possible add-ons.

4.5 CGT Task 5

This task focused on the Concept Mapper scaffold app, and – for the first time with the CGT – asked them to use the PDot (Participatory Design online tool) to record their observations. The PDot environment for the study was prepared, together with logons and passwords for all the CGT members. A document was produced to explain how to use PDot and provided as an attachment to the email which described the task and provided a link to the PDot environment, which itself linked to the Concept Mapper. The findings are summarised in tables 4.9a and 4.9b.

The key findings were:

- Of the 48 comments, 13 provided no usability insights (for Concept Mapper), and several more were non-specific accolades such as "It is easy".
- The usability-related comments were very diverse and most were mentioned by only one person
- All of the usability observations made by more than one person, and most of those made by just one person have been identified in earlier studies.
- There were some good suggestions for improvements (e.g. provide feature to select a number of the boxes and arrows and then click on trash icon to delete the selected ones). These were probably based on experience of other diagramming tools.

Methodologically, it was observed that a good number of participants, most of whom had never used Pdot, completed the task successfully and that there was some PDot-related confusion for 3 of the 9 participants.

CGT05 - Concept Mapper

Pdot analysis

9 participants, 48 comments (min 1, max 10).
12 likes, 19 neutrals, 17 dislikes.

PD artefacts

Startup difficulties - perhaps related to Pdot and the https / shield icon issue. (Two resolved the issue; one provided no further feedback.)	3
Difficult to interpret participant's text	4
No usability content in text	5
Pdot-induced confusion	1

Usability observations

Switching between arrow mode and drawing mode non-intuitive	2
Deleting arrow difficult or not achieved	4
Toolbar needs a redesign. Provide different sorts and shapes of arrows and arrowheads.	1
Need UNDO / REDO buttons	1
Need to be able to change fonts.	1
Need an option to use straight or curved arrows.	1
vocabulary on drop-down menu limited and changed since last week	1
online help needed	2
limited vocabulary on arrows	1
concept boxes translucent: reduces readability	1
ENTER key should be interpreted as user expects	1
Could not delete single concept	2
would like to be able to hand-draw relations (arrows) with mouse rather than system determine shape	1
too "heavy" / "abstract" for elementary students	2
add a facility to select the whole map and drag it to make room for another concept	1
should be possible to select some boxes then hit delete tool to delete all selected	1
the "change box colour" drop down menu is invisible when the box is near the right side of the drawing area	1
clicked in the wrong place and the whole box deleted	1
Save and reopen changes some aspects of the appearance.	1

Table 4.9a. Categorised user-submitted comments for CGT Task 5

Selected qualitative comments
"It's very easy to use. I like the pre-loaded concepts. That will be a great guide to our students."
"It is quite easy to built"
"too heavy for the elementary school"
"Last week I could not make any changes. Even now it seems to me that it does not work reliably. But maybe it's me, who is doing here a basic error? "
"I don't know where the concept map is saved. Is it on my computer or in the cloud or on the environment? My students won't work on the same computer every lesson so on the computer is not a good idea for my school."

Table 4.9b. Selected qualitative comments for CGT Task 5

4.6 CGT Task 6

This task requested teachers to explore the website http://www.wolframalpha.com/widgets/gallery/ which contains a large number of "widgets" of potential benefit in science teaching, some of which might usefully be incorporated into ILSs to enrich the lessons. Teachers were asked to freely explore the collection of widgets based on their teaching background and requirements, and to select a few which they found interesting, and provide feedback on:

a) Usefulness: Which of the widgets could add beneficial information or functionality to your ILSs and for which would you have no use respectively?

b) Usability: Which of the widgets can be easily used by you and your students and which are not working well? What did you especially like or which problem(s) occurred?

A spreadsheet was provided to structure the feedback. The findings from this study are presented in Table 4.10.

CGT06 - Wolfram Widgets

5 teachers responded, listing a total of 87 widgets (13, 4, 13, 6, 51). There were no obvious patterns in the choices of widgets they chose to evaluate. The main factors appeared to be the widget title and the teacher's curriculum speciality.

Useful	usable	T1	T2	T3	T4	T5	Totals
Y	**y**	10	3	0	5	29	47
Y	**n**	0	1	0	1	0	2
N	**y**	3	0	0	0	7	10
N	**n**	0	0	0	0	12	12
Y	**-**	0	0	6	0	3	9
not specified		0	0	7	0	0	7

Text responses to "usefulness" question: 74 (84%)

Text responses to "usability" question: 33 (37%)

This suggests more interest in usefulness than usability, though there were a higher proportion of blank responses when the usability was good than poor, so it may be that there was less to say. There were also some usability-related comments in the usefulness responses and vice versa, so this observation is tentative.

Table 4.10. CGT task 6 – Quantitative data

58 (73%) of the 80 widgets for which we have data were deemed useful. Since it can be assumed that teachers would only have selected widgets whose title suggested they would be useful, the 27% deemed "not useful" seems significant.

The most common themes from qualitative comments for non-useful widgets were:

- not relevant to my lessons
- not interesting / enjoyable enough
- not working - possibly "spam"
- trivial (e.g. application of a standard formula)

The most common themes from the comments for useful widgets were:

- gives useful / interesting information easily
- fits with my curriculum / lessons
- students can be drawn in to interacting with the widget

57 (80%) of the 71 widgets for which we have usability data were deemed usable. However 12 of the 14 deemed not usable were also deemed not useful, and from the qualitative data it appears that no specific usability problems were found. The two deemed useful but not usable both had usability problems concerned with language issues. The qualitative data on usability was, as noted above, rather limited. The only recurrent usability theme from this was to improve the graphics / visual appeal.

In summary, the Wolfram widget store seems to have some very appealing widgets in it, as well as some which appear not to work. Some teachers were enthusiastic about the better widgets, including comments about intending to use them, and also thoughts about how these might fit into an ILS. It is very clear from the way the comments about usefulness and usability were presented that - for these teachers at least - usability and perceived usefulness overlap and interact considerably.

5. Researcher-based Studies with Heuristic Evaluation

5.1 Overview

With the goal of identifying any major usability problems that may severely undermine user experience of using an interactive prototype, Heuristic Evaluation [HE] (Nielsen, 1994) is normally performed by usability specialists before rolling out the prototype to be tested with real end-users. We have applied the same principle for the WP3 PD work. Specifically, a team of four HCI researchers have conducted HE on several scaffolds, two Labs and associated ILSs – Electricity Lab: Ohm's Law and Splash: Sinking and Floating – and the ILS authoring process. HE focuses more on the design of user interfaces of the respective artefacts. Results of HE have been communicated to WP1, which meanwhile have conducted systematic reviews on ILSs, focusing more on the content while addressing some design issues at the same time.

ID	Date	Description
RH01	25/03/2014	Exploratory activities with scaffold apps in preparation for the PD Study Catalogue
RH02	08/05/2014	Exploratory activities with the ILS authoring toolset in preparation for PD workshops
RH03	06/06/2014	HE with Splash Lab, Sinking and Floating ILS, and accompanied scaffolds in preparation for PD workshops
RH04	20/06/2014	HE with Electricity Lab, Ohm's Law ILS, and accompanied scaffolds in preparation for PD workshops

Table 5.1. Researcher-based analytical evaluation studies in Go-Lab Year 2

In studies RH03 and RH04, two Go-Lab labs and associated ILSs with a different set of scaffolds were evaluated. The links to the versions, as specified by the date of evaluation, are as follows:

ILS Splash as of 6th June 2014:
http://graasp.epfl.ch/metawidget/1/783b7b75afb9aa5843ab2c54366220c5f728d3ea

ILS Electricity as of 20th June 2014:
http://graasp.epfl.ch/metawidget/1/27e621bed3c8e57d4aca8a97f8d494a6ec2ab556

The evaluators conducted a step by step walk through of the process, trying to imagine themselves in the role of teachers and of students and assuming very limited system knowledge. They also considered the role of the researcher(s) conducting a usability PD session using this ILS. Detailed notes were taken to identify:

- Preliminary outline ideas for what tasks and activities to give teachers for a usability study session;
- Any usability obstacles which a teacher may encounter; and
- Brainstorming of possible questions to ask teachers.

The main outcomes of HE are a list of usability problems identified and sorted in order of importance. When appropriate, improvement suggestions for individual usability problems are proposed. These analytic results have been documented and communicated to the respective partners who are the main developers of the artefacts. Consequently, some of the usability problems have been fixed prior to the PD workshops. In the following sections, we present the overall outcomes of two HEs (RH03, RH04).

As with the CGT studies, the findings from HE studies were shared with the development team by email and discussed in online meetings. The development team have been responsive in understanding the usability issues and using this as one input to their development priorities, though understandably they tend to give greater weight to studies involving real end users. The recent redesign of the Graasp ILS authoring environment is in very close agreement with the findings of RH02. In any case, many of the findings from HE studies were fixed before the Summer School end-user studies or reported by teachers in those studies.

For these reasons and for ease of reading we do not include detailed Development Team Responses to these studies in this already large document.

5.2 Overall outcomes of Heuristic Evaluations

In the following tables we report the outcomes of HEs on Splash Lab (Table 5.2), Sinking and Floating ILSs (Table 5.3), Electricity Lab (Table 5.4), Ohm's Law ILS (Table 5.5), and two scaffolds – Experiment Design Tool (EDT) (Table 5.6) and Concept Mapper (CM) (Table 5.7).

Obviously, there are overlaps between the empirical results of user-based evaluations (Section 3 and Section 4) and the analytic results of Heuristic Evaluations presented in this section. Instead of merging the results, we deliberately keep them separate in order to provide the pedagogical and technical teams with the distinct views of information sources, which can have different implications for decision making on the respective future work.

Table 5.2: **Results of Heuristic Evaluation on Splash Lab (SL)**

	Usability Observation	Importance	Recommended Response
SL14	The sliders are too sensitive to tiny mouse movements	Medium	Allow the user to type a number in the text box as an alternative.
SL15	The granularity of the sliders is too big (step size is 4.something). Thus you are physically not able to specify a mass of 200.00g.	Medium	Allow the user to type a number in the text box as an alternative.
SL16	The input boxes behind the sliders give the false impression of interactivity where they are only able to display the value. Changing it there does not affect the slider.	Medium	Allow the user to type a number in the text box as an alternative.
SL17	In the Archimedes tab the bottom third of the slider area can be reached temporarily but has no effect.	Medium	Make it consistent with the other sliders?
SL18	The bin button in the upper part of the results area is not working. Nothing happens if you press it.	Medium	Provide information to make the purpose and usage of these parts of the display clearer to the user.
SL19	In the Archimedes tab the sliders for density and fluid have the same units of measurement yet but the sliders are on a different scale.	Low	Use the same scale?
SL20	There are 5 tabs presented in the Splash lab, when only some are relevant for the current ILS. We think that the irrelevant tabs might distract and/or confuse the user.	Low	The lab should be configurable so that you could tell which of the tabs you want to be visible in your ILS.
SL21	The flask symbol in the results table is not working for the rows 7 and above.	Low	Make it work consistently or clarify to the user what it is supposed to do.

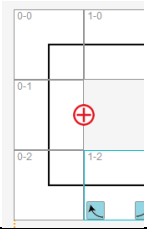
SL22	The meaning of the green star and the red down arrow in the results area is not clear.	Low	Provide clarification. We thought it might mean “floating” (green star) and “sinking” (red arrow). But that does not seem to be the case.
SL23	The 5 tabs seem to not “talk” to each other, so e.g., you cannot “design” an object in the density tab and then test its floatability in the floating and sinking tab. Instead you need to specify it again.	Low	If they are designed to be completely separate labs, make this clearer in the user interface.

Table 5.3: Results of Heuristic Evaluation on Splash: Sinking & Floating (SSF) ILS

	Usability Observation	Importance	Recommended Response
SSF1	Concern that Graasp issues warning message when used with Internet Explorer (9.0.8112.16421). Some PC users don’t have admin rights and cannot install other browsers.	High	Make it work on all standard browsers.
SSF2	There are no instructions or help given on how to interact with the concept mapper tool.	High	Provide online help via a HELP button.
SSF3	The first occurrence of the Splash lab only shows two error messages “{"exception":{"message":"You do not have sufficient rights to perform this action","name":"RuntimeError"}}”.	High	Make sure the lab is displayed correctly.
SSF4	Typos just above the videos: “W**h**atch the videos below to check if **you’re** right.” This might affect the teachers (and students) overall impression of the quality of the ILS and though have an impact on the perceived usability.	Medium	Check the whole ILS for typos and grammatical errors.
SSF5	Text refers to a notes tool/note tool. It is not intuitive that it is hidden in the tools bar.	Medium	Describe how to access the tool in more detail and/or make it more obvious.
SSF6	The notes tool is way too small to be used in the way described in the ILS.	Medium	Make it more obvious that you can widen and narrow the tool.
SSF7	To write in the notes tool you need to click in the text (and not the white space after it). That is not intuitive.	Medium	Make the white space clickable
SSF8	In the Discussion phase the report tool is currently missing.	Medium	Provide it.
SSF9	In the discussion phase the students are asked to save their “work, including all the steps” they took. We are not sure how they are supposed to do it and it is not explained in detail.	Medium	Clarify.
SSF10	Italic font makes it hard to read the introduction text of the orientation phase.	Low	Make the introduction text in orientation phase non-italic.
SSF11	American vs. British English: **Chips bags** / crisps bags? Color / **Colour**	Low	There should be consistent use of either US or UK English.
SSF12	No maximize icon on the youtube videos. The workaround with going to the youtube page and maximize it there might be distracting for the students.	Low	Offer the option to maximize the video right from inside the ILS, as it is very small to watch embedded.
SSF13	The table at the bottom of the page is cut on the right hand side in Chrome (e.g., “between these” and “come up” in the concept map definition are cut).	Low	Make sure all the information is displayed correctly.
SSF14	Lab doesn’t use the screen space fully and needs some scrolling.	Low	Provide a maximise option.

SSF15	To follow the instructions in the conceptualization phase (below the lab) the user needs to navigate and scroll a lot between phases (to look at the concept map) and between ILS and tools toolbar to look at the notes. This is a distraction from the actual learning.	Low	If the concept mapper is needed in different phases, it should be shown again or put in the tools section to be accessible from everywhere.
SSF16	There should be a connection between the experiment design tool and the splash lab to transfer data from one to the other. Currently you have to specify values twice using two different user interfaces.	Low	Automate if possible.
SSF17	In the experiment design tool you can lock values, which you cannot lock in the Splash lab. E.g., the latter automatically locks density.	Low	Pass information from EDT to Lab about which variables are locked, or pass information from lab to EDT to say which ones can or must be locked, or if neither possible, embed suitable advice in the ILS text.
SSF18	On the bottom of the page some text (about Hypothesis) is in light grey colour. This is hard to read. The reason is also not clear for the user.	Low	Change text format.
SSF19	In the Conclusion phase the lab is called “aquarium lab” when it has been called “Splash lab” before.	Low	Use consistent name for the lab.
SSF20	In several phases bold and underline style of a word seems to mean that it is defined or explained on the bottom of the page. Although these words look like a link (underlined) they are not interactive.	Low	It would be good, if the words were linked to the detailed information and/or a more recognizable footnote style (e.g., numbers) would be used.
SSF21	The title of the ILS (Sinking and floating) looks and changes the mouse cursor like a hyperlink, yet nothing happens when you click on it.	Low	Make the text look less like a hyperlink.

Table 5.4: Results of Heuristic Evaluation on Electricity Lab

	Usability Observation	Importance	Recommended Response
LE12	Scrolling is necessary to see all the content of the lab. Especially with the current size it is not very noticeable that there is more than the components, circuit board and meters.	Medium	The lab should be included in a way which makes all of it visible without scrolling “inside” the ILS. Avoid scrolling within scrolling.
LE13	The meters sometimes attach automatically to the border of the tile, not the wire: 	Medium	Make them attach only to wires.
LE14	The slider on the power meter is not working as you would expect. Instead of having fixed values on different positions the slider magnitude from the center influences the speed in which the voltage increases.	Medium	Use sliders consistently and only where the use of this interaction metaphor makes sense to the user (e.g., from real world or digital experience).

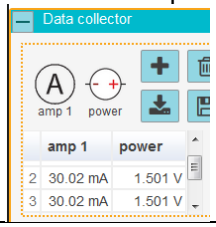
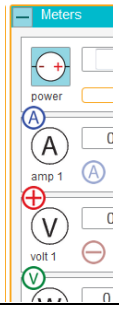
LE15	The table showing the collected data values shows only up to two rows. This makes it hard to work with and read/interpret the data. 	Medium	Allow more rows.
LE16	Sometimes the meter elements to be dragged on to the circuit board are not underneath the meter reading but in the top left hand corner. 	Low	Keep them in a consistent location relative to the meter.
LE17	Naming is not consistent: “Electric Circuit Virtual Lab” in the ILS text and “Circuit simulator with data viewer” in the lab.	Low	Use consistent naming for lab (and also for scaffolds).

Table 5.5: Results of Heuristic Evaluation on Electricity: Ohm’s Law ILS			
	Usability Observation	**Importance**	**Recommended Modification**
EOL1	In the experimentation phase the ILS asks the participant to “Keep record of your data in a table.” It is not clear which table is meant here (“offline” on a piece of paper, a table in the experiment design tool, a table in the ILS [where is this?]).	High	Make it clear what is meant and/or implement an interactive table to keep the records to the ILS.
EOL2	On the bottom of the hypothesis phase there is a second hypothesis scratchpad tool included in the ILS (Underneath “You can proceed to the next phase!”). Are the participants supposed to use it twice or is the second one redundant?	Medium	Remove the second scaffold if it is not needed to prevent confusion.
EOL3	In the Electricity Lab in the experimentation phase, data collector, data graph and circuits are only a distraction, because they are not needed to fulfil the task described in the ILS.	Medium	Remove unneeded material from the display (or – less usefully – add text to the ILS to say what to ignore)
EOL4	Loading of the orientation phase takes relatively long and the browser gets unresponsive.	Medium	Show loading animation until the whole page and content have finished loading.
EOL5	In the hypothesis phase the student is asked to use/relate the concepts noted down in the notepad during the previous phase. As the notepad is not shown, this will result in a lot of navigating back and forth to check the notes and apply them in the concept map.	Medium	Notepad tool should be shown above the concept mapper tool for easy and fast comparison/transfer.
EOL6	The data interpretation tool in the Data_Interpretation phase is missing.	Medium	Provide it, or modify ILS text so as not to create expectations.
EOL7	In the data interpretation phase the hypotheses are not transferred from the hypothesis	Medium	By implementing the vault or by other means, populate the

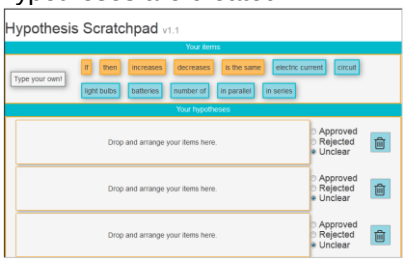
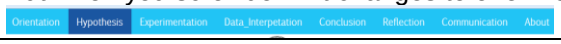
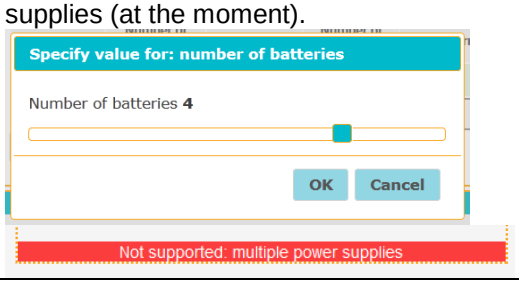
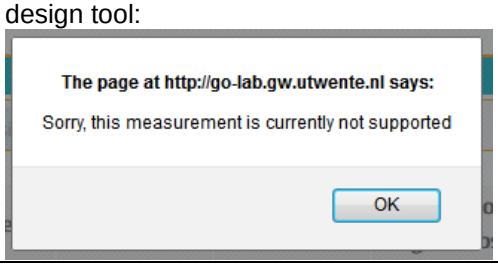
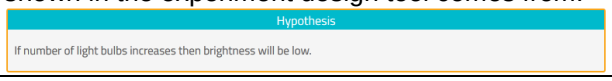
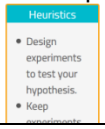
	scratchpad in the hypothesis phase, thus the participant has to re-create them. This is tedious and time consuming and could introduce errors.		hypothesis tool automatically from earlier phases.
EOL8	The hypothesis scratchpad in the data interpretation phase has no scrolling capabilities which would be needed if more than two hypotheses are created. 	Medium	Implement vertical scrolling and/or a larger window.
EOL9	In the conclusion phase the conclusion tool is again missing.	Medium	Provide a conclusion tool or edit ILS text so as not to raise expectations.
EOL10	In the reflection phase the user is asked to “write your thoughts”, but the rectangle provided does not allow text input/an appropriate tool is missing.	Medium	Make the box work for text input.
EOL11	In the communication phase the communication tool is missing.	Medium	To facilitate the session appropriately, at least the facilitators would need to know how this tool is supposed to work/look like.
EOL12	There are four tools at the bottom. At no point the ILS actually asks the participant to use one of them.	Medium	Remove any tools which are not of any use in this ILS.
EOL13	There is no sideways scrolling if you make the window too narrow to display all the content.	Low	Implement sideways scrolling
EOL14	The YouTube videos have text adverts covering nearly a third of the video. This may distract the participants.	Low	Use material without adverts where possible.
EOL15	When on top of the page, the tab navigation wraps:  But when you scroll down it changes to one line: 	Low	Have the tab navigation displayed in one line all the time (to resemble the tab metaphor).
EOL16	In the data interpretation phase the experimental designs are not visible. This causes a lot of “back-and-forth” navigation.	Low	Make this data visible by scrolling rather than clicking on tabs.
EOL17	The information presented in the about tab is irrelevant for the participant (at least the students or the teachers at this point, if they are working through the ILS).	Low	This information should be presented in the golabz repository (and/or the 2hidden” teacher notes or this ILS) and not as part of the actual/visible for students ILS.

Table 5.6: Results of Heuristic Evaluation on Experiment Design Tool (EDT)			
	Usability Observation	**Importance**	**Recommended Response**
EDT 21	The participant can design an experiment with multiple batteries, but the lab itself does not support multiple power supplies (at the moment). 	High	Alter Lab or Experiment Design tool (preferably lab) to be consistent.
EDT 22	Neither of the two “Measures” variables (Electric current or brightness) are currently supported by the experiment design tool: 	High	Implement this functionality.
EDT 23	For the participant it is not clear where the hypothesis shown in the experiment design tool comes from: 	High	Populate this automatically from the hypothesis scratchpad in earlier phase.
EDT 24	If you click on the Analysis tab right after designing your experiments, your experiment rows disappear.	High	Only allow students to do things in sequence (if this is required by the experiment tool to work).
EDT 25	We are not sure, what the run tab is for and neither the ILS nor the EDT tool is giving any information about that. E.g., are you supposed to give prediction in the run tab or keep track of findings after running the experiment in the lab?	High	Provide more online help in lab and/or guidance in ILS text.
EDT 26	It seems to be possible to specify inconsistent combinations of volume, mass and density.	Medium	Either show an error message and prevent going on with the experiment until it is corrected, or automatically adjust the variable to create a consistent set.
EDT 27	The instructions and heuristics are in very narrow boxes, which makes it hard to read where there would be enough white space to make them wider. 	Low	Provide wider box.
EDT 28	There is no default value for the density. Thus when selecting “Specify” it shows “density object **undefined** g/cm^3”	Low	Set a default value for the density as for mass and volume.

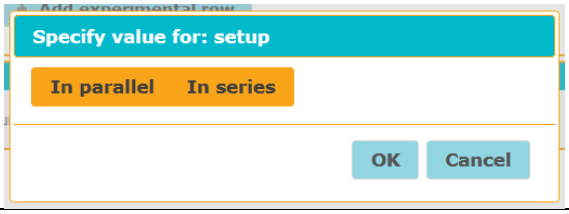
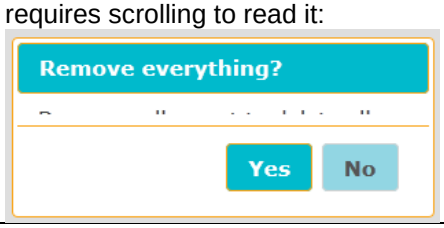
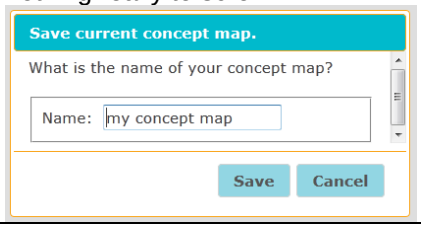
EDT 29	In the “specify value for setup pop up” both options can be visually selected so that they get orange. Still only the last one clicked on is transferred. 	Low	Make the colours behave as expected; i.e., when one turns orange make the previously orange one turn back to blue.
EDT 30	Delete button looks the same as Specify button, but one shows a pop-up to adjust the value, the other one directly deletes the experiment row. This is an inconsistent behaviour of interaction elements looking the same. (The same goes for the “Add experiment row” button.)	Low	Remove the upwards arrow from the buttons which don’t show a menu when clicking on them. From other user interfaces the user would expect something “popping out” when clicking on the arrow.
EDT 31	In the analyse tab the arrows in the table give the false impression that you could sort either ascending or descending, but it only works in ascending sequence.	Low	Implement ascending and descending sequences, or use a different shaped arrow.

Table 5.7: Results of Heuristic Evaluation on Concept Mapper (CM)

	Usability Observation	Importance	Recommended Response
CM25	User tried to add an arrow to concept map by dragging and dropping the arrow button. A blue “concept” box with the URL of arrow.png appeared instead.	High	Consider a more intuitive and consistent design to UI, or if not provide online help, and make this type of drag action do nothing.
CM26	If you try to delete the accidentally created arrow.png-URL box by dragging it onto the bin button, this is not working as the concept “title” is too long.	High	Make deletion method work the same for large components as small ones.
CM27	If you have relationships between concepts and create a new one by clicking on one concept first (which is then highlighted in orange) and click on a second one to create the relationship, all (other) relationships disappear.	High	Bug fix
CM28	If you save a previously saved concept map and don’t supply a new name, a new entry appears with the same name as a previously saved concept map, there is no warning message, instead the drop down list shows two (or more) entries with the same name. It is not possible for the user to distinguish between them.	Medium	Prevent saving of two different concept maps with the same name. Enhance “save” and “save as” dialogues to support expected functions.
CM29	When clicking on the bin button to remove everything, the warning message is not shown properly and requires scrolling to read it: 	Low	Use a bigger popup window with text showing.

CM30	The save dialog has a scrollbar although there is nothing really to scroll: 	Low	Don't display unnecessary scroll bar.

6. Implications for Design

A wide variety of PD events with teachers and sometimes also students have been conducted for WP3 in 2014. These activities have provided a steady flow of information back to the project leadership and thereby influenced system development direction and priorities. This section gives a brief overview of findings, and then followed by a discussion on possible systemic changes to be undertaken by the pedagogical and technical teams.

6.1 Synopsis of the Findings of PD Year 2

Most of the identified scaffold apps and some of the labs and ILSs have been tested, some several times and in several versions and levels of refinement. The most frequent ones have been:

- Hypothesis Scratchpad
- Concept Map Tool
- Experiment Design Tool
- Questioning Scratchpad
- Splash Lab
- Sinking and Floating ILS
- Electricity Lab
- Ohm's Law ILS
- GoLabz

Most of the data captured are subjective self-reports of teachers and students on their experiences, opinions, and, in case of Core Group of Teachers, performance (i.e., task completion time, number of error committed) after applying a specific Go-Lab artefact for a specified task. They are collected via paper-based/tool-supported surveys and facilitated discussions, and are complemented as well as supplemented by researchers' observations. In addition, the researcher-based analytic evaluations provide the third source of data. The multi-method triangulation is a worthwhile procedure that enables us to draw a clear picture about the qualities of the Go-Lab design artefacts developed in Year 2.

The findings converge to an overall observation that teachers and students are generally highly persuaded of the potential benefits of Go-Lab. Nonetheless, they do find some of the components in their current form and some aspects of the integration of the components not intuitive to use. In other words, in interacting with the Go-Lab Portal, the users face some usability obstacles, which need to be removed before they can fully exploit its potential. Above all, a frequently occurring theme in anecdotal data is that teachers appear to be concerned that a tool should fit in with and support the pragmatic aspects of their work – classroom management, assessment and marking, limitations of computer infrastructure. Encouraging is that the users have provided many constructive comments and improvement suggestions. Apart from user-based inputs, the PD team discussed the data to derive recommended responses for the other teams to consider when deciding on the strategies for the work in the next phase.

6.2 Summary of Specific recommendations

Based on the most common themes identified in the usability data detailed above, we make the following broad recommendations, which could improve the usability of the Go-Lab Portal. Such improvements are necessary to make the Portal more attractive, engaging, efficient and effective for a wide range of students and teachers to acquire and use, and they will also help build an engaged and motivated community of users in the coming years of the project.

	Recommendation	Notes
SR01	Adopt consistent usability paradigms across the system, wherever possible based on intuitive or well established interaction modes.	It is recognised that some components (e.g. labs) are external and cannot be made to conform. Nevertheless consistency for all in-house components would be beneficial. A suggested approach to this is provided at 6.3 below.
SR02	Provide appropriate (informative and engaging) online help facilities for all components.	Where a component is not completely intuitive for all users, provide a means for them to find out what it is for and how to use it. For complex components this may take the form of a narrated video.
SR03	Scaffold tools which contain vocabulary should be customisable by the teacher or ILS author.	Some of the tools tested seemed too specific to a particular lesson or an area of science. The new app builder (not yet usability-tested) may remedy this.
SR04	Make user-supplied data consistent between components and flow naturally between them.	The implementation of the Vault (not yet usability-tested) may cover most of this.
SR05	Implement auto-save throughout the system, and UNDO and REDO features.	This is regarded as a normal expectation by many of the users.
SR06	Provide exemplary scientific rectitude in all material.	For instance, graphs should always have labelled axes; measures should always specify units.
SR07	Clarify workflow; e.g. don't provide tabs if they have to be accessed in a particular sequence. Do not provide controls and visual features which are irrelevant to the current task.	If some features are irrelevant in a particular ILS, there is no need to show them. Students may try them and be distracted. If features have to be accessed in a particular sequence, show them when they become relevant. Note: This may suggest splitting up some of the more complex labs or scaffolds so different parts can be presented at different stages in the ILS.
SR08	Provide facilities to ensure students cannot accidentally or deliberately access or change someone else's work.	The current logon by nickname without any password is thought by some not to be sufficiently strong.
SR09	Make all components available in all languages for target demographic.	
SR10	Consistency across browsers.	Make all components work consistently in all mainstream browsers (IE, Chrome, Firefox).
SR11	Provide useful outputs	Teachers and students have asked for facilities to save and print their work or send it as attachments for various purposes.

Table 6.1. Summary of specific recommendations for improving the Go-Lab Portal

There were a few suggestions that accessibility features (e.g. customisable font sizes or colours) should be added to provide users with variable capabilities. Some countries have legislation concerning accessibility of websites. An accessibility review is therefore also recommended.

There were also some suggestions that the scaffold app toolset could be rationalised. The calculator tool may be unnecessary because students have access in other ways, and that there seem to be multiple note-taking tools which appear to have very similar and simple function, and could possibly be replaced by a single one. Some other tools and apps seem to have been used to

a small extent so far, and seem to add little value, as perceived by the students and teachers involved the PD events. However, such tools may be appropriate for specialised purposes, so no firm recommendation can be made.

6.3 Style Guide

In addition to the fine-grained recommendations documented in Section 3, 4 and 5, in this sub-section we propose a more systemic approach to designing the Go-Lab system. We suggest that a **style guide** be required for developers of Go-Lab owned scaffolds and labs, covering the following aspects: Visual design, Content, Interaction design, Data handling, and Online help.

This style guide is not intended as a straitjacket but rather to ensure a consistent and predictable and seamless experience for users. The purpose is to deliver a user experience which feels like a coherent whole system rather than a mixture of assorted components. Some elements of the style guide could form a basis for a tutorial on how to use Go-Lab artefacts. ILSs could also follow some elements of the style guide. The following paragraphs do *not* attempt to be a style guide, but attempt to show some of the areas arising from usability studies which the style guide can include as its constituent sections.

- ***Visual Design***

Visual design could cover fonts and colours, but also a lexicon of icons (so e.g., trashcan or open or save are consistent across apps) and also a convention to distinguish interactive elements (apps and labs), pedagogical text, instructions and usability hints. Also, whether video links are presented as an image or as a link, and how the user interface is rendered on different screen sizes. Ideally this could be authored by someone with skills in graphic design or website aesthetic design.

- ***Content***

This mainly concerns the semantic, scientific and pedagogical impact of text and other components in an ILS, but the design of other components such as labs and scaffolds should ideally be coherent with this design. This could perhaps be authored by pedagogical experts.

- ***Interaction Design***

“Interaction Design” could usefully cover such things as dragging and dropping behaviour for Go-Lab components in which students build diagrams using a toolbox, consistent recycle bin behaviour including the provision of warnings, consistent typing behaviour (including use of e.g., ENTER key to indicate complete, and standard text styling facilities), provision of help text accessed with a “?” icon in a specific place, and access to video tutorials for any component over a certain level of complexity. This section of the style guide could also specify whether or not UNDO and REDO functions are provided. Notification of success (e.g., ‘save’ in Hypothesis Scratchpad) and warning before delete takes effect. Also, policy on keyboard shortcuts (if any) should be specified. It could also cover aspects of workflow. Ideally this section could be authored by interaction design specialists.

- ***Data handling***

This section of the style guide could specify policy for automatic and UI-triggered data storage and retrieval, and also provide consistent data model and data communication architecture (e.g., Vault) for passing data from one component to another. Access rights would have to be considered to

satisfy user concerns about plagiarism and also accidental or deliberate modification of another user's data. Ideally this could be specified by technical specialists, informed by user requirements.

- ***Online help***

This has frequently been requested, though many different ideas proposed of how to deliver it (e.g., pop up text, YouTube videos). Multimodal help, if resources allowed, seems optimal as it could address different needs and preferences of different user groups. Irrespective of the option(s) taken, there should be a consistent way to find the help.

7. Discussion

7.1 Interplay between Evaluation and (Re)Development

It is important to remember that the purpose of Participatory Design is formative evaluation, specifically on matters of usability and user experience, but also taking note of the overlapping area of (perceived) usefulness and benefit of the system. The aim is to ensure the development team are aware of and respond to user requirements for improved usability and user experience, and also to ensure they are aware of and respond to any new or changed functional requirements which arise as a result of using the software.

It is therefore necessary for the programme of user studies to have a close interaction with the development team, to understand what components or mock-ups exist and what important questions they have, and also to ensure they understand and respond to findings from usability studies. It is recognised that there will be many other influences and constraints on development effort, so usability studies can never seek to dictate priorities but it is crucial that they inform priorities.

Throughout this document we have included the development team's responses to usability findings (except where findings are too recent or not yet sufficiently proven to expect a response).

In general the development team have understood and accepted the usability findings, and the ongoing work of development appears to be remedying many of the usability findings which have been reported. The clearest example is the recent redesign of the ILS authoring process, which – though not yet formally usability-tested – appears to resolve a very large proportion of the usability concerns and make visible improvements to overall user experience. There are also clear examples of incremental improvements having already been applied to some of the more long-standing (and usability-tested) scaffold tools such as Hypothesis Scratchpad and Concept Mapper.

7.2 Findings from different user populations and study methods

As discussed at Section 3.1 above, the study programme has engaged with a very wide and diverse range of teachers and students from across Europe. The recruitment and sampling approach has targeted diversity rather than any strict notion of "representativeness" since it would be impossible to represent every perspective fully in every study without unacceptably extending timescales. The face to face PD activities have been supplemented by the remote PD activity methods involving Core Group Teachers, and by Heuristic Evaluations conducted by HCI researchers.

Several of the results tables in Section 3 specify which usability findings came from teachers, which from students, and which were identified by both groups. Additionally we have compared the findings from the Heuristic Evaluations with those from face-to-face studies when they covered similar ground. The varying circumstances of the studies (e.g. varying amounts of time available) mean a quantitative comparison would be misleading, but when going through the data in detail a number of qualitative observations (or strong impressions) are apparent.

It appears from our face-to-face studies so far that:

- Some students are very tolerant of what teachers perceive as poor usability features
- Students appear not to report usability inconsistencies between different parts of the system

- Students seem to want high interactivity and sophisticated graphics, and are less tolerant of long passages of text
- Older students seem to identify far more usability issues than younger students
- Students seem to focus on the immediate, whereas teachers seem also to notice many broader issues – e.g. lesson dynamics, pedagogical issues such as engaging with creativity, or concerns for not subverting inquiry paradigm
- Teachers seem to be more aware than students of issues around the overall flow of a session – e.g. how data is passed from one part of an ILS to another, or whether it has to be retyped.
- Teachers also tend to consider how the system might work with students of other ages
- Teacher seem more concerned than students that online help should be provided
- Teachers want the system to be scientifically exemplary – e.g. in having units on all measures, and labels on axes of graphs.
- Observer notes made by researchers add a lot of value to student sessions, since students don't always report usability issues which they encounter.

When comparing the HE studies with later end-user studies accessing similar Go-Lab components, some apparent trends are:

- HE studies identify a number of the usability issues quite quickly.
- HE studies may lead to 'false alarms', sometimes finding flaws which might not be discovered in short end-user studies, and sometimes ones which might never be discovered in practice.
- End-user studies provide far more practical perspective based on classroom practice.
- End-user studies give a good understanding of how the system may be used in practice and therefore possibly better information on prioritisation.
- End-user field studies are subject to a wide range of external influences and confounding factors - e.g. timetable constraints, distractions, IT infrastructure issues etc.
- End-user studies often also provide a wide range of non-usability-related findings.
- Findings from end-user studies are sometimes difficult to understand and categorise.

Finally, as one way of mitigating the risk that our sampling of teachers and students may be non-representative (e.g. teachers with poor IT skills being less likely to engage; teachers who engage more than once no longer identifying "first time use" experiences) we have sometimes asked teachers to usability questions three times: "for me", "for other science teachers" and "for my students". (E.g. CGT task 4, table 4.8a above).

Comparing responses "for me" with "for other science teachers" confirms that some of our teacher groups recognised that they probably were able to use Go-Lab facilities more fluently than other science teachers might, and this therefore reinforces the importance of remedying usability issues if we wish the system to be attractive to a wider market.

Comparing responses "for me "with "for my students", and also considering some of the qualitative comments provided, it appears that at least a few teachers (perhaps older ones) believe their students are better with ICT than they are themselves. This is not a proven conclusion but worth considering as we plan PD activities for the future.

8. Conclusion and Outlook

8.1 Achievements

A series of 21 user-based Participatory Design (PD) events have been successfully conducted involving teachers and students from nine countries across Europe in Go-Lab Year 2. Among these events, the Go-Lab Summer School 2014 was most extensive where three formal studies with a large and diverse group of teachers were conducted, as well as deep involvement with the teachers engaged in Go-Lab tasks throughout the week which resulted in a wide range of observational data, interview data and ethnographic style immersion and insights.

Furthermore, the HCI research team of WP3 have also carried out several heuristic evaluations, feeding relevant input on usability and user experience issues to other work packages, also covering system functionality, robustness, and other qualities. The usability findings from heuristic evaluations seem to have correlated closely with usability findings from teacher evaluations; however the teacher evaluations have provided more insight into contextual and practical matters and also perceived usefulness.

Additionally, a core group of teachers has been established to provide rapid online feedback on specific Go-Lab-related questions. This proves to be a highly effective method for rapid engagement.

All these PD activities are valuable sources of feedback on a wide range of Go-Lab system components, some based on non-interactive mock-ups and others on working interactive prototypes.

The findings have been communicated to the pedagogical and technical teams on an ongoing basis, and many have already been addressed in their respective design and development work.

8.2 Challenges

Clearly, there have been different challenges in implementing the PD events. The most critical one was the recruitment of teachers, students and schools. The recruitment process has been more successful in some countries but more challenging in others, due to cultural and political differences, varied professional styles, and the workload and level of work-related stress experienced by teachers in some countries. Even when a session is organised and agreed, in field studies there are many possibilities for unplanned and uncontrollable factors (e.g. attendance, timetabling issues, distractions, technology issues) to influence the outcomes. Good planning is beneficial but researchers still sometimes have to improvise.

Within Go-Lab, like many RTD projects involving public users, there is a concern for the reputation of the product, and for it to be shown in its best possible light, partly to create a good reputation for the product and partly so that teachers who have engaged in participatory design sessions would find them useful and would be willing to continue to be involved. This means that there are sometimes compromises and trade-offs to be made when considering asking teachers to help evaluate software in their classes, especially in this early phase of development when the system is – quite understandably – not yet market ready. The Heuristic Evaluations and the use of the Core Teacher Group have helped us considerably in meeting these challenges.

8.3 Limitations

As has already been mentioned, recruitment of schools and particularly getting access to live classes with students has been limited in some countries, and we have therefore not been able to be as purposeful and systematic about when and where we conduct studies as we might have liked. As a result, some of the system components – especially the more recently created ones – have been evaluated by a relative small number of participants, perhaps not including every perspective to be found in the target domain.

Feedback received from participants has sometimes been difficult to interpret, especially when they have not been responding in their native language, or when the material has been translated. Also in some cases different participants have very different usability needs, preferences and ideas so elements of feedback can contradict each other.

The workflow defined at Section 2.5.4 has been most effective when partners have been able to negotiate a PD session well in advance and give the WP lead partner plenty of notice to prepare or adapt a protocol. However, occasionally PD opportunities occur ad hoc, and protocols and materials have been prepared rather quickly and less optimally. We will therefore endeavour to prepare a toolbox of more flexible materials suitable for a wider variety of PD sessions with students of different ages, sessions of varied durations etc.

8.4 Outlook

In Year 3 and Phase B of Go-Lab, WP3 will be more ambitious to involve an even larger number of teachers and students to evaluate more mature interactive Go-Lab system components. With the ongoing support from WP7 (Pilot), we are optimistic that this rather challenging target number can be reached. Figure 8.1 illustrates that as the software becomes more functional and usable, and the in-house ILSs become more relevant to school curricula, teachers may find it more attractive, effective and productive to conduct a lesson in this way, and provide PD data. As the engaged teacher community grows we expect to have a reducing need for Heuristic Evaluations, but active involvement of Core Teacher Group in providing prompt feedback will continue.

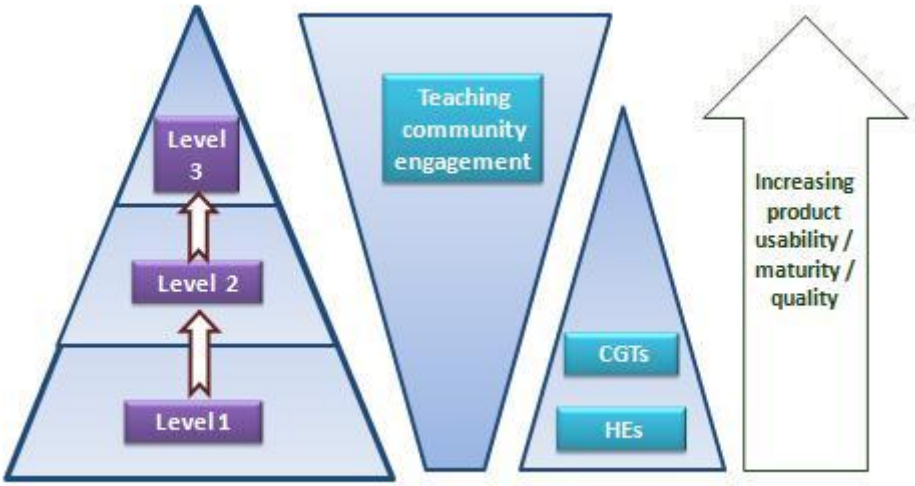


Figure 8.1. Changing communities of engagement

Furthermore, with prototypes of higher fidelity, more performance-based data, in addition to the perception-based, can be collected. One possible source of metrics which we will explore is the use of system-based measurements using web-analytic types of tools. This would require significant support from the development team, From the WP3 perspective, one intriguing research question is to understand the downstream utility of evaluation feedback to be collected with different approaches and instruments. In other words, we aim to assess the actual impact of such feedback on the design and development work by tracking how users' perceptions, attitudes, behaviours, performances, acceptance, and adoption will vary with the newer versions of the prototypes, which are redesigned based on the issues identified and associated improvement suggestions.

It is anticipated that WP3 will have even closer collaborations with WP1, WP4 and WP5 in the coming year, in order to prioritise studies and communicate findings even more effectively.

9. References

- Bargas-Avila, J. A., & Hornbæk, K. (2011). Old wine in new bottles or novel challenges: a critical analysis of empirical studies of user experience. In Proc. CHI'11 (pp. 2689-2698). ACM.
- Brooke, J. (1996). SUS – A quick and dirty usability scale. Usability evaluation in industry, 189 - 194.
- Garris, R., Ahlers, R., & Driskell, J. E. (2002). Games, motivation, and learning: A research and practice model. Simulation & gaming, 33(4), 441-467.
- Deterding, S., Dixon, D., Khaled, R., & Nacke, L. (2011, September). From game design elements to gamefulness: defining gamification. In Proceedings of the 15th International Academic MindTrek Conference: Envisioning Future Media Environments (pp. 9-15). ACM.
- Hartson, R., & Pyla, P. S. (2012). The UX book: process and guidelines for ensuring a quality user experience. Elsevier.
- Hassenzahl, M. (2008). User experience (UX): towards an experiential perspective on product quality. In Proceedings of the 20th International Conference of the Association Francophone d'Interaction Homme-Machine (pp. 11-15). ACM.
- Hassenzahl, M., & Tractinsky, N. (2006). User experience-a research agenda. Behaviour & Information Technology, 25(2), 91-97.
- Heintz, M. Law, E. L-C., Govaerts, S., Holzer, A., & Gillet, D. (2014).Pdot: participatory design online tool. In CHI Extended Abstracts 2014: 2581-2586
- Jamieson, S. (2004). Likert scales: how to (ab) use them. Medical education, 38(12), 1217-1218.
- Law, E. L. C., & van Schaik, P. (2010). Modelling user experience–An agenda for research and practice. Interacting with computers, 22(5), 313-322.
- Law, E. L. C., Roto, V., Hassenzahl, M., Vermeeren, A. P., & Kort, J. (2009, April). Understanding, scoping and defining user experience: a survey approach. In Proceedings of CHI'09 (pp. 719-728). ACM.
- Law, E. L. C., van Schaik, P., & Roto, V. (2014). Attitudes towards user experience (UX) measurement. International Journal of Human-Computer Studies, 72(6), 526-541.
- Marsh, T. (2011). Serious games continuum: Between games for purpose and experiential environments for purpose. Entertainment Computing, 2(2), 61-68.
- McCarthy, J., & Wright, P. (2004). Technology as experience. MIT Press.
- Muller, M. J., & Druin, A. (2010). Participatory Design: The third Space in HCI. Handbook of HCI. J. Jacko and A. Sears. Mahway.
- Nielsen, J. (1994). *Usability engineering*. Elsevier.
- Norman, D. A., & Draper, S. W. (1986). User-centered design. New perspectives on human-computer interaction. Lawrence Erlbaum.
- Simonsen, J., & Robertson, T. (Eds.). (2012). Routledge international handbook of participatory design. Routledge.
- Tuch, A. N., Trusell, R., & Hornbæk, K. (2013, April). Analyzing users' narratives to understand experience with interactive products. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (pp. 2079-2088). ACM.
- Tullis, T., & Albert, B. (2008). Measuring the user experience. Burlington, MA.
- Walsh, G., Druin, A., Guha, M. L., Foss, E., Golub, E., Hatley, L., ... and Franckel, S. Layered elaboration: a new technique for co-design with children. In *Proc. CHI'10*, 1237-1240.

Appendix A: Participatory Design Protocol (PDP) – Splash Sinking and Floating ILS

Go-Lab Work Package 3

Composite PD session for "Splash – sinking and floating" ILS

1. Objectives

- To obtain an impartial and unbiased assessment of the usability and user experience from teachers and students based on their use of this ILS.
- To provide creative ideas for refining the design where appropriate.

To achieve this we want to design a session which is attractive, engaging, enjoyable and has a good atmosphere so that teachers are encouraged to participate actively. We hope this will enable a productive and effective session and also leave teachers willing to participate again in the future.

2. Participants

This study is designed to elicit feedback from a group of 6 or more teachers at once, but can be used with minor adjustments with fewer or even just a single teacher. It could also be used by a group of students (with a teacher also present) with some minor adjustments.

As well as a facilitator to run the session, it is wise to have other Go-Lab staff to provide technical and system expertise and support especially in the main practical session. A ratio of at least 1 technical support for every 10 teachers is recommended.

3. Equipment & materials

A PC or laptop with web access and a browser (Google Chrome or Mozilla Firefox) should be provided for each participant (or if that is not possible, then one for every two participants, so they work in pairs).

The facilitator will need a PC and projector.

Camera

Printed materials:

- Feedback booklets (for annotation)
- Questionnaires
- Posters (for 3 stickers exercise)

4. Preparation checklist

1) Translate all materials into the language which will be used for the study
2) Conduct a pilot study if possible
3) Print / compile all materials and equipment
4) Test all technology to be used
5) Acquire PDot accounts / logins from Matthias Heintz (mmh21@leicester.ac.uk)
6) Prepare a Graasp space with links for use in class (see 7.4 below)
7) Prepare the room layout and visual aids

1

5. Methods

Always included:

- Questionnaires – paper or online survey ("Lime survey")
- PDot – Participatory Design online tool
- Answer booklet of screen shots for annotation
- Observer note taking

If time permits:

- "3 stickers on posters"
- Facilitated discussions

6. Session outline

1	Welcome and overview of the PD session	Talk	5 mins
2	Introduction to Go-Lab and the ILS	PowerPoint	10 mins
3	Concise questionnaire of first impressions	Online or paper	5 mins
4	**Brief** demonstration of how to get into ILS and how to navigate through it.	Projected laptop screen	5 mins
5	How to provide feedback: PDot and/or annotating the booklet	Talk + screen demo of PDot	10 mins
Brief break during which computers / browsers can be booted and checked			
6	Participants work through ILS, with technical or product support as needed, noting down any usability concerns either using PDot or booklet.	Computer session	60 mins
7	Elicit overall impressions (e.g. best features, worst features, what to add etc)	Questionnaire – online or paper	10 mins
8	(If time permits) Detailed assessment and critique	"3 stickers on posters", then facilitated discussion	20 mins
9	Final questions, thanks and session wrap-up	Talk	5 mins

7. Session script and materials

7.1. Welcome and overview of the PD session

The introduction should cover the following points.

- Welcome all the participants
- If appropriate get them to complete name badges
- Explain session length (and any breaks)
- Explain objectives
  - To obtain an impartial and unbiased assessment of the usability and user experience from teachers and students based on their use of this ILS.
  - To provide creative ideas for refining the design where appropriate.
- Mention the varied methods we will use and make it sound exciting.
- Emphasise that the aim is to help us understand how usable the system is and to give us ideas to change it to be more usable for a wide range of users. We are not trying to evaluate the computer abilities of individual teachers, and all data will be summarised and anonymous. We want genuine opinions and creative ideas to help us make the system as effective and usable as we can make it.

You may wish to cover these points in your own way, or base it on the sample script which follows:

Hello and welcome to this Go-Lab Participatory Design session, which will last approximately 2 hours. My name's xxxx and I'll be facilitating the session, assisted by yyyy & zzzz who will be available to help out in the practical work. Participatory Design is a process to involve potential users of systems to be involved in and influence the design of those systems.

We'll be doing a whole range of activities, some individual on computers or answering questionnaires and some more interactive and creative group sessions, so we hope you'll find it enjoyable as well as useful. We'll also have a very brief break about half way through when you can check mobile phones etc.

The aim of the session is for us to get a better understanding of how usable various parts of the Go-Lab system are at the moment, and how best we can improve it, as parts of the system are not yet finished and can still be improved, and we want the system to be as usable as possible. We'll be asking you to try out part of the system and provide us with feedback. We'd like you to be completely frank about your opinions and experiences of the system; telling us what you find helpful, what you find unhelpful and if possible giving us ideas of how to improve anything which you think would benefit from improvements. The session is to assess the system, not you as an individual user. We'll only report trends and patterns or perhaps using anonymous quotes. If you'd like to make any completely anonymous comments, please write them on the sheet of paper in front of you and leave them on my table at the end.

We promise to take all of the information you provide us with seriously, and to use them to refine the system design, but we can't promise to satisfy everyone with the final design.

A couple of administrative points:

3

- *We'll all get the most from this session if we switch off our mobile phones and other devices and keep our attention on Go-Lab as much as possible.*
- *It will help me if you'd write your name (whatever you'd like me to call you) on the card on front of you and place it where I can see it.*

Do you have any questions before we get started?

7.2. Introduction to Go-Lab and the ILS

Use the PowerPoint presentation "Go-Lab WP3 General Intro" provided. There are speaker notes in English for the presenter. If the audience are already familiar with Go-Lab then it can be used quite quickly as a reminder.

7.3. Concise questionnaire of first impressions

Questionnaire provided in Appendix 1. We suggest printing this in advance and handing them out at this point. Questionnaires on paper add a bit of variety so the session isn't all listening and computer work.

7.4. Brief demonstration of how to get into ILS and how to navigate through it.

Appendix 1 explains how to create a Graasp space with a simple url in advance to help in providing participants with links.

At this stage, using the facilitator's screen and projector, demonstrate how to get the url for the ILS and show just the first few screens. Remind the participants that we are interested in feedback on usability, and that we know the system isn't perfect yet, so they should feel free to make comments, and also any constructive ideas for improvements.

7.5. How to provide feedback: PDot and/or annotating the booklet

Use the following script, or adapt it for your own needs.

PDot is a tool designed to gather feedback online. To access it you click on the link we provided for you.

Open the following URL in the browser of the demonstrator PC (projected for the participants to follow): https://campus.cs.le.ac.uk/tomcat/PDtoolPlayground/?locale=en&eventId=26 . Then login to PDot using the credentials 21 and demo.

You can login to PDot using the credentials provided.

In this demo, PDot will be used to annotate the picture of a phone. However you are going to use it later to annotate the inquiry learning space we are evaluating today.

After logging in, the PDot tool will be shown in the upper left hand corner of the screen {point it out on the screen and/or projected image}. *Short instructions what to do to perform the evaluation with PDot are given in the PDot instructions on the top of the screen* {point them out}, *in this case we are asked to "Please give feedback on the front of the phone. Once you are done, click next".*

*So we now look at the phone and as soon as we notice something, we want to comment on, we click on the Give Feedback button in the PDot tool. For example, we notice, that it is helpful to have large digits for the clock, so that it is easy to read. Thus we click on the give feedback button. You will notice that the mouse cursor now changes to a little sticky note. This is to point out the spot we want to give feedback on. We click on the clock {*click on the watch on the screen of the mobile phone*} to indicate that we are commenting on it. You will now notice, that the content of PDot changed, allowing us to enter a textual comment and specify our feeling. So we go ahead and enter the comment "I like that the digits of the clock are this large, because it makes it easy for me to read them." And click on the Like button underneath it. Then we save this feedback, by clicking the save button {*perform the described actions*}.*

Secondly we notice that the on/off switch is on the right hand side of phone, but we would like it to be on the top. Thus we click on the switch to create a second sticky note there. Then we use the mouse to draw an arrow from the current position of the switch to the desired one on top of the phone {perform the described actions} and add a textual comment "I think the on/off switch would be more convenient to press if it would be on top of the phone.". We press the Dislike button to indicate that we do not like the current location of the button. Then we press the Save button to save our comment.

5

Thirdly we would like to have a slider to unlock the screen. Thus we click on the position where we would like to have the slider to create the little "location sticky note" there. Then we use the pen to draft the slider. Because we do not like the first draft, we use the Undo button to remove it and start over again. Alternatively we could also have pressed the Cancel button. As we think the slider would be a good alternative, but we are also fine with the current unlock mechanism we press the Neutral button before we save our feedback
{Perform the described action while explaining them}.

6

As we are currently done with giving feedback we click the Cancel button to leave the feedback mode. This hides our comments so that we can look at and interact with the original element to evaluate undisturbed. As we cannot think of any additional feedback regarding the front of the phone we click on the Next button in the PDot instructions {Perform the described action while explaining them}.

This leads us to the next step in the evaluation process, where we are asked to give feedback about the back of the screen. We look at the phone and notice that we would like the speaker to be on the other side of the phone as we are afraid we might cover it when holding the phone with our right hand. Thus we press the Give Feedback button. We notice that our comments regarding the front are no longer displayed as we are now looking on the back of the phone and entered the next step in the PDot instructions. We click on the speaker, to create a "location indication sticky" there and draw an arrow using the pen to indicate our wish to move it to the other side of the phone. Then we write a short comment explaining our reason "I am afraid I would cover the speaker with my right hand while holding it. Thus I recommend to move it." Then we save our feedback by pressing the Save button. {Perform the described action while explaining them}.

In this demo the evaluated content changed automatically when pressing the Next button in the PDot instructions. For technical reasons this is currently unfortunately not possible in the ILS we are going to evaluate, thus you will have to first click on the Next button in the PDot instructions and then switch to the appropriate tab using the ILS navigation.

Please delete all your comments after giving the demo to your teachers so that the demo is reset for the next facilitator.

7

Do any of you have any questions regarding using PDot to give feedback?

Do you all feel you'll be able to use it?

We'd like at least half of you to use this method, but we have a paper and pen alternative which some of you may prefer.

The Feedback booklet is supplied separately. Have some of these printed and stapled in advance. Show them a copy of the Feedback booklet.

I can give some of you a copy of this booklet which contains screenshots of screens you may encounter as you work through the ILS. You can write your comments and suggestions for the product in the booklet in whichever page seems the most relevant. Feel free to use it in your own way and write or draw whatever you think is the best way to communicate your ideas to us.

(Hand out Feedback booklets to those who will be using this method)

The partner conducting the study may choose to have a short break at this point, so the session isn't too demanding for the participants. You may wish to use the time to ensure all PCs are working and have a usable browser, and provide each participant with a link to the "Splash – floating or sinking" ILS.

7.6. Participants work through ILS.

Introduce this session in your own words or as follows:

We now come to the main part of our session.

You should all have a link to a pre-written Go-Lab ILS which provides a lesson in buoyancy entitled "sinking or floating".

I want you each to follow the link, and then work through the ILS trying out all the features that a student would use when actually doing this lesson. You will see the five learning phases of the Inquiry based learning module. You will find a mixture of information to read and actions to do.

You do not need to do everything a student would do in full detail; your purpose is try out all features to let us know how well designed they are, and especially how easy to learn and use they are. So for instance in the Orientation phase you will find links to four youtube videos. There's no need to watch all four in full, but we'd like you to try watching at least one of them for a while to understand how easy it is to get into them and to finish watching.

As you are working through the ILS, we would like you all to record your experience of the usability – any points good or bad, and if you have time any suggestions for improvements. Some of you will be doing this using PDot; others using the feedback booklets. Please try to

do this independently rather than helping each other, so your comments reflect your own personal and subjective experience. Remember that in this session we are mainly concerned with issues of usability and your overall experience of using the system. There are no right or wrong answers, and we won't be making any judgements of your individual capabilities; we want to know how good the system design is for a range of teachers and students and whether it needs improving.

Please feel free to ask me or any of the go-Lab team here if you have any questions or would like some help or advice, or experience any technical problems.

We will also be taking notes during this period. These will be about the product, not individuals.

During this session, mingle informally with the participants, to check they're getting on okay. Help them out if they ask or if they seem to need it. Encourage them to record their usability observations on the booklets or in PDot. Take notes of any relevant observations, especially about what parts of the system needed the most help, and why. Watch how the time is progressing and when appropriate, encourage the group that they should soon be moving on to the next phase.

7.7. Elicit overall impressions

Questionnaire provided in Appendix 2. We suggest printing this in advance and handing them out at this point. Questionnaires on paper add a bit of variety, so the session isn't all listening and computer work.

7.8. Detailed assessment and critique (if time permits)

Hand out three stickers to each participant: one green, one yellow and one red (if you do not have such stickers, coloured sticky notes will also do). Have a set of large (A3 or larger) printouts of each phase of the inquiry learning space either on the walls or on flipcharts. (You can produce these by printing separate pages of the Feedback booklet enlarged). Ask the participants to stick their green sticker on something they liked best, their red sticker on the thing they liked least and their yellow sticker on something they would like to discuss.

Ask participants to attach their stickers and return to their seats. Then use the stickers to facilitate the discussion about the ILS (e.g. point out areas with a lot of stickers, discuss some of the yellow stickers, discuss some of the green and red stickers and the reasons behind putting them there).

Again you may wish to explain this in your own way, or base it on the sample script which follows:

Thank you for all the feedback you've provided us with so far. We would like to finish this session with a group discussion. First though, we are going to use the "stickers on posters" technique to facilitate our discussion. We will now hand out three stickers to each of you; one green, one yellow and one red. Please stick these on the posters of the Go-Lab screens. Use the green sticker to indicate the part you liked best, the red one on what you liked least, and the yellow one to indicate one thing you would like to discuss.

The facilitated discussion depends highly on the placement of stickers and group dynamics. Thus we can only give general advice in this case.

We would start with discussing the two or three elements (depending on the time even more or less) with the most yellow stickers. What did the teachers like to ask or discuss regarding these elements. This could be followed by a discussion of the one (again depending on the time also some more) element or area with the most red stickers. To end the session on a positive note we would then suggest finishing with a discussion about the area with the most green stickers. Why have they been put they. What exactly did the teacher like? Could this also be used in other parts of the inquiry learning space?

Please keep an eye on the time, as this can be tricky during discussions and keep enough time for the final questions, thanks and wrap-up.

To document the results of this session, please take photographs of the printouts with the stickers and take notes of the discussion.

7.9. Final questions, thanks and session wrap-up

In the final few minutes:

- Ask them if they have any final comments about the system which they haven't already provided
- Ask if they liked using PDot. If positive, ask whether they think they could conduct a PDot session with their students?
- Ask if the session was interesting and enjoyable and felt worthwhile? Ask how it could be improved?
- Thank them all for their great input and wish them lots of success using Go-Lab.

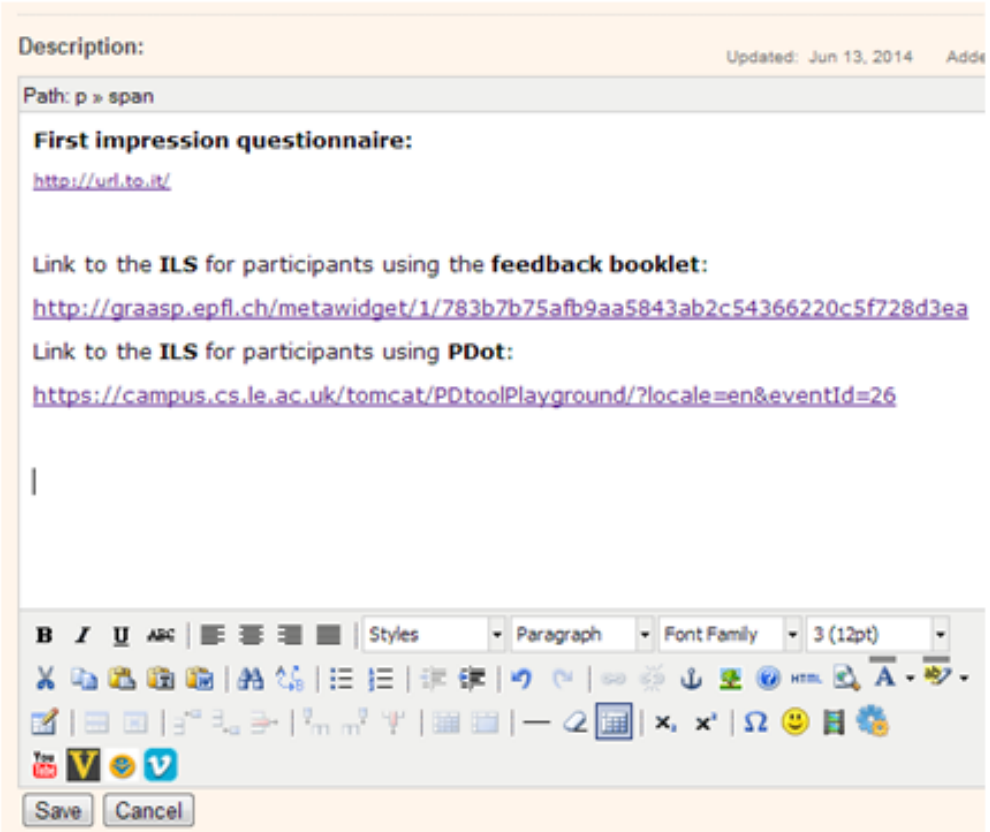
Appendix 1 – Preparing Graasp space with links

In preparation of the PD session please create a Graasp space containing all the links used during the session (to the ILS, to PDot, to any online questionnaires).

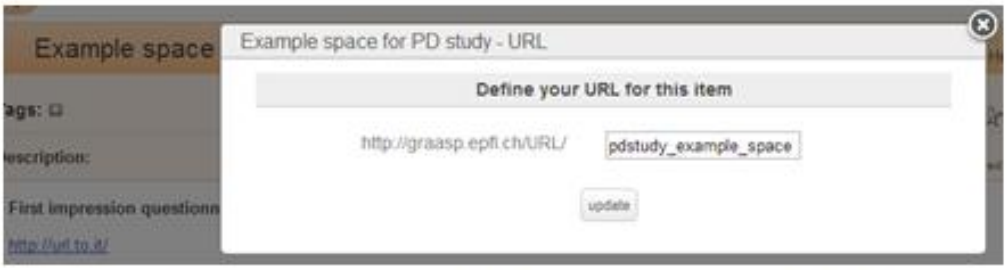
You could create this space either in the pad of your personal Graasp space.

To create the space click on the + button in the parent space () and select new space (

). Give it a name and create your space by pressing the Create button then access it by selecting it from the available spaces in the Pad. To provide the information for your participants double click on the description text and add the links you are going to use during the session in there (preferably with some structuring descriptions, e.g.:)



To ease the access to this space for your participants, please give it an easy to type URL by clicking on the URL link and specifying it, e.g.



Then on the day of the PD session you could just ask your participants to insert the URL in a browser (e.g. https://graasp.epfl.ch/#url=pdstudy_example_space) and access the resources from there.

12

Appendix 2 – "Early impressions" Questionnaire

Go-Lab Participatory Design

Questionnaire J01 – "Early impressions"

Date (DD/MM/YYYY) ……

Name or participant number (optional) ……

1. In total, how much time have you spent learning about or using Go-Lab so far?

   < 1 full day / 1 to <3 full days / 3 to < 6 full days / 6 or more full days

2. What do you think of Go-Lab so far?

3. How do you feel about the possibility of using Go-Lab in your school?

4. I think the Go-Lab system is easy to use?

   VERY EASY / QUITE EASY / SOMEWHAT DIFFICULT / QUITE DIFFICULT / VERY DIFFICULT

5. I think it will be quite easy for me to learn how to use Go-Lab fluently?

   STRONGLY AGREE / AGREE / NEITHER AGREE NOR DISGREE / DISAGREE / STRONGLY DISAGREE

6. Any other comments or thoughts at this stage?

14

Appendix 3 – Questionnaire after using ILS

Go-Lab Participatory Design: Questionnaire J02 – "After ILS usage"

Date (DD/MM/YYYY) ……

Name or participant number (optional) ……

Which ILS have you just used? ……SPLASH…-….floating…or…sinking……

Thinking of the features of Go-Lab which you have just used, especially of how easy or difficult they were to use, please respond to the following questions.

1. What was the best part of the system, and why?

2. What was the worst part of the system, and why? Can you provide any brief suggestions to improve the experience of using it?

3. Can you suggest any changes or extra features which would make the system easier to follow and to use?

16

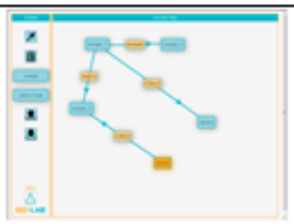
Again, **thinking of the features of Go-Lab which you have just used, especially of how easy or difficult they were to use,** please respond to the following statements, indicating your level of agreement or disagreement by putting an X in the relevant column.

Statement	Response				
	Strongly agree	Agree	Neutral	Disagree	Strongly disagree
I feel Go-Lab is an excellent system.					
Computer based labs can never be as good as real labs.					
The Go-Lab system could work in practice in my school.					
I felt the system lacked an overall design.					
I hope to be able to use Go-Lab when it is formally launched.					
The system needs a lot of work before I would be prepared to use it.					
I think that I would like to use go-Lab frequently.					
I found the system unnecessarily complex.					
I thought the system was easy to use.					
I think that I would need the support of a technical person to be able to use this system.					
I found the various functions in the system were well integrated.					
I thought there was too much inconsistency in this system.					
I imagine that most people would learn to use this system very quickly.					
I found the system very awkward to use.					
I felt very confident using the system.					
I needed to learn a lot of things before I could get going with this system.					
I think that I would like to use this system frequently.					

17

Finally, **thinking of the features of Go-Lab which you have just used, especially of how easy or difficult they were to use,** please consider each of the 5 main components and respond to the following statements, indicating your level of agreement or disagreement by putting a number 1 to 5 in each cell.

Strongly agree	Agree	Neutral	Disagree	Strongly disagree
1	2	3	4	5

	This tool was easy for me to use	There is a need for such a tool	I would use this tool in its current form	This tool needs improvement
 Concept map tool				
 Notes tool				
 Sinking and floating lab (1)				
 Sinking and floating lab (2)				
 Questioning scratchpad				

 Experiment Design tool				

Finally, **thinking of the features of Go-Lab which you have just used, especially of how easy or difficult they were to use,** please consider each of the 5 main components and respond to the following statements, indicating your level of agreement or disagreement by putting a number 1 to 5 in each cell.

Strongly agree	Agree	Neutral	Disagree	Strongly disagree
1	**2**	**3**	**4**	**5**

	Concept map tool	Notes tool	Sinking and floating lab (1)	Sinking and floating Lab (2)	Questioning scratchpad	Experiment Design tool
						
1. This was easy for me to use						
2. There is a need for such a tool						
3. I would use it in its current form						
4. This needs improvements						

Appendix B: Participatory Design Protocol (PDP) - Electricity Ohm's Law ILS

Brief:

There will be 27 students aged around 14, and one teacher, for 3 classroom lessons of 50 minutes. The first lesson, today, will not use Go-Lab technology but will introduce the ideas and concepts.

The other two will involve students in hands on activities including the Electricity Lab http://go-lab.gw.utwente.nl/production/circuitSimulator/build/circuitSimulatorDataViewer.html and the use of the embedded Data Viewer tool.

The students have some prior experience of Go-Lab using the Splash lab, the Hypothesis scratchpad and the Experiment design tool.

The objective of the sessions from the Go-Lab project perspective is to get responses from students about usability, problems they face, things they like and to get suggestions about improvements.

The teacher's main objective however will be for the students to learn about electricity.

The opportunity is somewhat opportunistic and therefore the researcher will mainly be observing, not running the classes, but there is an opportunity to administer a questionnaire of 10-15 minutes after at least one of the two lessons.

ULEIC are asked to provide questionnaire(s).

Assumptions:

We assume the Electricity Lab tasks will consist of:

- Creating a simple circuit diagram – perhaps just a battery, bulb and switch
- Adding some meters to the circuit
- Testing the effect with different voltages
- Collecting this data in the data tool
- Using the graph part of the tool to view the data

Suggested observations

Please note down any relevant observations, and also any comments or questions they might ask, especially about usability or interaction style.

Some specific points to look out for:

- Did pupils start before they had read everything?
- Did they notice and read the “Hints” at the bottom of the screen before they started?
- Did they attempt to hover the mouse pointer over anything in the hope of getting some help?

Suggested Interview

If you have chance, we would suggest having an interview or discussion or email exchange with the teacher after the second class. Point to cover could include:

- Did she notice the students encountering any particular problems with using the system?
- Would she feel it necessary to teach students about the system before they could use it? Is this a problem?
- If she was using this system with a class without any Go-Lab staff present, does she think she would have to provide extra support because of the usability of the system? Would this be a problem?
- Does she think other teachers would be more or less comfortable using it? And why?
- Does she think students could use this without supervision or support – e.g., for a homework assignment?

Suggested Questionnaire

The following pages contain a questionnaire we have drafted and believe will be useful. We suggest printing this out and giving a copy to each student straight after the have used the system.

We have not piloted this questionnaire with a school aged student so we do not know for sure how long it will take, but we have aimed to try to get the most useful feedback in the available 10-15 minutes.

We’re not quite sure how the use of the system will be split between the two lessons, so you may need to adapt this as appropriate.

Please encourage them to use the text boxes to give us deep insights, not just the multiple choice responses.

Go-Lab WP3 – Questionnaire after using Electricity Lab and embedded data tools

Date: ……

Age: ……

Please circle the appropriate answer or fill in the text boxes.

Did you have enough guidance from the teacher on how to use this lab?

NOT ENOUGH / ENOUGH / MORE THAN ENOUGH

Do you think you could use this lab as part of a homework assignment?

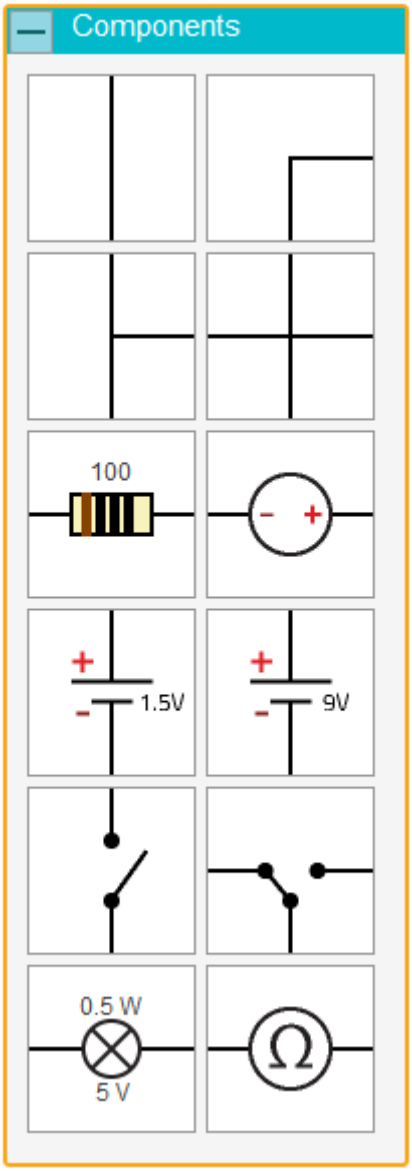
Do you think there is enough information included in the lab about how to use this lab?

Do you think the overall layout of this lab is good?

Can you suggest any way of improving the overall layout?

Did you find any of the symbols difficult to recognise?

1. COMPONENTS



a) Did you use this? NO / A LITTLE / A LOT

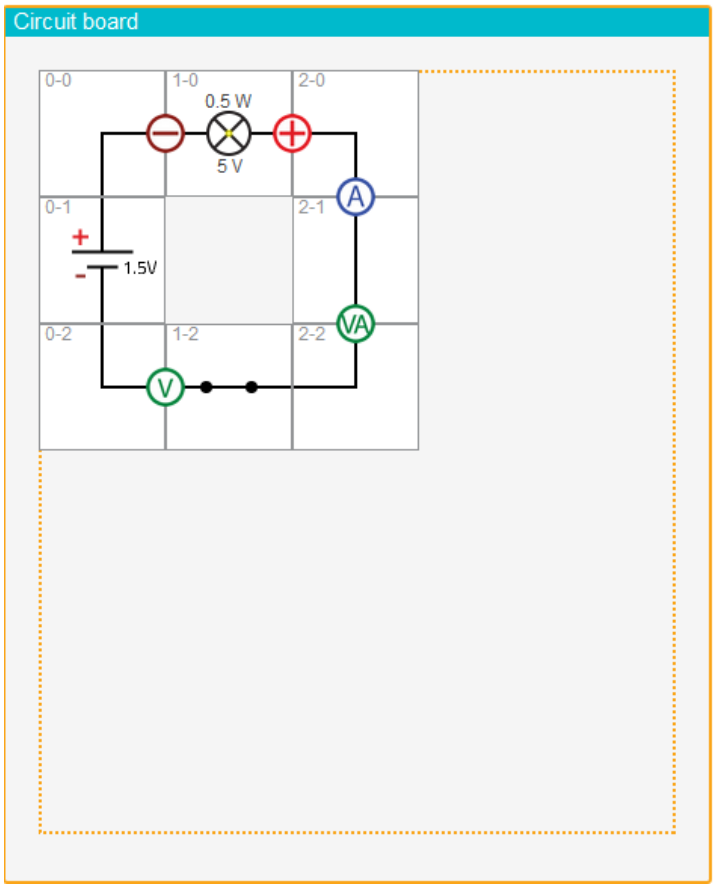
b) Did you find this easy to use?

VERY EASY / EASY / NOT EASY

c) Why?

d) How could it be improved?

2. CIRCUIT BOARD



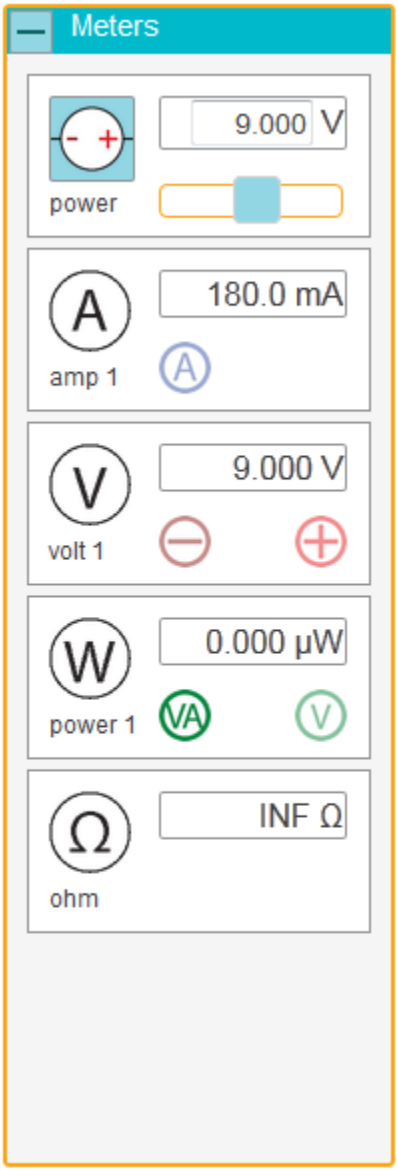
a) Did you use this? NO / A LITTLE / A LOT

b) Did you find this easy to use? VERY EASY / EASY / NOT EASY

c) Why?

d) How could it be improved?

3. METERS



a) Did you use this? NO / A LITTLE / A LOT

b) Did you find this easy to use?

VERY EASY / EASY / NOT EASY

c) Why?

d) How could it be improved?

4. DATA COLLECTOR

a) Did you use this? NO

/ A LITTLE / A LOT

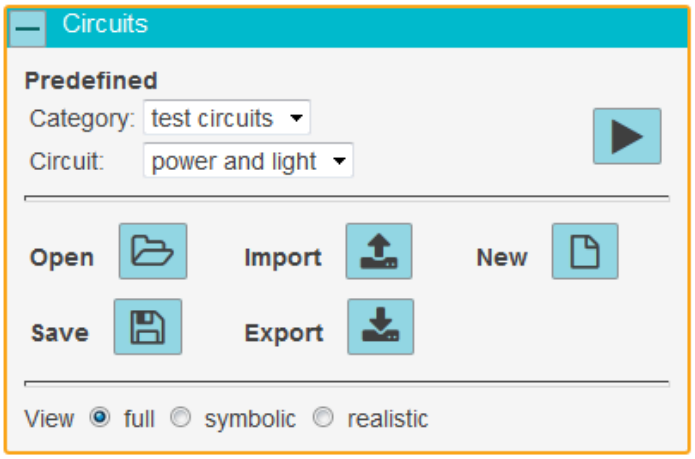
b) Did you find this easy to use?

VERY EASY / EASY / NOT EASY

c) Why?

d) How could it be improved?

5. CIRCUITS



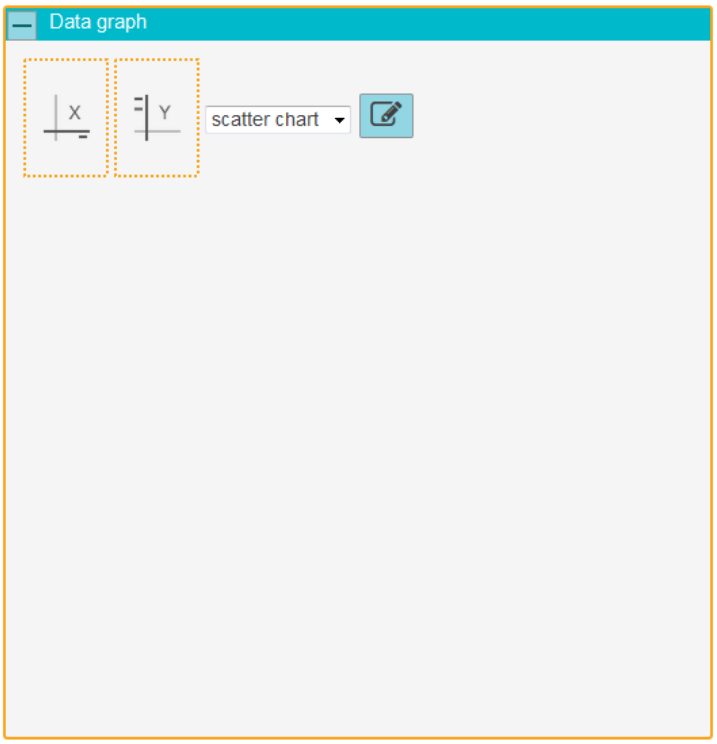
a) Did you use this? NO / A LITTLE / A LOT

b) Did you find this easy to use? VERY EASY / EASY / NOT EASY

c) Why?

d) How could it be improved?

6. DATA GRAPH



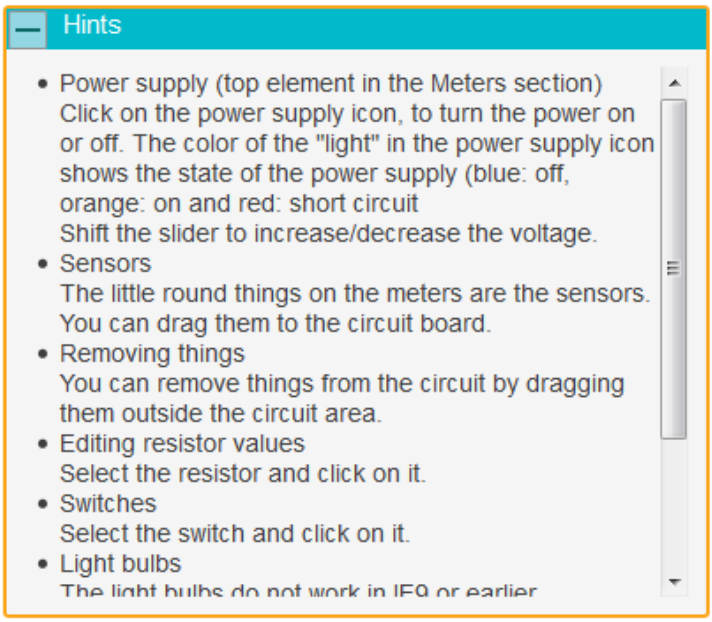
a) Did you use this? NO / A LITTLE / A LOT

b) Did you find this easy to use? VERY EASY / EASY / NOT EASY

c) Why?

d) How could it be improved?

7. HINTS



a) Did you use this? NO / A LITTLE / A LOT

b) Did you find this easy to use? VERY EASY / EASY / NOT EASY

c) Why?

d) How could it be improved?

Appendix C: Sample Results of Electricity Ohm's Law ILS

Date	16/7/2014 (Summer School)
ILS	Electricity ("Ohm's Law")
Duration	2 hours
Researchers	Rob E-W, Matthias H
Participants	35 teachers

Notes:

- Conducted in School computer lab.
- Teachers worked in pairs due to limited number of computers
- All feedback - questionnaires, Pdot and feedback booklets, represents two teachers
- The session felt highly motivated and energised, and questions afterwards suggest teachers enjoyed it
- Technology and internet connection worked well.
- It was difficult to explain what was going to happen and specifically to make them focus on usability, not educational effectiveness of ILS content. This took a lot of time and meant some felt rushed towards the end of the session.

Data collected:

Paper questionnaires, observations, discussion notes, feedback booklets, PDot material, annotated posters.

Questionnaire J02:

Q1	**best part of system**	**Why**
P1	Hypothesis & Experimentation	very amazing for students to come up with many hypotheses and with many setups
P2	Experiment Design tool	clear, understandable. "It fixes the number of variables, not necessarily electrical variables."
P3	Interaction	because you can work directly
P5	Experiment Design Tool. Electrical circuits virtual lab	
P6	Interactive	
P7	Virtual Lab	
P8	All parts are very good	designed to engage student in inquiry learning
P9	The possibility of having all the lessons and tools in one page	
P10	Experimentation	
P11	Maybe Experiment Design tool	It ensures students don't make some mistakes, like varying two variables at the same time
P12	Consistency with other labs	It helps the user to perform an activity
P13	Experimentation	it is well designed. It could have been more appealing and not so abstract

Q2	worst part of system	Why
P1	Concept Map	confusing. You can't change the description / content of a field
P2	Concept map tool	Implement concepts and delete others obliges restart. To make one ILS ittakes too much time.
P3	(general comment)	"It's difficult to work all phases in 1-2 hours of time for these lesson. I suggest to be more easy to work."
P5	Data Collection	It was not obvious how to use
P6	(general comment)	It was not working properly. Not enough variety of tools.
P7	(general comment)	we are not familiar with the system so as to work with it professionally
P9	The virtual lab	needs to be improved. We spent a lot of time trying to do things that should be easier to do.
P10	Concept map	hard for pupils to use
P11	Concept map	the possibility of UNDO. Not easy to delete a single wrong concept.
P12	New tools - e.g., hypothesis scratchpad, experiment design tool	Not clearly explained. Additional Help is necessary.
P13	Orientation	because it presented in a wrong way the electric current in connection to the electron flow. So change the videos.
P14	general	It's not very intuitive

Q3	suggested changes or extra features
P1	Feature for teachers to see what students are doing or have done.
P1	replace video with teacher demo of lab, so pupils can ask questions
P2	provide an UNDO function with an arrow
P3	"To be more operative; it is so complicate to use in school for a lesson. It is a way to work in a lesson, so much to spend a time."
P7	"when student touches a tool there could be directions for the tool. Or there could be an example to see how tools work first. Like an editorial."
P9	"create shortcuts in the keyboard. Improve the labs".
P10	online help. Dictionary.
P12	"This help system I mention above should work as each similar - with search tools etc"
P13	More instructions and help buttons, especially in the hypothesis phase
P14	A video to explain better how to use the lab

Questionnaire J02 Question 4											
		Str Agree	Agree	Neutral	Disagree	Str Disagree		n=	Agree score	Statement positivity	Positivity score
Q4a	I feel this could be an excellent system	4	10	0	0	0		14	1.29	1	1.29
Q4b	I felt the system lacked an overall design	0	1	3	8	1		13	-0.69	-1	0.69
Q4c	The system could work well in my school	3	9	2	0	0		14	1.07	1	1.07
Q4d	The system needs a lot of work before I would be prepared to use it	3	3	4	4	0		14	0.36	-1	-0.36
Q4e	I hope to be able to use the system in the future	7	7	0	0	0		14	1.50	1	1.50
Q4f	I think that I would need the support of a technical person to be able to use this system	0	4	4	6	0		14	-0.14	-1	0.14
Q4g	I thought the system was easy to use	0	6	4	4	0		14	0.14	1	0.14
Q4h	I thought there was too much inconsistency in the system	0	2	4	7	1		14	-0.50	-1	0.50
Q4i	I imagine that most people would learn to use this system very quickly	0	5	5	4	0		14	0.07	1	0.07
Q4j	I found the system very awkward to use	0	1	7	3	1		12	-0.33	-1	0.33

Qualitative comments (summarised)		
Q4a	I feel this could be an excellent system	
Q4b	I felt the system lacked an overall design	
Q4c	The system could work well in my school	(agree) if we were familiar with it
Q4d	The system needs a lot of work before I	Depends on personal attitude (of user)
Q4e	I hope to be able to use the system in the	
Q4f	I think that I would need the support of a technical person to be able to use this system	Sometimes we need interactive support / perhaps I need guidance for some "errors" / to start with
Q4g	I thought the system was easy to use	
Q4h	I thought there was too much inconsistency	
Q4i	I imagine that most people would learn to	
Q4j	I found the system very awkward to use	

Detailed Questions:

	This tool was easy for me to use	There is a need for such a tool	I would use this tool in its current form	This tool needs to be improved
Viewing youtube video	1.71	1.00	1.50	-0.50
Notepad tool in toolbar	1.50	1.00	0.88	0.00
Concept map tool	0.90	1.00	0.78	0.60

Hypothesis scratchpad	1.44	1.38	1.11	0.00
Experiment Design Tool	0.70	1.33	1.13	0.40
Electric circuit virtual lab	1.63	0.88	1.43	0.40
Data Viewer tool	1.20	0.83	1.17	0.44
Conclusion tool	1.40	0.67	1.14	0.13
Calculator tool	1.75	0.86	1.71	-0.78

Summary of qualitative comments

Note - comments about usability are included. Comments about content, pedagogy, science, classroom mana

Viewing youtube video	
Notepad tool in toolbar	
	It didn't work
	Doesn't work
	It did not work
	Didn't see it.
Concept map tool	
	Don't see difference between concept map and hypothesis scratchpad. Do we need both?
	It needs an "OR" branching function with different routing for True or False evaluations of supplied condition
	(unreadable P8)
	"the students need for adaptation for this tool"
	add a print function
Hypothesis scratchpad	
	provide in other languages
	add an example
Experiment Design Tool	
	provide in other languages
	add an example. Also add help and print functions.
Electric circuit virtual lab	
	Add capacitor and coil components, then can conduct experiements involving magnetic interactions
	provide in other languages because not all students are taught in English
	Provide a more attractive appearance.
	Provide more help in the usage of measurements.
	Didn't have time to use it
Data Viewer tool	
	Didn't get time to try it out carefully
	Couldn't use it (N.B. this user did not answer the quantitative questions)
	No time to analyse it
	Didn't have time to use it
Conclusion tool	
	Didn't get time to try it out carefully
	No time to analyse it
	Didn't have time to use it
Calculator tool	
	students may prefer to use an existing calculator
	perhaps add come buttons like x squared, square root, sine, cosine, reciprocal etc
	haven't used it
	Didn't have time to use it

The usability comments were categoried as P (positive evaluation), N (negative evaluation) or I(improvement sugg

	P	N	I	
Concept map tool	2	7		not able to delete individual boxes, just the whole thing
				difficult to use; students need learning time
				not clear how it works, and how to prepare a concept map
				needs guidelines / help
				not clear what the difference is between "concept" and "select or type" buttons
				the word provided on arrows is translucent and difficult to read
				on "select or type" button, the words provided will influence and bias students
				difficult to remove parts of a concept map without losing the whole thing
hypothesis scratchpad	3	1		not clear where these are stored and how to load them
Conclusion tool	1	2		Couldn't formulate or load a hypothesis and no examples provided
Experiment design tool	1	3		data has to be retyped
				not clear how to start and then proceed
				difficult to remove parts of the circuit
data viewer tool		2		can’t find own data
				after trying simulation, Greek letters and comments appear
Electricity circuit virtal lab	1		1	provide more space for larger circuits?
		1		difficult to run and to prepare the right circuit
				"Run" tab doesn’t seem to work
ECVL demonstration video		2		not clear how to use with Experiment Design;
				not instructive

PDot data analysis

12 sets of feedback received (based on user id). Most of these will represent the views of two teachers.	
115 comments provided, (max 16, min 2), categorised as:	
U - Usability	60
F - Functionality	6
C - Content	37
B - Insubstantial / uncategorsable	8
I - Unreadable	4
(N.B. Where comments included both usability and another category, they have been counted as usability)	

The usability comments were categoried as L (like), D (dislike) or N (neutral)				
	L	D	N	
Logon screen			1	What the nickname is for ? What's the use for it, must it always be the same ?
Orientation	2	4	2	It s good that you have this sentence but maybe it could be organized like that...if student is not reading it through and watching video linke he can not proceed to next stage. For example teacher can put timing to each page.
				what does this symbol mean ?
				why this video is given asa a link and in not embedded as the others ?
				This video should be also embedded inside the ILS and not shown in youtube.
				many "possible" phases that i can choos - not just one "protoype"
				It is better to have the link made on a picture of the video like video 2 and 3
				Its a great tool to use. Its easy and clear.
				Good app that order the process of writing an hypothesis. Thank you
Hypothesis	6	12	1	Sometimes deleting one concept does not work, if you do not target exactly in the center of the dustbin .
				It must have a undo button to avoid erase all the map by mistake
				teacher should be allowed to change the concepts that appear inthe tool in the same blue color.
				the use of this completely open (empty) concept map requires that student have already been trained to use them !
				why the two possibilities to write the hypothesis (scratchpad or free) are not given with the same priority ?
				there is no notification thah i have saved a hypothesis
				it is not easy to delete parts of the concept map and the arrows.
				We didn't easily find the + symbol to add a second hypothesis
				students maybe don't know what concepts they should use...
				students should be guided while giving them some basic concepts as a start (for the weak students)
				"Remove everything?" Hints are not readable!
				Concept map is not explaining itself!
				Its a great tool to use. Its easy and clear.

						Good app that order the process of writing an hypothesis. Thank you
						Good that this concept map appears dierectly not with a weblink
						Good that this is added maybe the text could be somewhere at the menu so all teachers could add the same one. For example a button add definition
						It is difficult
						Hypothesis Scratchpad is very useful! My students will love it ! Next time I will do it this way and use this ILS.
						students should be given an example from daily life
Experimentation - EDT Plan tab						
			6	6	1	put "vary this variable"
						why connot we go back ?
						Why only 5 ? Why not going with multiples of two ?
						we did not understand this window here. what is it for. What to do. Students need more time or orientation to work with this tool. We did not have enough time to explore it
						All these tools have to be translated in other languages
						more information needed to use this tool, maybe another example or video
						it is not clear what the students should measure...
						The "Confirmation page" is not amazing when it's coming up every time you are droping down an object.
						it is a very good idea to put this video here insted of written instructions ...
						Easy to use.
						This video is useful in order to introduce the students to the vitrual lab.
						It is important that students organize their experiment before executing.
						very good to demonstrate the use!
						The students will love it, but I'm sure, they are playing around with it!
						Why pre -digest these information ? Students should be caarefull to their work and fill them
Experimentation - EDT Design tab						
			1	2	0	hard to understand what is to be done
						How am i supposed to change the number of batteries on this lab? This lesson plan was designed in order to use the power source.
						Easy to use.

Experimentation - EDT	0	1	0	Here it is not clear what to do. Should the electric circuit virtual lab be used to run the experiment?
Experimentation - EDT Analyse tab				
	1	4	0	Drag and drop of the icon to the data collector is not possible because the window of the circuit is too big.
				It is not clear that it is necessary to drag and drop the icon to the data collector.
				The hints should be at an earlier stage.
				it is not intuitive to measure with drag & drop of the voltmeter etc.
				The realistic view is very useful.
Interpretation	2	2	0	Response is in Greek!?
				I dont understand how it works. you should add info about it.
				usefull
				works as it should...
Conclusion	1	0	0	OK, to have all the things on one screen and have a look back at the hypotheses
Reflection	1	0	1	Hope that those tools will be saved as pdf or world so later students could save it and present a scientific report
Comunication	0	0	0	

Posters

Participants were asked to take 3 stickers, one red, one yellow, one green. Posters of screenshots were provided. Each participant was asked to place their green sticker on the poster which represented the part of the system which they believed was most usable, the red sticker on the one which was least usable, and the yellow one on one which they had questions about. If time permitted they could write their question on the yellow sticker.

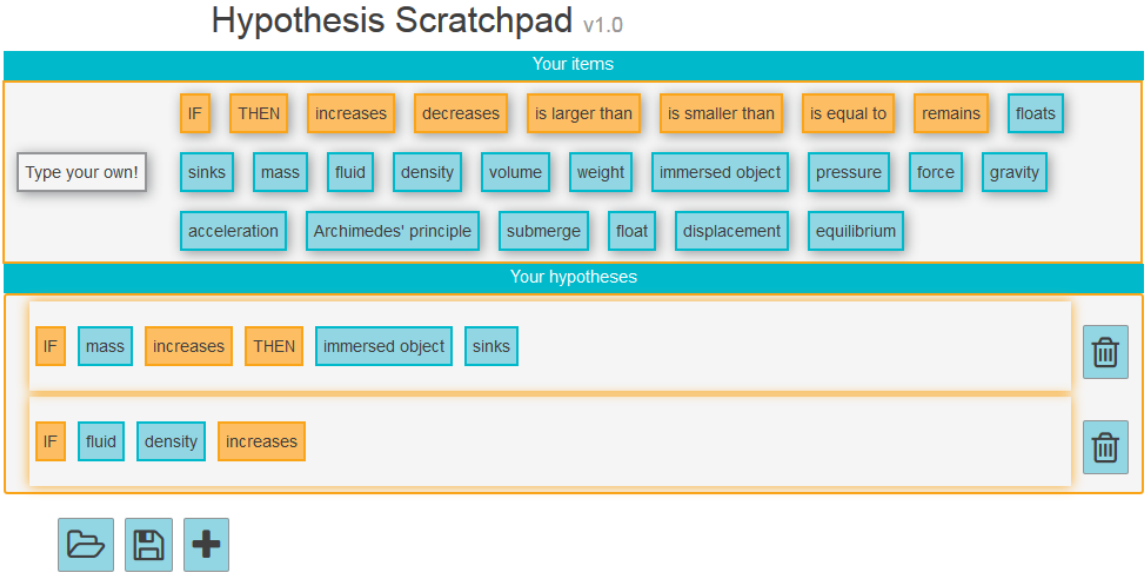
Screenshot description	green	yellow	red	text data
Concept map	4	3	3	(Y): Inquiry cycle. With a lot of driving questions we are cutting their ideas?
calculator tool	0	0	0	
Hypothesis scratchpad	0	3	0	
Experimentation phase, top part, including ECiVL demo video	2	0	0	(G): The purpose of the project and of the page itself is appealing
Experiment Design Tool - Plan tab	4	1	1	(G): "It was really interactive and helped me design my experiment". (Y): Why cannot results be saved from here?
Experiment Design Tool - Design tab	0	1	0	(Y): When are you going to translate into other languages?
Experiment Design tool - analyse tab	0	3	0	(Y): It is an operative system?
Electric Circuit Virtual Lab (top part - circuit design board)	4	0	8	(R): Lot of work (R): The online labs. We need more time to test them. (R): It needs improvement. Many tools are not working properly and not enough variety of tools available.
Electric Circuit Virtual Lab (bottom part - "cicuits" and "data collector")	0	2	2	(Y): Are there going to be shortcuts on the keyboard? (R): Guidance in data interpretation
Data Viewer tool	1	0	1	(R): The graph apps. We were not able to interpret data
Conclusion tool	1	1	1	(G): The relationship between hypothesis, experiment and conclusions. The Inquiry structure was clear.

Appendix D: Hypothesis Scratchpad - Questionnaire and Observation

Place: Twente
Date: March 2014

We like to hear your opinion about the Hypothesis scratchpad)

You can also give your comments by means of text in the picture.



1. What did you like about the HYPOTHESIS SCRATCHPAD?
   - It is easy to use
   - You can add your own words
   - It was simple to use. There were example words which were useful
   - It gives you the right terms to make a hypothesis
   - “Type your own”
   - The variation in things to choose. I also liked how you could type your own.
   - You could easily type your hypothesis there
   - It is clear and easy to use
   - That it gives some ideas to form a hypothesis. It also tells what kind of words you can use
   - It looks really well organised and structured. The saving and loading option works really well
   - Type your own. I also like that the scratchpad was giving you appropriate vocabulary to use

2. What didn’t you like about the HYPOTHESIS SCRATCHPAD?
   - You need more to choose from
   - You can’t remove one word at a time. You have to remove the whole sentence if something is wrong

- It was lots of work to fill in all words. It would be an improvement if you could just type and the words would be on the side to use for example
- There weren't words like "change" or "the"
- Some words which you did need weren't there
- The words there were pretty much useless. You could also just type it on your own
- There weren't enough boxes to choose from
- There were no words like "the" and "a" so you couldn't easily form a proper sentence. It would be much easier if you could just type it.
- The items could be organised into rows and columns
- Not enough langues: Sinks should also have sink. Fluid – fluids
- The way to write down the hypothesis. I didn't like all the blocks but prefer to write yourself

3. Were there certain aspects that didn't seem logical to you or that worked differently than you expected? If so, please describe why you didn't find this logical?
    - The EDT an de hypothesis scratchpad are not perfectly tuned to each other
    - I don't understand where my hypotheses are saved. On the computer or somewhere else?
    - The words. It isn't logical to use loose words instead of just typing your sentences.
    - Deleting words. You could also press the delete button on the keyboard. I would prefer clicking on words to add them to the hypothesis rather than swiping them.
    - If you drag as shown on the arrow (see figure) the box with the word will disappear
    - When you took an item and dragged it towards the bin sign, it disappears
    - Yes, the scratchpad seemed useless. The words can help you but making a sentence with it is hard
    - I expected there to be a clear question
    - I think I will never use the suggested items, it is just way easier to type your own
    .
4. Did you get stuck at some point? If so, how did you resolve this?
    - No
    - You made a typo, and had to start all over again. Solved by starting over.
    - No
    - No
    - No, I didn't
    - No
    - Yes, and I asked a student
    - At first I did not understand what I had to do. I did no know what my hypothesis had to be about.
    - No, it was really clear
    - When I no longer needed a word I clicked on rubbish bin and then the whole hypothesis was deleted; learned quickly

5. What would you change about the HYPOTHESIS SCRATCHPAD if you could?
    - Add the word: the object

- The add own words to sentence
- The name of the site
- Keep the words there but write it out your own
- Add some more boxes to choose from
- That you have to type everything yourself
- Make it into columns

6. Was there something you missed that we should add to the HYPOTHESIS SCRATCHPAD?
   - the word: the object
   - Add button to remove one word.
   - Explanation
   - The word “or”
   - No
   - Add a help button, maybe, to make it more clear to use
   - No
   - No, not really
   - Spelling check would be handy

7. How do you think the HYPOTHESIS SCRATCHPAD looked? Does it look attractive/ugly/simple/boring/… etc.?
   - It looked simple and nice
   - It looks clean, not to much nonsense.
   - Professional, someone with knowledge of computer programs made it
   - It looks a little boring, but it is very clear and doesn’t look messy
   - Simple, but that’s okay☺
   - I liked the way the blue and orange is used to see the difference between composing the sentence, and the items
   - Simple
   - It looks normal and simple, but not ugly
   - It looked simple but not boring
   - Quite attractive really. Simple as well
   - Simple

8. Is it easier to create a hypothesis using the hypothesis tool than without? Please explain why or why not.
   - It is easier because you can see the possibilities
   - It is because you have all “difficult” words, you don’t have to think them up yourself.
   - Without. It was hard and took longer to make normal sentences
   - Yes, it helps you with the words to use, but I prefer typing it with Word than using the tool.
   - Without, because some words which you did need weren’t there
   - I think it is easier because the items are already given

- The words might help you but it is easier to write it on your own because not all the words are there
- It is both easy, because with the tool you know what is expected of you
- No, it doesn't really matter
- No, I did not experience that
- In some ways it was easier

9. The next time you have to create a hypothesis, would you like to use the hypothesis tool or not? Please explain why or why not.

    - Yes, because it is easy to use.
    - Maybe only if it needs a lot of words. I don't know
    - No because it was hard and took longer to make normal sentences
    - I would use a word list, but type it myself because I think it takes too much time.
    - No, because some words which you did need weren't there
    - I would, because it's very easy to do
    - No, I can write it on my own
    - Depends on what the question is.
    - I would not. Because I can write myself and it doesn't make other things easier.
    - No, it takes up too much time
    - Not really, it's not about every type of science

Field Note observations

With regard to the Hypothesis Scratchpad, quite a few students used the "Type your own!"-element to enter words or even complete hypotheses. They did so because they felt the "pre-fab"-elements were too restrictive and/or did not allow them to form grammatically correct sentences. Those students did consider the "pre-fab"-elements helpful though. Looking at the "pre-fab"-elements helped them to think about the structure and contents of their own hypotheses, and they used the "pre-fab"-elements as a kind of "inspiration" for formulating their own hypotheses. Their own hypotheses were then entered into the tool by using the "Type your own!"-element. One of the teachers that was present during the session remarked that the "pre-fab"-elements triggered formulating hypotheses such as: "if density increases then…". She had seen one or more examples of that and remarked that density normally is not increased, but objects with different densities are compared. Perhaps, adding some more terms expanding the range of experimental possibilities is worthwhile considering.

The Experiment Design Tool (EDT) seemed to work fine. Perhaps unnecessarily, because of future connections between tools and labs, but it would be helpful if students could see (or even play around with) Splash before/while they use the EDT. Without having seen Splash, making the experiment design was quite abstract for the students. Having already an image, some idea about Splash, could make experimentation more concrete for them. That could make it easier for them to think about what they can do in the experimentation phase.

With regard to Splash, all students seemed to find it pretty easy to operate the lab and to find their ways through it. An issue that definitely needs to be solved however is how values can be entered. The sliders are coarse-grained: it is nearly impossible to enter the exact value you want to observe

by using a slider. As a result, most student use the keyboard of their computer to enter values in the boxes next to the sliders, but this is problematic because these values are not entered into the model unless the user types a semi-colon symbol after entering the value. This is not intuitive and gives rise to confusion and frustration. This issue definitely needs to be solved! A reverse slider/box enter problem can be observed with regard to fluid density. By using the slider, fluids with different densities can be selected (e.g., water, sea water, acetone). However, students try to use their keyboard to enter values into the box next to the fluid density slider (which obviously does not work). The slider suggests that there is a continuum of different fluid densities, but in fact, this is a confusing representation. In this case the slider is actually used to select qualitatively different fluids with different, but discrete, densities. I think it would be more clear and straightforward for the users if they were presented with a list of different fluids and their densities, from which they can choose (e.g., by clicking on a radio button or using a pull-down menu). The tables and graphs do not seem to be used much, if at all. The students that I spoke with, felt they did not need the table. They tested their hypotheses on the fly, just by observing what happened in the tubes. Of course, the table will be helpful, especially with more complex designs or with interaction effects, but it is worthwhile to notice that students tend to ignore the table. Last remark, the students that I asked about using the graphs, indicated that they did not understand the information the graphs provided. That is not really surprising, because this is a general problem with graphs. So, this seems to be more of a pedagogical problem than a design problem. Adding tubes, removing or moving tubes and running experiments did not pose any problems for the students.

All in all, the students really seemed to enjoy working with the tools and the lab.

Appendix E - Development Team responses to General Findings

Development Team Leaders have provided the following responses to the findings in 3.3.1 above.

Wish for facilities to monitor students' work and support assessment

Work is in progress on teacher dashboard tool to monitor student progress. Mock-ups have been produced and feedback received.

Sign-on security

A logout feature is implemented in the Month 24 release.

Storage and retrieval of work

ILS information added by students, such as text in input boxes, is saved as it is typed. The developers plan to have such a mechanism for all Go-Lab tools. (Externally provided labs may not provide this).
Student reports and other files should be saved in the vault. The Drop-file app provides one way to do this already; other such apps will follow.

Facilities to support group working

An app to support team working is under consideration. Basic infrastructure design is not provide advanced collaborative support.

Undo and Redo, Retyping

This is not trivial but developers will do when possible.

Scaffold toolbar visibility

Will work on app resizing feature.

Consistent interaction design

Hope to converge on a single drag-and-drop interaction. Will also try to apply similar aesthetic look and feel (though externally provided components – e.g. labs – cannot be forced to conform).

Help features

Can converge to a common design principle. As before, externally written components may be different.

Distinguishable interactive and non-interactive components

Issues concerned with distinguishing hyperlinks from other text will be / are resolved.

Multi-lingual components

New app composer feature allows teacher to translate apps. ILSs can be shared in different languages. (The language of the User Interface changes, not the content).

Usable scrolling

Agreed. We now resize apps to full size whenever possible. The user does not need to resize anything.

Browser issues

Working on it.

Appendix F – Development Team Detailed Responses to Some Specific Usability Findings

1.1 Scaffold Apps

- Hypothesis Scratchpad (HS)

Development Team responses:
HS1: Feature is already available through AppComposer Adaptor. Its usability will need to be evaluated. Needs to be integrated more comfortable into authoring. Work on refined authoring and configuration mechanisms in progress.
HS2: Single words can be deleted by dragging them over the bin icon - possible this wasn't intuitive enough. If the Style Guide - interaction proposes a different method, it can be implemented.
HS3: Already fixed.
HS4: Undo/redo is generally not a trivial feature, but since "model classes" have been introduced recently, there is a chance to introduce this feature.
HS5: Sorting of words should not change dynamically (i.e. by most recently used). Alphabetic sorting can be done through editing the list of words - see HS1. Storing a "lexicon" of own words can be realised when refined authoring and configuration mechanisms are in place, also see HS1.
HS6: Has been improved already, see also HS2.
HS7: Has been improved already, see also HS6 and HS2.
HS8: New, common load/save dialog has been introduced. Confirmations can be added there.
HS9: If a spell checker is requested in this tool, for consistency reasons we would one in all tools. Do we want this (as it is not a trivial feature…)?
HS10: Already fixed.
HS11: This is a resizing bug in the ILS, not in the HS.
H12: When Style Guide propose consistent interaction, it can be implemented.
H13: Okay - it is easy to change to one colour only.
H14: Would be fixed by HS4.
H17: Already solved by new common load/save dialog - see also HS8.
H18: Fixed already.
H19: Solved together with H13 -> one colour only.

- Experiment Design Tool (EDT)

Development Team responses EDT1: The EDT does not contain explanatory material yet. However, as pointed out by the reviewers users need to be able to retrieve guidance when necessary. This will be resolved in future versions. The presentation of the guidance needs be aligned across tools, and decisions will be made by the team on how to provide users with useful guidance on how to use tools.
EDT2: In the EDT users need to define all variables, because they need to understand to keep all variables the same except for the one variable they want to measure. This issue will be addressed by better communicating this idea to the user.
EDT3: The use of the padlock, as well as its functionality, indeed is not clear in the current version of **Splash**. The padlock is meant to function for interacting variables. For instance, density is comprised of mass and volume. If one variable changes, another automatically changes as well.

The padlock is placed at the variable that will not change if another variable is being manipulated. This functionality will be communicated to the user.
EDT4: The run tab will in the future work together with labs that can connect to the EDT. If users press the run-button, the experiment will be conducted in the lab. Students will need to fill out the results by hand in order for them to be actively involved in the results of the experiments.
EDT5: In future versions of the EDT restrictions between tabs will be communicated to the user. It will be made clear that it is not possible to go back, and why they cannot go back.
EDT6: The messages in the Plan-tab of the EDT will be further discussed and evaluated to find a solution in order to still provide users with the information without it causing frustration. The suggestion of the tick-box will be taken into consideration, as well as other means that do not disrupt the task.
EDT7: The values for the design-tab come from the lab that is connected to the EDT. In future vers ...input lab, not just EDT...
EDT8 – EDT20:

- Concept Mapper (CM)

Development Team responses:
CM1-4:
If usability and consistency is a problem, I'd propose to go for a style that's also common in other applications, e.g. Word or Powerpoint: You click an item in the toolbar (a concept or an edge), then you click to create, or drag to connect. Once done, the mode switches back to "move objects". Deleting objects would be a third mode, then.
An alternative way would be to build on dragging only, but this might be inconsistent with other tools.
As a developer of a tool, you often don't see those problems, so I'd be glad to see that covered in a Style Guide.
CM5: See comment HS4.
CM6: It's already there, but apparently not intuitive enough, see CM1-4.
CM7: Fixed already.
CM8: See HS1.
CM9: See CM1.
CM10: General request too all tools, I guess. Will be included.
CM11: List of concepts can be adapted, see CM8 and HS1.
CM12: Can be added.
CM13: Can be added.
CM14: Colours can be changed already, together with CM13 font size also.
CM15: A cross-button to delete can be added.
CM16: Agree.
CM17: [missing]
CM18: The order can be specified during authoring (see also HS1), a teacher can decide the order by himself.
CM19: The title of the saved concept map can be added in the title section.
CM20: Already fixed.
CM21: Not sure if we want this, see HS9.
C22: Can be fixed easily.

CM23: This request is a bit underspecified: Where to take the pictures from? Copy/pasting an external URL to a picture would work, but it cannot be guaranteed that this URL remains valid.
CM24: UNDO/REDO has been covered above. What does import do? What does the style guide say about Print and E-Mail?

- Data Viewer Tool (DVT)

Development Team response:
At the moment scales can be added through the edit graph button. But the known information (such as X-as label and the units) will be placed automatically. We are still working on a persistent storage of resources. At the time of the evaluation, we had only had a very limited version of the persistent storage of the resources.

- Questioning Tool / Questioning Scratchpad (QS)

Development Team responses:
QS1: Can be fixed.
QS2: Pre-defined terms can be authored through AppComposer, see comments for HS and CM.
QS3: ?

- Conclusion Tool (CT)

Development Team response:
How and where we will have display the online help is a point of discussion.

- Quiz Master Tool

Development Team response:

- Calculator

Development Team response:

- YouTube Widget (YTW)

Development Team response:
We have made it easier to insert videos in ILS and the video shows in a nice wide size, no need to be resized.

- NotePad / Notes tool (NP)

Development Team responses:
Agreed we need to make it clear what they do and what is the difference. typically collaborative vs personal. rich text, vs plain text.
NP1:
NP2: the size takes the whole width, the height is fixed for now
NP3: Done
NP4: to be discussed, we do not want to force passwords, we can think of a "welcome back message"
NP5: Done

1.2 Go-Lab Owned Labs

- Splash Lab

Development Team responses
SL1: See SL2.
SL2: Terminating edited values with the ENTER key is now possible.
SL3: See IMC YouTube movie.
SL4: The lock cannot be moved, clicking on one of the properties (mass, volume, density) sets the lock.
SL5: See SL6.
SL6:
(May have used an old version for evaluation). Graph is only visible to illustrate density as a relation between mass and volume. History table is indeed a little complex, could be made optional in a configuration.
SL7: Should be solved in latest version.
SL8:
Density is “semi-fixed” because it is also used as a colour of the liquid. Pull-down menu might work in some educational settings, but not in all. density = mass / volume is one of things students have to learn, and then changing mass or volume is a continuous value for density.
SL9: Depends on EDT.
SL10: There is not really enough space for units.

- Electricity Lab

Development Team responses:
EL1:
How to present online help is still a point of discussion. This could be placed in the tool or lab, or in the ILS. The circuit simulator is a bit special case, as it is a lab where the experiment has to be created first. One could say the student has to create the “real” lab first. A complete online help for the circuit simulator would only have global help and not address the support of the specific circuit the student is going to use to do the experiments.
We are working a persistent resource storage (a circuit is a resource). Such a storage can also also be used to store example circuits.
EL2:
We know the using of the sensors on the meter is not very. We are looking the improve this.
EL3 – EL11:
EL3: The size of the circuit board can be configured. But the maximum size of the lab is limited by the screen and ILS implementation. WE are searching a more efficient usage of the limited screen resolution.
EL4: fixed.
EL8: An earlier version used the common + and – minus buttons. But that implementation did not under android. A fix for this has be found and in the nearby future, the slider will be replaced by the common + and – minus buttons. The values can be types in. When adding the + and – buttons, we will try to make the UI more clear.
EL9: Hint will be placed in a popup window, to get more screen space for the other components
EL10: We are considering to make the circuit simulator also work for AC circuits using sinus shaped voltages. This would also allow the usage of coils and capacitors. But this means extending the circuit simulation engine, unfortunately this cost a significant amount of time.
EL11: There is a version of the circuit simulator, where data can be collected. This data can then be visualised in the data viewer. The designers of the ILS decided not to use that.

1.3 Go-Lab Designed ILS

- Splash Lab: Sinking and Floating ILS

Development Team responses:
cSSF1 – cSSF2:
SSF1:
SSF2: to be discussed
SSF3 – SSF9:

- Electricity Lab: Ohm's Law ILS

Development Team responses:
cEOL1: Difficult, if we have custom phases
cEOL2 – cEOL13:
EOL1 – EOL9:

1.4 ILS Authoring Toolset

Development Team responses:
General:
(Picture uploading complexity): pictures are now drag and dropped in the corresponding space.
("the text editor has too many icons, which represent some functionality not needed by the teachers": we have a simple 4 button description (italic bold link bigger).
AT1: will do for M24
AT2:

- we use app instead of widget
- we have a "write a description here" placeholder. to be discussed if further rewording is needed
- the teacher notes should be put in the about space which will have a crossed eye icon to show that it is hidden
- we have a big "standalone view" button in the share tab

AT4: done
AT5: done, no need to resize items.
AT6: done, drag and drop or upload image.
AT7: we are working on it (compatibility with contemporary browsers).
AT8: we have moved contextual function entities in a context component on the side with tabs, the "pad" is now central with little clutter.
AT9: done change name by just clicking the name
AT10: done, simple click on the text to edit it, this user experience is the same throughout the platform
AT11: we still have some hover over tooltips on space items as well as in the description, to avoid clutter when not used
AT12: removed, now we have a plus button to add things and d&d works everywhere
AT13: done, save as u type
AT14: the description is now completely new
AT15: same as above, new description
AT16: Done