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 D3.3

Formative usability and user experience evaluation report on the initial Go-Lab portal

Editor:	Effie Law
Date:	October 29 2015
Dissemination Level:	Public
Status	Final

The Go-Lab Consortium

Beneficiary Number	Beneficiary name	Beneficiary short name	Country
1	University Twente	UT	The Netherlands
2	Ellinogermaniki Agogi Scholi Panagea Savva AE	EA	Greece
3	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 South Wales	USW	United Kingdom
18	Institute of Accelerating Systems and Applications	IASA	Greece
19	Núcleo Interactivo de Astronomia	NUCLIO	Portugal
20	Cardiff University	CU	Wales

Contributors

Name	Institution
Effie Law, Rob Edlin-White, Matthias Heintz	ULEIC
Mario Mäeots, Leo Siiman, Margus Pedaste	UTE
Tasos Hovardas, Nikoletta Xenofontos, Zacharias Zacharia	UCY
Ton de Jong, Henny Leemkuil, Siswa van Riesen, Bas Kollöffel, Ellen Wassinkamp	UT
Evita Tasiopoulou	EUN
Eleftheria Tsourlidaki, Sofoklis Sotiriou	EA
Javier Garcia-Zubia, Olga Dziabenko	UDEUSTO
Denis Gillet, Adrian Holzer, Maria Jesus Rodriguez Triana	EPFL
Sven Manske	UDE
Rosa Doran	NUCLIO
Internal reviewers	
Diana Dikke	IMC
Yiwei Cao	IMC

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 system as highly beneficial and usable platform for use in science teaching. 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 3, WP3 is responsible to collect data on usability and UX 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 on 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 students or teachers 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) **Remote studies** involving 21 "Core Group" teachers from nine countries, 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) **Analytical Evaluations** conducted by usability researchers without involving end-users, to investigate a few artefacts; the results have directly been fed to the development team for improving the prototype before testing it with users.

From November 2014 to October 2015, the work package has designed, conducted, analysed and reported 46 face-to-face PD events with teachers and/or students (and has received, analysed and reported on usability data from 7 others), 13 remote events involving Core Group teachers and 4 analytical studies involving usability specialists. Altogether these PD activities involved over 550 students and over 450 teachers from 18 countries. The face-to-face PD events took place at schools, research/training centres, conference venues, 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, together with some performance measures such as timings, number of errors, etc.

Overall, the participating teachers and students are excited, enthused and positive about the potential benefits of Go-Lab, to a significantly greater extent than in Year 2. However many possible areas for improvement of usability or user experience have also been reported. The findings from the many studies have provided a steady flow of information back to the project leadership and thereby influenced system development direction and priorities. This document reports the year's findings, particularly areas of possible improvement. For all but the most recent studies, the findings have been communicated to the development cluster (WP4 & WP5), and their responses and progress are also documented herein.

Table of Contents

1. Introduction

1.1 Overall objectives of WP3 in Go-Lab

The purpose of WP3 in Go-Lab is to work towards making the system as usable and beneficial as possible to as broad a cross section as possible of real teachers and real students in real learning situations across Europe.

This is vital for the success of Go-Lab, since we want 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 find they can use it in a way which is effective, enjoyable, satisfying, efficient and beneficial to teaching and learning.

We therefore engage representative potential users of the final system – both teachers and students – in order to understand their objectives and requirements for such a system, to engage them in the design process and provide their own creative input, to provide them with access to prototypes of parts of or the whole system in increasing levels of fidelity in order to elicit their feedback on usefulness, enjoyability, intuitiveness, learnability, memorability, effectiveness, and other pragmatic as well as hedonic qualities, and finally to provide summative evaluation of the final system. Thus there will be user involvement throughout the project, feeding into a cycle of continuous refinement of objectives, requirements and design.

In Year 1, the main emphasis was on establishing user requirements. In Year 2 and Year 3 the emphasis shifts towards formative evaluation: testing prototypes in increasing levels of fidelity, to refine the requirements and the emerging system. In Year 4 the main emphasis will be on summative evaluation: assessing the usability and user experience and community engagement with the final system, conducted within WP8.

As argued at Section 2.3 below**,** many of these objectives are best achieved by engaging with real users in their workplace – teachers and students in schools.

The main outputs of WP3 are recommendations for software changes to extend or improve usability, usefulness or user experience. These are shared with the relevant technical cluster partners, and a response is agreed. WP3 studies have also identified occasional recommendations for software changes for other purposes (e.g. bugs) which are shared with the technical cluster; recommendations for selection criteria for labs which are passed on to WP2, and recommendations for ILS authors which are passed on to WP1. Go-Lab aims to implement the project's goals at a large scale in Europe. Fifteen countries have been selected for the main focus of the project this year. Stakeholders from these countries and others have been involved in a range of activities in WP3.

1.2 WP3 impact since D3.2

During Year 2 a wide range of end-user studies in usability, user experience and usefulness were conducted for WP3, communicated to development teams throughout the year, and finally summarised in deliverable D3.2. D3.2 provided very extensive and detailed tables of usability issues and recommended responses, and (in sections 3.3.1, 6.2 & 6.3) identified some major themes and general recommendations. D3.2 identified a large number of software changes which had been accomplished during Year 2 to mitigate the identified usability problems. Since D3.2 very considerable progress has been made on addressing the outstanding issues. The major themes are reviewed and progress noted at Section 5.2.2

below, along with further examples of exemplary development responses to usability findings during Year 3.

1.3 Overview of WP3 in Year 3

The overarching goal of WP3 in Year 3 is to evaluate usability and user experience of the various Go-Lab design artefacts, providing timely and clear feedback to development and pedagogical clusters to facilitate and encourage developments which enhance usability, usefulness and user experience. In Year 3 the main work is dedicated to Task T3.4 (Evaluating usability, user experience and educative experience) though tasks T3.1, T3.2 and T3.3 continue for artefacts which are new or newly changed. The artefacts under evaluation include small, self-contained, well-defined and relatively stable scaffold apps; more complex labs some of which are in-house and some provided by third parties; whole ready-to-deliver lessons (ILSs) which the authors can edit quite freely and which combine a selection of these and other online materials, and infrastructure elements such as the portal itself, the authoring tool, help facilities and community platform etc.

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 in order to conduct the most appropriate studies. Additionally WP3 has been collaborating closely with WP6 for community building, with WP7 concerning the recruitment of teachers/schools and with WP8 to scope the focus of the respective evaluation and validation activities.

The main outputs from WP3 are recommended software changes provided to the technical cluster (WP4 & WP5) throughout the year. Other outputs are ILS design recommendations for teachers, provided to WP1 and recommendations for Lab selection criteria provided to WP2**.** In addition, in Year 3 some more far reaching change recommendations are provided to WP9 as input to the sustainability agenda. This is partly because Year 4 is mainly concerned with evaluation, dissemination and community building, not software changes.

Compared to Year 2, some of the main differences are:

- The Go-Lab product and the range of resources have become and are becoming ever more comprehensive, mature, reliable, usable and useful, and are therefore more suitable for classroom use.
- The balance of participants in Year 2 was weighted towards teachers, whereas in Year 3 we are engaging more students.
- Studies and study methods in Year 2 were often quite exploratory and open to new requirements; in Year 3 we have moved towards more evaluative focus (e.g. usability

studies) for components which are well established, and using exploratory usability methods only for newer components or ideas.

- The portfolio of labs, apps and ILSs had increased dramatically (from around 80 a year ago to 356 at the time of writing), and some of the content has been quite fluid as ILSs can in principle be written, modified and deleted by any teacher, and external labs can be changed by their owners at any time. Therefore the WP has focussed on the most frequently used, and on delivering general as well as specific recommendations.

1.4 Resources and opportunities

Seven partners (UT, EA, EPFL, EUN, ULEIC, UCY and UTE) are formally engaged in this work package, and several others have contributed to varying extents. Different partners have had different amounts of time to dedicate to WP3 this year. ULEIC as Work Package leaders have a role in coordinating and of facilitating the work, and reporting on it. All WP3 partners work to engage teachers, students and schools, conduct evaluations and report them as well as contributing to wider aspects of WP3. Recruitment of schools is very 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. We have used field-based research with teachers and students wherever possible, supplemented by judicious use of analytical evaluations by Human-Computer Interaction specialists and remote evaluations by our Core Group of Teachers where appropriate.

1.5 Structure of D3.3

Section 2 presents the conceptual and methodological frameworks underpinning the WP3's activities. We articulate the role of UCD in the project. We describe three highly interrelated notions in the field of Human-computer Interaction (HCI): Usability, Accessibility and User Experience, and the methodological approaches of User-Centred Design (UCD) and Participatory Design (PD). It also describes at a practical level how the work has been managed, structured, organised, conducted, analysed and disseminated to study a wide range of Go-Lab artefacts at locations around Europe. It discusses the strengths and limitations of the three major types of PD approach employed for Year 3 work, namely the face-to-face PD approach, remote PD approach, and researcher-based analytic evaluation approach. It also outlines the policies and practices for targeted recruitment of schools, teachers and students, and finishes by describing how study protocols and materials were devised.

Section 3 outlines the study programme, listing all the studies – face to face, analytical or remote – which have taken place in Year 3, including details on number of participants, areas studied, duration, study methods, types of data collected etc. It provides information and profiling data on the participants (teachers, students) engaged. It describes how study findings were designed, conducted, analysed and reported.

Section 4 provides consolidated findings from all the WP3 studies in Year 3. The empirical data – particularly the data relating to areas of possible improvement - are analysed and aggregated, and then presented in the form requested by the development cluster – i.e. by design artefact. It provides user feedback and constructive design recommendations on a

wide range of infrastructure components, online labs, scaffold apps and 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. This section also includes a little of the considerable body of evaluative (quantitative) data collected

Section 5 provides a synopsis of the main findings of the PD studies in Year 3. It then 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, and the response of the development team to Year 3 Usability Findings and Recommendations, including work in progress and work planned. It also reviews the main Usability Recommendations from Year 2 and the considerable progress in these areas. Finally it reviews the relative effectiveness of the different methods of study deployed.

Section 6 concludes the deliverable by reflecting on the achievements and limitations of this work in Year 3, and the outlook for Year 4.

The **Appendices** provide a variety of supplementary and supporting material, including 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 Frameworks and Research Methodologies for Year 3

The methodologies and research paradigms for engaging users in this work have evolved since year 2, as predicted in D3.2. This chapter therefore discusses and provides a rationale for the various methods and systematic approaches we have deployed in Year 3. Some of this material was implicit or briefly described in D3.2 but is now more fully presented. We begin by stating explicitly the shared understandings in WP3 of usability, user experience and accessibility – the qualities we set out to measure and enhance. We then present the rationale for involving users in this process, and explain how user centred design (UCD) has been used in Go-Lab in Year 3. We describe Participatory Design (PD), and mention some related research methodologies, and show how PD and usability studies have been deployed in Year 3, and the strengths and limitations of the three approaches to user engagement which we have deployed, and how we recruited participants and utilised research opportunities. The chapter concludes with some detail about how we practically and organisationally conducted the work.

2.1 An overview of Human-Computer Interaction concepts in WP3

In this section we present several basic concepts in the field of Human-Computer Interaction (HCI) that are highly relevant to the work of WP3.

2.1.1 User-centred Design (UCD)

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, leading to a system with high levels of usefulness and usability and thus high user acceptance, adoption and ongoing engagement. Since its inception in the 1980s, UCD has been regarded as a cornerstone of the field of Human-Computer Interaction (HCI).

Systems which have low levels of user involvement can often fail to achieve their potential, in particular through poor usability. Don Norman (1998) and later Alan Cooper (2004) lamented that much technology – especially leading edge technology – tends to be designed by people with an enthusiasm for technology *per se*, and they tend to produce hugely versatile products and systems which present bewilderingly many options to the user, exploiting lots of exciting novel functionality. Such designers tend to produce applications (of technology) rather than solutions (to user problems), to be less concerned with ease of use, and to have limited awareness of the requirements, concerns, lifestyles and expectations of people who are different to themselves (Norman, 1998; Cooper, 2004).

Cultivating a rapport with user communities, through conducting field studies of the users, their tasks and environment early in the design lifecycle, or better still by including users in the design process, can help designers bridge the chasm, and understand the world, work and requirements of users, which can thereby lead to more accessible and effective products (Lindgaard et al, 2006; Muller, 2002). The value of this user engagement lies not only in the direct outcomes, but also in the process itself, in influencing the analyst and informing subsequent phases.

User involvement in User Centred Design (UCD) typically includes some or all of (Rogers, Preece, & Sharp 2007; Iivari & Iivari, 2011; Earthy et al, 2001):

- being consulted on their requirements or expectations,
- contributing to the creative work of designing solutions
- being engaged in formative evaluations of mock-ups and prototypes in gradually increasing levels of fidelity, and
- being engaged in summative evaluations of the finished product.

HCI literature provides many models of how this engagement between technical and end user communities can be conducted. Some are explored briefly below, and the approach adopted for WP3 is described and justified at Section 2.5 below.

2.1.2 Usability

Traditional usability is defined formally as: "*The extent to which a product can be used by specified users to achieve specified goals with effectiveness, efficiency and satisfaction in a specified context of use*." (ISO 9241-11:1998)[1]. In the context of HCI this could mean the ability to comprehend displays and use controls efficiently, fluently, effectively without errors, and without dissatisfaction.

Usability therefore depends not only on a product's user interface but also on:

- the capabilities and preferences and motivations and interests and energy levels of the users – including their experience and capability with other user interfaces;
- the environmental and social context in which users are using the system, including the availability of formal or informal sources of help and support;
- the tasks which users are attempting, together with related issues such as time pressure; and
- the other tools and aids in use for the tasks.

A gold standard in usability is to design an interface so that it is quickly apparent to even first time users how to use the system; this is sometimes described as intuitiveness. Jordan (1998) describes this as "Guessability", and identifies four other components of usability: Learnability (or Discoverability), Experienced User Performance, System Potential and Reusability (the latter relating to occasional use).

Because different users have different usability and accessibility needs, it may not be possible to create an "ideal" system which is perfect or completely intuitive for every possible user. To mitigate this problem, it is beneficial to provide support materials and structures to help users attain Experienced User Performance.

A specifically area of variability affecting Go-Lab is the different educational climate and paradigms across countries within Europe, with varied styles of teacher training, school and class sizes and organisation, varied curricula, varied levels of pressure from parents, head teachers, school inspectors, examination boards etc.

2.1.3 Accessibility

The notion of "accessibility" is similar to usability, but concerns the ability of users with very specific capabilities and needs to use a system. Accessibility is defined in ISO9241-171[2] by reference to usability as follows: "*usability … by people with the widest range of capabilities*"

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

[2] ISO9241-171: Ergonomics of human-system interaction -- Part 171: Guidance on software accessibility.

and in ISO TR22411[3] as: ''*design focused on principles of extending standard design to people with some type of performance limitation to maximize the number of potential customers who can readily use a product, building or service''*. The notion of accessibility reminds us that the design should be usable by a wide variety of teachers and students, not just by the most capable.

2.1.4 User Experience

User Experience (UX) has increasingly become recognised as a broader and more pertinent notion than usability. UX, broadly speaking, descends from the traditional UCD framework (e.g., Gould & Lewis, 1985; Norman & Draper, 1986), focusing on the experiential aspect of human-computer 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 & Tractinsky, 2006; Bargas-Avila & Hornbaek, 2011; Law et al., 2009; Vermeeren et al., 2010; Harrison et al, 2011). In the meantime, the "dated" notion of usability has been replaced by the then emergent UX, causing some confusion in the scope of research and practice, including job titles. Despite attempts to demarcate usability and UX (e.g., 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).

Table 1. Comparing usability with user experience

Usability	User Experience (UX)
Pragmatic quality	Hedonic quality
Do-goal (to find an e-book)	Be-goal (to feel competent)
Product: performance, task	Experience: emotion, affect
Reductionist	Holistic
Partly objective: quantitative measures	Highly subjective: qualitative narratives
Relatively persistent	Inherently dynamic
Standard usability metrics are widely used: *Efficiency, Effectiveness, Satisfaction*	Standard UX metrics are yet to develop; usability criteria are included[4]

With reference to the related literature (e.g., Hassenzahl, 2008; McCarthy & Wright, 2004; Norman, 2004), we list the major difference between usability and UX (Table 1). While the definition and metrics of usability are widely adopted in the HCI community (ISO 9241-11:1998[5]), UX is still plagued with definitional and measurement issues (Law & van Schaik,

[3] ISO TR22411: Ergonomics data and guidelines for the application of ISO/IEC Guide 71 to products and services to address the needs of older persons and persons with disabilities.

[4] ISO 9241-210 (2010): Note 3: Usability, when interpreted from the perspective of the user's personal goals, can include the kind of perceptual and emotional aspects typically associated with user experience. Usability criteria can be used to assess aspects of user experience.

[5] ISO 9241-11 (1998): Ergonomic requirements for office work with visual display terminals (VDTs) -- Part 11: Guidance on usability.

2010). Despite its crudeness, the standards definition, ISO 9241-210: 2010[6], which comes with three notes, is often referenced: *A person's perceptions and responses resulting from the use and/or anticipated use of a product, system or service.*

With the focus on users' 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 measure (Law et al. 2014). Furthermore, the UX evaluation methods (UXEM)[7] are largely drawn from the traditional usability evaluation methods (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). Nonetheless, both UEM and UXEM are applied in evaluating a variety of interactive systems, ranging from conventional work-oriented software to emerging mobile games.

2.2 Evaluation studies in Year 3

In Year 3, as in Year 2, the objective of the WP3 evaluation work 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, usefulness and user experience, to enable the development of a highly usable and attractive system.

Each study therefore seeks 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 using mock-ups.

The studies, individually and as a programme, aim to deliver project recommendations: e.g.

- to improve a component in specific ways
- to drop an aspect of functionality or even a component altogether
- to add new functionality or a new component
- to conduct further studies

These recommendations can – when taken up - affect development priorities and workloads in other WPs beyond the remit of WP3. They are therefore presented as recommendations to the project leadership and not directions to other WPs, though copies for information are usually sent to the relevant technical or pedagogical clusters.

The studies have mainly been conducted within the broad paradigm of UCD (Section 2.1) and the specific methodology of Participatory Design (Section 2.4).

2.3 Research methodologies

A huge variety of different approaches to HCI research exist. Martin and Hanington (2012) list 100 methods, though some are overarching methodologies and some are methods used

[6] ISO 9241-210 (2010): Ergonomics of human-system interaction -- Part 210: Human-centred design for interactive systems.

[7] http://www.allaboutux.org/

within them. Based on an extensive literature analysis, Iivari & Iivari (2011) began to map out this "jungle of user-centred design" (p. 127) and identified four clear dimensions of user-centredness in design:

- **User focus** – this dimension stresses that the design should take into account every user's capabilities and needs in relation to the system. Sometimes this is based on guidelines on what constitutes a usable technology, or on established knowledge about user capabilities (e.g. from cognitive psychology), or on average or exemplar or fictitious users.
- **Work centredness** – this dimension recognises that users (especially in workplaces) are acting in a particular organisational and social context, and that there are complex interactions between technology change, change in work demands, organisational change and changes in work performance.
- **User involvement in design** – this dimension (considered integral by most) originated in workplace systems, and involves workers (or sometimes others in surrogate or representative roles) in the design process as part of democratic or functional empowerment. Involvement can range from the superficial (user is subject of study) to the fully equal design partner (empowered participant).
- **Personalisation** – this dimension attempts to cater for the wide variability of users by allowing for various sorts of personalisation or adaptation or customisation of the system, either by user intervention or sometimes by the system automatically adapting to the user's characteristics and preferences, based on a user model (Iivari & Iivari, 2011).

Dimensions of user-centredness are not the only distinctions between UCD methodologies. Pathirage et al (2008) use a different categorisation (Figure 1), portraying a methodological continuum, positioning five indicative methodologies on a scale. Methodologies towards the left of this scale are more nomothetic, highly structured, and emphasise quantitative data and a deductive approach to research (i.e. starting with theory and creating data to support or challenge it), while methodologies towards the right are more ideographic, less structured, and emphasise qualitative data and an inductive approach to research (i.e. deriving theory from data).

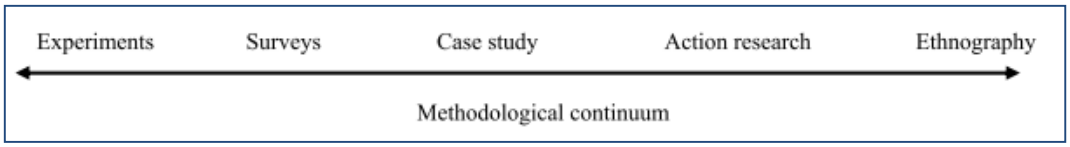


Figure 1. Methodological continuum (from Pathirage et al, 2008).

In the last decade or more, for a variety of reasons, the emphasis in HCI work has shifted from the more empirical to more social methods (Bannon, 2011).

2.4 Participatory Design (PD)

The predominant model of user engagement adopted for Go-Lab is Participatory Design (e.g. Muller & Druin, 2010, 2012; Simonsen & Robertson, 2012. In subsequent sections, we first describe some basic concepts of PD and then report how we have adapted and applied those concepts in WP3 evaluation activities.

2.4.1 Basic concepts

According to Muller (2002), the fundamental challenge of HCI methods is to bridge the chasm between the world of the software professional and the world of the user, each world having its own culture, language, space, and assumptions. PD involves respect for all users, recognition of workers as a source of innovation, a holistic view of systems involving users, technology and working practices, spending time in the workplace, listening to people and trying to improve workers' lives (Muller, 2002; Farrel et al, 2006).

PD has been used extensively and effectively as a way to involve user communities with distinctive requirements (including school children) in the design and evaluation of technology (e.g. Neale et al, 2002; Parsons et al, 2011; Millen et al, 2011). PD places a strong emphasis on involving users or their proxies 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 and 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.

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, priorities, capabilities, preferences and motivations. PD for educational technologies embraces the expertise of designers and pedagogical experts who engage collaboratively with end user communities. This hybrid team approach can 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.

2.4.2 PD approaches in Year 3

In Go-Lab, for Year 3, the study activities are increasingly based on usability studies, especially when evaluating well established artefacts. Year 2 PD studies mainly elicited detailed critiques of aspects of interfaces and the generation of creative design ideas to improve prototypes. Year 3 PD studies continue this aspect but also embrace more observational data and quantitative measures.

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.

For new or radically changed artefacts which continue to emerge, we continue to use the previously established pattern of studies, sometimes involving analytical evaluation before PD studies. We also continue to use remote studies for situations where this can yield valuable and timely information.

Face-to-Face PD Approaches: In general student events have been conducted in schools, and teacher events have been conducted in schools, on research facilities or in intermediate

spaces such as conference venues. In practice – especially for teacher workshops – we have used rather hybridised study methods, somewhat varied depending on the research opportunity and facilities and time available and other practical and educational constraints. Typical HCI methods such as observations, questionnaires, interviews and focus groups are used in our face-to face events.

Software-supported PD – PDot: The custom built *Participatory Design Online Tool*, PDot, enables the gathering of contextual usability feedback from users while they are using a system, either in remote engagement of participants in usability studies or in a face-to-face PD setting (Figure 2).

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 (Heintz et al., 2014):

- involving distributed users;
- disseminating results to distributed stakeholders within the project;
- addressing the limitations of existing online annotation tools

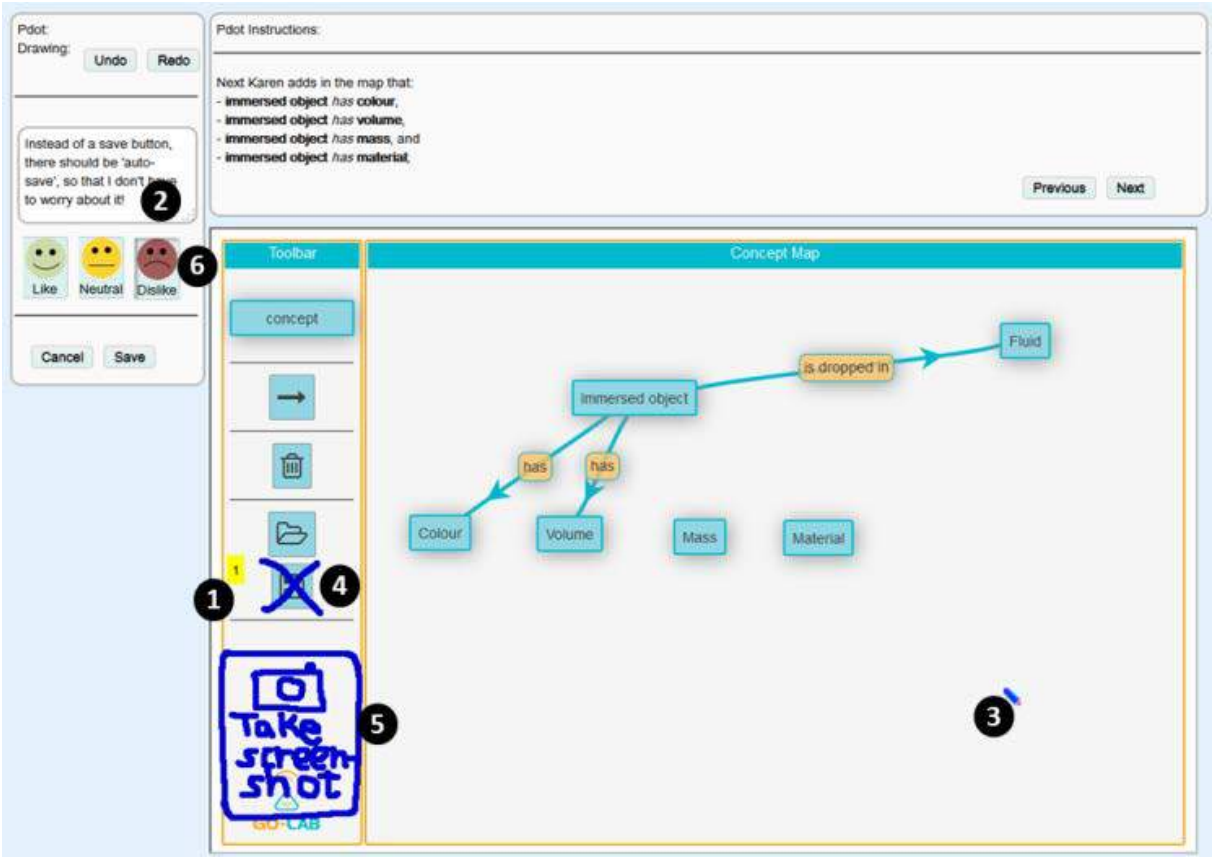


Figure 2. PDot for collecting feedback on Concept Mapper

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).

With the universal approach the Go-Lab project 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.

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. Nevertheless, PDot has been found very effective for gathering constructive feedback.

Co-operative evaluation: Broadly within a similar philosophy to Participatory Design and Action Research, Wright & Monk (1991) describe Co-operative Evaluation; a method in which the users are seen not as experimental subjects but as co-evaluators. This method emphasises targeted recruitment, well-selected and well described tasks for them to attempt, putting participants at ease by creating a supportive environment, emphasising that it is the system, not the user, which is under test, creating a good rapport using responsive rather than only scripted questioning, close observation and structured note taking supplemented by informal think-aloud methods and possibly recording. An advantage of this method is that it harnesses the skills and insights and self-awareness of the user, as well as their creative energies, providing a depth of insight into how they interact with the system, what aspects are problematical and sometimes how they could be improved.

Co-operative evaluation is normally conducted with a single user at a time, but can be used in group context without the think-aloud aspect. Many Go-Lab Year 3 teacher workshops included this method, with researchers mingling both observing and supporting during the practical sessions, and with worksheets provided to enable teachers to provide feedback.

At times researchers also asked scripted questions about the participant's experience, understandings, feelings or preferences with regard to the usability of the system. The worksheets also contained situationally appropriate specific questions. Task sessions were sometimes video-recorded or audio-recorded.

2.4.3 Remote PD approach: Core Group of Teachers (CGT)

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 from end-users on some partly formed idea or work-in-progress is through computer-mediated communication (CMC). With the support of WP7, in Year 2 we recruited a group of teachers, known as *Core Group of Teachers*, which currently consists of 21 teachers from 9 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.

2.4.4 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 usability specialists performed HE on several Go-Lab artefacts, which were subsequently evaluated with the teachers. Results of the HE were in the meantime communicated to both pedagogical and technical teams, and consequently the artefacts were improved to enable more productive use of teacher time and reducing any negative reputational impacts of exposing teachers to usability issues which the team could identify and eliminate.

2.5 Comparing the three main UCD approaches

In this section we describe each of the three main types of study used in WP3 in Year 3, specifying the strengths and limitations and typical structure of each, and the rationale for how and when they were deployed.

2.5.1 Face-to-face end-user studies

For most purposes, this is our preferred study approach, and the most frequently used.

Strengths and limitations

Face-to-face end user studies with teachers and/or students are a very effective method for eliciting findings based on genuine practitioners, rooted in in-depth experience of lesson delivery. A limitation with these studies can be that they are usually quite brief, and don't necessarily discover usability issues which might emerge when working with the technology for a longer time period.

Typical study structure – teacher-only sessions

- Presentation on Go-Lab – aims and objectives, concepts, resources provided etc.
- Live demo of portal
- (If time permits, hands on use of portal and questionnaire)
- Live demo of student experience of an ILS
- Hands on use of an ILS, Lab or App and feedback by PDot, observation, questionnaire, discussion or similar
- Live demo of ILS authoring
- (If time permits, get teachers to select a lab and begin developing an ILS for their own class, feedback by observation, questionnaire, facilitated discussion etc.)
- Conclusion – describe other features not yet seen + final feedback

Typical study structure – classroom session with students

- Briefing on what the lesson will contain – mentioning online labs etc.
- Live demo of how to use an ILS
- Hands on session – work through the ILS
- Feedback by observation, questionnaire, facilitated discussion etc.
- Later conduct teacher interview & debriefing if possible

Face-to-face end user studies conducted collaboratively with other WPs

Collaboration between Work Packages enables composite studies to be designed which deliver findings of interest to more than one work package – perhaps by using a questionnaire which includes some WP3 questions. This approach can maximise the usefulness of valuable participant engagement time. This method can broaden the number of assessments without large extra demands on recruitment and organisation of studies. They are usually most effective for summative style evaluation, and tend not to elicit creative design ideas in the manner of PD.

2.5.2 End-user remote evaluation studies

This method has been used for smaller evaluations, for evaluation of tentative designs, and when a response is needed quickly and it is impossible to schedule a face-to-face session.

Strengths and limitations

Remote end user studies are conducted mainly with the Core Group of Teachers (CGT). They are an effective method for eliciting findings from genuine practitioners with significant Go-Lab and teaching expertise. They can be scheduled reasonably quickly and easily to get answers to urgent questions. A limitation with these studies

can be that they are usually quite brief, and don't necessarily discover usability issues which might emerge when working with the technology for a longer time period. They are also rather dependent on self-perception and self-report; there is no possibility for observation, timing, or interviewing.

Typical study structure

- Email giving task name and deadline
- Description of a task to attempt with some Go-Lab component
- Feedback – often online questionnaire – subjective measures, + sometimes performance measures (timing, number of errors & retries, eventual success)

2.5.3 Analytical evaluation sessions

This method has been sometimes used as a predecessor to other methods, or when a very probing study is required. These sessions typically involve a team of usability specialists exploring one or more Go-Lab artefacts in depth while attempting to conduct an end-user task with it, and finding as many usability issues as possible. The Heuristic Evaluation approach has often been used in which the user interface is evaluated according to standard usability heuristics.

Strengths and limitations

This method can be used if an artefact is not yet suitably usable for a fluent end-user evaluation. It can be deployed without needing to gain access to end–users. Usability specialists can be very effective at finding many usability issues, and are less prone to distraction by content issues. However they may not have sufficient knowledge of how a real teacher or student would use the product in practice, and may sometimes highlight issues which an end user may never encounter in practice.

Typical study structure

- Usability specialists select typical end-user tasks with the artefact under study
- Each specialist attempts the task in every way they can imagine an end-user might attempt it, noting down any ambiguities, errors or faults
- List of usability issues is collated
- Usability specialists rank the issues independently and then form a consensus

2.6 Participant recruitment policy and practice

2.6.1 Policy

There are many areas of human and contextual variability which are likely to affect user experience and could therefore 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 and user interface paradigms. Student recruitment might cover similar factors, with a more specific emphasis on age. Anecdotal information from teachers suggests student response is also much affected by time of day, previous lessons, and time in the school year (e.g. before exams, new term, just before summer holidays etc.).

Recruitment policy is to work with 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. We aim to target diversity and inclusion of all

perspectives, but not necessarily to attempt “representativeness” in every study, since it would be impossible to represent every perspective fully in every study without unacceptably extending timescales.

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, and limited resources, there is necessarily a trade-off between covering as many Go-Lab components as possible and attempting to test each component with every possible category of user and context of use. Our recruitment decisions often prioritise broad coverage of system components, aware that this may limit the certainty and generalisability of findings.

2.6.2 Practice

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. In some countries very systematic and purposeful recruitment has been possible, whereas in others there has of necessity been a more opportunistic approach. ULEIC as WP leaders have worked to foster appropriately diverse, balanced and reasonably representative recruitment across the whole programme of studies. 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.

2.7 Organising and managing the PD Work

2.7.1 Practical context

A significant challenge for this project is to conduct PD sessions in multiple countries with different languages, involving teachers in different science subjects, with varying levels of technology aptitude and enthusiasm, and with different ages of pupils and different curriculum expectations. There are many Go-Lab design artefacts to be evaluated, and the portfolios of labs, apps and ILSs is frequently updated. In addition some individual artefacts go through multiple versions. This makes it impossible to test every Go-Lab artefact in every possible usage scenario. Hence, there is a need to prioritize the artefacts to be evaluated.

Additionally the number of teachers and students willing to undertake evaluation activities is constrained in some countries for various reasons. Teachers in different cultures and political and economic climates vary considerably in their work pressures and work patterns, their curriculum and timetable constraints, and have different motivations for use of online labs and for taking part in studies. School authorities in different countries have different attitudes to allowing researchers into their premises (from welcoming to discouraging) and to ethical requirements. Go-Lab partners in different countries therefore have different levels of access to, and rapport with, teachers, as well as different levels of HCI and PD experience.

Study opportunities vary considerably in duration, expectations and objectives of various stakeholders, ages and capabilities of students and teachers, technology available etc. General purpose study protocols are not always optimal, and often bespoke or heavily customised protocols and study materials are prepared for specific study opportunities.

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 PD protocols 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* and workflow was developed in Year 2 to manage the aforementioned challenges, and to optimise the match between supply (of study opportunities) with demand (for results). The version used in Year 2 was described in detail in D3.2. It was modified in the light of experience and expanded to include analytical and remote evaluations, and used in the modified form in Year 3.

2.7.2 Structuring and sequencing the study programme

As illustrated in Figure 3 below, usability tests of the Go-Lab artefacts 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 – perhaps using ILSs or mini-ILSs, or individual infrastructure elements; level 3 involves whole lessons or programmes of lessons delivered in the context of normal educational curriculum, and are summative evaluations (WP8 in Go-Lab). During the formative studies new requirements can arise leading to new components or ideas and therefore 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.

In Year 3, most of the studies have been at level 2 in this model.

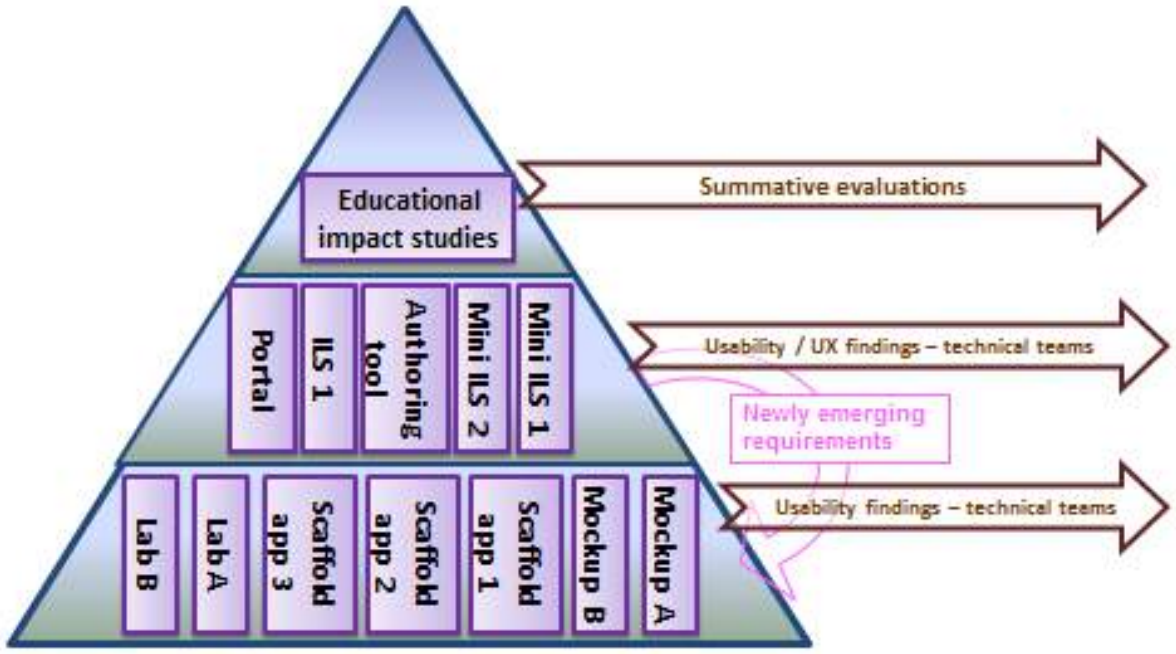


Figure 3. Study types and levels.

3. Year 3 Study Programme

In total the WP3 study programme in Year 3 consisted of 53 face-to-face PD studies (some of which involved multiple school visits), 4 analytical studies and 13 remote studies. 7 of the 53 face-to-face studies were conducted by other WPs but included WP3 research and delivered results to WP3. Several more WP6 studies also contributed quantitative data to WP3, reported at Section 4.7.

The engagement included over 450 teachers and over 550 students, from 18 countries. This represents well over double the activity and engagement levels in Year 2. Median contact time of a study was 120 minutes for teacher events and 105 minutes for student events.

Data collected included online capture of usability observations using PDot, observation reports written by researchers, online and paper-based questionnaires, system-generated performance data and screen capture data, interview material recorded in audio or handwriting, worksheets filled in by teachers or students, video recording, data written by participants on post-its or flipcharts, records of facilitated discussions etc.

In all, over 1100 artefacts (questionnaires, recordings, worksheets, field reports etc.) were collected, analysed and documented. The findings were disseminated to the consortium. This body of data is available to consortium members in Graasp. It has subsequently been summarised and collated and is presented in Section 4 below.

The study programme itself is reported in this section, followed by an analysis of the participants engaged.

3.1 Studies completed

As predicted a year ago in D3.2, for Year 3 the balance of studies has shifted towards Face-to-Face events, and with a higher proportion of student events compared to Year 2. The studies listed here were all conducted wholly or mainly for this work package. WP3 has also provided questions for and benefitted from input from a number of studies conducted primarily by other work packages, particularly WP6. These studies are not listed here but the findings are included in the next chapter.

3.1.1 Face-to-face Participatory Design studies

There were altogether 53 face-to-face PD sessions conducted during Year 3; 46 specifically or mainly for WP3 and 7 conducted primarily for another WP but providing supplemental data to WP3. Tale 2 shows an overview. Note that some of the WP3 played dual roles, collecting end-users' feedback and disseminating the project to them, and hence they are also reported in WP9.

Table 2. An overview of face-to-face PD studies in Go-Lab Year 3

Event ID	Date	Partner	Location	No. of teacher/ student	Comments / components covered
Y3PD01 *	Nov 2014	IASA	Greece: Athens	25S, 5T	ILS implementation activity
Y3PD02 *	Nov 2014	IASA	Greece: Korinth	35T	ILS implementation activity
Y3PD03	Nov 2014	UDE	Germany	13T	Authoring tools

Event ID	Date	Partner	Location	No. of teacher/ student	Comments / components covered
**					
Y3PD07	Dec 2014	UCY	Cyprus: Limassol	14S	Experiment design tool. Splash ILS., Splash Lab, EDT
Y3PD08	Dec 2014	UT	Netherlands	1T	Experiment Design Tool
Y3PD10	Dec 2014	UT	Netherlands: Groenlo	8S	Bond Lab
Y3PD12 *	Jan 2015	IASA	Greece: Argos	10T, 40S	ILS implementation activity
Y3PD14	Jan 2015	UCY	Cyprus: Nicosia	32T	Electrical Circuit Lab, series & Parallel circuits ILS, Splash Lab, Guppies lab
Y3PD15	Jan 2015	UCY	Cyprus: Limassol	11T	Osmotic Power lab, Bond lab, Electrical circuit lab, Series & Parallel ILS
Y3PD16	Jan 2015	ULEIC	UK: Leicester	3T	Portal, authoring, radioactivity ILS (demo only), Bond lab, Hypothesis tool
Y3PD17	Jan 2015	UT	Netherlands: Hengelo	30S	Bond Lab
Y3PD18 *	Feb 2015	IASA	Greece: Athens	5T, 50S	ILS implementation activity
Y3PD19 *	Feb 2015	IASA	Greece: Athens	12T	ILS implementation activity
Y3PD20	Feb 2015	UCY	Cyprus: Larnaca	12S	Hypothesis tool, Experiment design tool, Observation tool, Data viewer tool, Conclusion tool. Electrical Circuit lab, custom ILS.
Y3PD21	Feb 2015	UTE	Estonia	27T	ILS - Is it good to be beautiful (including questioning scratchpad and hypothesis tool); portal.
Y3PD22	Mar 2015	UTE	Estonia	9T	Portal. Splash ILS including Splash lab, Concept mapper, hypothesis tool and data viewer.
Y3PD23 *	March 2015	IASA	Greece: Crete	15T	ILS implementation activity
Y3PD25	Mar 2015	ULEIC	UK: Nottingham	3T	Authoring tool, portal, code compiler app
Y3PD26	Mar 2015	ULEIC	UK: Leicester	20S	Electricity ILS
Y3PD29	Mar 2015	UTE	Estonia	13T	Portal. Splash ILS. Splash lab. Apps: CM, Hypothesis Tool, Data Viewer.
Y3PD30 *	April 2015	IASA	Greece: Trikala	25S	ILS implementation activity

Event ID	Date	Partner	Location	No. of teacher/ student	Comments / components covered
Y3PD31	Jun 2015	UCY	Cyprus: Nicosia	32S	Hypothesis Scratchpad tool, Experiment Design tool, electrical circuit lab, Light fixtures ILS
Y3PD32	Apr 2015	ULEIC	UK: Bristol	21T	Radioactivity, portal (demo only), Bond (PDot), authoring
Y3PD33	Apr 2015	ULEIC	UK: York	20T	Portal (demo), Electricity (demo), Bond (PDot), authoring
Y3PD34	Apr 2015	ULEIC	UK: York	21T	Portal (demo), pH scale (demo), craters (PDot), authoring
Y3PD35	Apr 2015	ULEIC	UK: Leicester	4T	Portal (demo), Electricity (demo), Bond (PDot), authoring
Y3PD36	Apr 2015	ULEIC	UK: Birmingham	1T, 12S	electricity lab in class
Y3PD38	May 2015	EA	Bulgaria:Sofia	16T	Big Ideas workshop
Y3PD42	May 2015	ULEIC	UK: Leicester	24T	portal, craters ILS, authoring
Y3PD43	May 2015	ULEIC	UK: Leicester	8T	portal + authoring
Y3PD44	June 2015	ULEIC	UK: Birmingham	14T	portal, electricity lab(demo), Splash (Pdot), authoring
Y3PD45	June 2015	ULEIC	UK: Northampton	9T	portal, Bond lab, authoring
Y3PD46	June 2015	ULEIC	UK: Ealing	24S	electricity (series & parallel) ILS
Y3PD47	June 2015	ULEIC	UK: Ealing	13T	portal, electricity ILS, authoring
Y3PD48	June 2015	ULEIC	UK: Nottingham	5T	portal, authoring
Y3PD49	June 2015	ULEIC	UK: Leicester	30S	Craters ILS
Y3PD50	June 2015	ULEIC	UK: Greenwich	12S, 4T	Guppies and Craters ILSs + teacher demos
Y3PD51	June 2015	ULEIC	UK: Altrincham	21S	Gear sketch (use) and Craters (demo) ILSs
Y3PD52	July 2015	ULEIC	UK: St Asaph (Wales)	7T	
Y3PD53	July 2015	ULEIC	UK: Preston	6T	
Y3PD54	July 2015	ULEIC	Greece: Marathonas	28T	Observations & informal interviews
Y3PD55	July 2015	ULEIC	Greece: Marathonas	21T	"plenary" session - authoring + general
Y3PD56	July 2015	ULEIC	UK: Bournemouth	26S	Splash sinking and floating ILS. Splash lab. Hypothesis Tool. Conclusion tool.
Y3PD59	Apr 2015	UTE	Estonia	30S	
Y3PD60	Apr 2015	UTE	Estonia	24S	

Event ID	Date	Partner	Location	No. of teacher/ student	Comments / components covered
Y3PD61	Apr 2015	EA	UK: Cardiff	16T	Big Ideas workshop
Y3PD62	Jan-Mar 2015	EPFL	International	23T	
Y3PD64	May 2015	EUN	International	18T	
Y3PD66	Mar - July 2015	EPFL	Switzerland	2T	In depth study consisting of multiple visits
Y3PD67	July 2015	EA	Greece: Marathonas	16T	Big Ideas workshop
Y3PD69	Feb 2015	UT	Netherlands: Enschede	60S	Experiment with Splash Lab and EDT
Y3PD73	Jan-Mar 2015	EPFL	International	1T, 50S	
Y3PD74	Jan-Mar 2015	EPFL	International	1T, 50S	

Notes: * ILS implementation activities conducted by non-WP3 partners for other WPs but providing some usability findings for WP3.
** Study conducted by a non-WP3 partner specifically for WP3.

3.1.2 Analytical studies

There were altogether 4 analytical studies conducted during Year 3 (Table 3).

Table 3. An overview of analytic studies in Go-Lab Year 3

Event ID	Date	Partner	Components covered
Y3AN01	Dec 2014	ULEIC	App composer
Y3AN02	Feb 2015	ULEIC	ILS survey
Y3AN03	Feb 2015	ULEIC	Authoring tool
Y3AN04	Aug 2015	ULEIC	Conclusion tool

3.1.3 Remote studies

There were 13 remote studies conducted with the Core Group of teachers in Year 3 (Table 4)

Table 4. An overview of remote studies in Go-Lab Year 3

Event ID	Date	Partner	Components covered
Y3RS07	Oct-14	ULEIC + IMC	Tutoring platform
Y3RS08	Nov-14	ULEIC + IMC	Portal + Social Media
Y3RS09	Nov-14	ULEIC	Help and support facilities
Y3RS10	Jan-15	ULEIC	Portfolio of Labs
Y3RS11	Jan-15	EPFL	Requirements for Learning Analytics
Y3RS12	Feb-15	ULEIC	ILS Authoring tool
Y3RS13	Feb-15	ULEIC	App configuration facility
Y3RS14	Mar-15	ULEIC	Help & support revisited
Y3RS15	Mar-15	EA	ILS metadata
Y3RS16	Apr-15	ULEIC	Lab repository revisited
Y3RS17	May-15	ULEIC	Online support and help
Y3RS18	May-15	ULEIC	Access to ILSs from portal
Y3RS19	Jun-15	ULEIC	Go-Lab usage review

3.2 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, levels of experience 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 and statistically powerful sample of participants for each component; that is we prioritise system coverage over certainty and generalisability of findings.

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. For similar reasons, during part of one event, teachers were asked to assume the persona of a specific fictional student and respond in that persona.

As described in later sections, the face-to-face activities have been supplemented with remote studies using a Core Teacher Group, and with Heuristic Evaluations conducted by Human-Computer Interaction researchers, 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 in Section 5 below.

In all, over the course of the WP3 studies in Year 3 we have engaged with **over 450 teachers and over 550 students**.

3.2.1 Nationalities

Most of the formal PD events were conducted in the 7 countries represented in the Work Package partners (UK, Netherlands, Greece, Switzerland, Belgium, Estonia & Cyprus). However a number of events, including the Summer School events, recruited participants from a broader base. In all, participants were drawn from 18 countries.

3.2.2 Teachers

Based on the events where demographic details were collected, the teachers’ main subjects were: Physics (29%), Chemistry (16%), Biology (32%) and General Science or more than one subject (23%).

Their genders were 58% female and 42% male. Their ages ranged from 23 to 67 with a median of 40-45. Their teaching experience ranged from 0 to 37 years, with a median being approximately 12 years. The age of students taught provides less clear answers as many

teachers teach students of various ages. The type of school in UK studies (where primary = up to age 11; secondary means 11-18) was 30% primary and 70% secondary.

In many teacher events we asked “How much experience do you have of delivering lessons where student learning takes place using PCs”. The responses were “A lot” (20%), “A little” (69%) and “None” (11%).

3.2.3 Students

Based on the events where demographic details were collected, the students’ ages were as follows: 8-10 years (7%), 11-13 years (36%), 14-16 years (27%) and 17+ years (30%). Their genders were 49% female and 51% male.

3.3 Protocols and materials for Year 3 PD studies

Because of 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 wide range of other factors – there were many different protocols, some reused, some completely bespoke, and some bespoke including common reusable elements from the toolbox of resources (e.g. PowerPoints, questionnaires, worksheets) provided.

For reasons of space, we cannot include all study protocols below. However in Table 5 we provide an outline of the typical content of a teacher workshop of various durations. Appendix A and Appendix B show examples of materials for the main three different sorts of study. Materials for all studies are available to consortium partners in Graasp.

Guidance has also been provided to WP3 partners for analysing, translating and reporting findings from face-to-face studies. A slightly abbreviated form of this was used for studies conducted in English (mainly UK-based studies). The analytical and remote studies were all conducted in English by ULEIC. In most cases textual reports of studies were produced from this analysed data in a timely fashion for circulation within the consortium.

Table 5. Typical structure of a teacher workshop

Duration → / Activity type ↓	Short (1.5 hours)	Medium (2.5 hours)	Long (4 hours)
Introductions	Brief – facilitators only	Brief – facilitators and participants	Ice breaker - facilitators and participants
Presentation – Go-Lab PowerPoint	Brief version	Full version	Full version
Portal	Brief demo	Full demo + participants follow	Full demo + participants have specific task
Feedback 1	None	Audience response	Questionnaire
Student experience	Brief live demo of an ILS	Live demo of an ILS	Live demo of an ILS + participants follow online
Feedback 2	Brief verbal + optional worksheet questions	Brief verbal + optional worksheet questions	Detailed worksheet questions, then discussion
Detailed exploration of a specific lab, ILS or app	Observation or PDot session	PDot session + observation	PDot session + observation
Authoring demo	Demo by facilitators + participants follow online	Demo + participants follow online	Demo
Authoring practice	None	Test of functionality / usability	Test of functionality / usability + begin creating meaningful content
Feedback 3	Worksheet questions	Worksheet questions	Worksheet questions + facilitated discussion, or flipchart / post-it exercise in teams
Conclusion	Mention other features	Describe / demo other features	Describe / demo other features
Feedback 4	None	Worksheet questions	Worksheet questions + structured discussion

4. Consolidated Results and Recommendations

4.1 Overview

4.1.1 Scope

This section summarises the findings from WP3 studies (face-to-face, analytical and remote) conducted during Year 3, mainly relating to the usability, user experience and usefulness of the many design artefacts within Go-Lab: the infrastructure elements (portal, authoring facility etc), the labs, apps and ILSs. We do not list all observations and comments which require no action (generally good comments) since this would be superfluous in a formative evaluation report. There have been many, but just a few are included. Usability observations are accompanied by recommendations for possible improvements. Recommendations are shared with the development cluster and their responses so far are included below.

Given the volume of data collected and already disseminated from this study programme, it is not appropriate to present it all in full in this document, and a report organised study by study would not be useful to the development cluster. In accordance with their preference, WP3 formative evaluation findings below are collated and summarised by design artefact. Most studies have evaluated several Go-Lab artefacts (scaffolds, ILSs, labs etc), 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 preferred by the Go-Lab pedagogical and technical teams.

Every formative usability observation has been accompanied by a recommended response populated either directly from participant data or by WP3. In some cases recipients in the Development or Pedagogical clusters have suggested different responses to mitigate a usability issue, and these have been discussed and agreed. In the tables below the agreed response is listed, but – for the sake of brevity - not usually every responses which have been proposed and discussed.

This chapter collates the formative findings in four main sections: Infrastructure components, Labs, Apps and ILSs, and before them some general findings. It concludes by summarising one set of evaluative (quantitative and more summative) findings. Much more data in this area was gathered and shared but does not need to be presented in full in this formative evaluation report. A number of recurrent themes have been identified in the findings, and these are presented first.

Slightly different ways of summarising the findings are used for different types of design artefacts (e.g. labs, apps, ILSs) for the benefit of the different development approaches and responsibilities for those artefacts. Also slightly different approaches have been taken to studying the different types of artefact and this is explained at the start of each subsection below.

4.1.2 Component quality

As this is a formative evaluation report, most of the detailed comments recorded in this section describe areas which could be improved, and suggestions for how they could be improved. In many cases users have provided many positive and encouraging responses, but these are not collated here as this would serve no useful purpose in a formative evaluation report. The number of recommendations for a component is therefore not an

accurate indication of its design quality. For instance some components with few comments may be well-established ones which had usability improvements before Year 3, or relatively simple, or rather little used.

4.1.3 Coverage of component portfolio

Over the year, the number of labs, apps and ILSs has grown dramatically, and many have changed or even in a few cases been removed. At the time of writing (September 2015) there are 358 components on the portal (161 labs, 163 ILSs and 34 apps). We have endeavoured to cover an appropriately wide a range of these, but attempting to study every single component would have meant spreading our resources too thinly and providing rather superficial feedback. In several studies we have asked teachers to consider the whole portfolio and indicate which components they perceive to be most useful, and the portal also monitors which are most frequently used. This information has been used to help us plan our studies to cover the most important and useful components. The approach to prioritisation has been slightly different between labs, apps and ILSs, as described at Section 4.4.1 and Section 4.5.1. In these sections, as well as providing tables for improvements to individual components, we have provided emerging general guidance for selection of labs, design of apps and authoring of ILSs.

4.1.4 Findings shared with and shared by other Work Packages

The studies conducted for WP3 have sometimes produced findings of more interest to other work packages than our main audience – the technical cluster. This particularly applies to the large body of quantitative data, usually providing an assessment of how usable a particular component is deemed to be, but not giving detailed guidance on what aspect of the component could be improved. A selection of this data is presented here, but more importantly the findings have been passed on to the relevant WPs directly. WP3 has also occasionally received usability data from studies conducted by other WPs, and the relevant findings contribute to this section.

4.1.5 Development cluster response

In many cases the findings listed below have been identified and communicated prior to this report, sometimes resulting in discussion and agreement on a proposed course of action. 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 cluster are summarised in the sections below. The responses demonstrate that the WP3 does influence (but not dictate) development priorities. In some cases the development responses describe work in progress, or newly delivered, so it is not always possible to determine whether they fully satisfy the user issues. Progress is always ongoing. In Section 5.2.2 we show how the findings published a year ago in D3.2 have led to major improvements in Go-Lab.

4.1.6 Explanation of codes used in tables

The usability observation tables below use the following codes for the sake of brevity. *Frequency codes***:** these related to how often an issue has been experienced/ observed/ reported.

Table 6. Frequency codes

Code	Frequency	Numbers
A - Freq	Frequently	10+
B - Some	Sometimes	5 to 9
C - Occa	Occasionally	2 to 4
D - Rare	Rarely	1

*Category codes***:** these categorise the type of observation and/or its underlying cause.

Table 7. Category codes

Code	Description
UR	Usability Recommendation
FE	Functionality enhancement request
GF	Good feature - no change
BF	Bug fix
AP	Aesthetic preference
O	other, or no clear category

4.2 Major themes

Before providing the detailed findings, we first list some of the major recurrent themes and critiques which emerged during the programme of studies, with some suggested responses.

Table 8. Major themes

Id	Theme	Impact	Description	Response
Th01	**Usability - Invisible / appearing controls**	**Apps and labs - various**	Never have invisible controls which appear magically when you hover over nothing. This is not perceived as user-friendly by most users. Scroll bars which only appear when a scroll wheel is used or hovering are also problematical, especially on touchpad devices.	Include in guidance to internal and external suppliers of labs, ILSs and apps. Make internal artefacts comply when next modifying for other reasons.
Th02	**Usability - Icon recognition**	**Apps and labs - various**	Icons sometimes unrecognised (students and teachers) - especially novel icons - and are inconsistent across the system.	Encourage use of consistent, established and well-recognised icons in guidance to internal and external suppliers of labs, ILSs and apps.
Th03	**Engaging interactions**	**Guidance for ILS authors**	Some students skim over any material (pedagogical, scientific or UI guidance) which is not interactive or engaging - especially large amounts of text. (Especially for younger students)	Include in WP1 ILS design guidance.
Th04	**Scientific rectitude**	**Guidance for ILS authors**	Scientific and methodological correctness is very important to teachers. E.g. measures should always specify units, graphs should always have properly labelled axes and a title, one lab should say "nothing observed" rather than "nothing happens" etc.	Include in WP1 ILS design guidance.

Id	Theme	Impact	Description	Response
Th05	**Lesson duration**	**Guidance for ILS authors**	We need to be realistic about how much pedagogical time these lessons (ILSs) take	Include in WP1 ILS design guidance. Check figures on portal are good.
Th06	**Lesson planning - duration**	**Guidance for ILS authors**	Strong perception from many (not all) teachers that preparing lessons in Go-Lab will take a lot of time, and this may limit take-up	Include estimates for first time ILS authors and frequent ILS authors in teacher training and support materials.
Th07	**Use of videos**	**Guidance for ILS authors**	The practicalities of students watching YouTube videos at once in class need to be thought through - Wi-Fi loading; noise levels.	Include in WP1 ILS design guidance. Teachers can consider using video content in homework, or providing headphones, or deliver from the front.
Th08	**Suitability for younger students**	**Lab and app portfolio**	Common feedback from teachers that it's not suitable for primary school students	Investigate and publicise cases where it has been successfully used with younger students. Ensure ILS design guidance caters for the full target age range (10-18).
Th09	**More labs needed**	**Lab and app portfolio**	Need many more labs to satisfy some teachers.	A highly visible online button has now been added to request further labs.
Th10	**Labs with no logons and no booking**	**Lab portfolio**	Labs which require time slot reservation or which require an extra sign-up or logon are not considered practical by some teachers.	Pedagogical choice. Include advice in ILS design guidance. Possibly include need for logon in portal metadata.
Th11	**Teacher training need**	**Project rollout**	Some teachers perceive need for significant amounts of specialised training to use Go-Lab.	Covered in WP9 sustainability agenda. Community platform is useful. Ensure training need & opportunities are suitably visible.
Th12	**Sustainability**	**Project rollout**	Sustainability after the project is a concern for some teachers. Some have known EU projects end and leave them with nothing.	Covered in WP9. Make sustainability plans more visible.

Id	Theme	Impact	Description	Response
Th13	**Multiple languages**	**Project rollout**	Translation - need more material in multiple languages	Infrastructure in place (including app composer and customisation tool) to allow delivery in any language. As community builds, ready-made lessons will become increasingly available in various languages.
Th14	**Not yet ready**	**Project rollout**	Some teachers believe that Go-Lab is "a website in its infancy" or "will be worth looking at when it's nearer completion".	(Feedback mainly earlier in Year 3). Ensure site is clearly seen as maturing. Perhaps highlight some success stories?
Th15	**Marking and Assessment**	**System deployment in education**	Teachers across Europe absolutely want help with assessing students learning and progress as part of the system. (E.g. Y3PD62, many UK, Summer School, UCy, UTE email 17/4/2015 etc.) Must have reporting facility (preferably one click) so teachers can see and assess student's work, and demonstrate this. Assessment is a key part of teacher's job description and objectives and without it some may not consider Go-Lab.	Some features support this – e.g. Learning Analytics tools, several apps which collect student work. Plus teacher can access the ILSs of standalone viewers. Ensure this is understood by teachers, and investigate whether more is wanted.
Th16	**Organise student nicknames into classes**	**System deployment in education**	Several teachers are asking about organising nicknames by class, in case an ILS is used by children from different schools or classes within a school.	Recommended approach is to create a clone of the ILS for each class. Ensure advice is clear in training and support materials.

Id	Theme	Impact	Description	Response
Th17	**Many purposes and educational contexts for using Go-Lab**	**System deployment in education**	Many teachers suggested using this not only for main lessons but for homework, revision, detentions, excluded pupils, sick pupils, by teachers providing "cover" for an absent teacher (where otherwise pupils are often just asked to read textbook), lesson preparation (for real or online lab lesson), lesson reinforcement, catch-up work after a student absence, blended learning, distance learning etc. Go-Lab also appeals to colleges which are suffering cuts and have to deliver more teaching with fewer staff exposure hours (UK provision for 16-18 age group) as it can be used for independent study.	No action required. Provided for information of WP1 and project leadership.
Th18	**Ready-made curriculum materials**	**System deployment in education**	Some teachers want ready-made lessons for their curriculum, or at least very much stronger links to curriculum.	Expect more ILSs will be published and shared as the community grows in Year 4.
Th19	**Impact on Lab tech role**	**System deployment in education**	Impact on "Lab technician" role (in UK at least) needs to be considered and articulated. Some have felt their job is threatened by online labs.	Clarify that online labs complement, not replace, real labs. Be aware of possible concerns.
Th20	**Contextual help**	**System-wide**	Contextual help (e.g. F1 or "?" icon) is better than remote help (e.g. on portal). (N.B. student-oriented help ought to be accessible from ILS, not requiring them to go to portal)	Include in guidance to internal and external suppliers of labs, ILSs and apps.
Th21	**Usability distracting from science learning**	**System-wide**	Concern that IT issues and usability issues during lessons take time and focus away from science learning (for both teachers and students) / Concern that lesson time will be taken up with learning how to use ILS and the apps and labs within it rather than on science learning	Keep up the good work of focussing on ever improving usability; intuitiveness is even better than training & help & support.
Th22	**Online help and support**	**System-wide**	(Rarer than last year) need for help or step-by-step guide to using an ILS. Help, support and instructional material needs to be in all languages.	Templates, guidelines, forums and tutoring platform are addressing this.

Id	Theme	Impact	Description	Response
Th23	**Work on various devices**	**System-wide / technical infrastructure**	All components (especially those used by students) need to work usably on a variety of screen sizes in various browsers and devices. Some countries are beginning to provide BYOD (bring your own device) lessons. Everything should be designed for, and tested on, multiple devices and browsers.	Most in-house code is now responsive to screen size. Some external labs could be improved.
Th24	**Slow responses**	**System-wide / technical infrastructure**	Students are sometimes impatient and click on if things are slow to respond	Huge ongoing efforts to improve performance of Graasp and GoLab apps to improve loading speed even on poor connections.
Th25	**Usability - Varied ICT capabilities and preferences**	**Teacher facilities (e.g. authoring)**	Teacher sometimes believe their students have greater ICT capabilities than they do themselves.	General issue not limited to Go-Lab! Tutoring platform and training and support materials are mitigating this.
Th26	**Resilience with poor infrastructure**	**Technical infrastructure**	ICT infrastructure (devices, high bandwidth Wi-Fi) is limited and patchy, even in wealthier countries in Europe. Sometimes needs the teacher to book a (short) slot in computer room. Old kit, poor Wi-Fi; sharing devices etc.	Huge ongoing efforts to improve performance of Graasp and GoLab apps to improve loading speed even on poor connections. Also ILS design guidelines could mitigate the problem.
Th27	**Java issues**	**Technical infrastructure**	Java compatibility with browsers is a problem - e.g. for many biology labs.	Industry-wide issue affecting many websites, not just Go-Lab. Some external lab owners are working to eliminate Java dependency. Consider clarifying plugin dependencies in GoLabz portal.
Th28	**Window sizing and scrolling**	**Technical infrastructure**	Concerns about scrolling and making best use of screen area in both ILSs and authoring	New strategies for scrolling are being implemented.

4.3 General findings

These are findings which are not specific to any single lab, app, ILS or infrastructure component.

Table 9. General findings

Id	Usability observation	Frequency	Category	Agreed Response
Gen01	Loading time of the whole ILS and its different components is very long or components are not loading at all.	A - Freq	O	Done and ongoing. We have already dramatically reduced the loading time by avoiding the loading of all items at

Id	Usability observation	Frequency	Category	Agreed Response
				once, but only the ones visible. We also have optimized database queries. As an ongoing task we are monitoring performance bottle necks and improving them iteratively.
Gen02	Perceived as very time consuming to set up lessons - perhaps prohibitively so	A - Freq	O	In one or two discussions on this point, it seems that this is about the complexity of preparing web-based lessons (especially for someone bot used to it) rather than any usability issue with authoring tool. It would be worth further research (perhaps email to CGT or to Summer School 2015 teachers) to investigate. N.B. as community sharing grows, writing ILSs from scratch will be rarely necessary.
Gen03	Compatibility problems between Java websites (e.g. labs) and browsers gave many teachers - especially biology teachers - the impression that Go-Lab, or a lot of things in Go-Lab, are broken.	A - Freq	BF	Provide information / warnings on any labs which have special technical requirements not available on all mainstream browsers.
Gen04	Teachers need a facility to extract pupils' work for assessment and other purposes.	A - Freq	FE	An online facility now provided listing standalone users and giving teacher easy access to see all work done by the student on that ILS. Check whether this fully satisfies the need.
Gen05	Facility wanted to prevent plagiarism.	A - Freq	FE	Now provided (ILS author can require passwords)
Gen06	Occasional compatibility issues with Flash and browsers.	B - Some	BF	Provide information / warnings on any labs which have special technical requirements not available on all mainstream browsers.
Gen07	Wording is sometimes unclear (e.g. phase, “Tools” bar) for the participants.	C - Occa	UR	Advice for ILS authors: use student friendly wording throughout (learning content as well as apps, labs, …). No changes to infrastructure unless more evidence of widespread usability problems.
Gen08	Students should be able to access an ILS through google classroom so teachers can easily tell who's done the work and who hasn't. GC would also allow it to be stored and viewed by other teachers along with work done using other packages.	C - Occa	FE	No plans at present to integrate with GC as limited demand. Possible agenda item to consider this for sustainability release (WP9).

Id	Usability observation	Frequency	Category	Agreed Response
Gen09	Concern that an ILS with links in may not work as expected if external content changes.	C - Occa	O	ILS design advice – only link to stable, trustworthy, supported content.
Gen10	Provide facility for student to get quick preview or overview of ILS to help with budgeting time	D - Rare	FE	May be worth considering. Reflection tool contains much of the relevant functionality. Maybe students who use Go-Lab frequently will not need this. In some cases ILS author could put some text at the start to give an overview of the activity.
Gen11	Radioactivity lab doesn't scroll properly in its ILS; have to right click on SET and then open in a new window. (Example of several scrolling issues).	D - Rare	UR	Work in progress to make all scroll bars visible at all times.
Gen12	Would like more apps to support collaboration	D - Rare	FE	Unclear what apps they want. Investigate user need when priorities permit.
Gen13	Shouldn't need to log in more than once. Also concerns about booking labs.	D - Rare	UR	Reducing number of signups when booking labs by providing an integrated booking service.

4.4 Infrastructure components

A large number of studies addressed in Year 3 involved the use of the Portal and authoring tool, and – to a lesser extent – other infrastructure components. In general, the infrastructure appears much enhanced in usability and usefulness – especially the authoring tool and provision of extensive help and support materials. There have been difficulties with technology, including ICT infrastructure issues in schools, compatibility issues and some bugs. The latter have tended to be rapidly fixed.

4.4.1 Go-Lab Portal

Description

The Portal (www.golabz.eu) is the main entry point for teachers planning to use Go-Lab resources and facilities. It provides details of, and links to, all labs, ILSs and Apps, with appropriate search facilities. It has a wide range of instructional and help material (reviewed separately at Section 4.3.2) and links to the authoring facility. It also has material on the "Big ideas of science".

Screenshot

Figure 4. Go-Lab Portal

Overall evaluation

The Portal has been used in many teacher workshops with very few usability problems observed or reported. Three workshops focussed specifically on the Big Ideas of Science part of the interface, and on navigation by using metadata. It is generally felt attractive and easy to use. Quantitative research found its pragmatic qualities excellent, hedonic qualities good. Teachers tend to adopt varied and pragmatic approaches to searching for resources (e.g. for apps: exploring the app catalogue page – 36% / by exploring home page – 28% / by using search – 24% / by using filters – 12%). The text search might usefully be enhanced in

various ways. The social features which were added are generally considered useful. There were just three formative comments:

Detailed findings

Table 10. Portal findings

Id	Usability observation	Frequency	Category	Recommended Response
GP01	Preview feature (on Bond lab at least) led to a series of screen shots which looked so much like the lab that teacher was baffled by lack of interactivity	D – Rare	UR	Problem not reproducible. Preview works as expected. Perhaps user clicked on one of the screen images and got confused? No action.
GP02	If Portal is open for students to see, then don't show age-ranges for labs. Some students will find it stigmatising if they need to use labs aimed at a younger age range.	D - Rare	O	Maybe no action. Students won't normally use the Portal. This metadata element is useful for teachers.
GP03	Searching for labs doesn't work very well as you need an exact term.	D - Rare	UR	Search enhancement planned, so user can search among labs only, ILSs only or Apps only.

4.4.2 Help and user support

Description

Partly in response to D3.2, a considerable body of help and support materials and facilities has been added to the Portal under the “Support” tab. This includes video tutorials, “tips and tricks”, questions and answers, a community forum, a printable manual on authoring tool, a link to the tutoring platform and “MOOC” online courses.

Progress

In November 2014, the User Support Task Force (USTF) commissioned a remote study, Y3RS09 with the Core Group of teachers to investigate the level of need and elicit ideas and suggestions. This showed a unanimous feeling that more help should be provided (67% “Yes”, 33% “Possibly”, 0% “No”) with the most popular delivery media being Annotated Screenshots and Video Tutorials, but with support for a wide range of media. When asked about help and support for individual components, then brief textual help was considered most appropriate. The most popular way to access help was with an icon. The response from this and other activities was summarised in the following slide show in Figure 5:

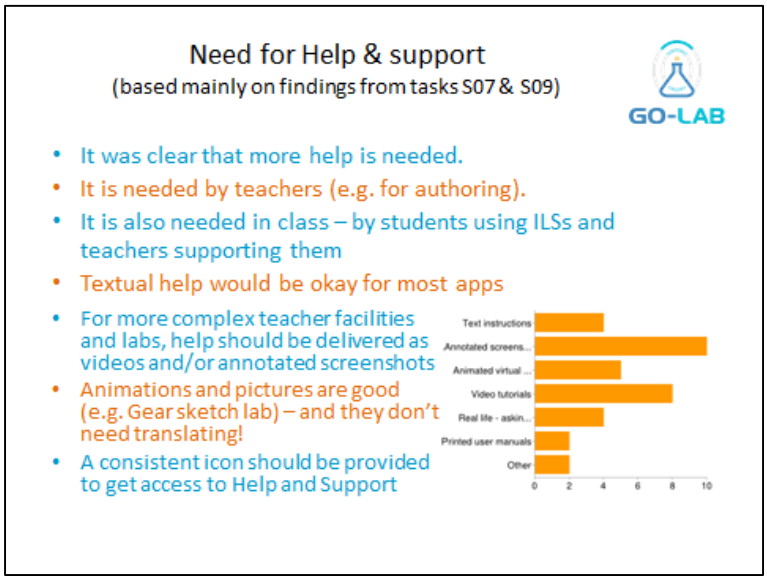


Figure 5. Need for help and support facilities.

This prompted a flurry of activity from the USTF, WP5 and others to provide a vehicle for delivery and also content. Some of the activity is summarised in Figure 6. Specifically, a new tab was added to the main Portal which leads to the screen show in Figure 7.

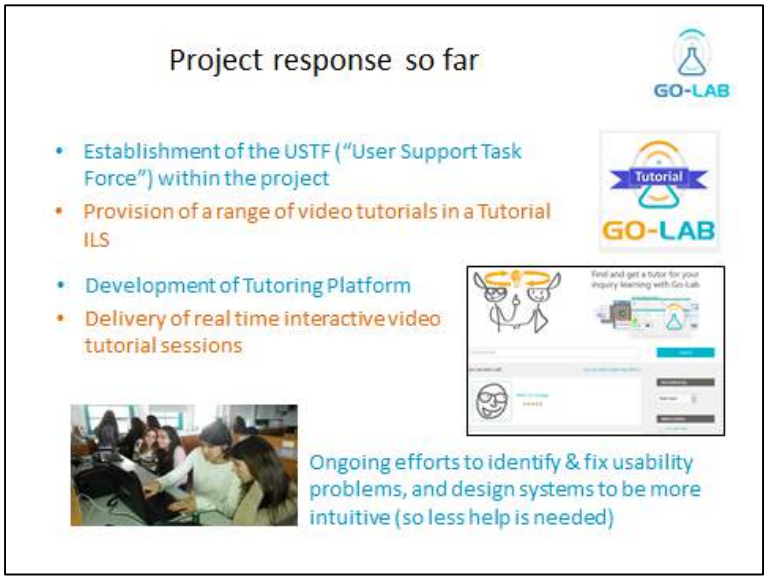


Figure 6. Project response.

This prompted a flurry of activity from the USTF, WP5 and others to provide a vehicle for delivery and also content.

Screenshot

Figure 7. Help and support screen.

Overall evaluation

In general, the feedback on this has been excellent. There are still a few requests for online help on specific details of specific apps, but it is much less of an overwhelming concern than last year.

Detailed evaluation

In March 2015, a second remote study, Y3RS14 with the Core Group of Teachers evaluated progress. The response was overwhelmingly positive, with also some suggestions for further refinements. The USTF has continued progressing this work and there are no further critiques to add.

4.4.3 Authoring tool

Description

The authoring tool is a web based facility based on EPFL's Graasp environment, which allows teachers to create and publish an ILS, either from scratch or based on cloning and modifying an existing ILS.

Screenshot of authoring tool

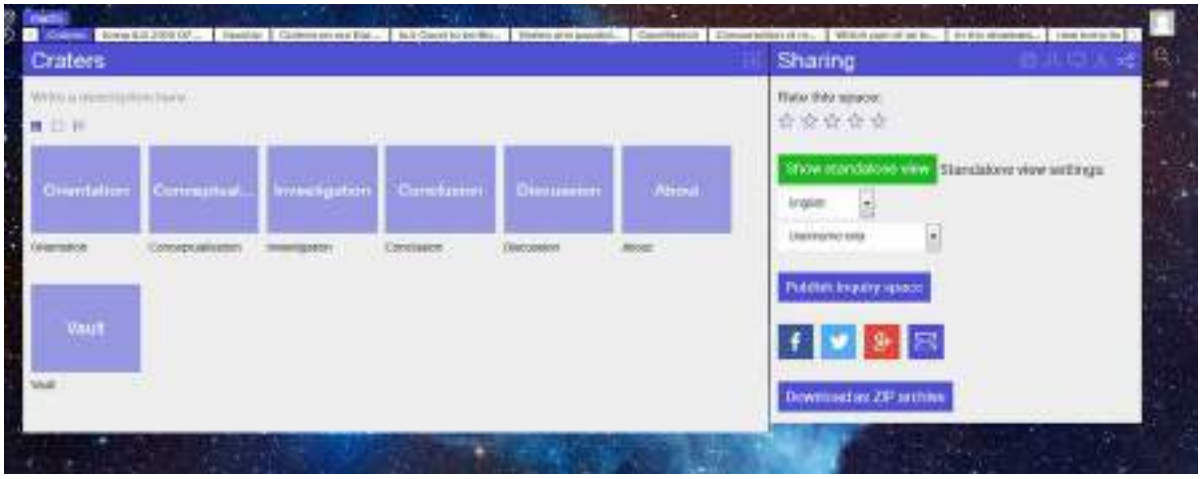


Figure 8. Authoring tool.

Overall evaluation of the authoring tool

The tool has been completely reworked since D3.2, and is now perceived by both teachers and project team members as much more intuitive and easy to use. There have also been further improvements, both in reliability and usability, during the year. Some teachers still some question why authoring needs a different environment and unfamiliar usability paradigm, and there is genuine scope for usability improvements, but these do not seem a major obstacle to most teachers using it.

Detailed findings

Table 11. Authoring tool findings

Id	Usability observation	Frequency	Category	Agreed Response
GAu01	Perception that creating an ILS would take a lot of time. E.g. "Would have to create lessons in Summer holidays, not when I have teaching duties".	A - Freq	O	In deeper discussion with a couple of teachers, it seems that the unfamiliar task of creating online lessons is what will take most of the time; not the user interface of authoring. Support platform and tutoring can assist. As community grows, fewer ILSs will have to be built from scratch and more will be available “off the shelf”.
GAu02	there were weird interactions during authoring - one person crates a text box and it appears in other people's ILSs	A - Freq	BF	Already fixed
GAu03	We need an app to display data in tabular format	A - Freq	FE	New table tool now provided. (Not yet tested for usability)
GAu04	Provide an "add text" button rather than forcing user to add a resource with no content	B - Some	UR	Done – button labelled “Add document”. Work in progress to simplify the UI to eliminate EDIT/VIEW mode.

Id	Usability observation	Frequency	Category	Agreed Response
GAu05	When adding a resource, if a URL is supplied but no name, nothing happens.	B - Some	BF	“Show a helpful error message.” “We have removed the need to type a name with a link” Clarity still being sought
GAu06	Users clicking on the "+" icon in the instructional text rather than the active "+" icon at the top of the screen. (Minor but quite frequent usability snag.)	B - Some	UR	Suggest: make the active "+" icon much more visible, not pale grey and tiny. Or - perhaps better - make buttons visible all the time without having to press "+" every time. Response: no action for now – cluttering the layout might cause as many problems as it solves.
GAu07	"Click on "+" is unclear	B - Some	UR	Suggest: Change the "+" to a more recognisable icon? Response: No action for now. The “+” is a standard UI affordance.
GAu08	When building an ILS it should be easier to re-order things and add text between other components. Can it be done in second or third view?	B - Some	UR	No action for now. The list view orders items alphabetically, so reordering there will be confusing; in the extended it might also be difficult to drag big components.
GAu09	Extend editor to provide (in some cases reinstate) facilities for subscripts, superscripts, and more complicated formulas	B - Some	FE	Formulas can now be added in descriptions and documents using standard Latex format inside tags. (Fully or partially satisfying the user requirement.)
GAu10	Need proper training for teachers	C - Occa	O	See GAu12
GAu11	Don't display "Create Space" button in contexts where it has no use.	C - Occa	UR	No software change – user learning issue.
GAu12	Needs an online tutorial	C - Occa	UR	Already done - pdf manual and video tutorial now available on Portal.
GAu13	Provide facility to include pictures or icons in a rich text document	C - Occa	FE	Under consideration
GAu14	Provide facility for teacher to scale a lab or app to fill more of the window seen by students.	C - Occa	UR	Ongoing work to make labs and apps respond to screen dimensions. This should mainly satisfy the requirement.

Id	Usability observation	Frequency	Category	Agreed Response
GAu15	ILSs can't safely be shared as students in different schools may use the same nickname.	C - Occa	FE	Similar to another request for a facility to create a structure of "classes" and have nicknames unique within a class. Current solution – clone the ILS so there's a unique instance for each class.
GAu16	Some ILSs on Portal don't have a student view or a "copy ILS" link. E.g. Impulse.	C - Occa	BF	Probably caused by a historic bug in publishing. Will not recur with the rewritten ILS publishing scheme.
GAu17	ILS publishing facility - provide a place for author to document any special browser requirements (e.g. plugins) which are needed.	C - Occa	FE	Done
GAu18	No facility for teachers to add their own mini-apps	C - Occa	FE	Unclear what is the problem. Seek clarity.
GAu19	Creating ILS from nothing is much easier than modifying one; e.g. adding text, or reorganising the material in a phase.	C - Occa	UF	Unclear why this is the case. Possibly duplicate of GAu08
GAu20	"Standalone view" should be renamed "Pupil view"	D - Rare	UR	Rare – no action
GAu21	the 5 funny little icons have wrong luminosity levels; bright when selected, greyed out when not. Greyed out normally means unavailable.	D - Rare	UR	Rare and only mildly problematical – no action.
GAu22	Would like a feature (app?) for ILS author to add a formula and have it displayed as a graph	D - Rare	FE	Rare – no action
GAu23	difficulty experienced trying to change a picture once added to the ILS	D - Rare	O	unclear and rare – no action
GAu24	Provide an app so teacher can create text with words missing for student to fill in.	D - Rare	FE	Useful possible enhancement, but rare so no action yet.

Id	Usability observation	Frequency	Category	Agreed Response
GAu25	Provide facility for teacher to create a table with headings, for student to complete. Also - even better - allow teacher to partially fill table and leave other boxes for pupil.	D - Rare	FE	Useful possible enhancement, but rare so no action yet.
GAu26	Would be nice to have suggestions for what type of material to put in each phase	D - Rare	FE	Possibly useful enhancement. Pedagogical team may like to consider this alongside the "scenarios" work.

4.4.4 ILS delivery

Description

The student experience of an ILS is also used by teachers to check on the student experience, especially while writing an ILS with the authoring tool. In that context is called the “standalone view”. The findings here focus on the infrastructure through which an ILS is delivered. Matters relating to ILS content and style are at 0 below. However both tables should be read as there are some observations which could be categorised either way, and in some cases a usability finding can be remedied either by Go-Lab changes or by content changes.

Screenshot of ILS delivery

Figure 9. ILS delivery.

Overall evaluation of the ILS delivery

There have been a number of issues with the way ILSs are experienced. Some of these are the result of poor network capability at study venues, and some were bugs which are now fixed. Nevertheless there is still scope for further improvements in areas such as screen utilisation and scrolling, and working to ensure ILSs can be delivered on the very wide range of platforms and browsers (including old equipment) in schools and – for homework or BYOD (bring your own device) learning environments – in students' homes.

Detailed findings

Table 12. Findings – ILS delivery

Id	Usability observation	Priority	Category	Agreed Response
GDel01	Apps sometimes slow to load (e.g. 2 minutes), or endlessly waiting	A - Freq	O	Field-based performance testing; tuning; improve reporting of performance problems. Development effort has been and is being devoted in Graasp and Golabz apps to improve the app and ILS loading. Much improved already.
GDel02	Flash and Java plugin compatibility issues between browser and code - sometimes fixable with techy intervention, sometimes not depending on browser	A - Freq	BF	Flash and Java apps and labs will present incompatibilities from now own due to the lack of support provided by certain browsers. Portal changes now highlight dependencies.
GDel03	Too much text - students get bored and/or skim	B - Some	UR	WP1 ILS design guidelines: design all ILSs to be suitably engaging and interactive for the target audience.
GDel04	Components not presented as intended on various devices (e.g. iPad 2)	B - Some	UR	Sizing and scrolling improvements are work in progress. See also GDel02
GDel05	Having no passwords means students could plagiarise too easily	B - Some	FE	Now fixed. ILS author can select from 3 accessing mechanisms: 1) anonymous, 2) with just nickname, 3) with nickname & password.
GDel06	Many resources (YouTube videos, labs, apps) within an ILS are displayed in small areas of the screen, and are hard to view and use because of size, when large parts of the screen are unoccupied or little needed.	B - Some	UR	. wherever possible enhance design to make best use of screen space available . Provide facilities to resize or maximise windows Partially done. Pictures, documents, videos (and other resources) and apps are set to use the maximum width unless the definition is enough. In the case of pictures, the teacher may specify if it's necessary to provide a small, medium, big size. The option of maximising windows is currently out of scope
GDel07	The ILS itself does not occupy the full width of the screen. Too much of the	B - Some	UR	Under discussion. Recommended: "Make the ILS content occupy more of the

Id	Usability observation	Priority	Category	Agreed Response
	background image, too little of the content.			screen, and the background image less." Response: "This is a design decision that follows layout patterns and good practices that promote that generation of responsive ILSs suitable for multiple screens."
GDel08	Entering backspace in various places can wreak havoc and lose your work. It sometimes logs you off completely.	B - Some	O	This is an inherent problem with most browsers. However it can be very disconcerting and disruptive in a complex interaction such as an ILS. Problem can be mitigated by designing so that it's always clear to user whether they're in text editing mode (in which case backspace behaves as expected) or not (in which case backspace key acts as browser back button). N.B. Mitigation - To avoid users losing their work, many Go-Lab apps & labs now save the users work automatically.
GDel09	YouTube videos can't be maximised. {The maximise button seems to be available sometimes and not others - perhaps depending on browser version?)	B - Some	UR	Same response as GDel04 & GDel07
GDel10	Working on computers isn't collaborative and good science lessons should be collaborative (several students in Y3PD49)	B - Some	O	Interesting perspective but probably no action at this stage. "CSCL is a full research area about collaborative learning supported by computers that shows the benefits of this approach. Teachers may argue that Graasp doesn't provide much collaboration support but, as in any real lab, students may work together with the same computer/device".
GDel11	Sometimes apps (e.g. Hypothesis Scratchpad, Concept Mapper) which have been configured to have a specific vocabulary get displayed in uncustomised form in student view.	B - Some	BF	Fixed.
GDel12	Suggest providing a comment / chat / discussion / feedback system to allow students to feed back comments to the teacher online	C - Occa	FE	Sounds like a potentially great idea. Out of scope for project lifetime but suggest put on agenda as future work.
GDel13	Sometimes scrolling impossibilities presented -	C - Occa	UR	Ongoing programme of improvements to display

Id	Usability observation	Priority	Category	Agreed Response
	e.g. a lab which you can't scroll to the bottom of.			presentation.
GDel14	YouTube video in ILS has rapid spoken English with a US accent which is hard to understand	D - Rare	UR	WP1 design guidance: Where possible use videos in appropriate language (and dialect). Use videos with subtitles.
GDel15	The use of "nickname" rather than "logon" or "username" may be seen as patronising by older pupils.	D - Rare	AP	Not worth changing on the basis of a single user comment.
GDel16	Concern about embedding links (e.g. YouTube) in an ILS: might the content change between teacher selecting it and student accessing it? If so then teacher has to check every ILSs just before lesson which is impractical.	D - Rare	O	WP1 ILS design guidance: use links which are reliable and stable and supported.
GDel17	In Y3PD56, YouTube videos would not play at all on some PCs but would on other PCs	D - Rare	BF	Bug not reproducible. Perhaps a school had a very out of date browser.

4.4.5 App composer

Description

The App composer provides a sophisticated and comprehensive set of facilities to make Go-Lab apps much more customisable by teachers, to match their specific linguistic and pedagogical needs. This addresses some frequently mentioned usability and functionality needs expressed by teachers in Year 3. The software also provides a range of facilities to allow teachers to share their customisations, make them visible to others and take on aspects of editorial responsibility.

Some of the functionality has now also been provided in the configuration facility built into the authoring tool, and so the app composer has been used mainly for creating and storing translated versions of apps.

Screenshot

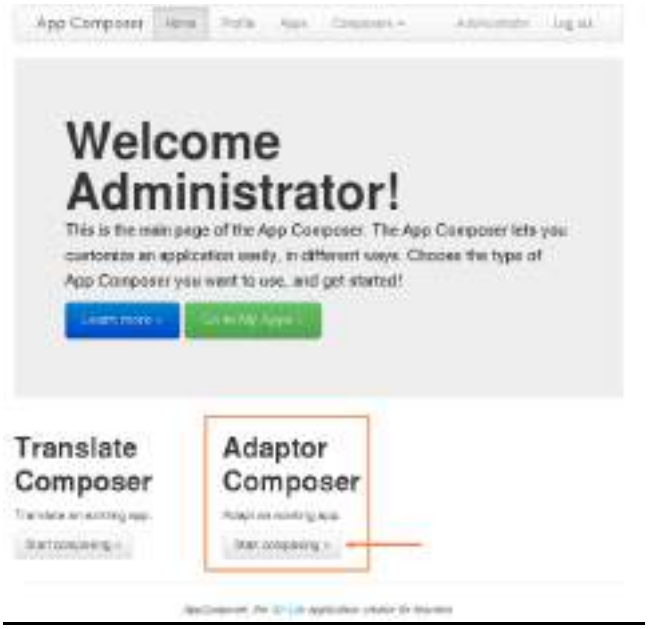


Figure 10. App Composer front page

Overall evaluation

The App composer delivers the required functionality effectively. However it appears to have a rather complex design requiring the user to have a sophisticated mental model of how it works.

Detailed findings

Note that most of the findings on this tool are based on remote evaluation, and therefore follow a somewhat different format to most of this chapter. As explained above, some of the functionality is now provided (more easily) in the authoring facility, and therefore feedback from teachers has so far been limited. The evaluation report including tables of findings is rather long to include here so it can be found in Appendix B.

4.5 Labs

The portfolio of labs on the Go-Lab Portal has increased dramatically this year. At the time of writing there are 161 labs. Many of these are externally supplied and well established, and have undergone scrutiny and quality checks before being admitted onto the Go-Lab repository. Some are also very specialised and may not be widely used. It therefore seemed most appropriate to focus WP3 resources for formative evaluation on the labs which are most frequently used, and also on labs which are in house and still being refined, where feedback is most likely to be effective. The Portal has a feature to monitor lab usage and an option to display them in order of popularity[8].

In addition, some studies evaluated the whole portfolio of labs, or asked teachers to select labs of particular interest to them. Therefore this section presents feedback on labs

[8]At the time of writing the five most studied labs rank 2, 6, 1, 4 & 3 respectively in the portal popularity rankings.

generally, then feedback on the whole portfolio of labs, and finally feedback on some of the most popular and useful labs in the current portfolio.

4.5.1 Labs – general & portfolio

Detailed findings

Table 13. General Labs findings

Id	Usability observation	Frequency	Category	Agreed Response
LGen01	Need to make labs which use Java workable on all browsers	A - Freq	O	This is a problem with infrastructure, (compatibility of modern browsers with Java), not a problem with our labs, but it makes some labs unusable on some platforms. Suggest highlighting dependencies on browser plugins in a highly visible way in the lab repository. Now done. There is a metadata field called "technical requirements" on the lab page, so that such information about java is given. We will also think of give a big label to those java labs to warn the users.
LGen02	Need to dramatically increase the number of labs to meet {UK} curriculum requirements.	B - Some	O	The portfolio has now grown to 161 labs, but some teachers want more and more specialised labs. There is now a button for teachers to propose a lab in Golabz.
LGen03	Will never use a lab which needs an extra logon or registration	C - Occa	O	Booking info nbow available on Portal, facilitating teacher choice. (WP2 may like to try to find more easily accessible labs, or ones with a free preview).
LGen04	Labs which need to be booked in advance are not of much practical value - simulations are better	C - Occa	O	try to find more easily accessible labs
LGen05	user wants to create own labs	C - Occa	O	perhaps provide some information on Portal about how to do this (and what skills needed and quality criteria)
LGen06	Need a facility for findings from a lab to automatically populate a results table in the ILS from which graphs etc. can be created.	D - Rare	FE	Some teachers would want this; others not, depending on learning objectives. Possible criterion for lab selection for WP2 and lab design for any future in-house labs..

Extra labs requested / suggested by teachers

Both in response to questions and often spontaneously when exploring the Portal, teachers have requested or suggested a huge number of extra labs. The list of such requests from WP3 studies is presented in Appendix C for the attention of WP2.

4.5.2 Electrical Circuit Lab

Description

This in-house lab allows students to create and manipulate a range of electrical circuits using standard components.

Screenshot

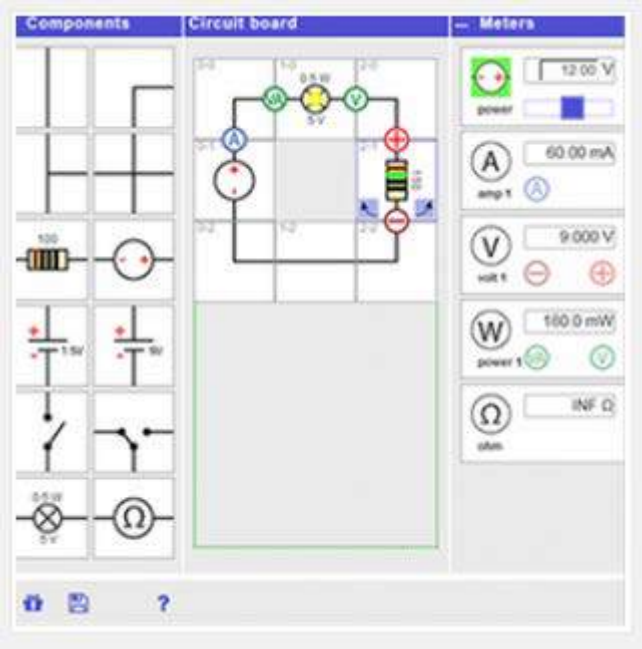


Figure 11. Electrical circuit Lab.

Overall evaluation

The lab is very popular with teachers and students, and effective. It is not obvious to all users how to use meters, and there are some minor but valuable improvements possible.

Detailed findings

Table 14. Electrical circuit lab findings

Id	Usability observation	Frequency	Category	Recommended / Agreed Response
LEC01	Arrows on power supply go up in millivolts, so no practical use.	A - Freq	UR	In discussion. Recommended: “Make them go up in steps of e.g. 0.1V” Response: “The user just has to keep the arrows pressed and the change speed will increase. How should that be made clear to the user?”

Id	Usability observation	Frequency	Category	Recommended / Agreed Response
LEC02	Provide text labels for components for students who don't know symbols	B - Some	FE	Possible enhancement ... But it may do more harm than good. Some teachers like the way this teaches the proper circuit diagram symbols. No action.
LEC03	Lab doesn't support multiple batteries	B - Some	FE	Useful possible enhancement / correction. No action – overly complex for limited benefits.
LEC04	Volt Meters and Amp meters should be dragged to the circuit in the same way as other components. In the Electric Circuit Lab it wasn't obvious that users need to drag the meters down to this part.	B - Some	UR	This interaction could be made more obvious (e.g. by labelling the area to drop meters accordingly or by showing an animation). “Dragging the volt and amp meters to the circuit (as the components) would also means wiring the meters inside the circuit. The meters would then be a part of the circuit. As a result the circuit would become more cluttered and placing the meters in another location is a lot more work”
LEC05	Should be able to replace a component by simply dropping another component over it, rather than having to remove the old one first.	C - Occa	FE	Minor possible usability enhancement – under consideration.
LEC06	Would be great to add extra components - buzzers, thermistors, diodes, LDRs (light dependent resistors), capacitors and coils to be useful for a wider range of classes and activities. (LDRs and thermistors would create a need for some way to control the ambient temperature and lighting levels).	C - Occa	FE	Possible enhancement being considered. ”Adding new components means also enhancements of the circuit analysis code. Capacitors and coils are being considered as it is a limited enhancement of the circuit analysis code and would make the lab a lot more useful.”
LEC07	Add a tool to measure the brightness of a bulb rather than subjective impressions?	C - Occa	FE	Extend the feature set of the Data Collector tool in the Electrical Circuit Lab to also allow to measure brightness. Possible enhancement being considered. ”Adding new components means also enhancements of the circuit analysis code.”
LEC08	Saving and loading was not possible in the Electrical Circuit lab (“it hangs on loading”).	C - Occa	FE	Done

Id	Usability observation	Frequency	Category	Recommended / Agreed Response
LEC09	Request for having a help button in the Electrical Circuit Lab.	C - Occa	UR	Make existing help facilities more obvious, as they seem to get overlooked.
LEC10	Data Collector tool perceived as unusable or not working	C - Occa	O	Dev trying to find a design solution to make this easier.
LEC11	Descriptions for the components the diagrams represent in the Electrical Circuit Lab are missing.	D - Rare	FE	Could possibly provide this as an option selectable by ILS author. (Some teachers like it the way it is, as it teaches students about standard circuit diagram symbols).

4.5.3 Sexual Selection in Guppies

Description

This well established biology lab provides the student with facilities to explore genetic changes over many generations of guppy breeding, giving them facilities to control factors such as the number of spots on male guppies, the prevalence of predators, and selection preference of female guppies.

Screenshot

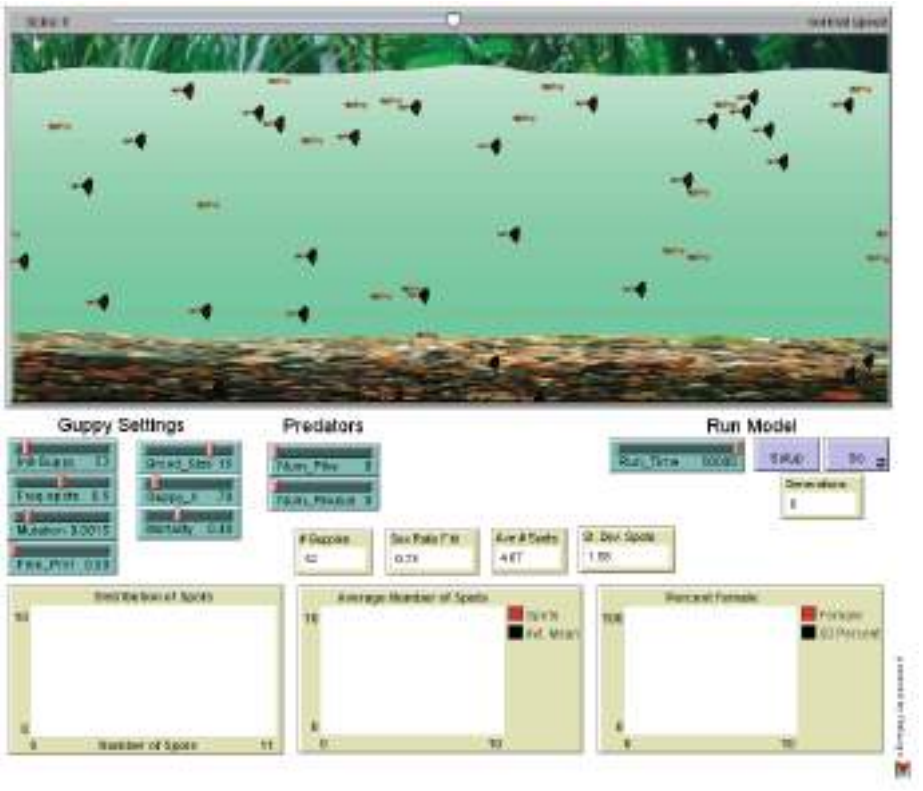


Figure 12. Biology Lab Guppies

Overall evaluation

This lab and an accompanying ILS “Is it good to be beautiful” were very much appreciated by teachers and students in two studies, and used effectively for learning. A number of biology teachers were very enthused by this lab, and didn’t report any formative findings.

4.5.4 Bond Lab

Description

The in-house written Bond Lab allows students to select two chemical solutions from a store cupboard of reagents, and to pour one of them from one beaker into a beaker containing the other, If a chemical reaction with an observable precipitate would take place for those two particular chemicals, there is a visual simulation of the reaction. The student is then expected to use a facility to document the reaction which took place in standard chemical (ionic) equation format.

Screenshot

Figure 13. Bond Lab.

Overall evaluation

The bond lab is highly popular with teachers and students, and quite engaging. There are many comments and observations but these are not a major critique. There is a lot of enthusiasm to extend the lab. It could be much improved in usability by fixing just the first.

Detailed findings

Table 15. Bond Lab findings

Id	Usability observation	Frequency	Category	Agreed Response
LB01	Mouse target to drag beaker needs to be larger / more visible	A - Freq	UR	Create a larger and more visible "handle" - perhaps even depicted as a real handle. **Now done** – not yet tested by end users.
LB02	Present the full names of the solutions and not just chemical symbols	A - Freq	FE	Present the full names of the solutions and not just chemical symbols. Being considered, but there are space constraints.
LB03	It is not at all obvious how to manipulate the equation - which bits drag and drop and when, and finding invisible arrows. Not at all intuitive. {Some teachers added that finding out how to use it would distract attention from the science learning}.	A - Freq	UR	Suggest: Either add more detailed instruction text to say how to do it (or an animation), or - far better - make it more obvious - by allowing drag before or after specifying charge and quantity, and making arrows visible. Dev willing but not yet planned.
LB04	Some animated precipitations - due to colour - were not very visible on certain screens / "colours a bit boring"	B - Some	UR	Suggest: Animation needs also to be as scientifically accurate as possible. Perhaps provide an option to change the background behind the beaker and rerun the experiment? Dev willing but not yet planned.
LB05	Perhaps provide a choice of store cupboards each focussed on a specific area of experimentation . (E.g. one on mixing oxidants with acids; one specifically on metals)	B - Some	FE	Possible enhancement - would probably benefit many teachers. Dev willing but not yet planned.
LB06	Animation was very impressive.	B - Some	GF	No action ... But see improvement suggestions too.
LB07	Could not find the arrows {to adjust the formula}	B - Some	UR	Make arrows visible at all times.
LB08	Appreciative comments about the appearance. E.g. "Attractive screen & easy to use. Appears to be pupil-friendly". "Like the chemical store".	B - Some	GF	No action
LB09	Have a button to mix solutions rather than drag action	C - Occa	UR	Have a button to mix solutions rather than drag action, or implement suggestion about bigger, more visible handles on beakers.
LB10	"Click on umbrella" message (which has an umbrella icon) leads some students to click on the umbrella in the help text instead of the active icon.	C - Occa	UR	Suggest: Add the word "above" to help text? Or make the umbrella icon a hyperlink? Dev willing but not yet planned.
LB11	Some students didn't discover solubility chart without help.	C - Occa	UR	Make it more visible, or provide guidance.
LB12	Underlying mistakes in	C - Occa	O	? It should be correct and

Id	Usability observation	Frequency	Category	Agreed Response
	chemical data.			reliable. However on one occasion where this was reported, investigation showed that the lab was correct.
LB13	Provide a wider variety of experimental results / observations (e.g. temperature, explosions, rates of reactions), not just precipitation.	C - Occa	FE	Possible enhancement - would probably benefit many teachers.
LB14	Enhance "nothing happened" message either to say why nothing happened or to ask student why they think nothing happened.	C - Occa	FE	Possible enhancement. Dev willing but not yet planned.
LB15	It's not clear to all users that the formula required is for the visible precipitate, and not the other formula for the ions which remain aqueous.	C - Occa	UR	?Add text to clarify the objective
LB16	Confusing {icon}. I thought the one that looks like the empty set {the circle with a diagonal line through it} would be a reset button. Also other comments about unfamiliar icons.	C - Occa	UR	The two arrow in a circle icon is increasingly recognised as "refresh" in browser based software but may have little recognition to users of windows based software. Probably not worth changing icons but tool tips recommended here and everywhere.
LB17	Move instructions to the top of the lab so people read them before using the lab.	C - Occa	UR	Possible improvement.
LB18	I liked the way it told me when I put in the wrong charge {in the formula}	C - Occa	GF	no action - though some teachers want to limit the number of guesses ...
LB19	I'm not a chemist but the correct formula can be found by randomly clicking. Could we have a facility to limit the number of attempts a student could have?	C - Occa	FE	The approach to this depends on the pedagogical style and purpose of the teacher. Some would find the suggested idea a bad one. It might be better to show the student a count of the number of "guesses" they have had, and to give teachers a facility to access this count too.
LB20	Instructions box is useful.	C - Occa	GF	No action
LB21	Can there be adequate support for less able pupils to use it and get the right answers independently while still giving sufficient challenge to the more able ones?	C - Occa	FE	Possible enhancement area. E.g. Have some hints and support available on request and its use visible to teacher.

Id	Usability observation	Frequency	Category	Agreed Response
LB22	Chemical formulas {on bottles} may be confusing for students. / "Add the words as well as the chemical symbols to bottles".	C - Occa	O	Suggest no change: ensure the target age range for the lab is appropriate. (Not much room for full names. Perhaps a tool tip?)
LB23	A reaction would have taken place but with no precipitate. The text says nothing happened. This needs to be reworded.	C - Occa	BF	To a scientist this is a bug and needs to be fixed. Dev willing but not yet planned.
LB24	Needs better help / instructions	D - Rare	UR	Perhaps add a "?" button to access guidance.
LB25	Precipitation animation rather slow to start and to run	D - Rare	UR	Start animation immediately. (This might reduce perception of slowness.)
LB26	Resetting the lab not quick and obvious. {Reload icon probably not recognised}	D - Rare	UR	No action.
LB27	Audio feedback during mixing or subsequent reactions would be nice	D - Rare	FE	Consider adding sound effects
LB28	"You should be able to adjust volume manually by typing the required value" {Probably means the volume of the two liquids}	D - Rare	FE	Possible functional enhancement
LB29	"Students pour more than one solution together. The lab then becomes very slow. Not sure whether it works correctly"	D - Rare	O	Unclear. Performance testing and load testing should be conducted.
LB30	Add the phases in chemical equations, with "(aq)" for an ion and "(s)" for solids	D - Rare	FE	Possible enhancement.
LB31	Needs a far larger store cupboard with many more reagents. / "Can we add a bromide please?"	D - Rare	FE	Done & tested.
LB32	Dragging of coefficients requires too much precision	D - Rare	UR	Make a bigger target area?
LB33	User trying to specify ion charges experienced problems with the invisible handles on the ions, and this led to confusion and self-blame.	D - Rare	UR	Make all handles and arrows visible.
LB34	"Why can I not pour the contents of two bottles into a single beaker instead of going through an intermediate step?"	D - Rare	UR	? Probably do nothing as only one mention and may be no consensus that it would be an improvement.
LB35	It would be good to be able to mix more than two solutions	D - Rare	UR	Suggest no change. Would probably be disliked by some other chemistry teachers.

Id	Usability observation	Frequency	Category	Agreed Response
LB36	Not clear to me that you had to put both solutions in the same beaker.	D - Rare	UR	No change needed. The Instructions seem very clear. See LB27
LB37	Liked the clear information on the bottles	D - Rare	GF	No action.
LB38	Would be good if you could choose the volume of each solution.	D - Rare	FE	Done & tested.
LB39	Nice animation of the reaction. Could it be slightly slower so students can take in what happens?	D - Rare	FE	Might be frustrating if it were slower, but perhaps a "slow replay" button would be a useful enhancement.
LB40	Enhance visuals so you can see the liquid being poured and even control how quickly and how much of it you pour.	D - Rare	FE	Possible enhancement / extension.
LB41	The animation is not an accurate representation of what happens	D - Rare	FE	Difficult criticism to satisfy.
LB42	Find some way to indicate (in the available solutions) which ones are likely to react?	D - Rare	FE	Probably best not to do this in the lab. Then ILS authors can give hints or guidance if they want, and leave it to pupil if they want.
LB43	Difficult to see which one is which between the two iron nitrates. {bottle labelling issue: the subscript "2" was tiny and on a line by itself)	D - Rare	UR	Clearer labelling would be helpful for this one chemical at least. Some chemical s with longs names seem to create a wider bottle to hold the name (e.g. NaCH3COO) but others split the name over two lines. Dev questioning the importance and practicality of this due to screen size and cluttering.
LB44	I like this feature as it is a long time since I worked with charges of ions, and it will be good for SEN children	D - Rare	GF	No action
LB45	After several attempts {at entering the formula} there should be a way of supplying the pupil with the right answer or some hints else they will get bored.	D - Rare	FE	Possible enhancement, but possibly detrimental to some, depending on teaching style. If implemented it should be a configurable option.
LB46	Available solutions bottles - could they be the same colour as the contents?	D - Rare	FE	Possible minor enhancement
LB47	Could molecule structure be shown?	D - Rare	FE	Possible enhancement

Id	Usability observation	Frequency	Category	Agreed Response
B48	The person sitting next to me put Fe2+ in the reaction and got it right. I put Fe3+ which was also accepted. We can't both be right.	D - Rare	BF	Investigate to see if there is a bug here. However it may be that one of them used Fe(NO3)3 and the other used Fe(NO3)2.
LB49	Consider making it ask the pupils to write the equation and predict the precipitation before trying it out?	D - Rare	FE	A pedagogically intriguing enhancement request. If implemented, it should be configurable by ILS writer.
LB50	Don't like the terminology used.	D - Rare	FE	?Not specific so impossible to know what to do.
LB51	{while entering formula} Not obvious that you have to click outside the counter to see if the charge selected is correct.	D - Rare	UF	An interaction style not familiar to users of typical windows-based applications in which you typically press on a button or icon or hit enter to get validation done. Perhaps leave as it is and hope familiarisation will suffice.
LB52	Selecting and mixing chemicals and observing reactions would be fun for all abilities, but the equations and balancing would be beyond some children.	D - Rare	O	Maybe make the equation part optional?
LB53	Some icons and some of the labels on solutions are very tiny and hard to read.	D - Rare	UR	?Increase fonts and icon sizes?
LB54	The empty {right hand} beaker, if lifted and poured into the beaker which had already reacted, causes reaction to be repeated.	D - Rare	BF	Sounds like a bug to fix.
LB55	Solubility chart - no information about AgF solubility	D - Rare	O	provide that information?
LB56	Ionic equations give the impression that the order matters, whereas writing it the other way around is perfectly valid	D - Rare	BF	Should allow both sequences since both are scientifically valid
LB57	The animation should also state textually what colour the precipitate is, for colour-blind students	D - Rare	UR	Possible accessibility improvement

Evaluative result

The lab got excellent ratings for usability from the 39 students in studies Y3PD10 & Y3PD17. This questionnaire is adapted from the System Usability Scale (Brooke, 1996).

Q1 - I think that I would need the support of a technical person to be able to use this lab.	1.87
Q2 - I thought the lab was easy to use.	1.59
Q3 - I thought there was too much inconsistency in this lab.	1.41
Q4 - I imagine that most people would learn to use this lab very quickly.	1.84
Q5 - I found the lab very awkward to use.	1.38
Q6 - I feel this could be an excellent lab.	1.81
Q7 - I felt the lab lacked an overall design.	1.73
Q8 - This lab could work well in lessons.	1.56
Q9 - The lab needs a lot of work before I would be prepared to use it.	1.22
Q10 - I hope to be able to use this lab in the future.	1.66
Q11 - I feel the lab should be more enjoyable to use.	1.28
Q12 - It was a pleasure to use this lab.	1.58

1=strongly favourable (disagree for odd numbered questions, agree for even numbered ones)
3 = neutral
5 = strongly unfavourable (agree for odd numbered questions, disagree for even numbered ones)

Development team response summary

Table 16. Response to Bond Lab findings

Response category	Observation ids
Software already changed and verified by end users	31, 38 (and some others), teachers can configure the solutions for their ILS
Software changed – not yet retested	1
Software change – work in progress	
In design	
Planned	
Willing	3, 4, 5, 10, 14, 23
Considering	2 (space available?)
Questioning importance	43 (and some others), cannot make the lab bigger (see also general observations about the use of screen real estate)
Do not intend to address the issue	In general, any item that suggests to add text to explain what students should have to do has been considered with great care.

4.5.5 Splash

Description

This well-established in-house lab provides a range of experiments concerning density of objects and liquids, sinking, floating, buoyancy and Archimedes principle.

Screenshot

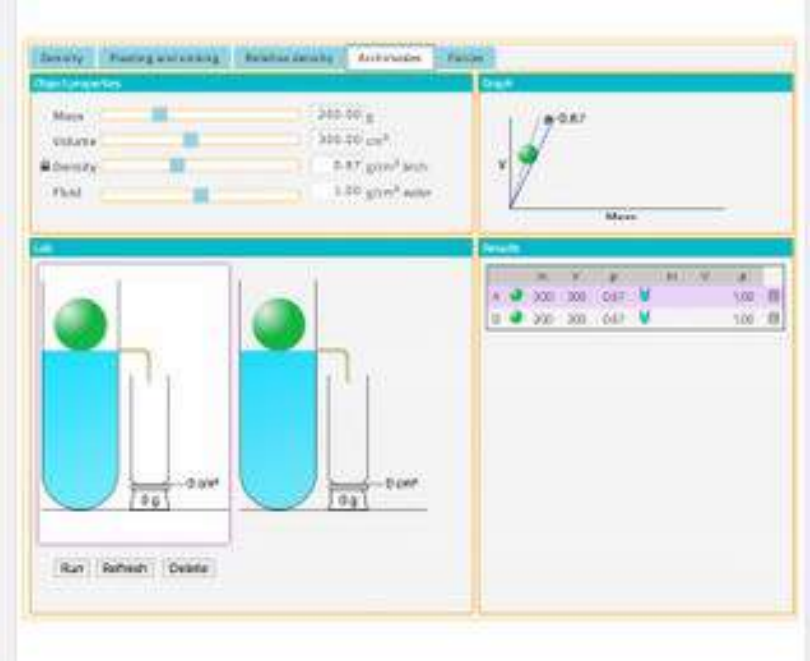


Figure 14. Splash Lab.

Overall evaluation

The lab could be improved by some of the range of usability enhancements listed below, but it seems to have enough interest for students to be rather tolerant of its usability foibles. It was much improved last year and as a result most of the comments this year are occasional or rare.

Detailed findings

Table 17. Splash Lab findings

Id	Usability observation	Frequency	Category	Recommended/Agreed Response
LSpl01	Had to type in a long URL	B - Some	O	No tech fix. WP1 ILS design advice: use URL shorteners or find an appropriate way to distribute links.
LSpl02	{S & F tab} Multiple tubes was a great advantage	B - Some	GF	No change required.
LSpl03	{Density tab} Not clear how to move padlock from density to mass or volume. Not even clear to some that you can.	B - Some	UR	Suggest: Find a clearer way to depict this. One way might be to have a padlock symbol above all three and to have radio buttons so that one can be selected. Dev willing.
LSpl04	Rather boring	C - Occa	AP	Perhaps consider more stimulating design. Perhaps this is partly dependent on teacher qualities. No change required.
LSpl05	{Density tab} The name of material (e.g. "Amber") to the right of the density figure needs to be clearer, else it might not be noticed.	C - Occa	UR	Suggest: Consider larger / brighter font. Dev willing.
LSpl06	{Relative Density tab} Labels needed on individual liquids in the tubes / clicking or hovering over a sphere or a tube should show the type of solid / fluid which it is made of / Use the colour of the ball or liquid to represent what it is.	C - Occa	FE	Useful possible improvement.
LSpl07	{Density tab} Suggest drop down box for material rather than a slider for density. / "It took me ages to find ebony. A drop down as well as a slider would be good"	C - Occa	UR	Would improve usability but have pedagogical implications. Suggest no action unless there's a clear consensus or a way of doing it which will leave existing option available too.
LSpl08	{Density tab} There is more than one value for the density of ebony!	C - Occa	BF	No plan to resolve – it selects the material with density closest to the specified value.
LSpl09	{Relative Density tab} Not clear how to select individual tubes / No instructions as to why there are several tubes and how we should be using them.	C - Occa	UR	Provide help button with suitable text?

Id	Usability observation	Frequency	Category	Recommended/Agreed Response
LSpl10	The two density sliders (for fluid and object) have different scales. It would be much better to have the same scale.	C - Occa	FE	Beneficial enhancement
LSpl11	{Density tab} The slider {for mass} is not precise enough.	D - Rare	UR	Provide a more precise slider, or some arrow buttons for fine tuning. (If the keyboard arrows provide this function then make it clear to the user). Or allow the user to type a value.
LSpl12	{Density tab} When the mass was meant to be locked, it could still be changed.	D - Rare	BF	If this is correct, then it was a bug. Suggest text and if it this user's claim is correct then fix it.
LSpl13	{Relative Density tab} Provide a feature to allow a change to mass, density or volume for all six tubes at once. Trying to repeat exactly the same values is tedious and error-prone.	D - Rare	FE	Possible enhancement.
LSpl14	{Density tab} Better to show the mass / volume / density calculation explicitly?	D - Rare	FE	Possible enhancement.
LSpl15	The results table is not very clear.	D - Rare	UR	Provide help button leading to explanatory text.
LSpl16	{Relative density} Reformat screen so that lab appears above sliders	D - Rare	AP	Suggest no action without establishing a consensus. The existing layout may be preferred by most.
LSpl17	{density tab} One student would like to be able to move mass and volume sliders independently (when density locked). {and presumably get feedback if the three were inconsistent}	D - Rare	O	Probably ignore. This would be a different functional and pedagogical design, and may disadvantage far more users than those who would want this.
LSpl18	Allow beaker size to be varied as well	D - Rare	FE	Possible enhancement.
LSpl19	Can only drop new elements on the end of the hypothesis - after which you can click again to resequence them	D - Rare	UR	Possible usability improvement: Allow elements newly copied from the list to be dropped anywhere in the hypothesis

Development team response summary

Except where noted above, and given resource constraints and priorities, most of these comments are treated as wish list but not yet planned.

4.5.6 Impact Calculator

Description

The well-established impact calculator lab allows a student to experiment with crashing astronomical projectiles (meteors or asteroids) on planet earth. The user can select the size

and density of the projectile, the velocity, angle of incidence, the type of surface (sea or land) of the impact site, and indeed may choose a specific impact site using Google Maps. The lab provides a range of visual imagery, scientific information and information on the crater size, the damage it would cause and the societal impact.

Screenshots

Figure 15. Impact Calculator.

Overall evaluation

This lab is extremely engaging and at the same time educational, and therefore much liked by students and teachers alike. There is some room for usability improvement regarding sliders and perhaps provision of more help.

Detailed findings

Table 18. Findings for Impact Calculator lab

Id	Usability observation	Frequency	Category	Recommended Response
LIC01	Sliders too sensitive - can't specify exactly 380 km from crash site, nor projectile diameter of 8000, nor a specific trajectory angle. Especially difficult on devices without a mouse - e.g. touchpad.	A - Freq	UR	Alter sensitivity, or provide a fine adjustment feature, or a numeric input box. (Keyboard arrow keys work but hardly anyone guesses this).
LIC02	Small fonts - difficult especially for visually impaired pupils	B - Some	UR	Some of the fonts (e.g. the ones used for "projectile diameter", "distance from site" etc.) could reasonably be increased.
LIC03	Leeds is not an option in the drop own menu! / More place names would be useful.	C - Occa	O	No action
LIC04	Would be good to be able to create much smaller projectiles - e.g. 5kg	C - Occa	FE	Possible enhancement
LIC05	Geography would be an issue / distraction for some pupils	D - Rare	AP	Other teachers and students loved this aspect and found it engaging, so suggest no action.
LIC06	Pupils may struggle using the map tool without guidance	D - Rare	UR	Google Maps is widely used both standalone and in other sites, so perhaps no action needed.
LIC07	Students may not know where Leeds is located in the country. {Instructions asked user to target the projectile on Leeds}	D - Rare	UR	No action for lab. ILSs using this lab need to be suitably generic so that target places can be found.
LIC08	projectile velocity box and slider inactive on some browser/ computer combinations and works in others.	D - Rare	BF	Already fixed? Reported Feb 2015, but cannot reproduce now.
LIC09	Target density box and slider disabled if water is selected.	D - Rare	UR	Probably working as designed, but some explanation for the student would be useful.

The development team have not yet had time to assess and respond to these recommendations.

4.5.7 Other labs used in WP3 studies

The following labs have also been used in WP3 studies (usually just one study) and produced minimal formative usability feedback:

- Osmotic Power lab
- Sexual Selection in Guppies
- Radioactivity lab
- pH Scale lab
- Gear sketch lab

4.6 Apps

There are currently 34 apps in the Portal. Some (e.g. Hypothesis Tool, Conclusion Tool) are pedagogical scaffolds, some (e.g. Periodic Table, Calculator) are reference material or learning aids, some (e.g. File Drop) are purely functional, and some (e.g. Online Users Visualisation) are for classroom management purposes. Some of these are well-established and some relatively new; some frequently used and some rarely if ever.

We have focussed most of our research efforts on the apps which are most frequently used and most crucial to the Inquiry-based Learning paradigm. The 5 apps most frequently studied this year were Hypothesis Tool, Concept Mapper, Experiment Design Tool, Data Viewer and Conclusion Tool and Reflection Tool. Their current rankings in terms of frequency of use (as recorded by the Portal) are respectively 1, 3, 2, 7 and 5. Their rankings when teachers were asked to nominate the apps most helpful for their teaching were respectively 1, 4, 2, 5 and 8.

We also present data from studies covering some less popular apps and covering the whole app portfolio.

4.6.1 Apps – general matters

Teachers asked to choose 3 apps most helpful for their teaching. Top apps were:			
	Number of times chosen	**%**	**Ranking (Aug 2015)**
Hypothesis tool	12	61	1
EDT	11	56	2
Quiz Master	11	56	4
Concept Mapper	10	51	3
Data Viewer	9	46	7
Function plotter	6	31	16

4.6.2 Concept Mapper

Description

The Concept Mapper tool is a diagramming tool typically used in the conceptualisation phase and allows students to select or enter some key concepts and the relations between them. It is sometimes used as an alternative to the hypothesis tool for younger children or in lessons where a more exploratory investigation is expected.

Screenshot

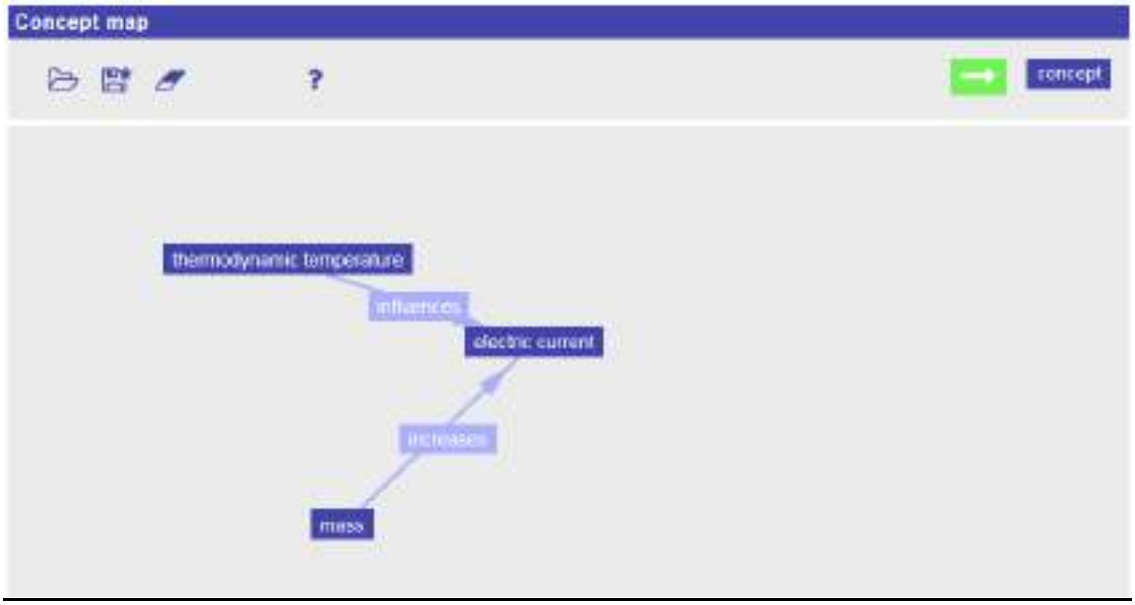


Figure 16. Concept Mapper.

Overall evaluation

The Concept Mapper was studied in year 2 and again in year 3, with some of the observations below echoing findings from year 2. It is a relatively simple application to use, but with some unexpected and mildly frustrating interactions. It would benefit greatly from some usability improvements.

Detailed findings

Table 19. Concept Mapper findings

Id	Usability observation	Frequency	Category	Recommended / Agreed Response
ACM01	Use of drag and drop for concepts but mode for arrows is confusing	A - Freq	UR	Make arrows drag and drop (like Microsoft Visio). **Dev considering.**
ACM02	When the Concept Mapper tool is in “add relation arrow mode”, labels of concepts cannot be edited.	A - Freq	UR	Make it clearer in which mode the user currently is. Allow leaving the “add relation arrow mode” without clicking on the arrow in the menu, e.g. by double clicking on a concept to edit its label. Or change the interaction to add relationships, e.g. to dragging and dropping arrows like for concepts and then connect the ends to the appropriate concepts. **Dev considering.**
ACM03	Switching from arrow mode to typing mode requires an unnecessary hit of arrow button.	B - Some	UR	Switch out of arrow mode automatically when a user attempts to type or to drag a concept (or better - see ACM01) **Dev considering.**
ACM04	The functionality of the	B - Some	UR	Add a label for each button (besides

Id	Usability observation	Frequency	Category	Recommended / Agreed Response
	buttons in the toolbar of the concept mapper is not clear.			the existing tooltips) and improve icons, e.g. by using standardized symbols where possible. **No action planned.**
ACM05	User can only use the arrow tool in the concept mapper if there are at least two concepts on the page. The flexibility of using this tool is restricted.	C - Occa	UR	Allow adding of relations even if there is only one concept (to enable a more flexible tool usage for more than concept maps) or at least provide a meaningful error message if that is not **No coding change planned**. (Would be counter-intuitive for many users)
ACM06	The erase button erases the whole concept map, not just one element.	C - Occa	UR	Make it more obvious how to delete single elements. **Done**
ACM07	The erase button caused the Concept Mapper to crash.	C - Occa	BF	Prevent crashes. **Done**
ACM08	Long concept labels that include white spaces are not visualized correctly in the Concept Mapper tool.	C - Occa	BF	**Done**
ACM09	The save button/feature doesn't seem to work to save the created concept map.	C - Occa	BF	**Done**
ACM10	The provided help is not easy and obvious to find.	C - Occa	UR	Make the help easy to find - e.g. large "?" icon. **Done**
ACM11	After creating a custom concept in Concept Mapper tool, if the user wants to edit the label again, the new text is just added to the existing one by default.	D - Rare	UR	When clicking on a concept, the existing label should be highlighted automatically and replaced when you start typing.
ACM12	The colour of arrow in the concept map is faded and not appealing.	D - Rare	AP	Try out different colours to find one that is appealing to most of the participants.
ACM13	There is no alignment or automatic formatting of concept maps.	D - Rare	UR	Add a feature to automatically tidy and format elements.
ACM14	The existing selection of colours to choose from when changing the colour of concepts in the Concept Mapper tool is limited.	D - Rare	UR	Offer a colour wheel to select the colour instead of a predefined list of fix colours.
ACM15	Participant was not sure what to write in his or her concept map.	D - Rare	UR	Add some examples for concepts and relations. (Questionable whether this should be in ILS text or concept mapper help).
ACM16	Map had disappeared on the page refresh.	D - Rare	BF	Already fixed?
ACM17	Search bar for concept blocks not particularly	D - Rare	UR	Implement the filter option differently, so that it filters from the start of the

Id	Usability observation	Frequency	Category	Recommended / Agreed Response
	efficient. Results returned tend to contain the letters typed, not begin with them, which is unlikely to be what the user wants.			word, not anywhere in the word.
ACM18	When user hovers over the colour options for the concept box the whole rectangle is highlighted blue, which brings confusion on what colour has actually been chosen.	D - Rare	UR	Use a different way to highlight the current colour selection, e.g. with a frame around it.

4.6.3 Hypothesis tool

Description

The Hypothesis Tool allows students to formulate hypotheses to guide their following experiments and data analysis. Hypotheses are created with this app by dragging items (either pre-written words provided by ILS author, or the student's own words in boxes) and dropping them one after the other to form the desired sentence.

Screenshot

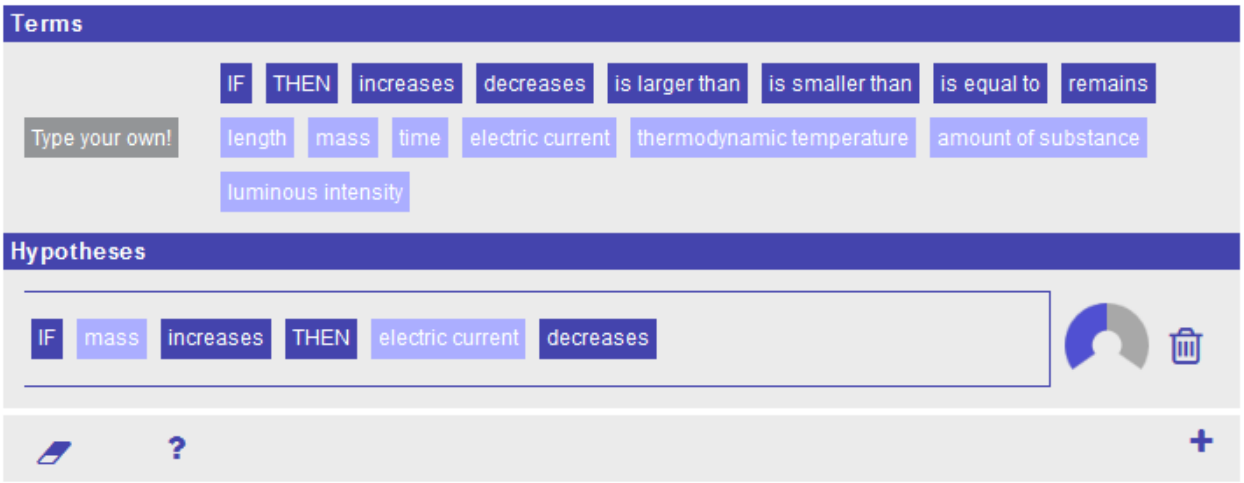


Figure 17. Hypothesis tool.

Overall evaluation

This tool has an overall appeal as a diagramming tool, but has some features which are not used in the way some users (e.g. those familiar with diagramming in Visio, PowerPoint etc) would find intuitive. Most people can use it reasonably well after a bit of trial and error. Applying some of the following usability improvements would be very beneficial for this frequently used tool.

Detailed findings

Table 20. Hypothesis tool findings

Id	Usability observation	Frequency	Category	Recommended / Agreed Response
AHT01	Users write whole hypothesis in a single "type your own" box / "much easier to type own in full rather than use drag and drop system" (7 of 15 in YPD46 did this)	A - Freq	UR	Consult WP1 to see whether this is desirable. If not, perhaps implement a word limit? **Done**
AHT02	users seen to be deleting whole hypothesis because "delete element" feature is not apparent	A - Freq	UR	Implement some more guessable way to delete elements - e.g. right click menu, or select feature. If not, then provide instructions how to do it. **Done**
AHT03	Difficulty finding out how to delete a single element from a hypothesis	A - Freq	UR	As AHT04: Implement some more guessable way to delete elements - e.g. right click menu, or select feature. If not, then provide instructions how to do it. **Done**
AHT04	Users drag a "type your own" box and then find they can't type into it	B - Some	UR	Possibly enable typing in box before and after it's dragged to hypothesis area. **Planned**
AHT05	Deleting element difficult even when you know how	B - Some	UR	Requires too much precision in dropping on the bin icon. Make it more generous, especially for long elements (e.g. type your own)
AHT06	Provide an example hypothesis so students know what to attempt. (5 students out of 15 in Y3PD56)	B - Some	UR	? Possibly add a help button which shows an example and explains how to use it. **Done**
AHT07	Clicking on dustbin deletes whole hypothesis without warning.	B - Some	UR	Add a confirmation dialog before deleting the whole hypothesis. Maybe also include guidance on how to delete an individual element.
AHT08	Difficulty dragging a "type your own" box into the hypothesis area. Dropping text blocks in the Hypothesis Scratchpad tool doesn't work properly and user has to order them afterwards.	C - Occa	O	(From Y3PD20 & Y3PD26) - ?firstly clarify the problem ... **Done**
AHT09	Difficulty using drag and drop within a hypothesis to change the sequence	C - Occa	UR	Make it require a bit less precision? **Done**

Id	Usability observation	Frequency	Category	Recommended / Agreed Response
AHT10	Cannot edit words in ready-written concepts - important in some languages where word endings have to be adjusted to make grammatical sense	C - Occa	UR	Allow editing of words in ready-written concepts once they have been dragged into hypothesis area **Planned**
AHT11	Ignoring the confidence meter	D - Rare	UR	? No need for change?
AHT12	Use of "+" button to add another hypothesis not recognised	D - Rare	UR	? No need for change?
AHT13	User didn't realise that after typing a "type your own" they had to click outside the box before dragging it to the hypothesis construction area	D - Rare	UR	? No action? This is standard MS Windows interaction style.
AHT14	Enable ILS authors to disallow "type your own" in some ILSs	D - Rare	FE	Provide a feature in app configuration menu to do this.
AHT15	A user had tried to add a second hypothesis in the Hypothesis Scratchpad tool by using space or new line, which didn't work. Seems like the functionality of the "+" button to add hypothesis is not clear.	D - Rare	UR	Add a better or detailed description of how the Hypothesis Scratchpad works. Reposition the "+" button underneath the existing hypothesis to make it more obvious.
AHT16	There could be hidden answers for the questions asked in the Hypothesis Phase, on which the user could click to reveal them.	D - Rare	FE	Could add a feature / app that allows this kind of content and interaction.
AHT17	It is not explained/clear where the hypotheses are saved.	C - Occa	UR	Add a description of how the Hypothesis Scratchpad works. **Planned**

4.6.4 Experiment Design Tool

Description

This well-established in-house scaffold app allows students to plan empirical research by first nominating which variables they wish to control (keep equal), which to manipulate (independent variables) and which to measure (dependent variables), and secondly by specifying the values they will use for controlled and independent variables.

Screenshot

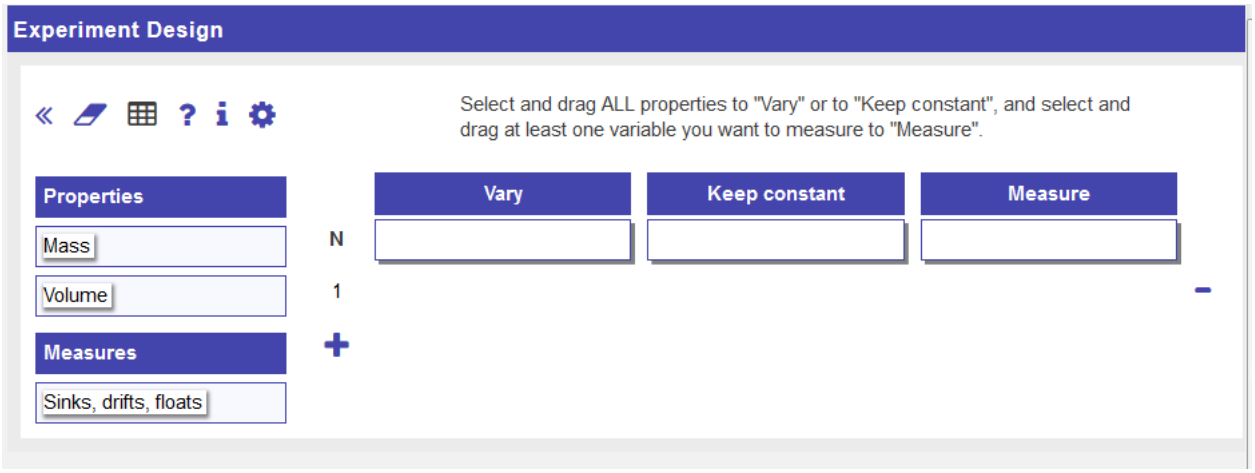


Figure 18. Experiment Design Tool.

Overall evaluation

This tool has been considerably improved since D3.2 and is more intuitive for many users. There are still a couple of areas of both comprehension and control which could usefully be improved.

Detailed findings

Table 21. Experiment Design Tool findings

Id	Usability observation	Frequency	Category	Recommended/Agreed Response
AEDT01	"Difficulty in the insertion of a measure because of an incomplete design" (7 students in Y3PD20).	A - Freq	O	Unclear what this means. Unable to reproduce so no action.
AEDT02	Dragging variables into the header is not intuitive when there is an empty box beneath it which looks like a destination.	A - Freq	UR	Either provide an online cue for destination, or make boxes active destinations. **Dev considering**
AEDT03	Sliders not easily controlled.	A - Freq	UR	Make more controllable; consider arrow icons for micro-adjustments (or keyboard arrow keys).
AEDT04	Error message when dragging a variable to the "Measure" column when in fact it was a valid variable to measure.	B - Some	BF	Allow appropriate variables to be dragged to measure column.
AEDT05	Sliders don't always have appropriate ranges for the experiment (e.g. volt slider goes up in millivolts).	B - Some	UR	Make them appropriate, or provide an authoring customisation feature so the ILS author can. **No action planned.**
AEDT06	It is confusing that the Properties and Measures are only draggable at the part of the boxes that contain text.	B - Some	UR	Fulfil the user expectations triggered by the box-design and make the whole box draggable, not only the text. **Dev considering**

Id	Usability observation	Frequency	Category	Recommended/Agreed Response
AEDT07	Sliders don't always have consistent ranges even when they refer to the same type of variable (e.g. two sliders for mass of different objects have different scales).	C - Occa	UR	Maybe no action; maybe make them consistent; maybe have a scale displayed. **No action currently planned.**
AEDT08	Students skip over EDT, and perhaps return to it later.	C - Occa	O	A matter for teachers / ILS authors rather than software fixes.
AEDT09	The Experiment Design Tool needs clearer instructions.	C - Occa	UR	Provide clearer instructions in/for the Experiment Design Tool. **Dev willing**
AEDT10	Not obvious how to assign values to "keep constant" variables. Workflow requires user to use add button to add an experimental trial.	D - Rare	UR	Provide online hint
AEDT11	Using a tool described as "Experiment design" for recording results of an experiment is counter-intuitive	D - Rare	UR	Consider using a separate app to record results, and remove this functionality from EDT. **No action planned.**
AEDT12	When you click on "view the experimental trials you have conducted", a black "X" symbol appears on a purple background with too little colour contrast to be seen on most screens.	D - Rare	UR	Display the X in a more contrasting colour to make it readily visible. **Change done – not yet user-tested.**
AEDT13	the drop-target boxes are perceived as text input boxes that are not working.	D - Rare	UR	Re-design the drop-target boxes to make their purpose clearer (e.g. add an arrow pointing downwards to indicate dropping something there). **Dev considering**
AEDT14	Clicking on a button in the Experiment Design Tool scrolled the page down for some reason for one of the participants.	D - Rare	BF	Doesn't appear to be reproducible. (Fixed? Illusory?)
AEDT15	If the hypothesis created in the Hypothesis phase could be seen in the Experimentation phase, would be better.	D - Rare	UR	Show hypotheses created, maybe also add according hints to do so to the teacher guidelines. **No action planned – would need a broad consensus.**
AEDT16	One participant reported a bug with the Experiment Design tool. After putting the “Measures” Brightness in the wrong box, he or she was not able to move anything anymore.	D - Rare	BF	Doesn't appear to be reproducible. (Fixed? Illusory?)
AEDT17	The functionality of the icons in the experiment design tool is not clear.	D - Rare	UR	Make use of labels or a brief description above this tool to explain how it works.

Id	Usability observation	Frequency	Category	Recommended/Agreed Response
AEDT18	It is not clear that the circle icon in the Experiment Design Tool resets everything.	D - Rare	UR	Use a different icon. Other than that the tooltip and popup message already are good ways to prevent accidentally deleting work.

4.6.5 Observation tool

Description

The observation tool allows students to record observations made while preparing, conducting and analysing experiments. Observations, together with data analyses, can later be retrieved in the conclusion tool as a basis for drawing conclusions.

Screenshot

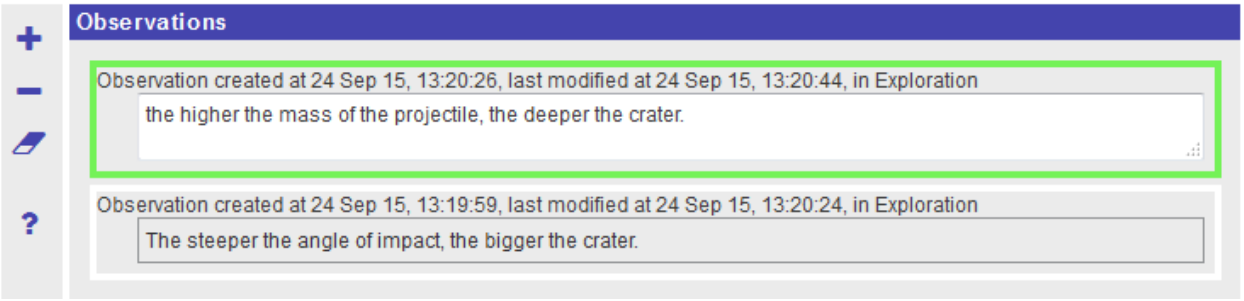


Figure 19. Observation tool.

Overall evaluation

This relatively simple tool has been used in several studies with no significant usability or related problems reported.

4.6.6 Data Viewer

Description

The Data Viewer tool 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.

Screenshot

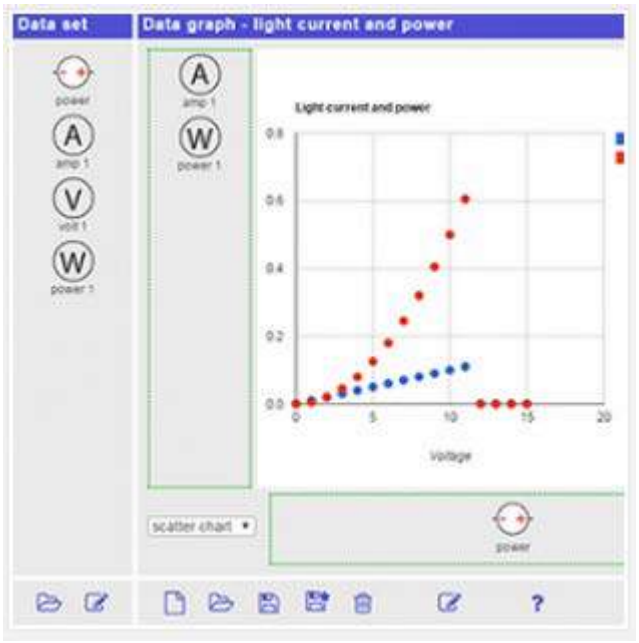


Figure 20. Data Viewer tool.

Overall evaluation

The tool seems to provide important and valued functionality. However at least some users find it difficult to use. It would benefit from some online help – perhaps a video demo – or, better still, some usability improvements. (The limited frequency counts below may be partly because some users did not explore it fully.)

Detailed findings

Table 22. Data Viewer tool findings

Id	Usability observation	Frequency	Category	Recommended / Agreed Response
ADV01	Difficulty understanding how to create a new data set	B - Some	UR	Provide clearer instructions. **Work in progress.**
ADV03	Unclear how to extract data from Splash Lab into Data Viewer tool. Clicking on "Load Dataset" does not seem to work.	C - Occa	UR	Make it clearer whether this is possible and if so how to perform it.
ADV04	Data from lab not available as they were not saved. (Y3PD20, electric circuit lab)	C - Occa	UR	Consider making it clearer to the user how and why they should store lab outputs.
ADV05	Confusion between the open button in the data set and the data graph space	C - Occa	UR	Provide clearer instructions **Under consideration**
ADV02	Saving in the Data Viewer tool took very long or did not work at all.	C - Occa	BF	This issue needs to be fixed (if it can be reproduced). **Already fixed.**
ADV06	Creation of a wrong graph (same variable on both axes)	D - Rare	O	Teachers may want to allow this as part of learning from mistakes. **No action planned – students**

Id	Usability observation	Frequency	Category	Recommended / Agreed Response
				learn through mistakes.
ADV07	One participant could not retrieve his or her experiment data in the Data Viewer tool and gives that it was not saving properly before as possible explanation.	D - Rare	BF	This issue needs to be fixed (if it can be reproduced).
ADV08	Additional functionality is needed for the Data Viewer tool (e.g. “Would be good if things like regression lines and such can be added on top of the data. As Well as giving more information for things like r^2.”)	D - Rare	FE	Possibly add additional features to the Data Viewer tool. **Under consideration**
ADV09	Because one of the earlier tools did not work for him or her, one participant could not use the Data Interpretation tool.	D - Rare	O	This issue needs to be fixed (if it can be reproduced).
ADV10	One of the features of the Data Viewer tool did not work, the participant assumed maybe because of too many users using it at the same time.	D - Rare	BF	Obscure what this means. No action suggested.
ADV11	The red colour of the error message in the Data Viewer tool clashes with the purple colour of the tool.	D - Rare	UR	Seems like this has been fixed in the meantime (error message is no longer red).
ADV12	Error message in the Data Viewer tool is only partly visible and cannot be moved into view.	D - Rare	BF	Seems like this has been fixed in the meantime (error message is no longer red).
ADV13	It is not clear if the interpretation entered in the textbox is saved (or that it is auto-saved).	D - Rare	UR	Make it clear for the user, that his or her entry is instantly saved, e.g. by showing a small auto-save animation in the corner of the tool when the text is changed.
ADV14	Users cannot add a title to their graph in the Data Viewer tool.	D - Rare	FE	Add a feature to the Data Viewer tool that allows the user to specify a title to their graph. **Under consideration**
ADV15	Table of Data Set is squashed in the Data Viewer tool and hard to use.	D - Rare	UR	Make the Data Set area in the Data Viewer Tool larger. **Work in progress.**
ADV16	One participant could not find his or her data set.	D - Rare	O	This issue needs to be fixed (if it can be reproduced).

4.6.7 File Drop tool

Description

This app allows students to upload files, e.g., assignment and reports, to the Inquiry learning Space. The app also allows teachers to download the uploaded files.

Screenshot

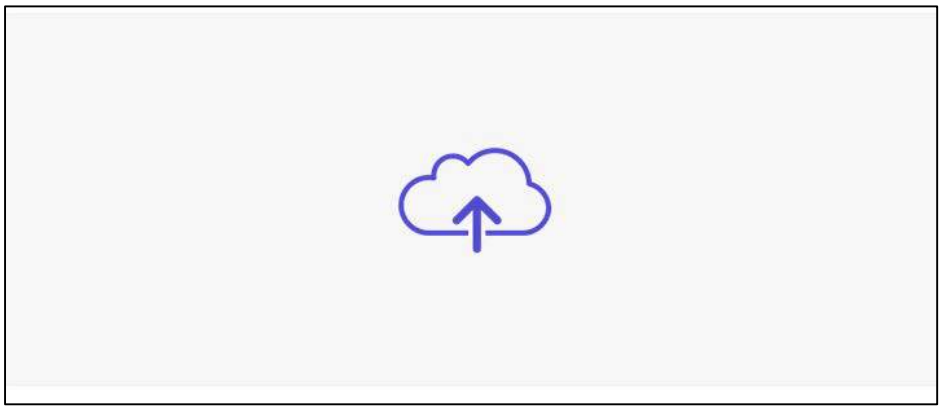


Figure 21. File Drop tool.

Overall evaluation

This tool fulfils an important function, and seems to have very few usability issues. Some of the observations below seem to have arisen from a temporary problem or perhaps an earlier version. The new version of the app resolves some of the observations.

Detailed findings

Table 23. File Drop tool findings

Id	Usability observation	Priority	Category	Recommended Response
AFD01	It is unclear what to do with the File Drop in this phase (upload something instead of download conclusion?).	B - Some	UR	Maybe separate File Drop in two different scaffold apps, one to upload and one to download files. **Resolved in new version**
AFD02	Content of textbox can sometimes not be changed, after submitting a file.	D - Rare	BF	Observation unclear and rare – no action at present.
AFD03	ILS learning content states that the expert conclusion can be found in the File Drop, but it is not there.	D - Rare	O	Observation unclear and rare and not reproducible – no action at present.
AFD04	To access the resource in the File Drop, the participant is asked to log in to the system (to Graasp?).	D - Rare	O	**Fixed**
AFD05	There are no restrictions on the file type that can be uploaded using File Drop.	D - Rare	FE	File Drop should restrict the file types that can be uploaded. **No action – working as designed.**
AFD06	One participant suggests to remove the sub-heading “Resources in the space:” if there are no resources.	D - Rare	UR	Separating the upload area from the empty resources area seems to be valuable to structure the tool. Collect additional feedback from users on this matter before changing

Id	Usability observation	Priority	Category	Recommended Response
				anything. **No action – working as designed.**

4.6.8 Conclusion tool

Description

The Conclusion Tool allows learners to check whether the results of experiments in the form of data graphs and/or observations support their hypotheses from the hypothesis tool, or are relevant for the questions posed in the question tool.

Screenshot

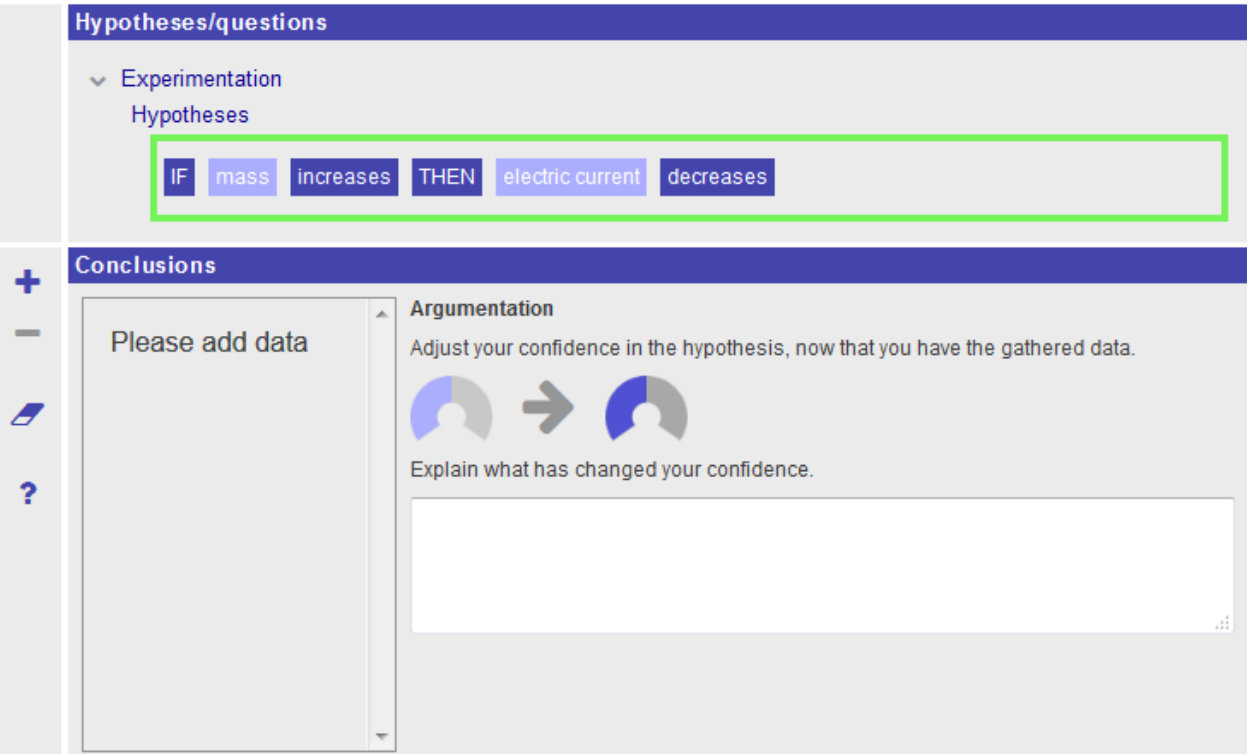


Figure 22. Conclusion tool.

Overall evaluation

This tool is valued pedagogically by teachers, and was found to be relatively simple and intuitive to use, though with slight scope for improvement. Feedback was slightly limited as it comes at the end of an ILS and not all users reached the end in timescales available.

Detailed findings

Table 24. Conclusion tool findings

Id	Usability observation	Frequency	Category	Recommended / Agreed Response
ACT01	User didn't realise that there were different hypotheses under different ILS phases (conceptualisation and experimentation), as the first	C - Occa	UR	Rather than opening one of the phases initially have them both collapsed so that neither draws the eye and the user simply chooses which one they want.

Id	Usability observation	Frequency	Category	Recommended / Agreed Response
	one was open and my eye was drawn to it.			**Dev work planned.**
ACT02	User didn't initially realise that different hypotheses had entirely separate sections of evidence and conclusions.	C - Occa	UR	Perhaps number the hypotheses and conclusions, so we see hypothesis 1 and hypothesis 2 and that we are working on conclusion 1. **Dev work planned.**
ACT03	The "Play" symbol on the button to load a Hypothesis in the Conclusion Tool, does not match the action it performs.	C - Occa	UR	Use a different symbol, e.g. a check mark, instead of the play symbol. In fact Go-Lab would be improved if the play symbol were eradicated throughout except in video playback. **Fixed and tested.**
ACT04	Even in the help for the conclusion tool, the fact that observations are not universally applied to all the hypotheses is made clear at the end!	D - Rare	UR	Rather than "Select a hypothesis" try "Select the first hypothesis which you have studied." Rather than "add one or more data graphs and/or observations…" try "add one or more data graphs and/or observations pertinent to the current hypothesis…".
ACT05	Hypotheses are listed as "unnamed hypotheses" but there doesn't seem to be a way to name them	D - Rare	UR	If there is no way to name them, remove this quantifier. If there is, make it more obvious / put it in the hypothesis scratchpad help dialogue
ACT06	It is possible to completely cover words in the hypothesis with the tooltip, which is not telling us anything of value	D - Rare	UR	Improve location of tooltip
ACT07	Locating saved hypotheses in the Conclusion Tool takes quite a long time.	D - Rare	UR	Make it more obvious, how to find and load saved hypotheses in the Conclusion Tool. **Fixed and tested.**
ACT08	Current way to retrieve a Hypothesis is disliked, a combobox to select from is proposed as alternative solution or making it feel more like a button.	D - Rare	UR	As the current way fits the overall design of the tool, collect additional feedback from users on this matter before changing anything. **Fixed and tested.**
ACT09	Hypotheses are not always shown in the load list.	D - Rare	BF	Cannot reproduce. Possibly fixed (or mis-reported). **Fixed and tested.**

4.6.9 Reflection tool

Description

The Reflection Tool gives feedback to students about their use of an Inquiry learning Space (ILS). The tool displays the percentage of time a student has spent in the various inquiry phases compared to a norm set by the teacher (see image). Students are prompted to reflect on their ILS use by a number of questions.

Screenshot

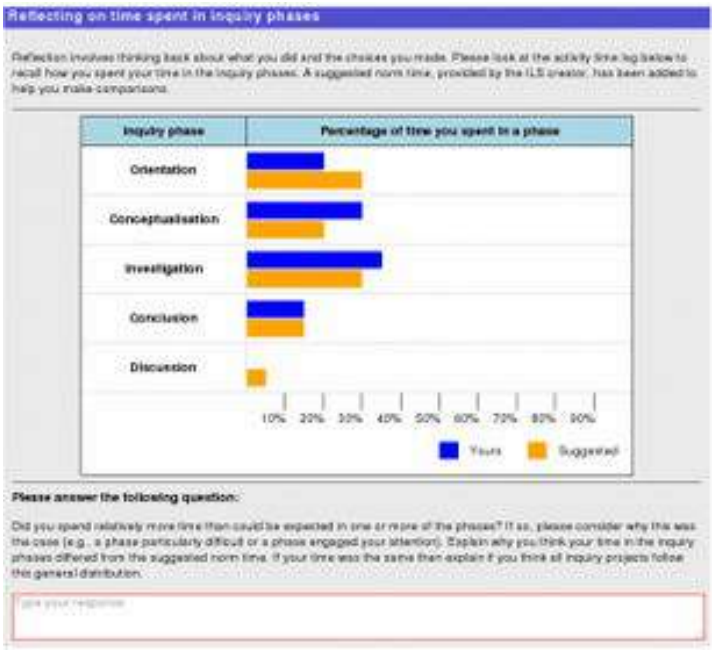


Figure 23. Reflection tool.

Overall evaluation

This tool attracted several rather disparate observations, little enthusiasm from students and slightly more from teachers. There is no obvious pattern to the findings. (During studies, because this is used at the end of an ILS, in some cases the session was getting rather hurried by this stage and feedback may therefore have been rather limited).

Detailed findings

Table 25. Reflection tool findings

Id	Usability observation	Frequency	Category	Recommended Response
ART01	The Reflection Tool takes too long to load or does not load at all.	C - Occa	O	This issue needs to be fixed (if it can be reproduced).
ART02	"This may be a technical glitch but it didn't allow me to type a reflection"	D - Rare	BF	? Not reproducible. Possibly a one-off glitch? Check this works okay. If so, no action.
ART03	“Refresh arrow” in the Reflection Tool is not noticeably clickable to access data.	D - Rare	UR	Make it more obvious that the refresh arrow image is a button, or refresh the content automatically, when the user enters this phase.

Id	Usability observation	Frequency	Category	Recommended Response
ART04	The bars in the reflection tool sometimes are longer than 100%.	D - Rare	BF	This issue needs to be fixed.
ART05	It is not clear for the participants, where the suggested times comes from.	D - Rare	O	Provide information, where the suggested times come from.
ART06	Reflection Tool does not show the time spent for all the phases (only for one, not for orientation and investigation, …).	D - Rare	BF	This issue needs to be fixed (if it can be reproduced).
ART07	Reflection Tool should explain why more time should have been spent in different phases.	D - Rare	O	Provide information or explicitly ask the student to also reflect on that.
ART08	One participant lost the text entered in the Reflection Tool (“I typed my response and then pressed something else and it disappeared so I had to type it again”).	D - Rare	BF	Not reproducible. Possible user error?
ART09	Colour scheme of Reflection Tool is not aesthetically pleasing.	D - Rare	AP	As the current colour scheme matches the Go-Lab colour scheme, no change appropriate unless extra feedback becomes persuasive.
ART10	The suggested times for each phase in the Reflection Tool make the students feel like they did something wrong, if they have not spend the right amount of time on a section.	D - Rare	O	Provide information or explicitly ask the student to also reflect on that.
ART11	Reflection Tool lacks indication if saving is necessary and how it is done.	D - Rare	UR	Make it clear for the user, that his or her response is instantly saved, e.g. by showing a small auto-save animation in the corner of the tool when the text is changed.
ART12	The times shown in the Reflection Tool seem to be inaccurate at times.	D - Rare	O	Unclear what action to take. Retest first.

No responses have yet been received from the development team to these (mostly rare) findings.

4.6.10 Quiz Master tool

Description

This tool allows teachers to prepare a multiple choice quiz using a Google spreadsheet, and incorporate it into an ILS, so students can respond to the quiz.

Screenshot

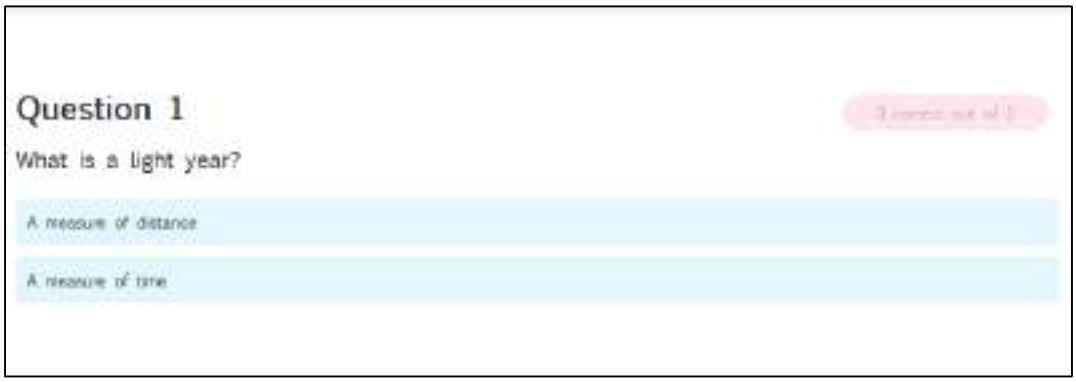


Figure 24. Quiz Master tool.

Overall evaluation

The student experience of the tool is generally very satisfactory. Some teachers would prefer a simpler interface for creating quizzes, and also possibly a more versatile tool. Some have prior experience of other online quiz tools such as HotPotatoes, Moodle (SCORM), Google module, Quiz Faber, Flubaroo etc. There were also occasional concerns about installing Google Drive and using unfamiliar spreadsheet software.

Detailed findings

Table 26. Quiz Master tool findings.

Id	Usability observation	Frequency	Category	Recommended / Agreed Response
AQz01	Should be much easier for teacher to use, without having to go into Google apps and use confusingly similar to but different from Excel	B - Some	UR	If possible, a purpose-designed tool without needing a Google logon would be beneficial. **Done – not yet user-tested.**
AQz02	Should support different sorts of quizzes, not just multiple choice.	C - Occa	FE	Useful possible enhancement. **Under consideration but beyond current scope.**
AQz03	In the Orientation phase in the Quiz tool it would be beneficial to have instructions informing a user to click on the correct event. Students tend to click on the first one and the other ones remain in the shade.	D - Rare	O	Unclear (UTE, 16/4/2015) and cannot be reproduced. **No further action for now.**

4.7 Inquiry Learning Spaces (ILSs)

The number of ILSs published on the Portal has grown to 152 this year, and the portfolio is constantly changing as teachers create or modify ILSs. It would clearly be impractical and also of limited value to study every ILS, since it would spread resources very thinly yielding only superficial results. Additionally most of the findings on specific ILSs tend to be about specific pedagogical content. Any usability issues tend to apply to all or many ILSs, not just specific ones.

Furthermore, in many cases WP3 partners have created special purpose ILSs (often based on published ILSs but customised) for use in their studies, to maximise the research benefits from the time available, to explore specific research questions and to suit the time available, the class available and their learning needs. Such comments on detailed aspects of particular ILSs have been fed back directly to the author and would unnecessarily complicate and extend this document.

The formative evaluation findings for ILSs are therefore summarised in two categories:

- ILS delivery – already presented at 0 above.
- ILS general design guidance below.

The ILS delivery feedback is of relevance to the development cluster. The ILS content / ILS design feedback is passed to the pedagogical cluster to contribute towards ILS design guidelines for teachers. However some comments could be resolved by either design or development, so both tables should be read by both audiences.

The ILSs which have been most commonly used in studies which contribute to these findings include:

- Craters on earth and other planets
- Electricity – an alternative approach of Ohm’s law
- Splash Inquiry Space
- Is it Good to be Beautiful - Understanding Evolution through Natural and Sexual Selection
- Series and parallel circuits
- Splash – sinking and floating
- How are the light fixtures in a house connected?
- Is Radioactivity always harmful for humans?
- GearUp

or in some cases translated versions of these. Additionally, in some events teachers have been invited to explore ILSs of interest to them and to provide feedback.

4.7.1 ILSs – general design feedback

Results are displayed in Table 27.

Table 27. ILS Design findings

Id	Usability observation	Frequency	Category	Recommended Response
IDes01	Large chunks of text inappropriate for some classes	A - Freq	O	ILS Design guidance - make suitably engaging for target audience
IDes02	Java required for ILS completion but not compatible with some browsers	A - Freq	O	ILS Design guidance - use only features which exist on all standard browsers, OR document any special browser / plugin requirements when publishing ILS
IDes03	Background image is currently not good: - distracting and blurry	A - Freq	AP	The background picture needs to be not distracting and of high resolution. The default background image should comply to this as well and an option for “no background image” (plain white background) should be available. As background image can be selected by the teachers, the teacher guidelines/help should ask them to pick one that is not distracting and of high resolution, if they like to change it from the default one.
IDes04	Playback of “Video 1” has been disabled on other websites than YouTube, thus the student has to leave the ILS to watch it. They would prefer to stay in the ILS instead.	A - Freq	O	Include only YouTube videos into the ILS, that can be played back inside the ILS (add this suggestion to the teacher guidelines/help material as well).
IDes05	Spelling mistakes	B - Some	UF	correct them
IDes06	It needs to be clearer where to find the note taking tool and how to access it.	B - Some	UF	ILS instructions should be clear on where to find tools when it is located in a pull down bar at the bottom of the screen. Alternatively the pull down tool bar should be much more visible. Instead of “Tools” the label could actually tell which tools can be found when opening the “Tools area” (e.g. use a label like “access note taking tool and calculator” instead) or an icon could be used their and to reference to it in the ILS learning text.
IDes07	Some ILSs are slow or ineffective at displaying YouTube videos in some browsers due to compatibility issues with Shockwave	B - Some	BF	Fixed? Browser compatibility issue. Reported Feb 2015; not occurred recently.
IDes08	Broken links & "404" messages	B - Some	O	ILS design advice - only use links to locations with stable trustworthy content
IDes09	Plugin compatibility issues	B - Some	O	ILS design advice - document any special browser requirements when publishing.

Id	Usability observation	Frequency	Category	Recommended Response
IDes10	ILS using inappropriate language for the age of students.	B - Some	O	Design guidance
IDes11	Participants think that the text layout of sub-headings could be emphasized.	B - Some	UR	One possible improvement would be to make sub-section headings stand out more, e.g. by underlining them.
IDes12	Tabs on top of the ILS should not need and have scroll buttons to the right and left	C - Occa	UF	Teacher guidelines should advise them to limit the number (and labelling) of phases to an amount that can be displayed without scrolling if possible. (The scrolling option currently implemented seems to be the most flexible one regarding customization.)
IDes13	Colour scheme should be improved.	C - Occa	UF	Try out different colour schemes to find one that is appealing to most of the participants. Teachers could get a feature to specify a custom colour scheme. Collect additional feedback from users on this matter before changing the colour scheme.
IDes14	Concern that an ILS with links in may not work as expected if external content changes.	C - Occa	O	Can't think of a feasible remedy to this, except to advise ILS authors to link only to stable trustworthy supported content.
IDes15	Inappropriate vocabularies in apps (e.g. hypothesis scratchpad, concept mapper etc.)	C - Occa	O	Design guidance - always configure apps before publishing ILS
IDes16	Font type and text colour are currently not appealing to some of the participants and regarded as unprofessional. It makes them want to stop reading rather than continuing.	C - Occa	AP	ILS design guideline - make text appealing to the expected users
IDes17	Having only tabs on top does not introduce the page enough.	C - Occa	UR	Add an introduction tab (as the first tab) or text above the tabs.
IDes18	The current titles of the videos (“Video 1/2/3”) are not very meaningful.	C - Occa	O	Add a meaningful title behind the labels “Video 1/2/3”.
IDes19	Description about the video and its content is missing, e.g. in case it cannot be loaded.	C - Occa	UR	Add a description of the content of every video to the ILS.
IDes20	“Tools” bar is perceived as too big.	D - Rare	AP	It could be a little bit smaller.
IDes21	ILS title tooltip states that the ILS description can be shown/hidden by clicking on the title, but nothing happens.	D - Rare	BF	If there is a ILS description show/hide it on click, if there is none, either remove the tooltip.
IDes22	It is not always clear what the tools in the toolbar are used for.	D - Rare	UF	Tools in the toolbar should have a title and/or description explaining their purpose and usage.

Id	Usability observation	Frequency	Category	Recommended Response
IDes23	Spacing between paragraphs is quite big, leaving a lot of white space	D - Rare	UF	Change text layout to have smaller gaps between paragraphs.
IDes24	One participant did not think that the note taking tool is a “tool”.	D - Rare	UF	The participant suggested to put it straight on the bottom of the page, not in a “tool area”.
IDes25	"Could we have access to ILSs written by other teachers?"	D - Rare	FE	Now provided (Publishing ILS to Portal). No action required.
IDes26	Could we add some way for student to communicate with teacher - particularly when they are using remotely?	D - Rare	FE	Possibly worth considering as an extra feature. Many teachers envisage remote use of ILSs.
IDes27	ILS usage requires student to download and install a software item	D - Rare	O	ILS design guidance - be aware that some users may not have admin rights to install software. If requiring this, make it clear when publishing ILS.
IDes28	ILS using old version of certain apps	D - Rare	O	New versions of Apps should be compatible and use the same url as the old one.
IDes29	Hypotheses should be closer to lab so you can see them while you do the experiment	D - Rare	UR	May be a possible usability enhancement.
IDes30	It is not clear enough where and how to keep notes of the results of the brain storming about concepts.	D - Rare		Make description in the ILS more precise.
IDes31	Too many videos can get boring.	D - Rare	O	Put more information in text, instead of adding too many videos.
IDes32	Content is perceived as insufficient, e.g. one participant did not know what he or she is supposed to do in this phase.	D - Rare	O	Make description in the ILS more detailed.
IDes33	It was not clear how and where to add notes.	D - Rare	UR	The notes tool should be directly underneath the respective video.
IDes34	One participant stated the interface could be made more navigable.	D - Rare	UR	He or she suggested to use audio effects to make the interface more navigable.

Response

These findings are shared with the pedagogical cluster to consider incorporating into the ILS design guidelines for teachers.

4.8 Sample quantitative findings

The main purpose of WP3 in Year 3 was to advise development teams on recommended areas to improve usability, as presented above. However the studies also included a certain amount of quantitative evaluations, both of components and of broader areas of acceptance. These provide ongoing advice to the project. As before, there has been far too much to present in full in this document, but some sample material from studies designed, conducted and analysed in collaboration with WP6 are presented here.

4.8.1 Overview

An integrated WP3-WP6-WP8 questionnaire has been administered to teachers participating in Practice Reflection Workshops (PRW) organized by the Go-Lab National Coordinators in their respective countries. The questionnaire was aimed to capture the teachers' opinions of different components of the Go-Lab Portal. Among the 14 questions, Q3 and Q4 (Figure 25) were used to evaluate how teachers perceived the usability and user experience of the Go-Lab Portal. Results of analysing Q3 and Q4 are presented in this section and results of the other questions are included in the deliverables of the other WPs (e.g., WP6).

Note that PRW were organized throughout Year 3 while the Go-Lab Portal and its components were evolving. Hence, participants of PRW held at different times interacted with prototypes of different levels of maturity. Furthermore, *not* all participants of PRW responded to the integrated questionnaire. Hence, the total number of PRW participants differs from the total number of respondents of this questionnaire.

3. In the table below you will find 10 pairs of contrasting attributes. Where you place your choice between two attributes indicates your view about the quality of [the Go-Lab Portal]

For example:

Disagreeable ○ ○ ○ ○ ◉ ○ ○ Likeable

This choice tells us that the Go-Lab Portal is somewhat likeable, but there is still room for improvement (Note: *There is no right or wrong answer. Your personal opinion is what counts.*)

3.1 Confusing	1	2	3	4	5	6	7	Structured
3.2 Practical	1	2	3	4	5	6	7	Impractical
3.3 Predictable	1	2	3	4	5	6	7	Unpredictable
3.4 Simple	1	2	3	4	5	6	7	Complicated
3.5 Dull	1	2	3	4	5	6	7	Captivating
3.6 Stylish	1	2	3	4	5	6	7	Tacky
3.7 Cheap	1	2	3	4	5	6	7	Premium
3.8 Unimaginative	1	2	3	4	5	6	7	Creative
3.9 Good	1	2	3	4	5	6	7	Bad
3.10 Ugly	1	2	3	4	5	6	7	Beautiful

4. For each of the following statements, please indicate your extent of agreement by circling the number of choice

Statement	Strongly Disagree						Strongly Agree
1. [The Go-Lab Portal's] capabilities meet my requirements.	1	2	3	4	5	6	7
2. Using [the Go-Lab Portal] is a frustrating experience.	1	2	3	4	5	6	7
3. [The Go-Lab Portal] is easy to use.	1	2	3	4	5	6	7
4. I have to spend too much time working with [the Go-Lab Portal].	1	2	3	4	5	6	7

Figure 25. **Two Questions on Usability and User Experience**.

4.8.2 Instruments

Q3 is an abbreviated version of AttrakDiff2 (Hassenzahl et al., 2003) with 10 items being extracted from the original 21. The psychometric properties of both versions have been established (Hassenzahl & Monk 2010). The instrument AttrakDiff2 has been built upon Hassenzahl's (2003) user experience model. Accordingly, *features* (i.e., content, presentation, interaction and function) of a system contribute to users' perceptions of its hedonic quality (HQ), pragmatic quality (PQ), aesthetic quality (Beauty) and overall quality (Goodness). As shown in Figure 25, the first set of four items are for evaluating PQ, the following four are for HQ, one item for Beauty (ugly-beautiful) and one for Goodness. A 7-point semantic differential scale is used for each pair of contrasting adjectives. The items are mixed in terms of their positive/negative anchors, and reversal is required for further analysis.

Q4 is based on a more recent scale called Usability Metric for User Experience (UMUX) by Finstad (2010), who argued that it could be an effective proxy of the System Usability Scale (SUS) (Brooke, 1996) – a scale widely used in usability research and practice. UMUX is a standardized instrument claimed to conform better to the ISO definition of usability (Finstad, 2010) than SUS. We have slightly adapted the original 4 items by replacing the word "system" with "Go-Lab Portal".

The rationale of including the abbreviated AttrakDiff2 and UMUX is that the former focuses on the users' perceptions of the quality of the system whereas the latter focuses on the user's perceptions of their performance and feeling resulting from interacting with the system. Both perspectives are closely related as well as complementary.

4.8.3 Results

Altogether 229 participants responded to both Q3 and Q4. Based on the activities they were involved in PRW, participants were classified in one of the three groups:

A: I am only searching for online labs on the Lab Repository to use them in my lesson.
B: I am using existing Inquiry Learning Spaces (ILS) provided by the Go-Lab team or other teachers.
C: I am an experienced Go-Lab user. I am creating new Inquiry Learning Spaces (ILS) for my lessons.

Table 28 shows the distribution of participants per group per country. 45.9% of the participants are in Group B whereas 26.6% and 27.5% are in Group A and Group C, respectively.

Table 28. Distribution of participants per country per group

Group	Belgium	Cyprus	Estonia	Germany	Greece	Netherlands	Portugal	Spain	Other*	Total
A	29	12	0	4	2	2	7	3	2	61
B	8	36	49	2	0	2	4	3	1	105
C	3	17	1	4	12	1	8	16	1	63
Total	40	65	50	10	14	5	19	22	4	229

***** Other = 1: Israel, 2: Macedonia, 1: Russia

AttrakDiff2

Table 29. Descriptive statistics of AttrakDiff2 constructs (N = 229)

Construct	Item	Mean	Std. Dev.
Pragmatic Quality (PQ)	Confusing - Structured	5.08	1.276
	Impractical - Practical	4.63	1.632
	Unpredictable - Predictable	4.33	1.282
	Complicated - Simple	4.28	1.493
Hedonic Quality (HQ)	Dull - Captivating	5.06	1.130
	Tacky - Stylish	4.56	1.469
	Cheap - Premium	4.94	1.315
	Unimaginative - Creative	5.38	1.068
Goodness	Bad - Good	5.05	1.631
Beauty	Ugly - Beautiful	5.31	1.160

The ratings for the four items of PQ and those for HQ were aggregated and averaged over all participants:

PQ: Mean = 4.58, SD = 0.97, CI-95% = 4.45 - 4.71;
HQ: Mean= 4.99, SD = 0.82, CI-95% = 4.88 - 5.10.

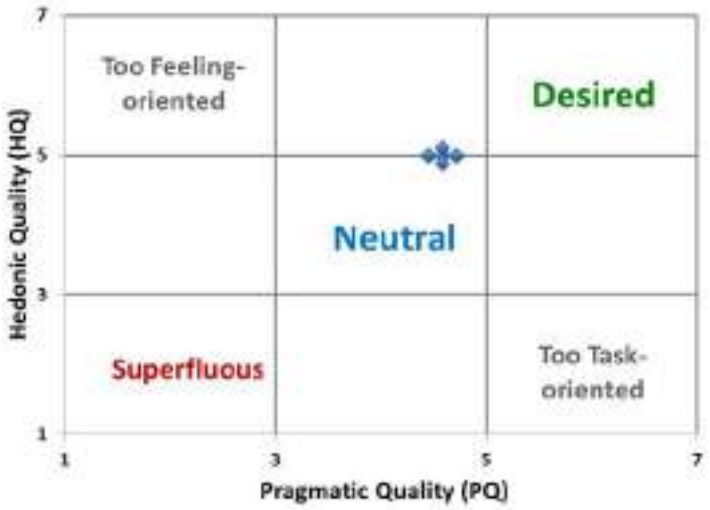


Figure 26. Perceived hedonic and pragmatic quality of the Portal

According to the standard analysis of AttrakDiff (http://attrakdiff.de/index-en.html), in plotting HQ vs. PQ, we can position the system under evaluation. The upper rightmost quadrant "Desired" represents the most ideal status of a system with a well-balanced PQ and HQ, enabling users to fulfil their goals with the system and enjoy the usage at the same time. The Go-Lab Portal is close to this status with its position at the upper end of the middle quadrant "Neutral", bearing in mind that the responses collected were averaged over the prototypes of different maturity.

We were interested in finding out whether participants in Group A, B and C differed significantly in their perceptions of the Go-Lab Portal. Presumably, Group C had the deepest level of interaction with the system than Group B, which in turn had a deeper level of

interaction than Group A. ANOVA with the four constructs as dependent variables and group as independent variable was performed (Table 30). There were significant difference in Beauty ($F(2,226) = 3.05$, $p<.05$) and Hedonic Quality ($F(2,226)=3.96$, $p<.05$). Independent t-tests show that there are significant difference in Beauty between Group A and C ($t(122) =2.47$, $p<.05$), suggesting that Group C appreciated the aesthetic quality of the Portal more. There are also significant differences in Hedonic Quality (HQ) between Group A and C ($t(122) = 2.64$, $p<.01$) and between Group A and B ($t(164)=2.45$, $p<.05$), implying that it was more likely that participants found the Portal enjoyable to use when they interacted with it at a deeper level.

Table 30. Average ratings of the four constructs of AttrakDiff2 for the three groups

Group	Goodness	Beauty	Pragmatic Quality	Hedonic Quality
A	5.05	5.03	4.57	4.74
B	4.92	5.33	4.55	5.07
C	5.25	5.54	5.64	5.10

UMUX

The four items of UMUX measure ease of use and three prototypical usability metrics.

Item 1 – Effectiveness
Item 2 – Satisfaction
Item 3 – Ease of Use
Item 4 – Efficiency

As shown in Table 31, the average ratings of Efficiency, Satisfaction and Ease of Use were above the neutral value of 4.0 whereas the average rating of Efficiency was below 4.0. The findings suggest that participants could achieve their goals with the Go-Lab Portal with ease and satisfaction, but they needed to spend quite some time in accomplishing them.

Table 31. Average ratings of the four constructs of UMUX for the three groups

Group	Effectiveness	Satisfaction	Ease of Use	Efficiency
A (N=61)	4.57	5.34	4.43	4.24
B (N=105)	4.74	4.56	4.45	3.59
C (N=63)	4.63	5.14	4.92	4.14
All (N=299)	*4.66*	*4.93*	*4.57*	*3.92*

We were also interested in finding out if participants in Group A, B and C differed significantly in the four aspects. Results of ANOVA show that there are significant differences in Satisfaction ($F(2,226) = 5.25$, $p<.05$) and in Efficiency ($F(2,226) = 5.05$, $p<.05$) among the groups. Independent t-tests show that Group A and B have highly significant differences in Satisfaction ($t(164)=2.88$, $p<.01$) and Efficiency ($t(164) = 2.85$, $p<.01$) and that Group B and C have also significant difference in these two constructs, though to a less extent ($t(166) = 2.19$, $p<.05$; $t(166) = 2.41$, $p<.05$). The observations imply that among the three groups, Group A had a strongest tendency to be satisfied with the Portal and to feel that they could complete their tasks efficiently. It is consistent with the fact that in general the time they spent in using the Portal was less than the other two groups, given the relatively lower level of interaction.

5. Discussion

As reported above, a wide variety of PD events with teachers and students have been conducted for WP3 in Year 3. 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. We also include a few reflections on the effectiveness of our study programme and study methods.

5.1 Synopsis and summary of the findings of PD Year 3

The Go-Lab infrastructure – much extended and improved since Year 2 – has been thoroughly tested in multiple studies. The hugely increased and changing portfolio of artefacts (Labs, Apps and ILSs) on the Portal have been tested in varying levels of scrutiny, focussing most effort on those which are most frequently used or which teachers believe they are most likely to use frequently.

The huge body of data which has been collected has been very diverse in content as well as in type. The main types of data have been subjective self-reports by students and teachers (e.g. on online or paper questionnaires, flipcharts, worksheets, facilitated discussions, interviews and collected online with PDot etc). We have also collected more performance based data by observation, analysis of video recording or system-captured performance data, timing of tasks etc. Our measures are not of uninformed opinion, nor merely measuring responses to well-rehearsed presentations; all studies involve the participants in practical hands-on use of the system before responding to questions.

The appropriate use of three complementary approaches (field-based study, remote studies and analytical studies) as well as the sampling across age groups, subject areas and nationalities, has given a very broad base to the findings.

In general the findings have been that the infrastructure is very much improved and largely of good quality in functionality and usability. Some big areas of infrastructure improvement since Year 2 have been authoring, configuration facility for apps, app composer (for translating apps), the Portal, the provision of resources for Help, Support, Tutorials and Tutoring, During the year the Go-Lab infrastructure has also been improving in terms of reliability and quality, and also portability to different browsers and devices, and it is expected that these areas of improvement will continue.

The portfolio of resources (labs, apps, ILSs) on the Portal has improved dramatically in volume and therefore usefulness, and most of the well-established resources have also improved in usability and user experience. A more consistent user interface and interaction design would be even better, to help students focus on learning and not technology, though it would be hard to achieve with components from different providers.

The findings for Year 3 suggest that teachers not only see lots of potential in Go-Lab but also that they want to use it, and are eager to see further improvements and extensions to the system.

There is evidence in the data that teachers are beginning to think of Go-Lab seriously as a vehicle for delivering lessons rather than as an interesting visionary idea, and this makes them more demanding and more imaginative in their responses and questions. They ask if

it's free of charge and will always be free, and always available. They are increasingly keen on the practicalities of using this in a variety of real school and homework situations. They make suggestions for how the system could be improved rather than just criticising it. The tone of their responses is often requesting improvement with a sense of urgency, rather than as a detached intellectual critique.

In particular there is evidence that teachers would like the system to have more resources (ILSs and labs) available "off the shelf" and closely aligned to their curricula; they would like all aspects of the system to be rendered well on a variety of devices and browsers, utilising screen space well and supporting effective resizing and scrolling; they would like even better usability with consistent use of interaction paradigms and icons across all components; they want facilities to assess/mark their students' work.

Some challenges - including some beyond the control of the project - remain. Some schools across Europe have limited IT facilities – e.g. shared computer rooms which have to be booked, poor or older computers, computers with a "locked down" software portfolio, erratic internet connections etc. Some teachers also have limited experience of delivering lessons using PCs, and a level of fear about it and worry about how to manage a classroom and keep children focussed when they are using PCs. Teachers often say they want everything to work instantaneously and speedily. They want systems to be secure and private but many are averse to any extra logon processes. Some new challenges have arisen for web-based software during the year as browsers have become less tolerant of Java and some other plugins, and this has meant some of the externally supplied labs do not run on all platforms.

5.2 Effectiveness of formative evaluation programme

The purpose of formative evaluation is to influence development priorities leading to a more usable and useful product and better user experiences. In this section we revisit some of the main findings reported a year ago in D3.2 and note progress made; we give some examples of progress made during the year, and we summarise the responses to the usability recommendations recently published in precursors to this document.

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.

5.2.1 Interplay between evaluation and (re)development in Year 3

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.

Findings have regularly been reported more frequently but less formally than this document, and influenced development priorities during the year. Appendix D contains an example of an email response from the development cluster to a specific area.

5.2.2 Review of main recommendations from Year 2

A very considerable body of development work has taken place since D3.2, much of which has provided great improvements to usability, usefulness and user experience. Some of the main infrastructure changes which have improved the system from WP3 perspective have been:

- Implementation of new authoring facility which is dramatically more intuitive and accessible, and makes it much easier to demonstrate and "well" Go-Lab convincingly
- Provision of a wealth of Help and Support and Tutorial materials and strategies
- Implementation of the vault
- Implementation of App Composer enabling apps to be translated and shared
- Implementation of app configuration tool enabling apps to be customised by the ILS author
- Feature allowing ILS author (optionally) to require students to use password based log on method
- Consistent use of Autosave
- New range of "learning analytics" for teacher to monitor progress of class through an ILS
- Extension of support to include all three main windows-based browsers
- Dramatic increase in the range of labs available

There have also been many changes to individual components which have also contributed.

In the rest of this section we revisit some of the main recommendations from D3.2, published a year ago, and demonstrate the excellent progress which has been made, and in a few cases note scope for further progress.

Table 32. Progress on general findings from D3.2 Section 3.3.1 October 2014

Finding title	Summarised description	Progress by September 2015
Overall impression	Users believe system 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.	General perception that system is much more refined and fuller with content, though not market ready.
Other tools reported as used	It has become apparent that many teachers in various contexts use other online tools in teaching situations.	Increased awareness of alternative products within team. No specific action required.
Sign-on security	Some concerns about ILS signon security – privacy, potential for plagiarism or inadvertent corruption of someone's work – exacerbated by a bug.	New feature to allow ILS author to require password protection. (Not yet tested for WP3). Bug no longer reported.
Storage and retrieval of work	Frequent questions on where a user's work is stored and how it can be retrieved were raised. 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.	Vault is considerably improved, and auto-save implemented. There is still high demand from teachers in tools to extract student work from vault for marking purposes.
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	Fuller implementation of Vault has made big improvements to data flow and to reduce retyping, though still not all data passes automatically from tool to tool. UNDO and REDO only partially implemented.
Scaffold toolbar visibility	Pull down toolbar not easy to find unprompted. Also some tools in the toolbar appear too small for use in some circumstances. The facility to make them wider is not visible or not recognisable.	Fewer reports of this though no design change noted. Sizing and scrolling of components on various devices should still be improved.
Consistent interaction design	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.	There appears to be some convergence but some remaining inconsistencies (e.g. arrows in Concept Mapper). Difficult to get all external labs to conform to a single usability style quickly.
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. Ideally use consistent and well recognised icons throughout, with either text labels or too tips.	Very considerable progress on help and support, following establishment of user support task force. (And tool tips widely implemented too).
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, and to extract data for student assessment (marking).	A good range of online classroom progress monitoring tools have been added, as well as a consolidated concept map. There is still a perceived need for facility to extract student work for marking purposes.

Finding title	Summarised description	Progress by September 2015
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.	No major technical changes to address this, but the pedagogical cluster's work on scenarios will help.
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 some "decorated" text in ILSs appeared to be hyperlinks but wasn't	This has in part been addressed by improved ILS design (and ILS design guidelines). Fewer reports this year. A further possible improvement would be for every lab or app to have very clear visual boundaries – e.g. be in a box.
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.	Average learning time has been added to ILS description on the Portal.
Multi-lingual components	A frequent comment has been that tools are not currently available in the participant's own language.	App composer and also app configuration tool now allow translation of most elements.
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,	This still needs to be addressed. Made more difficult as this year more diverse devices (e.g. tablets, very old computers) have been encountered.
Browser issues	Some components of the system seem to be 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 I.2LS) would take the user all the way back to the start of the ILS and cause the user to lose her work.	Code enhanced to properly support recent versions of Internet Explorer as well as FireFox and Chrome. Some bugs eliminated. A new issue has appeared concerning compatibility of browsers with Java-based and (to a lesser extent) Flash-based labs.

Table 33. Progress on summarised specific recommendations from D3.2 Section 6.2

	Recommendation	Progress by September 2015
SR01	Adopt consistent usability paradigms across the system, wherever possible based on intuitive or well established interaction modes.	There appears to be some convergence, but also some remaining inconsistencies (e.g. arrows in Concept Mapper). Difficult to get all external labs to conform to a single usability style quickly.
SR02	Provide appropriate (informative and engaging) online help facilities for all components.	Very considerable progress on help and support, following establishment of user support task force. This includes an online manual, narrated demonstrations on a video channel, tutorials, a user support forum and availability of online tutoring sessions. ILS authors have a facility to customise the Help text for the tools they include in their ILS.
SR03	Scaffold tools which contain vocabulary should be customisable by the teacher or ILS author.	Fully addressed by app configuration facility. Very considerably improvement.
SR04	Make user-supplied data consistent between components and flow naturally between them.	The implementation of the Vault has provided most of this. Externally supplied labs are more difficult.
SR05	Implement auto-save throughout the system, and UNDO and REDO features.	Autosave implemented everywhere. UNDO and REDO in places.
SR06	Provide exemplary scientific rectitude in all material.	There have still been a few comments about graphs without properly labelled axes, or measurements without 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.	The EDT app has been much improved, and this may have been the main source of these comments as they have not re-emerged in Y3.
SR08	Provide facilities to ensure students cannot accidentally or deliberately access or change someone else's work.	Feature added to allow ILS author the choice of whether to require or not require passwords.
SR09	Make all components available in all languages for target demographic.	App composer and configuration tool allow apps to be localised. ILS authoring allows any language. The Portal now has ILSs in far more languages. Many labs are largely visual, but some are language specific.
SR10	Consistency across browsers.	Full support has been extended to include all recent versions of the main three Windows browsers (Chrome, Internet Explorer and 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.	Implementation of the Vault has laid the foundations, but extract facilities also needed.

5.2.3 Summary of development responses to Year 3 findings

Observations in the tables of usability observations have been analysed. Where an observation was made more than once (i.e. not "rare") and a response from the Technical Cluster seemed appropriate, the actual response was classified into whichever of the following categories best matched the response.

Table 34. Development team response codes

Code	Response category	Description
SCV	Software already changed and verified by end users	Component already changed to remedy the usability observation, and remedy verified by subsequent user studies
SCN	Software changed – not yet retested	Component already changed to remedy usability observation; remedy not yet verified by subsequent user studies
SCIP	Software change – work in progress	Component change to remedy this usability observation is in progress; not ready for testing.
SCP	Planned	Change to remedy this usability observation is planned.
SCW	Willing	Development teams agree that a change should be implemented to remedy this usability observation.
SCC	Considering	Development teams recognise the issue and are considering whether or how a change could be designed and implemented to remedy it.
QV	Questioning value or importance	Development teams are not yet decided whether the work involved in remedying this usability observation is justified. Include issues where user requirements needs clarifying or corroborating, or not reproducible issues, and issues which need more consensus before making a change.
NA	Do not intend to address the issue at this stage	Development teams believe the usability benefits of remedying this observation do not justify the costs. E.g. minor problem, low usability benefits, may do more harm than good, functional extensions beyond scope etc.
AS	Non-technical alternative solution planned instead	Dev cluster believes the issue is best addressed without software fixes - e.g. by learning material, help material, or ILS design.
II	Infrastructure Issue	Problem is inherent in infrastructure elements beyond the scope of the project (e.g. browsers, plug-in support, hyperlinking)
SA	Sustainability Agenda	The issue is beyond the scope of current project but worth considering for inclusion in a sustainability agenda
Disc	Still in discussion	Not yet agreed how or whether to resolve the issue

The frequencies are presented in Table 35. These are based on usability observations which were reported more than once, for which a change in software was the recommended response and which have been presented to the development team in time for them to provide a response.

Table 35. Development team response frequencies

Code	Response category	%
SCV	Software already changed and verified by end users	20.6
SCN	Software changed – not yet retested	9.5
SCIP	Software change – work in progress	5.6
SCP	Planned	4.0
SCW	Willing	8.7
SCC	Considering	14.3
QV	Questioning value or importance	8.7
NA	Do not intend to address the issue at this stage	12.7
AS	Non-technical alternative solution planned instead	4.8
II	Infrastructure Issue	2.4
SA	Sustainability Agenda	2.4
Disc	Still in discussion	6.3

This is further summarised in Figure 27.

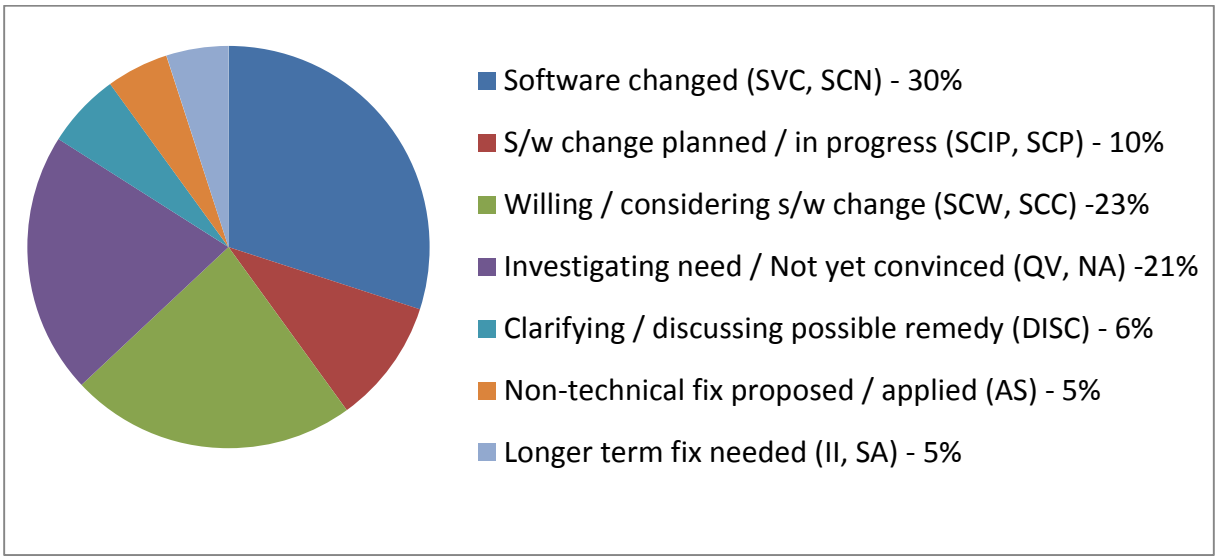


Figure 27. Summary of development responses

This is another illustration of the extent to which usability studies are helping development to improve the usability of the system. Discussions of these matters have always been cordial and productive.

5.3 Analytical evaluation – methodological reflections

While the evaluation team were conducting the evaluation and also discussing usability scoring, we recognised that – although we try to imagine the usability experience of a teacher rather than our own – the process is somewhat subjective, and partially influenced by our own capabilities and preferences for interaction styles and paradigms. The more time we have spent observing science teachers attempting to use technology, the more insight we have into their experience and the less influenced we are by our own capabilities and preferences.

We identified the following quite subjective factors which to varying extents influenced the usability issues we identified and the importance assessments:

- The user's task, purpose, objectives, support structures, age, prior experience, capabilities, preferences, context of use, motivation levels, sense of urgency, distraction levels, frequency of use etc. Also how resilient, persistent and motivated we think the user is; whether they prefer traditional or novel interaction paradigms; how important specific issues are for them – e.g. privacy and confidentiality, aesthetically pleasing design, personal productivity, etc.
- Whether we empathise with those experiencing usability difficulties or whether we believe they should become more technologically literate.
- How seriously we consider different potential user responses to usability obstacles (e.g. briefly slightly puzzled, mild frustration, leave it and try again later, abandon the interaction, abandon the product)

We also identified some more pragmatic factors which influenced our varied assessments of an issue's importance:

- Expected data volumes. In some cases (e.g. whether and how scrolling or paging should be provided) the data volume is a key factor.
- How frequently an issue might be encountered, e.g. difficult to understand text in a rarely encountered error message would be less important than difficult to understand text in the main screen.
- How important a function is to the user's main purpose in using the system, e.g., a problem with adding tags to an artefact to help others find it may be less important than a problem creating the artefact.

5.4 Findings from different user populations and study methods

In the Year 2 studies and analysis, we compared the effectiveness of different study methods, and sampling from different populations, and reported findings in D3.2. This informed the methods we adopted and the way we deployed our resources in Year 3. Based on a sample of the raw data from Year 3, it appears that the observations we made in Year 2 can be reiterated with little modification.

5.4.1 Comparing findings from teachers and students

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

- Some students are very tolerant of what teachers perceive as poor usability features, especially novel user interface elements
- 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.

5.4.2 Comparing analytical evaluations with end user studies

When comparing the analytical evaluation (or Heuristic Evaluation) 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.
- 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.

5.4.3 Broadening our sampling – secondary perspectives

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".

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.

6. Conclusion and Outlook

6.1 Achievements

A very considerable body of field-based end-user research activities have taken place this year, engaging a very diverse sample of over 450 teachers and over 550 students from 18 countries in 50 studies – some of which involved multiple visits. Complementing this, we also conducted 13 remote studies with an engaged Core Group of Teachers, and 4 analytical evaluations.

Recruitment of teachers and schools has been considerably more successful this year than last year, perhaps due to the increasing quality and credibility of the system as well as the developing relationships with teachers' communities.

This has generated a huge body of data on the usability and usefulness of the system and the user experience, as well as, inter alia, some data on related areas such as performance, reliability, bugs, aesthetic preferences etc. These data have been translated into English where necessary, analysed, and shared with the Development cluster throughout the year, resulting in many system improvements. This data, including some new findings from recent studies, is also summarised by component for this document.

The main outputs from WP3 are software change recommendations provided to the Technical Cluster (WP4 & WP5). Other outputs are ILS design recommendations for teachers, provided to WP1 and Lab selection recommendations provided to WP2**.** In addition, in Year 3 some more far reaching change recommendations are provided to WP9 as input to the sustainability agenda.

All change recommendations are discussed, and proposed remedies agreed with the relevant development teams. The underlying data from studies is stored on Graasp and shared with consortium partners.

6.2 Challenges

Recruitment of schools continues to be challenging, though as noted above, much more successful this year than last year.

Field study is very worthwhile as it provides more realistic and profound insights into how the system would be used in situ, but it also means less control of the agenda. Good planning of sessions is beneficial but researchers still sometimes have to improvise when the unexpected occurs. Technical problems – often but not always caused by the school's infrastructure – can be one source of such challenges. So can issues such as pupil behaviour and mood – e.g. just before exams or holidays.

Thirdly, in designing and conducting a study, there are balances to be struck between a variety of stakeholders and agendas; the teachers, the students, the HCI researchers and also the reputation of the product and project. For instance, teachers may be concerned for education and curriculum; students for enjoyment; HCI researchers for unbiased feedback and the wider project for positive dissemination and enhancing the reputation of the product and project. There is sometimes a trade-off between showcasing the best parts of the system (for maximal reputational benefit) and allowing free exploration and an openness to critique

(to maximise usability feedback). The Heuristic Evaluations and the use of the Core Teacher Group have helped us considerably in meeting these challenges.

6.3 Limitations

As noted earlier, the portfolio of labs, apps and ILSs on the Portal has grown huge and been in constant flux. As a result, some of these components – especially the more recently created ones – have been tested by relatively few participants perhaps not including every perspective to be found in the target domain. (The Go-Lab infrastructure and the more established or frequently used components have been thoroughly tested by a great many participants).

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 co-ordination of studies 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 have used materials which are generic and easily adapted to cope with such situations.

6.4 Outlook

The outputs from WP3, including but not limited to this document, continue to influence the Technical Cluster's agenda, leading to a system of increasing usefulness and usability and with an increasingly positive user experience. The responsiveness of the Technical Cluster to WP3 findings, the progress so far and the plans for further improvements suggest that the system will continue to improve in usability and usefulness and provide ever more engaging and beneficial user experiences.

Some of the more far reaching, visionary or complex recommendations are passed on to the sustainability programme – specifically WP5 deliverable D5.7 and WP9.

Findings on ILS usability are also feeding in to the guidelines for ILS Designers provided by WP1. Findings on the lab portfolio and on general lab usability are passed on to WP2 to be used as part of their lab search and selection criteria.

Given the huge, fluid and growing body of components on the Portal, and the multiple authorship and ownership, and policy to allow open publishing of ILSs by teachers, the content cannot be fully policed forever. Instead WP5 are developing "social" approaches for high quality artefacts to be clearly visible, and in some cases gaining online badges of approval. Using this approach, coupled with guidelines on writing good content, it is expected that high quality ILSs will be incentivised and will gain highest visibility on the Portal.

The formative evaluation agenda of WP3 is now completing, and in Year 4 the perspective will shift towards dissemination, community building and summative evaluation. Facilities are increasingly being provided for users to provide feedback online – e.g. using DISQUS. Effort will be made to encourage the teaching communities who have been engaged in WP3 to continue to use Go-Lab as part of their teaching practice, and to continue to engage with other Work Packages.

7. References

Ballabio, E and Moran, R. (1998). Addressing the needs and potential of older people and people with disabilities in the information society: an RTD Approach for the European Union, Brussels: EC DGXIII, Telecommunications, Information Market and Exploitation of Research, 1998

Bannon, L. (2011). Reimagining HCI: toward a more human-centered perspective. *Interactions, 18*(4), 50-57.

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 *Proceedings of CHI'11* (pp. 2689-2698). ACM.

Bassey, M. (2008). *Case study research in educational settings*. Maidenhead, Open University Press.

Boehm, B. W. (1988). A spiral model of software development and enhancement. *Computer*, 21(5), 61-72.

Brooke, J. (1996). SUS – A quick and dirty usability scale. Usability evaluation in industry, 189 - 194.

Cooper, Alan. *The Inmates Are Running the Asylum: Why High Tech Products Drive Us Crazy and How to Restore the Sanity*. Sams Pearson Education (2004)

Crabtree, A. (1998). Ethnography in participatory design. In *Proceedings of the 1998 Participatory Design Conference*, 93-105.

Earthy, J., Jones, B. S., & Bevan, N. (2001). The improvement of human-centred processes—facing the challenge and reaping the benefit of ISO 13407. *International Journal of Human-Computer Studies, 55*(4), 553-585.

Eveleens, J. L., & Verhoef, C. (2009). The rise and fall of the chaos report figures. *IEEE Software* (1), 30-36.

Farrell, V., Farrell, G., Mouzakis, K., Pilgrim, C. & Byrt, P. (2006). PICTIOL: a case study in participatory design. In *Proceedings of OzCHI*. Sydney, Australia, Nov. 191-198.

Finstad, K. (2010). The usability metric for user experience. *Interacting with Computers*, 22(5), 323-327.

Gould, J. D., & Lewis, C. (1985). Designing for usability: key principles and what designers think. *Communications of the ACM, 28*(3), 300-311.

Harrison, S., Sengers, P., & Tatar, D. (2011). Making epistemological trouble: Third-paradigm HCI as successor science. *Interacting with Computers, 23*(5), 385-392.

Hartson, R., & Pyla, P. S. (2012). *The UX book: process and guidelines for ensuring a quality user experience*. Elsevier.

Hassenzahl, M., & Monk, A. (2010). The inference of perceived usability from beauty. *Human–Computer Interaction, 25*(3), 235-260.

Hassenzahl, M., & Tractinsky, N. (2006). User experience-a research agenda. *Behaviour & Information Technology, 25*(2), 91-97.

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., Burmester, M., & Koller, F. (2003). AttrakDiff: Ein Fragebogen zur Messung wahrgenommener hedonischer und pragmatischer Qualität. In *Mensch & Computer 2003* (pp. 187-196). Vieweg+ Teubner Verlag.

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

Iivari, J., & Iivari, N. (2011). Varieties of user-centredness: an analysis of four systems development methods. *Information Systems Journal, 21*(2), 125-153.

Jordan, P.W., 1998. Human factors for pleasure in product use. *Applied Ergonomics, 29* (1), 25–33.

Landauer, T.K.,1997. Behavioral research methods in human-computer interaction. In: Helander, M., Landauer, T.K., Prabhu, P. (Eds.), *Handbook of Human Computer Interaction*. Elsevier Science, Oxford, pp. 203e227.

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.

Lindgaard, G., Dillon, R., Trbovich, P., White, R., Fernandes, G., Lundahl, S., and Pinnamaneni, A. (2006). User Needs Analysis and requirements engineering: Theory and practice. *Interacting with Computers, 18*:47-70

Martin, B., & Hanington, B. M. (2012). *Universal Methods of Design: 100 Ways to Explore Complex Problems, Develop Innovative Strategies, and Deliver Effective Design*. Rockport Publishers.

McCarthy, J., & Wright, P. (2004). *Technology as experience*. MIT Press.

Millen, L., Cobb, S., & Patel, H. (2011). A method for involving children with autism in design. In *Proceedings of the 10th International Conference on Interaction Design and Children* (pp. 185-188). ACM.

Muller, M. J., & Druin, A. (2010). Participatory Design: The third Space in HCI. *Handbook of HCI*. J. Jacko and A. Sears. Mahway.

Muller, M., and Druin, A. Participatory design: The third space in HCI. In: Sears, A., and Jacko, J (eds.) *The Human-Computer Interaction Handbook* (3rd Edition). Lawrence Erlbaum, Hillsdale, 2012, 1125-1154.

Muller, M. J. (2002). Participatory design: The third space in HCI. In J. A. Jacko and A. Sears (Eds.), *The Human Computer Interaction Handbook: Fundamentals, Evolving Technologies and Emerging Applications* (pp 1051–1068). Mahwah, NJ: Lawrence Erlbaum.

Muller, M.J., Wharton, C., McIver, W.J. & Laux, L. (1997). Toward an HCI Research and Practice Agenda Based on Human Needs and Social Responsibility. In *Proceedings of CHI '97 Human Factors in Computing Systems*, Atlanta.

Mynatt, E. D., & Rogers, W. A. (2001). Developing technology to support the functional independence of older adults. *Ageing International, 27*(1), 24-41.

Neale, H., Cobb, S., & Wilson, J. (2002). A front ended approach to the user-centred design of VEs. In *Proceedings of Virtual Reality*, *IEEE* (pp. 191-198).

Nielsen, J. (1994). *Usability engineering*. Elsevier.

Norman, D. A., & Draper, S. W. (1986). *User-centered design. New perspectives on human-computer interaction*. Lawrence Erlbaum.

Norman, D. A. (1998). *The Invisible Computer*. MIT Press, Cambridge, MA.

Norman, D. A. (2002). *The design of everyday things*. Basic books.

Norman, D.A. (2004). *Emotional Design*. New Books, New York.

Parsons, S., Millen, L., Garib-Penna, S., & Cobb, S. (2011). Participatory design in the development of innovative technologies for children and young people on the autism spectrum: the COSPATIAL project. *Journal of Assistive Technologies, 5*(1), 29-34.

Pathirage, C., Amaratunga, R. D. G., & Haigh, R. P. (2011). The role of philosophical context in the development of theory: Towards methodological pluralism. *The Built & Human Environment Review, 1*.

Radermacher, H., L (2006) Participatory Action Research with People with Disabilities: Exploring Experiences of Participation. School of Psychology. Accessed 10 March 2008 http://wallaby.vu.edu.au/adt-VVUT/uploads/approved/adt-VVUT20060904.160428/public/02whole.pdf, Victoria University.

Rogers, Preece & Sharp (2007). *Interaction design: beyond human-computer interaction*, (2. Ed.). West Sussex: John Wiley & sons, Ltd,

Roto, V., Law, E. L., Vermeeren, A. P., & Hoonhout, J. (2010). User experience white paper. Retrieved from http://www.allabout.ux

Simonsen, J., & Kensing, F. (1998). Make room for ethnography in design! *The Journal of Computer Documentation, 22*(1), 20-30.

Simonsen, J., & Robertson, T. (Eds.). (2012). *International handbook of participatory design*. Routledge.

Sixsmith, A., & Sixsmith, J. (2000). Smart care technologies: meeting whose needs?. *Journal of telemedicine and telecare, 6*(suppl 1), 190-192.

The Standish Group International Inc (1994). Chaos. Technical report, The Standish Group International Inc.

Thüring, M., & Mahlke, S. (2007). Usability, aesthetics and emotions in human–technology interaction. *International Journal of Psychology, 42*(4), 253-264.

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.

Vermeeren, A. P., Law, E. L. C., Roto, V., Obrist, M., Hoonhout, J., & Väänänen-Vainio-Mattila, K. (2010, October). User experience evaluation methods: current state and development needs. In *Proceedings of the 6th Nordic Conference on Human-Computer Interaction: Extending Boundaries* (pp. 521-530). ACM.

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.

Wright, P. C., & Monk, A. F. (1991). A cost-effective evaluation method for use by designers. *International Journal of Man-Machine Studies, 35*(6), 891-912.

8. Appendices

Appendix A: Example PD protocols and study materials

A.1 Protocols and materials for a face to face teacher workshop

The materials below were created for a Scientix conference held in York, UK in April 2015. Two workshops were available of 2.5 hours each, and with 20-30 teachers expected at each. The "running order" document was for use by researchers only.

University of Leicester Computer Science
Scientix Conference – York April 2015 - Go-Lab Teacher Workshop

Objectives

- Showcase Go Lab
- Enthuse teachers and network with them to **gain hot contacts**
- Collect usability & related data:
  - To improve Go-Lab
  - To show in deliverables
  - PDot data for M PhD

How

- Distributing promotional material (flyers, pens etc)
- Networking - talk to teachers about Go-Lab at every opportunity
- Workshop:
  - Present brief PPT on purpose, role & relevance of Go-Lab
  - Live demo of portal (& exploration by teachers?)
  - Demo Student experience of a whole ILS
  - Teachers experience part of an ILS – a lab + give PDot feedback?
  - Demo authoring (+ simultaneous hands-on?)
  - Discussion
  - Present possible next steps & gather contact details & willingness to take it further

Data collection

- PDot
- Worksheets – including tick boxes for further involvement (& contact info)
- Observer notes (including perhaps approximate count of "traffic light" questions)
- Photos
- Post-its & flipchart data from group working
- Others?

Resources

- PowerPoint for opening intro (memory stick?)
- Laptop with appropriate browsers etc installed for demos
- Worksheets – 60 – with team numbers
- Pdot passwords and links
- Promotional materials: Go-Lab standard
- ULeic flyers
- Pre-written flipchart sheets for discussion section
- Prints of PowerPoints with speaker notes

Preparation:

If possible, get to the room in advance and do as much of the following as possible:

- Check out all technology including projection works with your laptop
- Organise seating etc. as you like it
- Provide each place with a worksheet and possibly also a flyer, traffic lights, post-its, free Go-Lab pen etc.
- Boot laptops, connect to Wi-Fi if possible, check out browser works with the hands-on activity; maybe even open PDot links in a tab
- Organise pre-prepared flipcharts for discussion section in an appropriate place – if there's enough room then in a separate place to where people will be seated.
- Brief one of the team to use a clock to keep presenters to the schedule.
- Ask in advance if it's okay to take our own photos.

Roles:

- P1 & P2 – present as noted below
- Someone else – keep an eye on clock and indicate to presenters when they are 1 minute from end of slot
- All – tech support / advice & containment especially in practical sessions
- All – support a group in discussion sections
- All – encourage completion of last section of worksheet
- All – take observer notes, especially of traffic light questions
- All – soon afterwards write down remembered observations & reflections

Agenda

1. Arrivals + brief PPT on purpose and role of Go-Lab – P2 - **10 mins**
2. Live demo of portal – P1 – **5 minutes?** + some brief opportunity for live feedback (traffic lights)
3. Student experience – demo some whole ILS — P2 – **10 minutes +** traffic lights?
4. Hands-on + feedback: - P1 – **30 minutes**
    a. Pdot intro
    b. Teachers use a lab and provide feedback with PDot
5. Demo authoring (teachers may follow in real time?) – P1 – **15 minutes +** traffic lights?
6. Discussion – P2 and P1 - in groups + feedback– **15 minutes**
    a. Maybe 4 groups of 10? Each with a ULeic facilitator
    b. Maybe hand each group a flipchart page pre-written with 4 boxes: what was best, what was worst, what questions remain, any other comments
    c. Get individuals to write >= 1 post-it for each box
    d. Group discuss / prioritise
    e. **Collect & photograph flipchart sheets**. (To ensure they stop this task and focus on the next one)
7. Conclusion: – **P2 – 5 minutes**
    a. Mention a couple of themes from priorities comments
    b. Remind them that we can offer further free sessions, online support and workshops in their own schools or our campus including CPD and INSET days
    c. **Get them to complete last section of worksheets (further contact) and collect as they leave (+ facilitators reinforce the message)**
    d. Thank for participation

Detailed Running Order

11:00-11:10 / 13:30 – 13:40 Introduction to this session and to Go-Lab concepts - P2

Detailed speaker notes in PowerPoint. Also:

- Go-Lab is EU initiative to support science teaching across Europe by providing online laboratory experiences packaged into a pedagogical environment with supportive scaffolds.
- Software is work in progress but much of it is ready to use.
- Growing body of teachers across EU planning and beginning to use it to deliver science lessons.
- Our objective in sessions like today is to let teachers see it, work out whether and how it could be useful to you, and let you provide feedback as to what you would like to see in it so we can make it more relevant and useful.
- We hope you will be inspired to adopt this technology as part of your teaching practice, but as a minimum we are sure you'll be able to make an informed decision.
- We also like to engage with students in classes delivered using Go-Lab.
- We have a worksheet with various questions which you may answer as we go on.
- If we're okay to take our own photos, ask teachers if they're happy to be photographed.

11:10-11:15 / 13:40 – 13:45 Live demo of Go-Lab portal - P1

Show them:

- the url
- three main tabs: labs, apps, ILSs + mention that there are searching, sorting and filtering options
- the "drill down" feature to actually see the lab / app / ILS
- point out but don't explore the tutorial tab and DIY icon

Ask some of the traffic light questions (Set 1)

11:15-11:25 / 13:45-13:55 Demonstration of a Go-Lab ILS – P2

- **Electricity (a.m.) / pH scale (p.m.)**
- Mention that different phases are provided to support a pedagogical model of inquiry based learning
- Ask some of the traffic light questions (Set 2)

11:25 – 11:55 / 13:55 – 14:25 Teachers try out a selected lab and provide feedback – P1 + team

- Explain – we are going to ask you to use one of the online labs and also to use an online tool called PDot to provide feedback on the lab and suggest any problems or possible improvements. First I am going to explain how to use PDot and show you how to use it to critique the design of a mobile phone.
- Intro to PDot
- **Lab: "BOND" (a.m.) / "Craters" (p.m.) / Splash (backup in case of problems)**
- **All four researchers help!**

11:55 – 12:10 / 14:25 – 14:40 Demo of Authoring – P1

Depending on how long this takes in rehearsal, you may be able to invite the teachers to follow this in real time.

- Explain – for lessons delivered on Go-Lab, you can either use a prewritten ILS, clone and modify one, or write your own from scratch! We will now show the latter method.
- From GoLabz, click on DIY icon
- Use Graasp screen to create a new logon (mention that something called Graasp is part of the infrastructure)
- Use + symbol and click on create inquiry space
- Explain the general concept of a learning space having several phases represented by tabs, each phase being composed of different sections of text, pictures, videos, apps, a lab, ...
- Then demo how to get into a space, how to add different types of things:
    - photo (upload a file from your PC)
    - text (Add resource without specifying a link gives you the option to enter text),
    - YouTube video (paste the URL to the video to be included, then change the description automatically generated from YouTube content),
    - apps (mention but don't demonstrate that they can be customised by teacher at this stage)
    - labs (Explain how to identify an appropriate lab on the GoLabz website first and then find it in the pull-down list of labs),

  and mention that other things can be included too.

- Show how to get a url to distribute to students so they can use ILS to learn
- Mention that you can also publish your ILS on portal if you'd like other teachers to have access to it

Ask some of the traffic light questions (Set 3)

12:10 – 12:25 / 14:40 – 14:55 Group discussion – P2 & P1 + team

- Create teams – suggest pre-written number 1 to 4 on everyone's worksheet
- Get them to gather round flipchart sheets (also numbered)
- Ask them each to complete at least one post-it for each of the four questions (best , worst, questions, comments)
- Discussion – people pool post-its in each category on to flipchart and try to prioritise
- **Collect & photograph flipchart sheets**. (To ensure they stop this task and focus on the next one)

12:25 – 12:30 / 14:55 – 15:00 Conclusion – P2

- Thank them for being engaged throughout the session
- Two more quick items before we conclude:
  - Future opportunities: state that we can offer free workshops in your school or our campus; online support, help you build ILSs, visit your classes the first time you deliver using Go-Lab, and make improvements to the resources and facilities. This can include CPD (Continuing Professional Development) sessions and also support for school INSET days. **Please complete the last section (further engagement) of the worksheet. Whole team mingle and reinforce the message.**
  - Thank them for feedback on flipcharts and – if possible – mention a couple of themes already identified, or answer one of the top questions
- Thanks again and do keep in touch.

Collect & label all materials, worksheets, flipcharts etc

The following worksheet was given to participants to complete during the session and collected afterwards.

Team:

University of Leicester Computer Science
Go-Lab Workshop – 18th April – Scientix event

Teacher Worksheet

GO-Lab – Global Online Science Labs

Welcome to this Go-Lab workshop.

The programme we envisage is outlined below, along with some questions which you may choose to answer as we go along, and some instructions for the practical sessions. We will happily answer any pressing questions you have as we progress. Other questions can be noted on the last page of this worksheet and contribute to the discussion at the end.

We would like to keep these forms at the end as one of the methods we will use to understand how to make Go-Lab better for yourselves and other teachers.

It would help us if you would answer the following two questions:

What ages are the pupils you teach? ……

What subject(s) do you teach? ……

1. Introduction to this session and Go-Lab concepts – P2

2. Live demo of Go-Lab resources portal - P1

Questions

Do you have experience of delivering lessons in which student learning takes place through PCs?
A LOT / A LITTLE / NONE

Do you believe this can be an effective way to deliver science learning? YES / POSSIBLY / NO

Do you believe online labs can provide benefits for science teaching? YES / POSSIBLY / NO

Do you believe online labs can provide benefits for your own lessons? YES / POSSIBLY / NO

If no, then why not. If yes, then what benefits and in what circumstances?

3. Demonstration of a Go-Lab ILS – P2

Questions

Do you think that this ILS environment, with several phases of learning, and a mixture of text, pictures, videos, labs and scaffold apps, can be an effective way to deliver science learning?
YES / POSSIBLY / NO

Please note down any comments, concerns or questions arising

4. Exploring the student experience of a lab – P1

This will be the hands-on part of the workshop. P1 will explain what to attempt. If you have any questions or difficulties, please ask. Our team will be happy to help.

5. Creating a Go-Lab lesson with the Authoring Tool – P1

Creating an ILS is like creating a lesson plan but using online tools. For lessons delivered on Go-Lab, you can either use a pre-written ILS, clone and adapt one, or write your own from scratch! We will now show the latter method.

We can offer workshops or support which focusses specifically on this area. This session is a brief demonstration only. The 5 phases presented are designed to support Inquiry-based learning paradigm. When creating an ILS teachers can choose whether or not to use all these phases.

P1 will show how to add each of the following types of entity:

- photo
- text
- YouTube video
- app
- lab

Various other components can be added, and there are facilities for the teacher to customise some of the apps to be more relevant to a particular lesson plan.

Questions

Would you envisage writing your own ILS using these facilities? YES / POSSIBLY / NO

Do you believe these facilities are sufficiently user-friendly for your own use? YES / POSSIBLY / NO

Do you believe these facilities are sufficiently user-friendly for many science teachers to use effectively? YES / POSSIBLY / NO

If not, how could we improve them? (Note down the highest priority things)

Do you have any comments, concerns or questions about this process?

6. Reflection / Discussion – P2 & P1

We are interested in whether you would like to use Go-Lab either soon or later? Are there any practical obstacles you can foresee? Also do you think other science teachers would like to use it? And finally: What we can do to make it more useful and attractive?

- From what you've seen today, please write a post-it about each of the four areas:
  (a) (GREEN post-it) What was best / most encouraging / most useful looking?
  (b) (RED post-it) What was worst / most discouraging / difficult?
  (c) (YELLOW post-it) What questions are you left with?
  (d) (BLUE post-it) What comments do you wish to make?
- Find the flipchart sheet with the team number indicated on the front page this worksheet, and work with others in that team to collate your responses and try to prioritise them.

7. Conclusion and next steps - P2

Future opportunities. We can offer further workshops and events of various sorts, free of charge, in schools, other venues or on our campus. These can contribute to teachers continuing professional development. Some schools have used them on INSET days. We are also very keen to see how this works with students and are happy to provide support in preparation and delivery of lessons. We can also keep you up to date with news about further developments of Go-Lab.

If you would like to be kept up to date or engage further with Go-Lab, please supply contact details.

Urgent request: We are urgently looking for a class of 24 science students aged 15 or over to undertake a 1.5 hour Go-Lab based science learning activity this Summer – perhaps after their exams. **If you are able to offer such an opportunity or know of someone who can, do please let us know.**

Contact details (optional) email:.....

Phone: .....

Please indicate whether you would like:

to attend a workshop ☐

to have a workshop at your premises ☐

to attend an online tutorial ☐

to register for fortnightly emails ☐

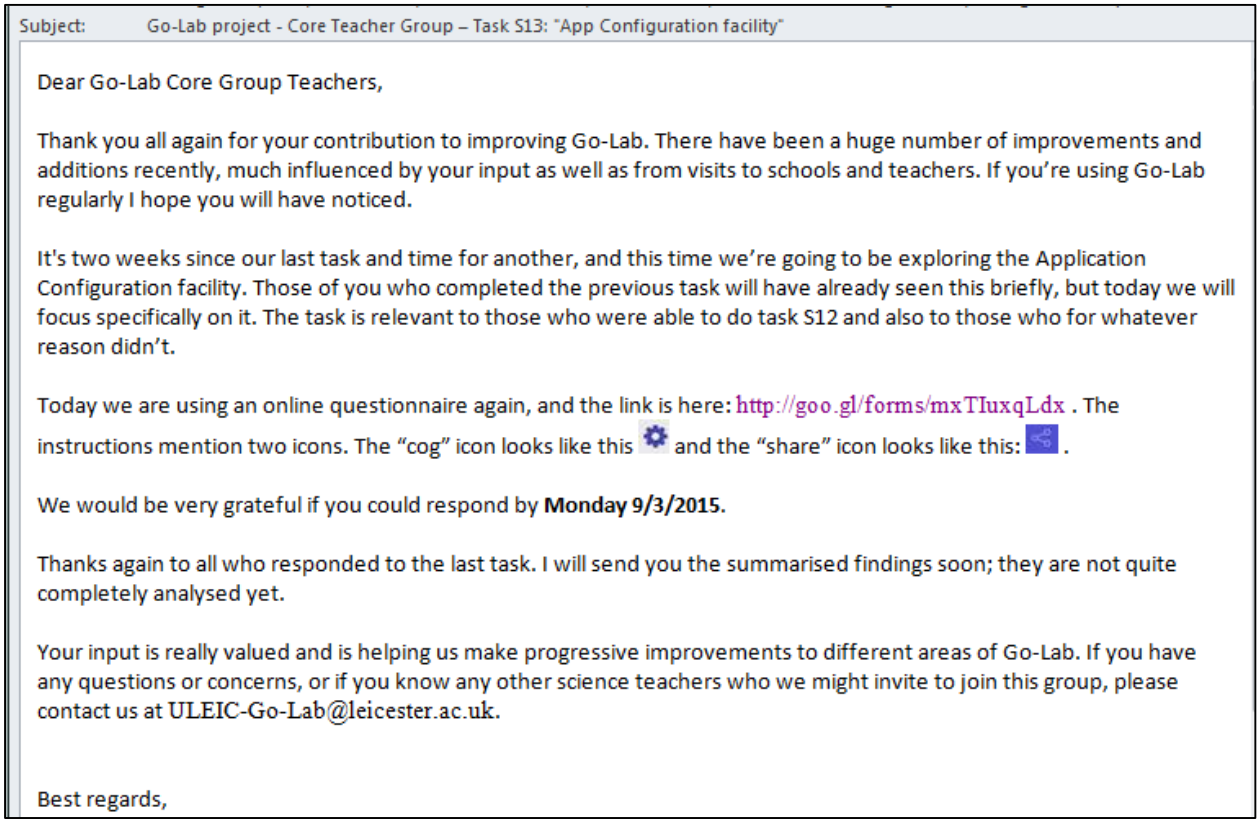
further information on ongoing improvements to Go-Lab ☐

Can you help us to locate a class of 15-18 year old students this summer for our proposed activity? YES / NO

The session also used a PowerPoint presentation and some prepared flipcharts which are not included here.

A.2 Protocols and materials for a remote study

The documents attached below show a typical Remote study.
The participants receive the following email:



Subject: Go-Lab project - Core Teacher Group – Task S13: "App Configuration facility"

Dear Go-Lab Core Group Teachers,

Thank you all again for your contribution to improving Go-Lab. There have been a huge number of improvements and additions recently, much influenced by your input as well as from visits to schools and teachers. If you're using Go-Lab regularly I hope you will have noticed.

It's two weeks since our last task and time for another, and this time we're going to be exploring the Application Configuration facility. Those of you who completed the previous task will have already seen this briefly, but today we will focus specifically on it. The task is relevant to those who were able to do task S12 and also to those who for whatever reason didn't.

Today we are using an online questionnaire again, and the link is here: http://goo.gl/forms/mxTIuxqLdx . The instructions mention two icons. The "cog" icon looks like this and the "share" icon looks like this: .

We would be very grateful if you could respond by **Monday 9/3/2015.**

Thanks again to all who responded to the last task. I will send you the summarised findings soon; they are not quite completely analysed yet.

Your input is really valued and is helping us make progressive improvements to different areas of Go-Lab. If you have any questions or concerns, or if you know any other science teachers who we might invite to join this group, please contact us at ULEIC-Go-Lab@leicester.ac.uk.

Best regards,

The link in the email provides the following online questionnaire:

App Configuration

In today's task we would like you to explore the App Configuration feature. Please follow the instructions in the task section and then answer the questions below. Feel free to contact us if the instructions are not clear, not easy to follow or if you have any questions.

Task

1. Sign on to Graasp (www.graasp.eu), and create an ILS for temporary use. (Use the "+" icon to find the "Create Inquiry Space" button).

2. Insert the Concept Mapper app into the Conceptualisation phase. (Open the Conceptualisation space, use the "+" icon again to show buttons, click on "Add App", use the "choose from Go-Lab" feature and select "Concept Mapper").

3. Click on the "cog" icon to edit the Concept Mapper configuration.

4. Explore the facilities to alter the Concept Mapper Options, Help and Resource Options. (For this particular app there is only one "Resource Option").

5. Use these features to create a customised version of the Concept Mapper which might be useful for one of your own lessons. Finish by using the Save button. (You may wish to do this only partly, to test the features rather than to create a completely finished customiation).

6. Check out the student experience of your customised Concept Mapper (Click on "share" icon and then "Show standalone view").

If his has sparked your interest and you would like to explore more, then please do. Most of the apps now have customisation facilities of this sort.

Questions on App Configuration facility

Does the facility to configure apps seem useful to you?

- ○ Very useful
- ○ Somewhat useful
- ○ Not very useful
- ○ Not at all useful

Did you encounter any problems or difficulties while configuring the app?

- ○ Many
- ○ Some
- ○ A few
- ○ None

Please describe the problems or difficulties you encountered:

Do you believe the App Configuration facility would benefit from some improvements?

- ○ Improvements are essential
- ○ Improvements would be very beneficial
- ○ Improvements would be somewhat beneficial
- ○ No improvements are required

Please describe any improvements which you would suggest:

Questions about Concept Mapper

Is the Concept Mapper now a useful app?

- ○ Yes
- ○ No

Has the Concept Mapper been made more or less useful by the provision of this configuration capability?

- ○ Much more useful
- ○ A little more useful
- ○ About the same
- ○ Somewhat less useful
- ○ Much less useful

If you were preparing a lesson using Go-Lab, would you include the Concept Mapper?

- ○ Yes
- ○ Possibly
- ○ No

Please explain why, or in what circumstances you would include the Concept Mapper.

If you had created a customised app, would you like a facility to be able to store it independently of the ILS, so it could be incorporated into other ILSs?

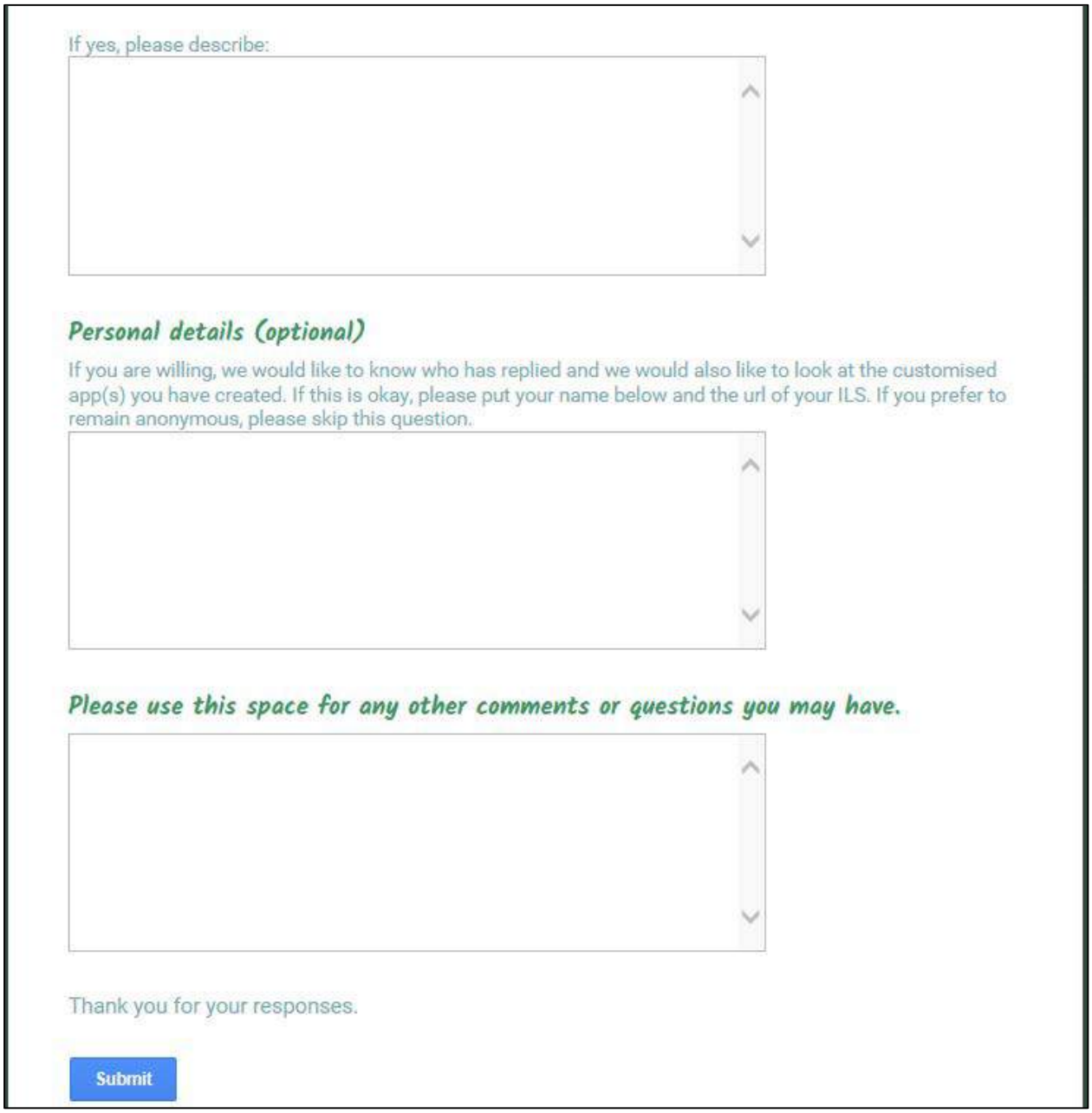
- ○ Yes
- ○ No

Please explain.

Today's task

Was there anything in today's task which you might not have achieved easily without our detailed instructions?

- ○ Yes
- ○ No


If yes, please describe:
Personal details (optional)
If you are willing, we would like to know who has replied and we would also like to look at the customised app(s) you have created. If this is okay, please put your name below and the url of your ILS. If you prefer to remain anonymous, please skip this question.
Please use this space for any other comments or questions you may have.
Thank you for your responses.
Submit

Appendix B: Analytical evaluation of App Composer

Go-Lab WP3 - Participatory Design – Heuristic Evaluation Report

App Composer

1st -2nd December 2014

Aim of session

As requested by technical project partners:

- To explore the new App Composer functionality
- To evaluate for usability and user experience

Method

Based on the link http://composer.golabz.eu/, four HCI specialists with a minimum of prior information explored the new software facilities – which are broadly to allow teachers to produce and share their own translations or customised versions of Go-Lab apps. Two approaches to exploring the system were used:

- screen driven – i.e. based on what we could see, and
- process driven – i.e. based on what we believed a teacher's goals would be and how they would try to achieve them with the software.

This took place over two substantial workshop sessions, during which the review team imagined themselves in the teacher role, aware of a wide range of computer expertise among the teaching population, and assuming limited knowledge of this new functionality. Detailed notes were taken to identify any usability obstacles, which might limit the ability of the teachers to achieve their goals effectively, efficiently, quickly, accurately and enjoyably. Some concepts closely related to (and correlated with) usability such as the aesthetic and affective factors were also taken into account.

After the two sessions, the reviewers' notes were compiled to create a consolidated list of usability observations. This was circulated around the review team so that all could independently assign importance levels (high / medium / low) to the issues:

- Low importance for issues which would be noticed by at least some users, and might affect their overall sense of the quality of the interface, but would not hinder them significantly in achieving their objectives.
- Medium importance for issues which would be noticed by users and may confuse, delay or distract them briefly and temporarily.
- High importance for issues which would be an obstacle to some users, either preventing them from achieving their goals, or causing some significant delay, disruption, confusion or annoyance.

Finally a meeting was held to discuss any discrepancies in "importance" scores and achieve consensus. For many of the issues this also resulted in considerable clarification of evaluators' understandings and the issue descriptions or recommended modifications.

Equipment

- Two laptop computers running Windows 7
- Screen resolution: 1366 x 768 (both)

- Browser: Mozilla Firefox (30.0) and Google Chrome (39.0)

N.B. All screenshots are taken with Mozilla Firefox

Overall Assessment

- **Provides relevant and needed functionality**

The software we explored provides a sophisticated and comprehensive set of facilities to make Go-Lab apps much more customisable by teachers, to match their specific linguistic and pedagogical needs. This addresses some frequently mentioned usability and functionality needs expressed by teachers last year when evaluating apps. The software also provides a range of facilities to allow teachers to share their customisations, make them visible to others and take on aspects of editorial responsibility.

- **Non-trivial design**

The software achieves this using some non-trivial and non-obvious structures and processes, leading to a number of areas in the user interface which seemed rather complex. Because of this complexity we believe some users might not immediately understand some of these features and underlying models.

- **Mainly successful, but also improvable**

Most of the functionality we explored worked well and seemed to be well designed for user experience and functionality. In the text and tables which follow we have not recorded all the good aspects as that would make the document over-long and jumbled. Rather we have focussed on aspects and areas where we believe software changes could make a contribution to improving the user experience of usability and pleasure in use. In some cases, and for some users, usability problems can have a big effect on their adoption and use (or non-use) of a product. Therefore we focus on areas where we believe usability improvements could increase product adoption and engagement.

- **Identified issues and suggested improvements**

We encountered a number of usability and usability-related issues with the system which we believe would pose significant difficulties to teachers, and might have a negative effect on their overall impression of the system. We have listed these issues in tables below and categorised them as high, medium or low importance based on the scenario of a teacher using it for the first time without support or prior training. We recommend addressing at least the most important of these issues before much end user testing is conducted.

We also provide a few more far-reaching questions about the level of complexity and conceptual model and how this might compare with the expectations of teachers. If the current model is retained in full, we believe there should be some help material to make the overall functionality and underlying models clearer. We also believe it would be worth checking some of these models and underlying design assumptions with a few teachers using mock-ups or descriptions or demonstrations.

Summarised usability findings

The detailed tables of usability findings contain a wide variety of material. This material is not easily summarised as it is very diverse, but three themes can be identified:

Issue	Suggested Approach
Non-standard logon process.	This should be streamlined to make it quicker, easier and clearer.
Somewhat misleading name for the product ("App Composer") and some of its main functions.	Adopt names which more closely match the actual functions.
Displays sometimes include computer jargon and elements of code (e.g. HTML). This could give some teachers the impression that it's a product for "techies" and not for them.	All material displayed should be readily understandable by science teachers who have no programming knowledge.

We also here provide three general observations:

- The underlying design model creates a variety of roles and responsibilities for teachers, including hierarchical aspects, as well as certain models about workflow. E.g. some teachers can have editorial control over translations, and the ability to transfer responsibility to another person (without their agreement). There also seem to have been design decisions made about the privacy/visibility/copy-ability of artefacts.

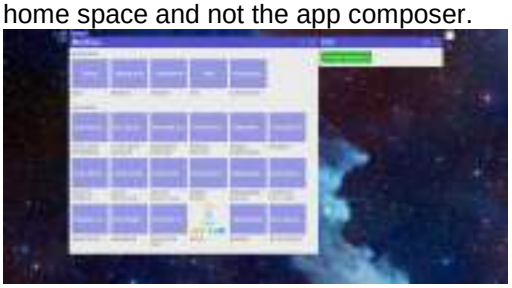
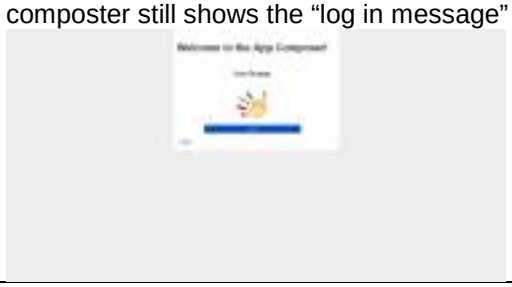
  We wonder whether these design decisions are based on teacher requests, or are working assumptions. If the latter, we suggest testing these assumptions. It may be that a significantly simpler model – especially for translations – would be more appropriate.
- The system seems to have an underlying data model in which one base application can have several translations, each of which can have several adaptations (though we did not find out exactly how to adapt a translation). The user interface also in one place gave the impression that a single translation "contained" many translations (see issues 92 & 93). It took the evaluators some time to understand this (and we are still not convinced we understand it perfectly). The system would be easier to understand and use if the model were simplified or better communicated to users, either through the interface itself, or some help material, or training or manuals. The best option – when possible – is through the interface itself.
- The "importance" scores are based on rather pragmatic criteria, i.e. whether the user can continue to meet their task objectives with a minimum of delay. There are some issues of a more aesthetic nature or affective impact (e.g. 28, 34, 43, 44, 45, 49, 55 & 66) which have relatively low ratings individually but which, when combined, might have a rather negative effect on some users' judgement of, and attitudes towards, the software. This should be considered by the developers when assessing the usability observations reported and their importance rating.

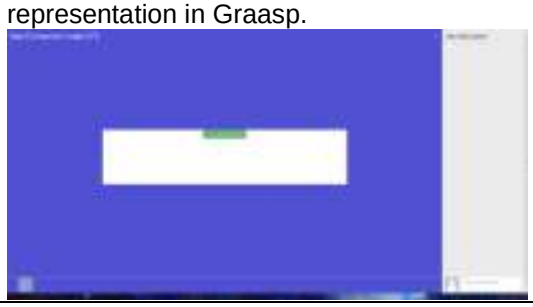
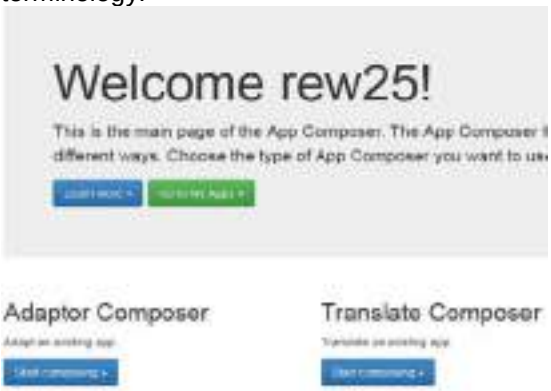
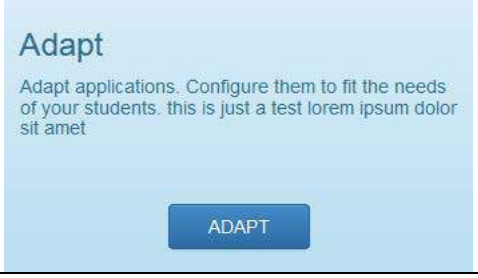
Examples of the Findings

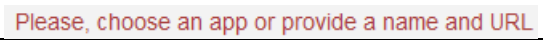
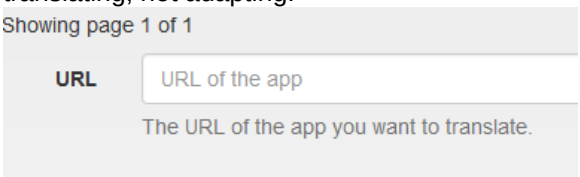
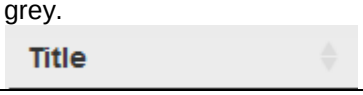
Here below we present some examples to illustrate how we have categorised and reported the findings of the analytic evaluation.

1. General usability issues encountered

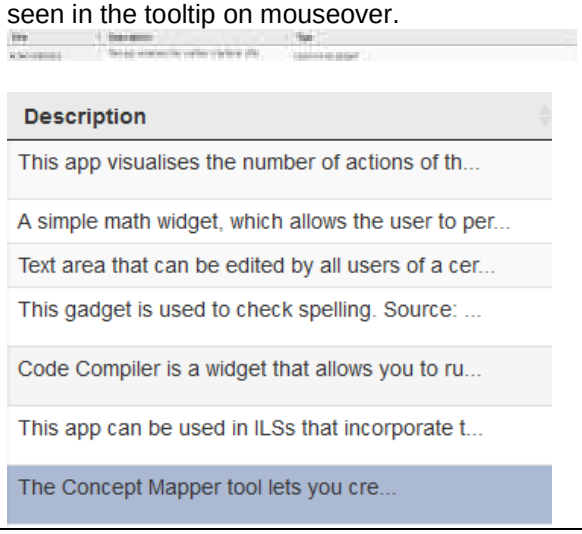
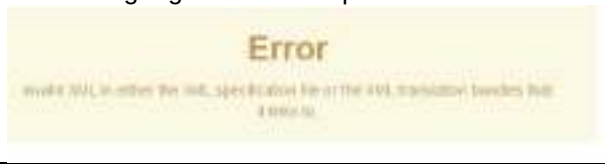
	Usability Observation	Recommended Modification	Import-ance
1	Does the name "App Composer" lead to mistaken expectations?	Consider renaming as App customiser?	Medium

	Usability Observation	Recommended Modification	Import-ance
2	"Super-cool" sounds a little unprofessional 	Perhaps replace with "customise Go-Lab apps for use in your lessons"?	Low
3	Not clear whether this is a "button bar" or tabs that don't quite look like tabs: 	Make them look more like tabs?	Low
4	Problem getting on to app Composer screen in Chrome ("Page unresponsive" error message)	Think this solved itself – maybe no action required?	High
5	Admin tab seems – confusingly – to do the same as "Use it" button.	Label the tab more appropriately, e.g. "Compose" / "Customize".	Medium
6	Screens accessed by "About" and "Contact" tabs have text missing and no help/additional information is provided.	Provide relevant text	High
7	After logging in, the user gets to his Graasp home space and not the app composer.  Instead of 	Should be automatically redirected to the app composer (logged in). Also make App Composer findable from one or more of Graasp home page, Portal and Authoring facility	High
8	(Removed - duplicate)		
9	If the user is already logged in, the app composter still shows the "log in message" 	Should be directly redirected to the App Composer (logged in).	Medium

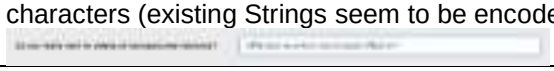
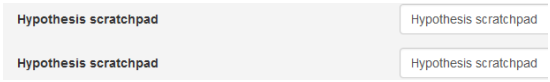
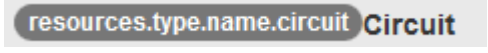
	Usability Observation	Recommended Modification	Import-ance
10	After clicking on the login button (when already logged in) the user is shown an oversized space representation in Graasp. 	Should be directly redirected to app composer page (logged in) [what happens when pressing the “Go to App Composer”]. 	Medium
11	“Learn more” button on App Composer start page (logged in) does nothing when clicked. 	Provide suitable help material	Medium
12	Slightly confusing misleading use of terminology. 	Suggest product is called App customiser, and this screen says “choose the type of customisation you wish to do” with buttons for “Customise app functions” and “translate app”. Also consider whether it would be beneficial to use different terms for base apps and customised apps, since they are treated differently by the system. Suggest also abandoning the use of “adaptation” and using “customisation” throughout.	Medium
13	On the Apps page “Translate Composer” and “Adaptor Composer” are not aligned. Also the previous screen has them in the opposite order. 	Line them up straight. Have them in the same order in both screens.	Low
14	The text inside the Adapt button is not all correct. 	Tidy it up!	Low

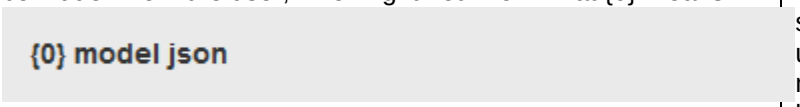
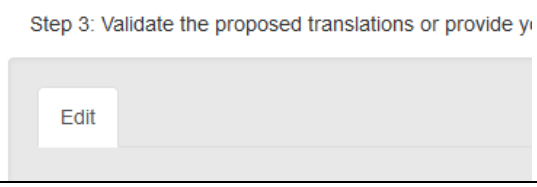
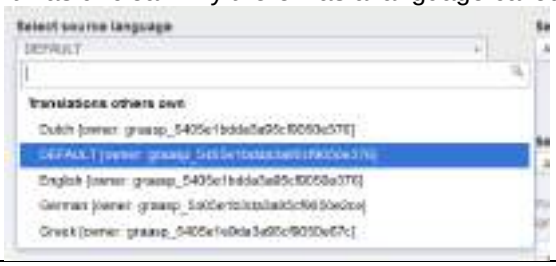
	Usability Observation	Recommended Modification	Import-ance
15	The “highlight box” labelled “No applications.” on the Apps page looks very similar to a button, but isn’t one. 	Make all buttons look more like buttons, and all non-buttons look less like buttons.	Medium
16	On the “Home” screen there are two buttons with the same label “Start composing >>”, but different intentions: 	Should be labelled differently; not “start composing” but e.g. “Adapt an existing app” and “Translate an existing app”.	Low
17	“Log out” does not log you out. Credentials are still retained.	Make it do what it says.	High
18	In the “Adapt” tab the second half of the error message does not make sense when not in advance mode, because you can not specify name and URL. 	Make the message contextual depending on the advanced mode button.	Medium
19	It is not entirely clear whether the “Advanced mode” button has some long term effect or whether it merely affects a single visit to the current screen.	Find some way of communicating this. If it is simply an alternative display for this screen, perhaps have it as a tab at the top of the screen? If it has long term persistence and relevance to other screens, then perhaps make it visible at all times and in all screens?	Low
20	In the Adapt tab, one message suggests you are translating, not adapting: 	Change “translate” to “adapt”	Medium
21	Sorting arrows in the table of apps in “Adapt” tab are nearly invisible, because they are grey on grey. 	Change colour / intensity to increase visibility	Low
22	Headings underneath the table are confusing. 	Remove them.	Low

2. Translate functionality

	Usability Observation	Recommended Modification	Import-ance
23	In table of apps, “Description” column is too narrow, especially when there is lots of blank space. Whole description can only be seen in the tooltip on mouseover. 	Resize columns for improved usability.	Low
24	In the table of apps, description should not contain the computer programmer terminology “OpenSocialGadget”	Replace with something understandable by science teachers if this column has any use	Low
25	Apps that cannot be translated should not show up in the table or at least be marked as “not translatable”. 	Adjust display.	Medium
26	Select language sometimes produces this error: 	Bug fix?	High
27	Clicking on the “Step 2 – select language” button when no app is selected does nothing.	Provide a suitable message asking user to select an app. “Greyed out” state of button should be more obvious.	Low

	Usability Observation	Recommended Modification	Import-ance
28	Mouseover of description has non-meaningful text (html) in it. 	Present only meaningful text, understandable by teachers, not formatting codes. Also eliminate references elsewhere in the UI to “shindig” and “json” and “original XML” which are programmer jargon.	Low
29	Message below looks like an error message (it’s in red) but appears as soon as you reach the screen. Please, choose an app or provide a name and URL	If it’s an instruction on how to proceed it should be above the list of apps and not in red. If it’s an error message then only display if an error has been made.	Low
30	When an app is selected and then the paging buttons are used to navigate through the list of apps, the selection is still active but not visible.	Display information about the selected app above the “continue” button.	Medium
31	The 3 steps are displayed and referred to inconsistently and confusingly. Step 1 - App creation Step 2 - Select language » Step 3: Translate « Go Back to Step 1 Step 3 - Translate » **Step 3: Validate the proposed translations or provide your own. Save to apply the changes.**	Label each of the three steps and the matching navigation buttons consistently. E.g. Step 1 should be “app selection”, not “app creation” “Step 3 - translate” should be “Step 2 - select language” + use consistent fonts	Medium
32	(Removed – duplicate)		

	Usability Observation	Recommended Modification	Import-ance
33	Usernames displayed in the translation selection are weird. You want to translate the app from English [owner: graasp_5405e1b3da3a95cf9050e2ce] (for anyone) into German [owner: graasp_5405e205da3a95cf9050e958] (for anyone).	All text displayed should be understandable to a typical science teacher and avoid computer programmer jargon.	Medium
34	(Removed – duplicate)		
35	Translation does not always cope with German special characters (existing Strings seem to be encoded incorrectly). 	Extend the font capability	Medium
36	The same word needs to be translated several times (e.g. Save). 	Either create a scheme to present each word only once, or (if translations need to differ depending on context) provide some context.	Medium
37	Adding the translated version of the app to an ILS results in an error message “Unable to retrieve spec for http://composer.golabz.eu/composers/translate/app/F-wvl4pnSTuMT2it0W0P/ALL/languages/ALL_ALL.xml. HTTP error 404”. 	Bug fix? Error seems to be fixed, but App is still not in the selected language (seems like the URL created when publishing is for ALL and not language specific)	High
38	“Show identifiers” changes the layout. 	Make the layout of the translated column stable.	Low
39	Show identifiers also displays programmer text not understood by teachers. 	If there is a purpose for teachers to be able to see these identifiers, make them understandable. If not then eliminate this feature.	Medium

	Usability Observation	Recommended Modification	Import-ance
40	In the list of phrases to translate, implementation details should be hidden from the user, who might not know what {0} means. 	Supply {0} "behind the scenes" as the user should not need to be bothered with it.	High
41	On the translation page there is a tab navigation with only one tab "Edit". 	Don't show the tab bar if there is only one tab	Low
42	Users can see all proposed translations. 	Consider whether there should be privacy and also whether only accepted translations should be shown,	Medium
43	It was unclear why there was a language called DEFAULT. 	Consider whether and why this is necessary, and if there is a need, make the purpose of DEFAULT clearer	Medium
44	Relabel "group" as it is not a "group of languages" you are selecting from this list. 	E.g. you could call it "level of language".	Low
45	What's the point of having the URL of the original XML (as a link). 	Remove this unless it has a purpose and if it has, make the purpose clearer.	Low

Appendix C: Areas of Suggested Labs by WP3 Participants

Here below is the list of areas where some WP3 participants have suggested:

- Titrations
- Health and safety
- Rates of chemical reactions
- Thermal output of chemical reactions
- Electrophoresis
- PCR (polymerase chain reaction)
- Chemistry: titrations, reactivity series/displacement, electrolysis
- DNA
- Lab to educate on the effects of taking illegal drugs, including simulations to demonstrate the effect on user's capabilities, and
- also simulated video journey through the user's body to see the effect on their internal organs!!! (Pupil suggestion)
- solar and lunar eclipse
- levers
- circulatory system
- heat measurement and transfer
- classification of animals and plants
- systems of human body
- reproduction in plants and humans
- add some biology apps which don't need Java
- more things in Arabic language
- Could we have a lab which consisted of a video of the inside of a human body in which the student could navigate into different organs etc.
- combustion
- lenses & optics
- PCR, electrophoresis, genetic engineering, ELISA, blood tests etc.
- digestion and GE {I think GE might mean gross energy}
- kidneys, photosynthesis, respiration
- more chemistry labs
- blood analysis
- more for computer science
- isomerism
- periodicity
- Simple harmonic motion
- Buffer solutions, mole calculators, redox titrations, fractional distillation, NMR (nuclear magnetic resonance I think), a time app
- Enzymes, DNA building, protein synthesis

Appendix D: Example Responses of the Technical Cluster to Mid-year Findings

In the following text, the plain text is from a WP3 report, and the bold text is the Technical cluster response to some *mid-year findings*.

Bug fixes

1. Text entered into a rich text field by one user sometimes appears in another user's ILS. **Fixed**
2. CONNECTING & DISCONNECTED messages in bottom left corner of screen and system stalls. **This component shows if the connection is bad. This is standard in many online platforms (e.g. Google Docs, Sharelatex) we will keep it**
3. When you enter text into a "document" and go to standalone view and then back to Graasp view, the text sometimes seems lost. It can be recovered by clicking on EDIT and VIEW alternately a few times. **On tentative todo list**
4. Customised version of app created by one user sometimes appears in another user's ILS. **On tentative todo list (Bugs)**
5. Some ILSs on the Portal do not have the "Copy ILS" button – possibly caused by a bug when the ILS was published. For others, the COPY ILS button is visible but does nothing at all. **On tentative todo list (new publish scheme)**
6. A phase with lots of apps and labs can fail to load, and it's fixed by splitting it up into multiple phases. **On tentative todo list (improve performance)**
7. When typing in a nickname to use an ILS, you have to hit ENTER twice. **This only happens if the nickname already exists, the first click is to select the nickname in the list the second to log in. No action.**
8. When adding a url without specifying a name, nothing happens but no error message either. **On tentative todo list (Bugs)**

Usability issues

1. Reinvent scrolling to be user-friendly, i.e. familiar for Windows users (**On tentative todo list details to be discussed**), and possible on devices such as touchpads:
    - Eliminate invisible / disappearing scroll bars,
    - no very narrow scroll bars
    - no scrolling within scrolling **(for pdfs we could use another viewer that requires to press next instead of scrolling, would this solve the issue?)**
    - no impossible scrolling
    - preferably no sideways scrolling (if **this happens in apps, it's out of our control**)
    - preferably all vertical scroll bars at right edge of screen (if maximised) and all horizontal ones at the bottom – including app customisation screens.
2. "Windows" – anything which looks like a (Microsoft) window – e.g. blue bar at the top – should behave like one, fully. **Unclear. No action.**

3. Constantly having to click on a tiny and pale "+" button to get the icons to be visible is unnecessarily tedious. Better to have the icons always visible. Also some users have been observed clicking on the "+" button in the user instructions. **Most content can be put by drag and dropping, for the rest we feel that the suggested solution is unnecessarily cluttering the UI.**
4. Contents of vault are incomprehensible. **On tentative todo list (List view for vault with more details about creators)**
5. Maybe rename spaces as folders, rename "create ILS" as "create online lesson" **(core project vocabulary)**, rename "add document" as "add text box" **(this is not a text box, it is a doc or document (either rich text or html))**, rename "standalone view" as "student view"**(core project vocabulary)** , and remove "create space" unless it has a purpose **(it can be used to store content)**. **No action.**
6. Provide better facilities for entering formulas (both chemical and mathematical) – preferably easier than LaTeX, or provide learning materials for LaTeX. **Community cluster can provide help in the tutoring platform, maybe a link to this. No action.**
7. Provide better facilities for cloning / copying an ILS – e.g. when using for another class – without having to publish it on Portal. The clone needs to be identical in all ways. **On tentative todo list**
8. Difficult to change the sequence of items within a phase unless the title matches the content well - heavy working memory demand. Could drag and drop be implemented in the full view? **Not practical to drag and drop big components in the expanded view, changing view seems a better option. No action**
9. There are two ways of entering text into a lesson – at the top of a component (app, lab etc) and in a text box (insert document). Both have different characteristics and editing capabilities. Just one method would be better. (RWE-W personal thought: would it be more intuitive if the whole space could be text entry, with special buttons to add labs, apps, links, resources etc … more like MS Word or web-based WYSIWYG editor). **On tentative todo list (removing the edit and view options on the .graasp file to give it the same feel as the description) We will not implement MS Word.**
10. Please provide embedded online help. **We have added tooltips to provide guidance, further help can be added to the content of the tutoring/tutorial platform by the community cluster . No action.**
11. Components are always added at the end of a phase. Can we also have a facility to add components in a specific place? **Would need a UI redesign and does not seem practical when drag and dropping several files. No action.**
12. Why three separate views plus standalone view? There must be an easier way. Streamline? **Similarly to the directory viewer on a desktop different options are possible for users, and some have more or less options. Up to the users which one they want to choose. No action**
13. Having to give a name to text boxes (documents) seems unnecessary. **Since they are not text boxes but files, a name allows to identify them in the space and in the folder if they are downloaded. No action**
14. Not sure why we need EDIT and VIEW for documents. (Especially important as if a user presses delete arrow key when in view mode, most browsers navigate back in a disconcerting fashion). **On tentative todo list**

15. When creating an ILS, it scrolls to the bottom of your list of ILSs (and the scroll bar is invisible). **related to the scroll discussion above**

Function & usefulness issues

1. Add reporting facilities so teachers (or students) can extract students' work; e.g. for assessment. **On tentative todo list (add link on contextual users to their ILSs if not password protected)**
2. Allow us to include PDFs, Word documents and PowerPoints in an ILS easily. **They can be dragged and dropped. No action**
3. Get rid of "add space" and provide an "add text box" button.
4. Make everything work on a wide range of devices including old and new computers, tablets, maybe even phones. **Works even with phones. No action (unless we should make it work with atari :)**
5. If possible, design it to work as well as possible when there is a limited internet connection. **On tentative todo list (performance improvements)**
6. Lack of passwords on nicknames means plagiarism is too easy, or even (accidentally or deliberately) changing someone else's work. **Users can specify if they require nickname or nickname and password in the standalone view. No action**