ENhance **VI**rtual learning **S**paces using **A**pplied **G**aming in **E**ducation

H2020-ICT-24-2016

D5.1 - Pilot execution plan and evaluation protocol

Dissemination level:	Public (PU)
Contractual date of delivery:	Month 9, June 30, 2017
Actual date of delivery:	Month 10, July 24, 2017
Workpackage:	WP5 Pilot execution and evaluation
Task:	T5.1 Building virtual labs: Small-scale pilots addressing the educational scenarios T5.2 System test and evaluation
Type:	Report
Approval Status:	Final
Version:	0.7
Number of pages:	29
Filename:	D5.1_Pilot execution plan and evaluation protocol_Final.docx

Abstract

We present and discuss a pilot execution plan and evaluation methodology and protocol that will be followed to conduct a series of small-scale test implementations of virtual labs, accompanied by the authoring, analytics and visualization tools of ENVISAGE. The quantitative and qualitative feedback collection from users, both teachers and students, will be based on the herein proposed methodology and protocol of conduct. Complementary methods are to be utilized, including structured questionnaires, interviews, focus group discussions and on-site observations, according to a well-defined procedure of conduct and reporting. A schedule of activities related to piloting and evaluation during the school year 2017-2018 is also discussed.

Co-funded by the European Union

Copyright

History

Version	Date	Reason	Revised by
v0.1 (alpha)	09/06/2017	Table of Contents circulated to the consortium	Georgios Mavromanolakis & Spiros Nikolopoulos
v0.2	23/06/2017	Elaborated version circulated to all partners for collecting contributions	Georgios Mavromanolakis, Sofoklis Sotiriou, Benedikte Mikkelsen, Fabian Hadiji, Christoffer Holmgård, Giannis Chantas
v0.3	30/06/2017	Consolidated version incorporating contributions from all partners	Georgios Mavromanolakis
0. 4 (beta)	07/07/2017	Beta-version circulated for internal reviewing	Line Ebdrup Thomsen
V0.5	19/07/2017	Pre-final version incorporating the internal revision comments	Georgios Mavromanolakis
V0.6	24/07/2017	Version including the comments made by the coordinator	Spiros Nikolopoulos
V0.7 (final)	24/07/2017	Final version checked for consistency by the coordinator	Georgios Mavromanolakis

Author list

Organization	Name	Contact Information
EA	Georgios Mavromanolakis	gmavroma@ea.gr
EA	Sofoklis Sotiriou	sotiriou@ea.gr
AAU	Benedikte Mikkelsen	benedikte@hum.aau.dk
GIO	Fabian Hadiji	fabian@goedle.io
UoM	Christoffer Holmgård	holmgard@gmail.com
CERTH	Giannis Chantas	gchantas@iti.gr

Executive Summary

The current document aims to present a pilot execution plan and evaluation methodology and protocol that will be followed to conduct a series of small-scale test implementations of virtual labs, accompanied by the authoring, analytics and visualization tools of ENVISAGE. In the context of well-defined educational scenarios, the effectiveness of the developed technologies will be thoroughly tested and evaluated, with respect to their ability to benefit educational organizations utilizing the ENVISAGE solution towards the optimization of virtual lab design and functionality. The evaluation is foreseen to provide quantitative and qualitative feedback on the process of using the ENVISAGE authoring tool as a means for building virtual labs, on the offered analytics and visualization tools as means of support towards improving virtual labs and on the delivered virtual labs with respect to their effectiveness to meet the goals and expectations of both teachers and students in the learning process.

The process of using the authoring tool as a means for building virtual labs will be evaluated by teachers and e-learning experts as test users. Predefined tasks of gradual complexity will be given to them and feedback will be collected through various means and methods in order to evaluate the effectiveness, usability and functionality of the authoring tool as well the overall user experience. The evaluation outcomes will be utilized to develop, update and release an improved version of the tool.

With respect to the support offered by the analytics and visualization tools in the process of improving virtual labs, this will be evaluated also by teachers and experts. Similarly, these tools' effectiveness, usability, functionality and user experience will be assessed towards their improvement and upgrade.

The delivered virtual labs and the learning content will be evaluated by both teachers and students as well as experts. The evaluation process will focus on measuring to which extent the virtual labs have achieved to motivate and to engage students in the learning process, and also to which extent their utilization in the classroom facilitated the teaching and learning objectives set or expected by teachers.

In this context, a methodology and protocol of conduct is presented, which the quantitative and qualitative feedback collection from users, both teachers and students, is based on. Complementary methods are proposed and will be utilized, including structured questionnaires, interviews, focus group discussions and on-site observations, according to a well-defined procedure of conduct and reporting. A schedule of activities related to piloting and evaluation during the school year 2017-2018 is also laid out.

The document is structured as follows: a short introduction is given in Chapter 1. Chapter 2 presents the general objectives of piloting and evaluation work. In Chapter 3 we are discussing the methodology on which is based the quantitative and qualitative feedback collection from users. Complementary methods will be utilized according to a well-defined procedure of conduct and report. In Chapter 4 we propose a schedule of activities related to piloting and evaluation which coincides with the school year. The main points of the document are summarized in Chapter 5.

Table of Contents

1 Introduction

1.1 The ENVISAGE Concept

The concept of ENVISAGE is based on iterating the process of improving virtual labs through a structured and staged process. The first phase starts from the current available version of a lab that may already be in use in real classroom setting and the collection of preliminary analytics extracted from user data. The next step is to deploy machine learning methods to the obtained analytics. Then follows the integration of the concluded information and pedagogical insights into the authoring tool, which will be employed by teachers and educators to build an improved version of the virtual lab and accompanying educational scenarios. The above process can be iterated until the final version is reached. The whole approach of ENVISAGE is illustrated in Fig.1.1. Initial statistics on tracked data are collected to monitor the activity of users in order to model their behaviour and general practices. Further analysis with the application of machine learning methods follows in order to predict the expected behaviour of users. Both of these approaches are combined with visualization methodologies that will offer insights and guidance to educators and to lab developers on what features are important and what functionalities users expect to find in a virtual lab. These will allow the optimization of the design of the virtual lab, and will facilitate its implementation and finally lead to the enhancement of the learning process.

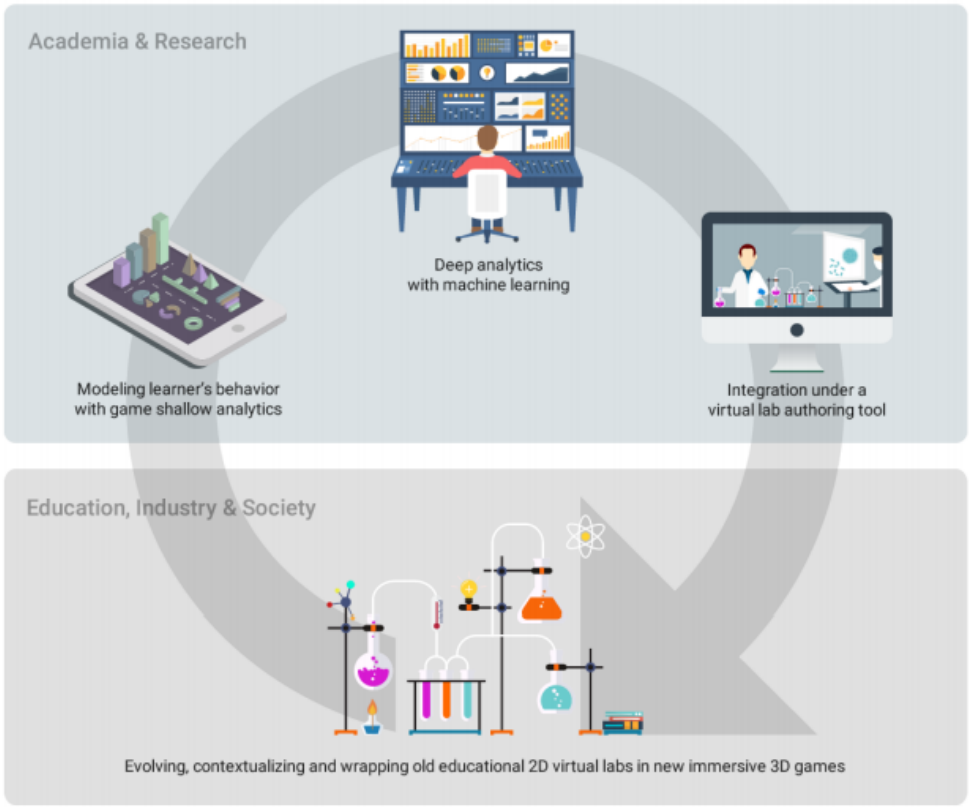

Figure 1.1: *The ENVISAGE virtual lab enrichment strategy through a staged and iterative approach.*

1.2 Scope of the current document

The current document aims to present a pilot execution plan and evaluation methodology and protocol that will be followed to conduct a series of small-scale test implementations of virtual labs, accompanied by the authoring, analytics and visualization tools of ENVISAGE. The document is structured as follows: in Chapter 2 we are presenting the general objectives of piloting and evaluation work. In Chapter 3 we are discussing the methodology on which is based the quantitative and qualitative feedback collection from users, both teachers and students. Complementary methods will be utilized, including structured questionnaires, interviews, focus group discussions and on-site observations, according to a well-defined procedure of conduct and report. In Chapter 4 we propose a schedule of activities related to piloting and evaluation which practically starts with the beginning of the school year and concludes at the end of it. The main points of the document are summarized in Chapter 5.

2 Objectives of piloting and evaluation

The general objectives of piloting and evaluation tasks are to conduct a series of small-scale test implementations of virtual labs, accompanied by the authoring, analytics and visualization tools of ENVISAGE, in the context of well-defined educational scenarios. This is done to evaluate the effectiveness of the developed technologies with respect to their ability to benefit educational organizations utilizing the ENVISAGE solution towards the optimization of virtual lab design and functionality. The evaluation is foreseen to provide quantitative and qualitative feedback on three separate conditions:

a. The process of using the ENVISAGE authoring tool as a means for building virtual labs.
b. The offered analytics and visualization tools as means of support towards improving virtual labs.
c. The delivered virtual labs with respect to their effectiveness to meet the goals and expectations of both teachers and students in the learning process.

In particular, the process of using the authoring tool as a means for building virtual labs will be evaluated by teachers and e-learning experts as test users. Predefined tasks of gradual complexity will be given to them and feedback will be collected through various means and methods in order to evaluate the effectiveness, usability and functionality of the authoring tool as well the overall user experience. The evaluation outcomes will be utilized to develop, update and release an improved version of the tool.

With respect to the support offered by the analytics and visualization tools in the process of improving virtual labs, this will be evaluated also by teachers and experts. Similarly, these tools' effectiveness, usability, functionality and user experience will be assessed towards their improvement and upgrade.

The delivered virtual labs and the learning content will be evaluated by both teachers and students as well as experts. The evaluation process will focus on measuring to which extent the virtual labs have achieved to motivate and to engage students in the learning process, and also to which extent their utilization in the classroom facilitated the teaching and learning objectives set or expected by teachers.

The piloting and evaluation work will primarily take place at EA schools in collaboration and with support from the other partners of ENVISAGE, namely CERTH on the authoring tool, AAU and UOM on the analytics and visualization components respectively. If necessary, a case that may be mostly depending on project time constraints and standard school year curriculum schedule, piloting and evaluation will also be conducted with teachers and students from other preselected schools in order to acquire the needed additional evaluation user data with higher statistics and to ensure objective results.

In the following we present and discuss the proposed methodology and protocol of conduct. Prior to that, and due to the fact that ENVISAGE is comprised by partners of diverse and complementary expertise and experience, we list the definition of certain typical terms for reference, better completeness and clarity.

Effectiveness

Effectiveness of a tool, system, service or product is defined as the capability of producing a desired result or the ability to produce desired output. In the context of the project, effectiveness is the accuracy and completeness of users' tasks while using the offered system or components, the degree to which objectives are achieved and the extent to which targeted or predefined tasks are executed. In contrast to efficiency, effectiveness is determined without reference to elements of cost, such as e.g. time needed to accomplish or complete a task.

Functionality

Functionality of a tool, system, service or product is defined as the set of features, properties and functions that it offers to users or/and also as a synonym of practicality, i.e. the quality of being suited to serve a purpose well. In this context, for example, the functionality of an authoring tool is the level at which its offered functions address the needs or expectations of a teacher for the purpose of building or improving a virtual lab.

Usability

Usability of a tool, system or service is defined as a quality attribute that assesses how easy their user interfaces are to use within an actual or realistic usage context. The term also refers to methods for accomplishing or improving easy-of-use during the design process. In the framework of piloting and evaluation of ENVISAGE components, the objective is to set up and conduct several usability tests involving users in realistic situations or usage scenarios. Therein the users perform a list of tasks using the tools being tested while observers watch and take notes and pre- and post-test questionnaires may also be used to gather feedback. The aim is to observe how actual users interact with the system, in a realistic manner, so that developers can identify non-anticipated use cases, usability violating issues causing errors, identify mental models etc.

User experience

Refers to a person's emotions and attitudes about using a tool, system, service or product. It includes not only the practical but also the overall experiential and affective aspects while interacting with it. Additionally, it includes perceptions of system aspects such as utility, ease of use and efficiency. User experience may be strongly considered subjective in nature to the degree that it is about individual perception. User experience is constantly modified over time due to e.g. changing usage circumstances or due to users acquiring better knowledge and experience. User experience design as a discipline is concerned with all the elements that together make up the interface, including layout, visual design, text, brand, and interaction. It works to coordinate these elements to allow for the best possible interaction by users.

3 Methodology and protocol of evaluation

3.1 Prerequisites and preselection criteria

The piloting and evaluation work will primarily take place at EA schools of primary and secondary education in collaboration with and with support from all partners of ENVISAGE. Both experienced and novice science teachers have been already identified to participate in the studies during the school year of 2017-2018, which starts in September 2017. If and when necessary, a case that may be mostly depending on project time constraints and standard school year curriculum schedules, piloting and evaluation may also be conducted with teachers and students from other preselected schools in order to acquire (if needed) additional evaluation user data of higher statistics and to ensure objective results. In order to ensure that this process will lead to comparable standards of conducted studies and collected feedback in terms of quality, we devised basic practical prerequisites and criteria. These are as follows:

- Teachers with good knowledge of English language, both reading and writing, as many of the tools to be tested, their end user interfaces and their instructions, will be provided in this language
- Teachers and school policy makers with interest in technology-enhanced science education approaches and methodologies
- Teachers and school policy makers with positive attitude towards innovative methods of teaching and learning
- Teachers with experience using computers for teaching and learning
- School equipped with classrooms or labs with personal computers connected to internet with minimum availability of 1 personal computer per 2 students

Any additional selected schools and their teachers should comply with these criteria, which are fulfilled in the case of EA.

3.2 Proposed methodology

The evaluation and piloting will be based on the following complementary methods of quantitative and qualitative feedback collection: structured questionnaires, interviews, focus group discussions, user testing and on-site observations. More specifically, each component of ENVISAGE will be tested and evaluated according to a combination of methods as follows.

3.2.1 Authoring tool

The process of using the ENVISAGE authoring tool as a means for building virtual labs will primarily be evaluated with standard questionnaires for user interface usability, perceived usefulness and ease of use of its front-end. Furthermore the evaluation will be elaborated with focus group discussions or interview with teachers where the main topics will be standard attributes of usability, such as learnability, memorability, efficiency, overall satisfaction etc. This will be conducted in the context of certain testing actions or sequence of tasks that teachers/users will be requested to perform. These are, in technical terms, as

follow: create a "game project", create a "new scene", edit an existing scene, create a "new Asset 3D", add an Asset 3D to a scene, save a scene, edit Asset3D, delete a "game project", delete a "game scene", delete an Asset3D, edit 2D scenes, assembly and compile of a "game project".

The above methods of feedback collection may be complemented by user testing and observation sessions by expert users or developers. This session can be followed-up by teacher interview or discussion. Its purpose is to collect direct feedback and compare user behaviour under two different conditions, namely when observers act also as facilitators and provide guidance and assistance, and when they do not intervene with user actions.

The standard structured questionnaires are depicted below (see Fig.3.1, 2 and 3) and can be in paper or electronic format [1-3]. A typical work-flow/agenda of user test and evaluation completed in a single session is shown in Appendix I. The proposed timeframe when these piloting and evaluation activities can take place is discussed in section 4 (see also Fig.4.1).

3.2.2 Analytics and visualization tools

The support offered by the analytics tools in the process of improving virtual labs will be similarly evaluated through teacher interview and discussion sessions. In these, general feedback is first collected using paper mock-up dashboards of visualizations, along with observation and interview for assessing teachers' intuitive ability to understand the displayed information of collected analytics within context. At a second iteration, actual user testing will take place where teachers utilize the analytics front-end interface. In this case evaluation is collected with standard questionnaires for user interface usability, perceived usefulness and ease of use.

The standard structured questionnaires are shown below in Fig.3.1, 3.2 and 3.3. A typical work-flow/agenda of a complete user test and evaluation session is shown in Appendix I. The proposed timeframe when these piloting and evaluation activities can take place is discussed in section 4 (see also Fig.4.1).

3.2.3 Delivered virtual labs

The delivered virtual labs and learning content with respect to their effectiveness to meet the goals and expectations of both teachers and students in the learning process, will be evaluated in actual classroom settings. The evaluation will be conducted using structured questionnaires for teachers and students, accompanied also by teacher interviews and discussions. Prior to this, teachers in collaboration with educational games developers will provide insights and assess various versions of proposed virtual labs, in particular the wind energy lab, the chemistry labs and their analytics and visualizations components.

The structured questionnaires for accessing the student learning outcomes and the teacher expectations and observations are shown in Appendix II and III. The proposed timeframe when these piloting and evaluation activities can take place within the school year is discussed in section 4 (see also Fig.4.1).

		1	2	3	4	5	6	7		NA
1. Learnability	bad	○	○	○	○	○	○	○	good	○
Comments:										
2. Efficiency	bad	○	○	○	○	○	○	○	good	○
Comments:										
3. Memorability	bad	○	○	○	○	○	○	○	good	○
Comments:										
4. Errors (Accuracy)	bad	○	○	○	○	○	○	○	good	○
Comments:										
5. Subjective Satisfaction	bad	○	○	○	○	○	○	○	good	○
Comments:										
		1	2	3	4	5	6	7		NA

Figure 3.1: *Proposed standard questionnaire, in printed paper or electronic form, for evaluation of attributes of usability [1].*

PERCEIVED USEFULNESS		1	2	3	4	5	6	7		NA
1. Using the system in my job would enable me to accomplish tasks more quickly	unlikely	○	○	○	○	○	○	○	likely	○
Comments:										
2. Using the system would improve my job performance	unlikely	○	○	○	○	○	○	○	likely	○
Comments:										
3. Using the system in my job would increase my productity	unlikely	○	○	○	○	○	○	○	likely	○
Comments:										
4. Using the system would enhance my effectiveness on the job	unlikely	○	○	○	○	○	○	○	likely	○
Comments:										
5. Using the system would make it easier to do my job	unlikely	○	○	○	○	○	○	○	likely	○
Comments:										
6. I would find the system useful in my job	unlikely	○	○	○	○	○	○	○	likely	○
Comments:										
PERCEIVED EASE OF USE		1	2	3	4	5	6	7		NA
7. Learning to operate the system would be easy for me	unlikely	○	○	○	○	○	○	○	likely	○
Comments:										
8. I would find it easy to get the system to do what I want it to do	unlikely	○	○	○	○	○	○	○	likely	○
Comments:										
9. My interaction with the system would be clear and understandable	unlikely	○	○	○	○	○	○	○	likely	○
Comments:										
10. I would find the system to be flexible to interact with	unlikely	○	○	○	○	○	○	○	likely	○
Comments:										
11. It would be easy for me to become skillful at using the system	unlikely	○	○	○	○	○	○	○	likely	○
Comments:										
12. I would find the system easy to use	unlikely	○	○	○	○	○	○	○	likely	○
Comments:										
		1	2	3	4	5	6	7		NA

List the most **negative** aspect(s):

1.
2.
3.

List the most **positive** aspect(s):

1.
2.
3.

Figure 3.2: *Proposed standard questionnaire, in printed paper or electronic form, for evaluation of perceived usefulness and ease of use [2].*

		1	2	3	4	5	6	7		NA
1. Overall, I am satisfied with how easy it is to use this system	strongly disagree	○	○	○	○	○	○	○	strongly agree	○
2. It was simple to use this system	strongly disagree	○	○	○	○	○	○	○	strongly agree	○
3. I can effectively complete my work using this system	strongly disagree	○	○	○	○	○	○	○	strongly agree	○
4. I am able to complete my work quickly using this system	strongly disagree	○	○	○	○	○	○	○	strongly agree	○
5. I am able to efficiently complete my work using this system	strongly disagree	○	○	○	○	○	○	○	strongly agree	○
6. I feel comfortable using this system	strongly disagree	○	○	○	○	○	○	○	strongly agree	○
7. It was easy to learn to use this system	strongly disagree	○	○	○	○	○	○	○	strongly agree	○
8. I believe I became productive quickly using this system	strongly disagree	○	○	○	○	○	○	○	strongly agree	○
9. The system gives error messages that clearly tell me how to fix problems	strongly disagree	○	○	○	○	○	○	○	strongly agree	○
10. Whenever I make a mistake using the system, I recover easily and quickly	strongly disagree	○	○	○	○	○	○	○	strongly agree	○
11. The information (such as online help, on-screen messages, and other documentation) provided with this system is clear	strongly disagree	○	○	○	○	○	○	○	strongly agree	○
12. It is easy to find the information I needed	strongly disagree	○	○	○	○	○	○	○	strongly agree	○
13. The information provided for the system is easy to understand	strongly disagree	○	○	○	○	○	○	○	strongly agree	○
14. The information is effective in helping me complete the tasks and scenarios	strongly disagree	○	○	○	○	○	○	○	strongly agree	○
15. The organization of information on the system screens is clear	strongly disagree	○	○	○	○	○	○	○	strongly agree	○
16. The interface of this system is pleasant	strongly disagree	○	○	○	○	○	○	○	strongly agree	○
17. I like using the interface of this system	strongly disagree	○	○	○	○	○	○	○	strongly agree	○
18. This system has all the functions and capabilities I expect it to have	strongly disagree	○	○	○	○	○	○	○	strongly agree	○
19. Overall, I am satisfied with this system	strongly disagree	○	○	○	○	○	○	○	strongly agree	○
		1	2	3	4	5	6	7		NA

List the most **negative** aspect(s):

1.
2.
3.

List the most **positive** aspect(s):

1.
2.
3.

Figure 3.3: *Proposed standard questionnaire, in printed paper or electronic form, for system usability testing [3].*

3.3 Protocol of conduct

3.3.1 General guidelines and ethics of conduct

During the piloting and evaluation of ENVISAGE tools, various information is collected about participants' profile, needs and abilities in what refers to their interaction with the virtual labs and their learning progress when carrying out the educational scenarios. In addition, the collected data may include, but is not limited to, personal information about the user such as: name, date of birth, interests, location or relations to other users. The data controller will be the EA organization and the data processor will be all the other partners of the consortium. In handling these data, both controller and processor parties will comply with national and EU legislation, as well as follow the best practice for ethics in Human-Computer Interaction (see also [4]). More specifically the general guidelines are, with respect to:

Data collection: any data collected, in paper, electronic or online forms, will be strictly anonymous or be anonymized. In all cases the personal identity of the data will be strictly protected from third parties and will only be used for testing purposes within the project. There may be cases where in-classroom activities will be video-recorded for evaluation purposes. This will solely happen after permission or appropriate consent procedures have been granted and followed in accordance with any specific rules that schools have and in addition to national and EU regulations.

Data usage and handling: the consortium is committed to maintain strict rules of privacy and security to prevent all personal data from being abused or leaked. The collected data will be used strictly for the purposes defined by the project objectives of piloting and evaluation. Under no circumstances, the consortium will provide, give or sell any information of its users to any third party. In this context data will not be used under any circumstances for commercial purposes and photographs or video recordings will not to be used for dissemination, will not appear at the project web site or/and will not be added to any project dissemination materials.

Data retention time: three months after the project completion all personal data that have been collected, stored and processed will be deleted, unless participants have been clearly informed otherwise and have provided appropriate consent accordingly.

Recruitment and consent of participants: all consortium members of ENVISAGE, according to expertise and experience, will take or contribute to all necessary steps to ensure that all participants, teachers and students, understand and are well informed about the objectives of the project and the processes employed during it to achieve them. In particular, recruitment of teachers to participate in the piloting and evaluation activities will be based on basic preselection and prerequisite criteria, as already discussed in section 3.1. Consent of participation will be: a) informed: given in possession and understanding of the principal, relevant information; b) voluntary: given freely and not as a result of coercive pressure (real or perceived); c) competent: given by somebody able, in virtue of their age, maturity and mental stability, of making a free, considered choice. The recruited participants can be freely choose to be withdrawn from the pilots at any moment and without any consequences

either through their own request, or through the advice of their school supervisor or principal. Upon withdrawal all data collected from the participant will be deleted and removed from the evaluation studies. Subsequently, the recruitment procedure will be re-initiated to select a replacement for the withdrawn participants, if this is required for the smooth execution of the project. Where applicable, and especially in cases of video-recording of in-school activities or other relevant instances of piloting and evaluation, where participation of students is involved, appropriate consent forms from school or school authorities and parents will be requested to be granted.

Executive research staff: all piloting and evaluation activities involving teachers and students will be performed, supervised, reported or observed by qualified staff and with the practical experience, which will guarantee a strict conformance with national and international ethics and regulations. The main purpose of research staff acting as facilitators or observers is to report accordingly the conducted activities and guarantee objectivity, integrity and overall quality of collected data without intervening to the actions of participants.

With respect to each proposed method of data collection, practical guidelines of conduct are briefly emphasized below for completeness.

3.3.2 Structured questionnaires

Collection of data from participating teachers and students using structured questionnaires in paper printed or electronic format will be anonymous or will be anonymized, e.g. nicknames or other predefined code ids may be utilized that do not reveal the true identity of participants.

Participation is always on voluntary basis without any imposed constrains or pressure e.g. instructions to answer all fields or questions, or within given time limit etc.

3.3.3 Interviews

Collection of data from participant teacher through structured or open interviews will also be anonymous or be anonymized unless written and signed consent is granted by the participant (see also Appendix I).

At the start of an interview session the participant will be reminded that her/his participation is always on voluntary basis without any imposed constrains or consequences, on understanding that she/he is free to leave at any time and that none of her/his opinions or thoughts will be shared with anyone outside the project, unless all identifying information is removed first.

The facilitator should ensure that an interview session is scheduled well in advance, especially in case the participant has to do some preparatory work, and that the participant is informed and familiarized with the procedure.

The interview may be conducted in the native language of participant, however transcripts or notes will be reported in English by the facilitator.

3.3.4 Focus group discussions

The method of focus group discussion may seem similar in procedure to an interview, thus all aforementioned apply here as well, however in this case the facilitator task is mainly to keep the focus of discussion among the participants without biasing or influencing their views. Follow-up sessions may be needed in case unforeseen significant topics of discussion or concern have arisen.

3.3.5 Observations

Collection of data through direct observation of user actions in an actual setting or situation will also be conducted during the piloting and evaluation of the project. The observing researcher should ensure that her/his presence does not affect, influence or bias the actions or behaviour of the participant user(s). The observer should not offer any technical guidance or assistance during the observation, but only prior to it, through clear instructions and description of tasks to be conducted in the test study.

Furthermore, and in compliance with the aforementioned guidelines, all researchers acting as observers involved in school-based activities will be required to:

- follow all national procedures for verifying fitness to access school premises,
- have school-verified identification and a school liaison person available at all times during school visits,
- be aware of essential health and safety issues concerning students on school premises,
- have granted or obtained parental consent or equivalent authorisation for the evaluation activity e.g. students' involvement in videos, photographs etc.

3.4 Reporting of activities

An integral part of the execution plan is the reporting procedure that should accompany the piloting and evaluation work. Proper and up-to-date reporting is vital to monitor the development of progress, to provide a standard logging of activities per project partner, to determine and to ensure overall progress and quality. The reports, if applicable, will be attached to the project's official deliverables, D5.2-Implementation of the educational scenarios and evaluation report (first phase) due in M12 and D5.3-Implementation of the educational scenarios and evaluation report (second phase) due in M21, respectively.

In general, soon after each piloting and evaluation event or series of relevant activities, the partner, or partners involved, is expected to produce a report (see Table 3 for the proposed activity report template). This should be sent to the consortium and be uploaded to a repository/common workspace of the project. The reports document basic information about the activity such as date/period held, location, number of participants, target group and type of activity along with a brief and comprehensive description of the activities and main tasks or findings. Also, any materials in printed or electronic format that are related to the activity are to be attached to the report as well (e.g., instructions materials handed to participants, educational materials produced specifically for the activity, photos or videos taken during the event, template of questionnaires, topics of discussions etc.).

The reports of activities are discussed and reviewed during regular online and face-to-face project meetings to assess progress and results. The goal and scope of the reporting procedure will not change over the lifetime of the project, but specific improvements and revisions or complementary actions might be added over time if needed.

Table 3: Proposed template of activity report

ENVISAGE Report Code	*[LLXX-YYYYMMDD] Please follow this format:* *LL= 2 letter country code, XX = partner id, YYYYMMDD = date*
Activity Title and Type	*Indicate one or more of these types* *Teacher induction/training workshop* *In-classroom activity with students* *Usability test/teacher interview/observation* *If other please describe*
Country, city/region	
Working language	
Start/End date	
Partners involved	
Facilitator name and email	
Name of pilot school	
Number and age of participants	*Provide the number, or an estimate, of participant teachers and students. In case of students involved please also indicate their age group*
Activity description	*Write a brief description of the activity (e.g., usability test of authoring tool, methodology used, main objectives, any weblinks etc.)*
Main findings and observations	*Give a short overview of main findings or observations*
Any other relevant material	*Attach here any other relevant material (e.g. paper mock-ups, notes, screenshots etc.)*

4 Proposed schedule of activities

The proposed schedule of activities related to piloting and evaluation is shown in Fig.4.1. It practically starts with the beginning of the school year and concludes at the end of it. The work is divided into the following consecutive tasks:

- Introduction and training of teachers. Participating teachers are given first introduction seminar and familiarize themselves with the scope, methodology and objectives of piloting and evaluation.
- Pre-event feedback collection. Prior to any actual in-classroom piloting, usability test or study, teachers provide feedback on available virtual labs and tools related to analytics and visualization through structured questionnaires or/and interviews.
- Implementation in classroom. During this period the implementation of virtual labs with students in science classrooms is taking place. On-site observations from expert partners are also conducted.
- Post-event feedback collection. After the in-classroom piloting, teachers provide feedback on virtual labs, analytics and visualization tools through structured questionnaires or/and interviews.
- Analysis of results. Partners analyse the gathered data. First conclusions are drawn towards the improvement of available tools by the design and development team.
- Usage of authoring tool. Teachers use the authoring tool to change themselves various features of a virtual lab. Observations from expert partners are also conducted.
- Feedback collection. This is with respect to the previous usage case and is conducted through structured questionnaires or/and interviews.
- Analysis of results. Partners analyse the gathered data. Main findings and conclusions are drawn towards the finalization of available tools.

	2017							2018					
	Jun						Dec	Jan					Jun
Calendar Month	6	7	8	9	10	11	12	1	2	3	4	5	6
Project Month	9	10	11	12	13	14	15	16	17	18	19	20	21
Training of teachers													
Pre-event feedback													
Implementation in classroom													
Data and feedback collection													
Post-event feedback													
Analysis of results													
Usage of authoring tool by teachers													
Post-event feedback													
Analysis of results													
Project Month	9	10	11	12	13	14	15	16	17	18	19	20	21

Figure 4.1: *Proposed time schedule of activities related to piloting and evaluation work.*

5 Summary

We presented a pilot execution plan and evaluation methodology and protocol that will be followed to conduct a series of small-scale test implementations of virtual labs, accompanied by the authoring, analytics and visualization tools of ENVISAGE. The evaluation is foreseen to provide quantitative and qualitative feedback on the process of using the ENVISAGE authoring tool as a means for building virtual labs, on the offered analytics and visualization tools as means of support towards improving virtual labs and on the delivered virtual labs with respect to their effectiveness to meet the goals and expectations of both teachers and students in the learning process.

In this context, the process of using the authoring tool as a means for building virtual labs will be evaluated by teachers and e-learning experts as test users. Predefined tasks of gradual complexity will be given to them and feedback will be collected through various means and methods in order to evaluate the effectiveness, usability and functionality of the authoring tool as well the overall user experience. The evaluation outcomes will be utilized to develop, update and release an improved version of the tool. With respect to the support offered by the analytics and visualization tools in the process of improving virtual labs, this will be evaluated also by teachers and experts. Similarly, these tools' effectiveness, usability, functionality and user experience will be assessed towards their improvement and redesign. The delivered virtual labs and the learning content will be evaluated by both teachers and students as well as experts. The evaluation process will focus on measuring to which extent the virtual labs have achieved to motivate and to engage students in the learning process, and also to which extent their utilization in the classroom facilitated the teaching and learning objectives set or expected by teachers.

We proposed a methodology and protocol of conduct, which the quantitative and qualitative feedback collection from users is based on. Complementary methods are to be utilized, including structured questionnaires, interviews, focus group discussions, user testing and on-site observations, according to a well-defined procedure of conduct and reporting. A schedule of activities related to piloting and evaluation, which practically starts with the beginning of the school year 2017-2018 and concludes at the end of it, was also laid out.

References

[1] J. Nielsen, *Usability Engineering* (1993), Academic Press, Chapter 2.2, p. 26.

[2] F. D. Davis, *Perceived Usefulness, Perceived Ease of Use, and User Acceptance of Information Technology,* Management Information Systems Quarterly, MISQ (1989), 13:3, pp. 319-340.

[3] J. R. Lewis, *Psychometric Evaluation and Instructions for Use,* International Journal of Human-Computer Interaction (1995), 7:1, pp. 57-78.

[4] ENVISAGE Deliverable D8.1, *POPD Requirement No 1*

Appendix I - Typical test procedure and script

Hello, [*subject name].* I am [*facilitator name]* and I will be the one who assists you through this session today. Before we begin the actual test, I have some practical information so you get an idea of what we will be doing here today. Let us go through the test session as a whole, so you know what activities we will be doing the next [*Estimated session time*].

1) The test will start with a short post session interview where I will ask you some questions about demographics, level of experience with similar systems and preferences within learning.
2) Then we move on to the actual test, where you will be given a task, which I will ask you to perform in the system. Throughout the test session I would ask you to, as much as possible, speak aloud the thoughts that you might have when performing the tasks provided. It can be hard to do, as it does not come natural for people but I will of course help remember this, by asking you questions doing the test. Therefore, you should not feel like you are doing anything wrong, if I ask you a question.
3) After the test session, we will end by talking loosely about how you experienced the system and hereby wrap up the test session.

In addition, I want to clarify is that we are testing the program and not you. Therefore, you do not have to be afraid to make any mistakes. In fact, any problems you encounter will help us improve the experience of the system and is thus a big help to us. Remember that you are the expert as a user and your experience matters to us. Also, do not worry about hurting our feelings by providing criticism, as we are doing this to improve the system and we would appreciate to hear your honest reactions.

If you have any questions throughout this session or feel uncomfortable in any way, please do not hesitate to speak out. We might not be able to answer the questions right away, since we are interested in knowing how people solves the task when they do not have someone sitting next to them for help but if you get completely lost, we will of course try to help you and to report any obstacles occur. If you still have, any questions after this session we will try to answer them. Also, if you need a break during the session, just let us know.

[If applicable] We would furthermore like to ask you, if we can record this session? The recording will be used to help us figure out how to improve the system, and will only be used by people working on the project. It also enables us to be more present, during the test, as we do not need to take as many notes during it.

If you can allow this, we would ask you to read and sign this simple permission form. It is just stating that we have your permission to recording and that only people working on the project will use the recordings.

- *Give test subject permission form and pen.*
- *Start recording while subject signs.*

Do you have any question so far?

Test procedure:

Pre-session interview

Demographic questions:

Now we just need the last couple of questions and then the test sessions can begin.

- What is your occupation?
- How old are you?
- How many years of experience do you have within learning/teaching?
- What is your main teaching subject and age range of your students?
- Do you have any prior experience with the tool [authoring tool, wind energy lab, chemistry labs, etc]?
- How often do you use ICT as part of your teaching practice in science classes?
- Any prior experience with making digital/non-digital learning material?
    - If yes, could you please describe it?
- Have you ever used games or virtual labs in teaching?
    - If yes, how was your experience with this?

Great! We are now done with the questions and can thus move on to the actual test session.

Any questions before we begin?

- *Ensure the participants are sitting comfortable and feel ready*
- *Ensure in advance that the computers are working properly and are on the starting screen/first page the users should see during the test.*

Test session

The test session now starts and again I just wanted to remind you to think aloud while performing the task and that you are the expert during the test

- *Ask participant to perform task 1 (or sequence 1 of tasks):*
- *Remember to ask what the participant is looking for and is thinking about while they perform the task*

If the task IS solved: Alright that was the first task.

If the task is NOT solved: Okay, let us go to the next task.

- *Ask participant to perform task 2 (or sequence 2 of tasks):*
- *Remember to ask what the participant is looking for and is thinking about while the perform the task*

If the task IS solved: Alright that was the second task.

If the task is NOT solved: Okay, let us go to the next task.

- *Ask participant to perform task 3 (or sequence 3 of tasks):*
- *Remember to ask what the participant is looking for and is thinking about while the perform the task*

If the task IS solved: Alright that was the third task.

If the task is NOT solved: Okay, let us go to the next task.

Post-session interview

Now you finished the last task I got for you, I would therefore like to ask you how the session was for you?

Post-interview questions:

- How would you describe your experience with the system?
- What do think about the system?
- Do you see yourself using a tool like this in the future?
    - If no, do you know other teachers for whom it might be relevant?
    - If no, is there anything that you would make to use it?
    - If yes, how?
- In your opinion, is there any features or functions the system is missing?
- Anything you would like to add?

Interview ends

Thank you for your help, we appreciate that you could help us with this project.

Do you have any questions that you want to ask?

Draft of Consent Form

The purpose of the [observation session or interview or group discussion] and the nature of the study and the evaluation procedure and questions have been thoroughly explained to me.

I consent to take part in it and give usability or other relevant feedback in the context of the ENVISAGE project.

[If applicable] I also consent to be tape or video-recorded during this session in case it is needed for the purposes of the project. I understand that this film or recording will not be shared with anyone outside the project and that I can request it to be withdrawn or to be permanently erased.

My participation is voluntary. I understand that I am free to leave at any time and that none of my opinions or thoughts will be shared with anyone outside the project, unless all identifying information is removed first.

Place and Date

Please write your name, job function and affiliation

Please sign on your name

Appendix II - Structured questionnaire for teachers for evaluation of virtual lab in educational setting

Below are several statements regarding the virtual lab in an educational setting.

Please indicate to what extent you agree or disagree with each statement.

	strongly agree	agree	neutral	disagree	strongly disagree
The content presented in the virtual lab is correct and well balanced					
The virtual lab fits well with the curricula					
The virtual lab presents the learning content in a relevant manner for the students					
The quality of the learning content did not meet my expectations					
It is difficult to integrate the virtual lab into a learning context					
The learning material is presented in structure and complexity that suits the students' competencies					
I found that the instructions for the virtual lab were good					
In general, the students found the virtual lab difficult to navigate					
The interface of the was easy to understand					
The virtual lab gave the student a better understanding of the topic					
The students preformed as I expected in the virtual lab					
The learning goals for the virtual lab are clear.					
It was hard for me to evaluate the student's performance in the virtual lab					
I had a good sense of how the students were working with the virtual lab					
The virtual lab supports differentiated learning					

I would use the virtual lab again in my teaching					
I would like to change part of the virtual lab to better support my teaching					
The students found the virtual lab engaging					
The students found the virtual lab challenging					
The students enjoyed using the virtual lab					
The virtual lab simulated the students' interest or curiosity in the subject					

Appendix III - Structured questionnaire for students for evaluation of virtual lab in educational setting

1. **Before you played the virtual lab, what did you think it would be like?**

2. **What did you learn from the virtual lab?**

3. **How much did you learn from the virtual lab?**

☐	☐	☐	☐	☐
Did not learn much	Learned a little	Learned some	Learned much	Learned a great deal

4. **How interesting did you find the virtual lab?**

☐	☐	☐	☐	☐
Extremely interesting	Very interesting	Interesting	Slightly interesting	Not interesting at all

5. **Did the virtual lab live up to what you thought it would be like?**

☐	☐	☐	☐	☐
Absolutely yes	Mostly yes	Neither yes or no	Mostly not	Absolutely not

6. **Did you find the virtual lab easy or hard to use?**

☐	☐	☐	☐	☐
Very easy	Easy	Neither easy nor hard	Hard	Very hard

7. **Did you find the virtual lab fun to use?**

☐	☐	☐	☐	☐
Not fun at all	Slightly fun	Fun	Very fun	Extremely fun

8. **Would you like to use the virtual lab again?**

☐	☐	☐	☐	☐
Would definitely not try again	Would probably not try again	Would maybe try again	Would quite likely try again	Would definitely try again