(Project Number: 945301)

DELIVERABLE D1.3

Report on evaluation of the new developed labs

Lead Beneficiary: LUH

Due date: 30/09/2023 Released on: 28/09/2023

Authors:	Vivien Pottgießer

For the Lead Beneficiary	**Reviewed by Work package Leader**	**Approved by Coordinator**
Vivien Pottgießer	**Vivien Pottgießer**	**Mojmír Němec**

Start date of project: **01/10/2020** Duration: **36 Months**
Project Coordinator: Assoc. Prof Mojmír Němec
Project Coordinator Organisation: CTU VERSION: 1.2

Dissemination Level		
PU	Public	**X**
RE	Restricted to a group specified by the Beneficiaries of the A-CINCH project	
CO	Confidential, only for Beneficiaries of the A-CINCH project	

Version control table

Version number	Date of issue	Author(s)	Brief description of changes made
1.0	28.09.23	Vivien Pottgießer	Draft version
1.1	16/10/2023	Jana Peroutková	MST check
1.2	17/10/2023	Mojmír Němec	Coordinator’s check and approval

Project information

Project full title:	Augmented Cooperation in Education and Training in Nuclear and Radiochemistry
Acronym:	A-CINCH
Funding scheme:	Coordination and Support Action
ECGA number:	945301
Programme and call	H2020 EURATOM, NFRP-2019-2020
Coordinator:	Mojmír Němec
EC Project Officer:	Kateřina Ptáčková
Start date – End date:	01/10/2020 – 30/09/2023 i.e. 36 months
Coordinator contact:	+420 224 358 331, mojmir.nemec@fjfi.cvut.cz
Administrative contact:	+420 245 008 599, cinch@evalion.cz
Online contacts:	http://www.cinch-project.eu/

“This project has received funding from the Euratom research and training programme 2019-2020 under grant agreement No 945301.”

EXECUTIVE SUMMARY

After the completion of the VR-Lab with three virtual hands-on trainings (vHoT), two online questionnaires were created to evaluate the VR-Lab and the learning success by the execution of the vHoTs, as well as to determine the usability according to Brooke and consider physical effects caused by the virtual environment. The questionnaires were used with four test groups: during the summer school in Cyprus, at a seminar for authorities and one test run each with students at POLIMI and LUH. In total, about 40 completed questionnaires were included in the evaluation. The age of the testers ranged from 16 to 60 years and their previous knowledge in the field of radiochemistry varied widely. The overall concept of the VR-Lab was well received, and participants positively commented on the playful nature of the vHoTs. Due to minor bugs and occasional crashes of the application, the usability was not rated as very good. However, the instructions guided users well through the vHoTs and many felt that the process was close to reality. Cybersickness was rather sporadic and no participant had to abandon the test because of it. Measuring the success of learning was difficult due to the small number of participants in each HoT, but most users stated that they had learned a lot during the test. For the experiments on the determination of the half-life of K-40 and the separation of Po-210 and Pb-210, the participants would have liked a more detailed description of the chemical and metrological background.

CONTENT

1 INTRODUCTION

The final version of the virtual radionuclide laboratory (VR-Lab) was completed in month 33. In addition to the structural expansion, which included several laboratory rooms, a measurement room and the airlock area, three virtual hands-on trainings (vHoT) and the entering and leaving procedure were developed. For the evaluation of the VR Lab and the vHoTs, two online questionnaires were created, which the users had to complete before and after performing one or more vHoTs. The pre-survey only asked about the users' prior knowledge. The same technical questions were asked in the post-survey to measure learning success. In addition, user data such as age, gender and education level were collected in the post-survey. Five questions were asked about possible cybersickness and the 10 standardised questions about usability according to Brooke. Users were also asked how they rated each vHoT and the VR-Lab as a whole. The surveys were created using the LimeSurvey tool and set up so that users could access the survey via a link without having to register. This prevents the pre-survey responses from being correlated with the post-survey responses, but a cumbersome registration process usually reduces the willingness to participate in a survey, so this was not done.
The VR-Lab was tested and evaluated at four different events: the seminar for authorities described in deliverable D1.6 (four participants), the test run of a flipped classroom event at LUH described in D1.5 (four participants), the summer school in Cyprus (20 participants) and two-day testing event at POLIMI (12 participants). In total, feedback was collected from about 40 people, although it should be noted that during the summer school in Cyprus, the participants did not have enough time to finish the vHoTs, so only the evaluation of usability, the overall impression of the VR-Lab and the resulting cybersickness can be used from these questionnaires. Learning success cannot be determined for this group. The group of participants in all evaluations is very mixed. In Cyprus there were pupils with almost no prior knowledge and students with some prior knowledge. The test at POLIMI was participated by Masters students in Radiochemistry and Nuclear Engineering and at LUH by students and volunteers who had just finished their high-school diploma. The age range of these participants was 16-31 years. In comparison, the participants in the seminar of the authorities were much older with an age range of 34-60 years. For the government seminar and the flipped classroom test, the participants completed the vHoTs independently at home, whereas for the other tests the participants were on-site and were supported by the tutor’s present.

2 EVALUATION

2.1 System Usability Scale

The System Usability Scale (SUS) was created by John Brooke in 1986 to evaluate the usability of a wide variety of digital products. Users are asked to agree or disagree with 10 statements about the application on a scale of 1 (=strongly disagree) to 5 (=strongly agree). The statements are the following:

1. I think that I would like to use this system frequently.
2. I found the system unnecessarily complex.
3. I thought the system was easy to use.
4. I think that I would need the support of a technical person to be able to use this system.
5. I found the various functions in this system were well integrated.
6. I thought there was too much inconsistency in this system.
7. I would imagine that most people would learn to use this system very quickly.
8. I found the system very cumbersome to use.
9. I felt very confident using the system.
10. I needed to learn a lot of things before I could get going with this system.

These statements are coded in the standard procedure with values from 0 to 4. The coding depends on the wording: In the case of positively worded statements, a full agreement response is coded as 4 and a completely negative response is coded as 0. In the case of negatively worded statement, a response of full agreement is coded as 0 and a completely negative response is coded as 4. The results of the SUS questionnaire are used to calculate a so-called SUS score, which can have a value between 0 (worst imaginable use) and 100 (best imaginable use): When the questionnaire is scored, the numbers obtained are added together - the sum is between 0 and 40 - and then multiplied by 2.5. Although the scores are between 0 and 100, they are not percentages and should only be considered in terms of their percentile ranking. Based on research, a SUS score above 68 would be considered above average and anything below 68 would be considered below average.

The scores for the four different user groups were as follows: Seminar for authorities: 46, test at IRS: 55, test at POLIMI: 67 and at the summer school: 69. When evaluating the results, it must be taken into account that the first two groups of participants were very small, which makes the score less statistically robust. The two groups also have in common that the users tested the VR-Lab alone at home. The lower score could therefore be related to the fact that the users did not have the opportunity to get short-term help with minor problems, but had to find a solution on their own. Which is sometimes time consuming and very frustrating. Such problems included imprecise tasks such as rinsing out the canister, which should be done via the "Rinsing" checkbox, which is marked red in **Chyba! Nenalezen zdroj odkazů.**, and not via the slider below it.

Figure 1: Screenshot from the task „rinse HDPE container“ which is only fulfilld by using the red marked checkbox.

However, the scores of the two groups that received on-site support are also just at the limit of what would be considered good usability. It can therefore be concluded that there is still room for improvement in the usability of the VR-Lab. This was also evident from the written and verbal feedback. Some users reported that the application crashed several times. Another point of criticism was the linear structure of the quest system, which was sometimes too detailed and sometimes not detailed enough. Users also had to repeat some steps because they were done at the wrong time, e.g. weighing samples in the wrong order in the K-40 experiment or disposing of pipettes too early in other experiments.

Despite the minor bugs and inconsistencies, most participants said or wrote that they enjoyed doing the vHoTs and enjoyed the playful nature of the VR-Lab.

2.2 Cybersickness

Cybersickness mainly occurs when using VR headsets, but this phenomenon can also be observed in 1st person computer games that are watched on a screen and controlled with a mouse and keyboard. Therefore, the users were asked how much the use of the laboratory had affected them physically. They were asked to rate how true the following statements were on a scale of 1 to 5, with 1 being strongly disagree and 5 being strongly agree:

- I felt nauseous during use
- During use I had eyestrain
- I felt vertigo during use
- During use I got a headache
- During the use, general discomfort occurred

The results for the four groups were as follows:

Figure 2: Results of the survey on the physical effects of using the VR-Lab.

The summer school participants in particular experienced fewer physical effects. However, this could also be due to the fact that the participants shared a computer in pairs and also spent less time in the virtual lab. Overall, the occurrence of physical effects can be classified as average.

2.3 Individual vHoTs

Participants were asked about the implementation of each vHoT and an attempt was made to measure learning. Unfortunately, the latter was not possible for two groups, the seminar for authorities and the summer school, as there was not enough time to complete the experiments in the summer school and 12 participants took part in the pre-survey for the authorities, but only 4 in the post-survey, so learning success cannot be measured here either.

For the vHoTs assessment, participants were asked to indicate their level of agreement with the following statements, where 1 is strongly disagree and 5 is strongly agree:

- The instructions guided me well through the experiment
- The execution of the virtual experiment was close to that of a real experiment
- I had problems with the language
- The workflow was smooth and without technical problems
- I learned a lot by carrying out the virtual experiment

As not all participants completed each vHoT, all groups were combined for analysis to give a more statistically significant result.

2.3.1 Half-live determination of K-40

Most of the participants carried out the experiment to determine the half-life of K-40. As can be seen in Figure 3, hardly any participants had problems with the language, although only a few participants spoke English as their first language. Most participants also rated the instructions and the realism of the VHoT as rather good. The workflow, on the other hand, is perceived as less good. In addition, most participants rated their learning gain as good. However, this statement should be interpreted with caution. Most of the answers come from participants of the summer school who were not able to complete the experiment, so that no statement can be made about the fulfilment of the actual learning objectives. In addition, there were many pupils among the participants who had little previous knowledge about radioactivity or working in a laboratory. For this group of participants, it is easier to have a higher learning gain through the VR-Lab than for students who already have more prior knowledge and practical experience in a real laboratory.

In the evaluation of the technical questions to check the learning success, no learning gain could be determined, which is due to the fact that almost all participants had already answered the questions correctly in the pre-survey. In the oral debriefing of the experiment, however, it became clear that the participants did not correctly capture the background of the experiment. They did not understand that it was not the gamma line of K-40 that was measured, but the beta decay. The self-absorption of the sample would play a negligible role in the measurement of the gamma lines, but the participants did not realise this because the evaluation of the measurement was done by the computer and they did not look at the measured values themselves. In order to improve learning of the correlation between the mass/thickness of the preparation and the counting rate, the information text in the instructions should be revised and a task for independent evaluation of the measurement results should be added.

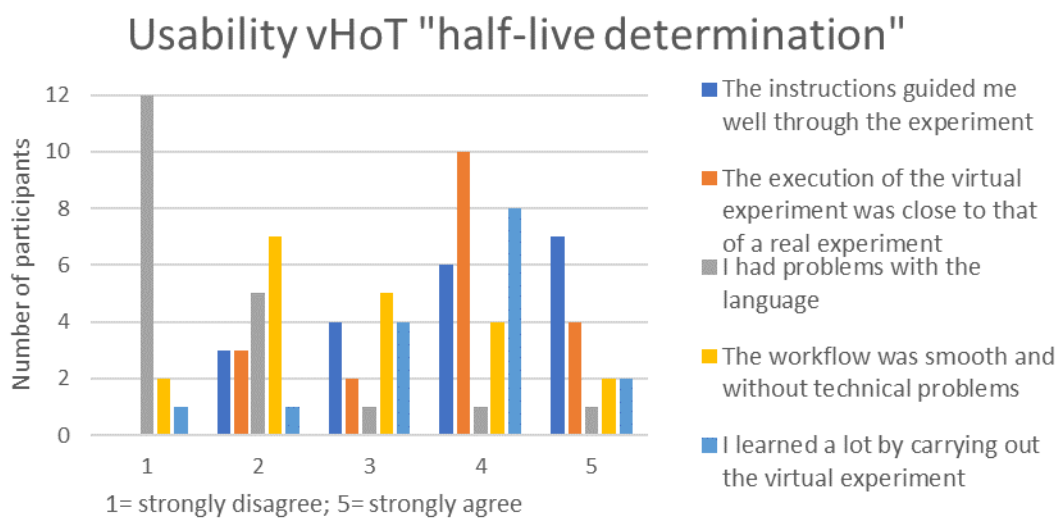

Figure 3: Evaluation of the vHoT to determine the half-life of K-40.

2.3.2 Seperation of Pb-210/Po-210

Figure 4 shows that also in the Pb-210 and Po-210 separation experiment, language was not a problem for most participants. The vHoT was rated as quite realistic and the instructions as good. The additional videos, which can be viewed at [I]nstructions, were particularly positively mentioned in the personal feedback. The workflow and the learning gain were rated as rather mediocre. This is also reflected in the written and oral feedback. The participants would have liked more explanations of why the individual steps are carried out and the evaluation of the measured values is no longer part of the vHoT. Unfortunately, the principle of column chromatography with a selective resin could not be explained to the participants. Nevertheless, the vHoT has given them a good insight into the steps required for the chemical separation of Po-210 and Pb-210 and the subsequent sample preparation of alpha and beta samples.

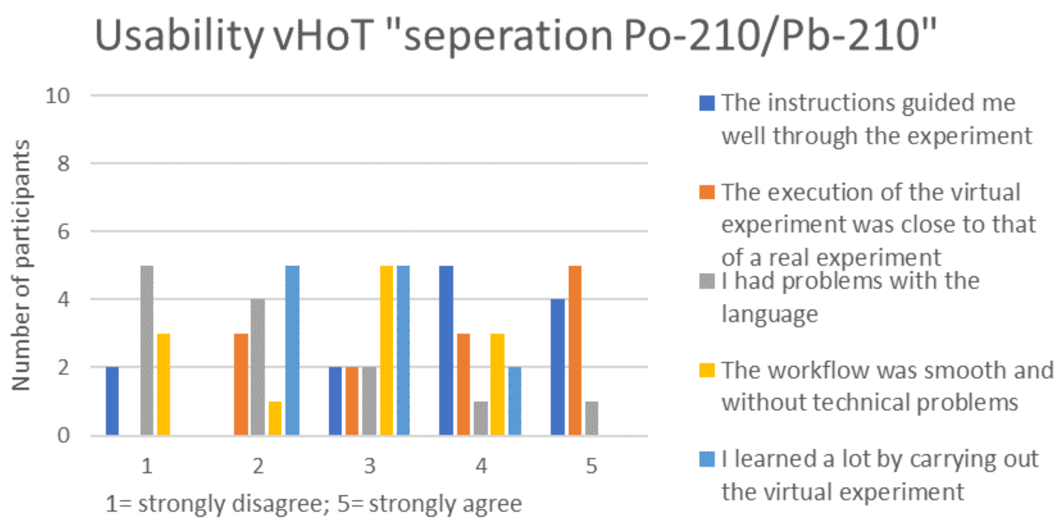

Figure 4: Evaluation of the vHoT on separation of Pb-210 from Po-210.

2.3.3 Superficial decontamination

The fewest participants chose the superficial decontamination and waste conditioning experiment, so the evaluation shown in Figure 5 is the least reliable. Nevertheless, it is clear that there were no problems with the language, the instructions guided well through the experiment and the execution felt realistic to the participants. As with the other vHoTs, the workflow is rated rather average. On the other hand, the learning gain was rated higher. This is probably due to the more detailed explanations under [I]nstructions. Unfortunately, the questions on learning gain do not support this, mainly due to the fact that many participants gave the correct answers in the pre-survey and some did not complete the post-survey. The video part of the vHoT was evaluated very differently by the participants. Some found it to be an interesting diversion, while others described the videos as disruptive to the flow of the vHoT.

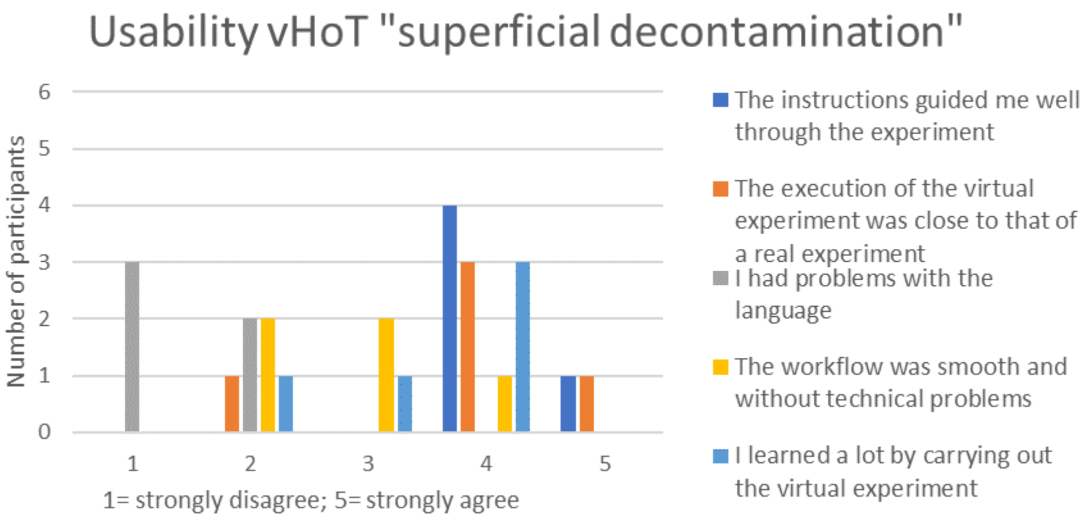

Figure 5: Evaluation of the vHoT about superficial decontamination and conditioning.

2.4 Overall feedback

At the end of the post-survey, participants were asked about their overall impression of the VR Lab. As can be seen in Figure 6, the majority of the participants enjoyed the application and it was able to motivate some participants to engage more closely with radioactivity.

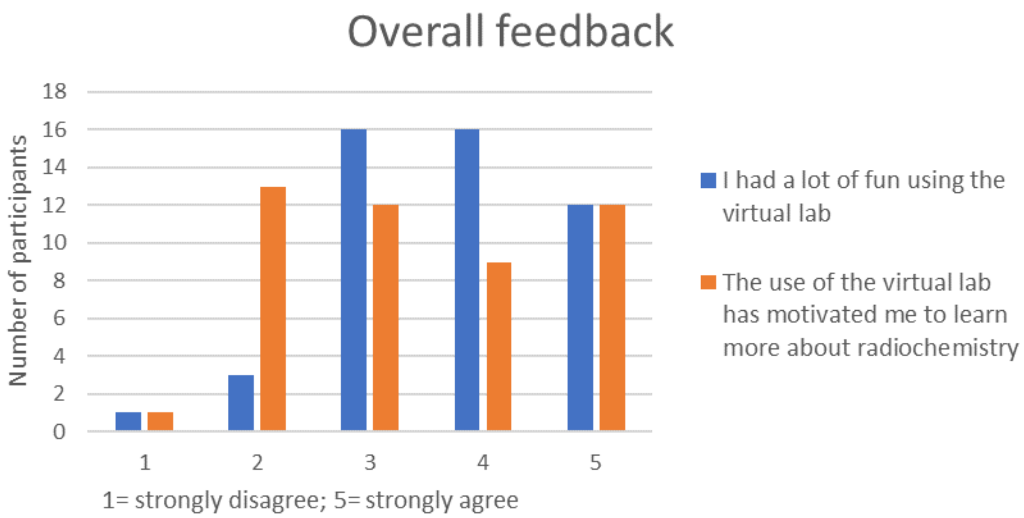

Figure 6: Evaluation of the whole VR-Lab experience.

In addition, there were three open questions: what the participants liked most or least, what could be improved and an option for further feedback. These responses emphasised that the general concept of the VR-Lab was good and that the playful nature of the vHoTs was fun. As expected, the minor bugs and the fact that the application is occasionally crashed were rated negatively. Suggestions for improvement were that frequently used items such as pipettes or syringe bottles should be kept at the workstations to save walking to the equipment storage cabinet.

3 CONCLUSIONS

In conclusion, the VR-Lab and the implementation of the vHoT was fun for most of the testers and the playful character was well received. The realism and the instructions were rated rather positively and even though the learning success could not be measured by the technical questions in most cases, many participants rated their learning success as good. There were two main problems with measuring learning through the technical questions: firstly, the number of participants was often not high enough, and often more participants took part in the pre-survey than in the post-survey, which made evaluation difficult. Secondly, the questions were not always optimally chosen, which was also due to the fact that measurable operators were not always used when formulating the learning objectives. Personal feedback indicated that participants would have liked more explanation of the chemical background and how to interpret the measurement data, especially for the VHoT on the half-life of K-40 and the separation of Po-210 and Pb-210. Small bugs and occasional crashes of the application partly affected the flow of the vHoTs, but most participants were still motivated to complete the tasks.