

D5.4 PRACTICAL RECOMMENDATIONS FOR ASSESSMENT METHODS

Project acronym: OTTER Project title: Outdoor Science Education for a Sustainable Future Call: H2020-SwafS-2018-2020



This project has received funding from the European Union's Horizon 2020 research and innovation programme under Grant Agreement No 101006482



Project no.	1010010082
Project acronym:	OTTER
Project title:	Outdoor Science Education for a Sustainable Future
Call:	H2020-SwafS-2018-2020
Start date of project:	01.09.2021
Duration:	30 months
Deliverable title:	Practical Recommendations for Assessment Methods
Dissemination level:	Public
Due date of deliverable:	29.2.2024
Actual date of submission:	29.2.2024
Deliverable Lead Partner :	University of Limerick
Work Package:	WP5
Keywords:	Assessment, EOC Assessment

Please cite as:

Kelly R., O'Neill D., McCormack O., Azevedo N.H., Kray, Zs. (2024). D5.4 Practical Recommendations for Assessment Methods. Ireland. 47 pages.





Name	Organization
Deirdre O'Neill	University of Limerick
Regina Kelly	University of Limerick
Orla McCormack	University of Limerick
Nathália Helena Azevedo	University of Groningen
Zsuzsanna Kray	European Science Foundation

History			
Version	Date	Reason	Revised by
01	8/01/24	First draft of deliverable	UL, RUG
02	15/01/24	Review of deliverable	ESF, UL
03	22/1/24	Review of deliverable	UL, RUG
04	9/2/24	Review of deliverable	UL
05	20/2/24	Review of deliverable	UL
06	28/2/24	Review of deliverable	UL, RUG, ESF
07	29/2/24	Final draft	UL





Table of Contents

1 OTTER project	7
2 Project Consortium	8
Executive Summary	9
3 Summary	10
4 Objective	10
Assessment Practices	11
5 Assessment practices in education	12
5.1 Purposes of assessment	12
5.2 Forms of Assessment	13
5.2.1 Feedback	13
5.2.2 Questioning	14
5.2.3 Formative use of summative assessments	14
5.3 Some considerations in terms of assessment	14
5.3.1 Curriculum coherence	14
5.3.2 Universal Design for Learning	16
5.3.3 Assessment of skills	16
5.4 Literature review findings about assessment of EOC	16
Feedback on assessments from pilots	18
6 OTTER pilot assessments	19
6.1 Country-specific assessment practices	19
6.1.1 Frequency and type	19
6.1.2 Role of the teacher in pilot countries	20
6.2 Reports of assessment practice that worked well in the OTTER pilot countries	21
6.3 Needs expressed by OTTER pilot countries	21
6.4 OTTER Lab adaptation opportunities in pilot countries	22
Proposed evaluation for EOC assessment methods	24
7 Proposed EOC assessment methods	25
7.1 Summary of overall assessment findings	25
7.2 Overall EOC assessment recommendations	26
7.3 Revised approach to OTTER	27
7.4 Student Focused EOC Assessment Rubric	27
7.5 EOC Principles and Suggested Principles in Practice	30
8 Reference List	34
Appendices	38
9 Appendix A	38



@tter

10 Appendix B	39
11 Appendix C	40
12 Appendix D	43
13 Appendix E	46





List of Acronyms

D	Deliverable
EOC	Education outside the classroom
STEAM	Science, Technology, Engineering, Arts, and Mathematics
WP	Work Package





List of Figures

Figure	1. Revised approach to	OTTER including reflect and	assessment throughout	27
Iguie	1. Revised approach to	OT TER Including reliect and	assessment infoughout	····· ∠ /

List of Tables

Table 1: Sample learning objective and appropriate assessment	15
Table 2. Methods of assessment and exemplary studies in EOC research	17
Table 3: Sample Instruments to assess skills throughout the OTTER Lab steps	29
Table 4: Principles and of Suggested Principles in Practice	31





1 OTTER project

OTTER is a H2020 funded project that aims to enhance the understanding of Education Outside the Classroom (EOC) approaches and how they can help improve the acquisition of scientific knowledge and transferable skills in students, specifically in the field of environmental sustainability and the reduction of plastic waste. It aims to increase interest in scientific topics among young people, while also contributing to the range of innovative educational projects and the increase of scientific citizenship within the EU.



OTTER aims to strengthen educational outside-the-classroom (EOC) **networks within Europe**, connecting experts from four different regions within the continent (**Finland**, **Hungary**, **Ireland and Spain**). The strengthening of these networks will be utilised to carry out a programme of EOC pilot schemes and analyse of the effect they have on the performance of participating students, including their levels of sophisticated consumption and scientific citizenship, to increase understanding of the effects of education outside the classroom on EU citizens. The pilot schemes will share a common theme revolving around issues of plastic waste and recycling in order to build upon recent momentum in tackling related global educational, social, and environmental issues and due to the close relationship between reducing plastic waste and the need for more sophisticated consumers.





2 Project Consortium





Geonardo Environmental Technologies (GEO)

European Science Foundation (ESF)



University of Groningen (RUG)





OLLSCOIL L

Learning Scoop - oppimisen osuuskunta (LS)

The Big Van Theory (TBVT)

Center for the Advancement of Research & Development in Educational Technology (CARDET)

Executive Summary

3 Summary

This report aims to provide some advice and recommendations regarding assessment within EOC, and also when evaluating EOC in this regard. The report is structured in four sections. Section one introduces the report and report structure. Section two explores the literature in terms of assessment practices in education and EOC. Section three provides feedback on assessment from the pilot countries, while section four proposes an approach to embed and evaluate assessment methods within EOC.

Drawing on research literature and findings from this project, emphasis is placed on supporting assessment practices that are coherent; draw on assessment for learning strategies; adopt principles of Universal Design for learning supporting choice and flexibility; support a range of learning outcomes; consider pre-existing assessment practices and support student and teacher assessment literacy.

Based on the research literature and findings from OTTER, this report recommends that flexible tools for assessment are implemented and embedded in EOC practices; that all and any approach recognises international best practice; that effort and priority is placed on developing teacher agency and assessment literacy and that both formative and summative approaches to assessment are necessary to support learning and alignment with existing practices.

Based on the research findings from OTTER, a revised approach to the OTTER labs is proposed within this report, wherein reflection and assessment are embedded throughout and at each stage of the EOC experience rather than as a separate and standalone approach, hence supporting coherency with the learning experience.

4 Objective

The purpose of this deliverable is to provide guidance regarding assessment methods for EOC, and to feed into the guidelines for accrediting EOC and the toolkit for practitioners. It will define the main quality parameters the assessment will have to be based on and will suggest the construction of a practical evaluation grid to assure a common ground for the harmonized evaluation of practitioner/programme compliance. These practical recommendations are based on ideas reflected over OTTER activities and on best practices used at the national level in different countries which are part of the OTTER consortium. The assessment methods are conceived as a mutual/parallel tool used to evaluate the educational activities which aim to complement the assessment methods in education and not to replace the "traditional" methods.

Assessment Practices

5 Assessment practices in education

Curriculum is not just the content to be covered but also includes the pedagogical and assessment approaches utilized during the learning process (Gleeson and Donnabháin 2009). Assessment, particularly high stakes summative assessment (assessment that takes place after the learning occurs), plays an influential role in the teaching and learning process, impacting on what is learned, how learning takes place and what is valued (Lange and Meaney 2012; Hopmann 2015). As a result, assessment becomes the tail that wags the curriculum dog (Hargreaves 1989). Therefore, it is important that assessment is considered during the planning process to ensure approaches to assessment align with the forms of learning one wishes to promote. Within this section we will explore the different purposes of assessment, some forms assessment can take and key aspects we need to consider when planning assessment.

5.1 Purposes of assessment

Assessment serves multiple purposes, evident within the many understandings of the research literature (Archer 2017; Schildkamp and Kuiper 2010; Newton 2007). In general terms, the differing purposes can be grouped into three main categories as identified by Archer (2017), which are listed, and briefly described below:

- 1. Assessment to Support Learning: assessment can be used to inform the learner (and teacher) about their learning to support, promote and improve learning. Assessment can identify gaps in understanding and areas that warrant additional attention, hence supporting the learner to improve their learning. Formative assessment, particularly assessment *for* learning, both of which are described in more detail in the next section, play a key role here. While not an exhaustive list, this may include criterion-based assessment, feedback (teacher/peer/self) and using formative assessment to study for summative examinations, i.e., traffic light approach to reviewing learning (Black et al. 2004).
- 2. Assessment for Accountability: such approaches to assessment are usually utilised at institutional or governmental level and can be employed to hold educators to account, particularly when justifying funding of educational institutions. This can entail using and comparing student grades against an agreed benchmark to show performance, or lack thereof. The No Child left Behind Act in the USA is one such example (Meier 2004).
- 3. Assessment for certification, progress, and transfer: in this instance assessment is used to indicate that a student has reached a specific standard or has acquired a particular level of competency, skills, and knowledge. Assessment can also be used to determine who can progress to the next level of education and what form that progression takes, i.e. to a university

or polytechnic. In an Irish context, for example, grades in the Leaving Certificate Established determine what, if any, university and/or college course students' progress to (McCormack et al. 2020).

While the three purposes are not mutually exclusive, our central focus here is on purpose one, using assessment to support, promote and motivate student learning. However, it is important to acknowledge that such forms of assessment and related learning occur within contexts where assessment for other purposes co-exist. This can create tensions within the learning process as focus can be placed mainly on learning to pass examinations to ensure progression, rather than on supporting deep and effective learning (McCormack et al. 2020; Au 2011).

5.2 Forms of Assessment

Assessment can be summative or formative and aligned with this can be considered assessment 'of' or 'for' learning (Black et al. 2004). Summative assessment occurs at the end of the learning process and indicates how much a person knows or can do on a specific topic. As a result, it is considered to be assessment 'of' learning, in that it assesses past learning at a given point in time. While it doesn't always have to be the case, summative assessment (assessment of learning) doesn't always improve learning or provide feedback on the learning process (Ninomiya 2016). Such forms of assessment merely tell a learner how much they knew about a given topic at a given time (McCormack et al. 2020). Such forms of assessment particularly align with purpose 2 and 3, as discussed above.

Formative assessment occurs during the learning process and is used to both inform and improve learning and student achievement (Elwood 2006). Formative approaches to assessment include such assessment for learning strategies as peer/self-feedback, teacher feedback, effective use of questioning and utilizing formative assessment to prepare for summative examinations, all of which are now briefly described.

5.2.1 Feedback

Assessment for learning supports the use of feedback on specific criteria, rather than grades. Grading rubrics are helpful here. Criteria should be discussed and agreed with students in advance, so they understand the criteria, know what to focus on and how they can succeed. Feedback can be provided by teachers, peers or can take the form of self-feedback. The approach adopted can depend on the age and experience of the cohort in assessment for learning procedures. Peer assessment/feedback has been shown to support curriculum coherence (Shiel and Murchan 2017), which is considered later in this section. Feedback should be constructive, identifying aspects students did well and areas for further improvement and consideration. The research literature suggests that when students receive both a numerical score or letter grade and feedback, the feedback tends to be ignored

therefore assessment for learning principles support the use of comment feedback only (Black et al. 2004).

5.2.2 Questioning

Research evidence indicates that teachers, when they ask a question, provide limited time for students to consider a response and/or answer the question themselves. Assessment for learning principles support the effective use of higher order questioning and 'wait time' in questioning to ensure that students have sufficient time to consider and provide an in-depth response. This could entail, for example, giving students a higher order question and allocating a set period of time for them to consider their response. Students can work individually, in pairs or small groups, depending on their age and experience of group work. Again, establishing and discussing success criteria in advance of the task is advised here (Black et al. 2004)

5.2.3 Formative use of summative assessments

Summative and formative assessment are not mutually exclusive and formative assessment can be used to support students to prepare for summative exams. Formative assessment strategies can support students to reflect on their learning and to identify areas of learning that warrant further time and attention e.g. students can 'traffic light' work as red, amber, green to highlight areas of serious concern, some concern, no concern respectively (Black et al. 2004).

5.3 Some considerations in terms of assessment

Along with integrating assessment for learning practices, there are other issues that warrant consideration when planning assessment. We will briefly explore three of these, which are coherence within curriculum, integrating Universal Design for Learning (UDL) principles when planning curriculum (and assessment) and assessment of skills.

5.3.1 Curriculum coherence

Coherence within curriculum focuses on the alignment within and between aims/objectives/learning outcomes¹, teaching and learning approaches and assessment. It can also focus on development coherence, where learning occurs in a coherent, development and sequenced way (Jin et al. 2019), and also on the coherence between classroom-based assessments and high stakes state assessment models i.e. does high stakes state assessment align with and support school and

¹ Terminology can differ across contexts.

classroom approaches to assessment (Shepard 2004). Education systems, schools or programmes that achieve curriculum coherence improve student performance (Newmann et al. 2001).

Learning objectives identify what a learner should ideally know, be able to do or value at the end of a learning experience. These learning objectives should ideally contain the cognitive, affective, and psychomotor domains inclusive of knowledge; skills; attitudes, values, and ethics and also include a range of higher order thinking.

Assessment, and pedagogical approaches, should be planned in alignment with these learning objectives to support curriculum coherency and to support the type of learning the curriculum is trying to promote. While there may be overlap between the approaches to assessment across learning objectives (i.e. the same assessment approach can align with several learning objectives), it is important that each learning objective has a suitable and appropriate assessment method to support curriculum coherency. Say, for example, we want to develop communication skills amongst students, and we identify a related and suitable learning objective to support this; we need to ensure that teaching and learning approaches within the curriculum enable students to communicate in a variety of different ways and we also need to ensure that we assess students capacity to do so in an appropriate way through, for example, a presentation, a debate or interview. Table 1 portrays further examples of this.

Learning Objective	Learning Experience	Assessment
Students will develop communication skills	Students will work in pairs and small groups; students will engage in conversation with external stakeholders in both written and oral forms; students will develop an awareness campaign	Self-assessment of ones communication skills based on agreed criteria administered at start, middle and end of term Teacher observation of individual students communication skills administered at 3 timepoints Group presentation on
Students will develop critical thinking skills	Higher order questioning by teacher; completion of a group artefact that reflects their learning	Self assessment on ones level of critical thinking Teacher observation and use of higher order questioning

Table 1: Sample learning objective and appropriate assessment

Collaborative environments, involving co-planning, that promote engagement with professional development are supportive contexts to enable coherency (Newman et al. 2001).

5.3.2 Universal Design for Learning

Universal Design for Learning (UDL) acknowledges the diversity of learners that exist within all educational contexts and places value on student differences and interests (Hymel and Katz 2019). The underlying principles of UDL support educators to consider and cater for the diverse learning needs of students (Carrington et al. 2020). UDL enables and supports educators (and students) to consider, plan and provide for multiple means of engagement, representation and action and expression (CAST 2018). UDL principles align, amongst other things, with choice, autonomy, collaboration, self-regulation, and self-assessment. UDL principles enable learners to display and present their learning in differing ways. At a practical level, adopting UDL principles may involve educators considering multiple modes of delivery of content as well as the inclusion of a diverse range of resources to achieve the same learning objectives. Offering students choice and autonomy is central. In terms of assessment, this could again entail providing choice for students both in terms of which assessment questions they answer, but also in terms of the mode in which they complete the assessment. This could mean, for example, that students when given an assessment time point, may complete the assessment through using a concept map, a podcast or an essay written response. Such flexibility has been shown to have a positive impact on students' learning experience (Balta et al. 2021).

5.3.3 Assessment of skills

A challenge that exists regarding the assessment of skills relates to the ambiguity around the description of a particular skill (Voogt and Roblin 2010). It is important that educators discuss the expected educational outcome for a particular skill at the specific educational level relevant to their students. Educators need to define each skill they aim to develop so that appropriate assessment can be developed. This may require additional professional support for educators (González-Salamanca et al. 2020). Exemplary skill assessment instruments should be shared within schools to support educators developing their own skill assessment.

Formative assessment procedures more so than summative assessments appear to support the assessment of skills (Voogt and Roblin 2010). Drawing on previous assessment considerations, it is important to adhere to coherence within curriculum, integrate Universal Design for Learning (UDL) principles and involve students in self and peer assessment when planning the assessment of skills.

5.4 Literature review findings about assessment of EOC

Assessment of EOC interventions can encompass a variety of methods and instruments aimed at evaluating their impact. Our comprehensive review of the literature (see D2.1) reveals a diverse array of assessment techniques employed across numerous studies, and Table 2 shows a summary of these results. There is a strong overlap within the research literature between using research instruments to evaluate an intervention and typical classroom forms of assessment a teacher might employ in their everyday practice. As our interest here is on assessment of EOC, rather than evaluating for research purposes, the focus is solely placed on assessment approaches and methods within this section. Readers can review D2.1 for further information on research tools for evaluating EOC.

Several studies opted for assessments or exams to evaluate the knowledge acquired in specific subjects, as evidenced by the inclusion of 15 studies in our review. Additionally, interactive quizzes, assessing student work (artefacts) and observation protocols were frequently used to assess student learning.

Methods of assessment	Studies
Interactive Quizzes	Ying et al. 2019
Analysis of Students' Work/Drawings	Bhattacharya et al. 2021; Ghadiri Khanaposhtani et al. 2018; Giamellaro 2014; Harris and Bilton 2019; Kärkkäinen et al. 2017; Kermish-Allen et al. 2019; Puttick and Tucker-Raymond 2018
Observations	Affeldt et al. 2015; Ariosto et al. 2021; Dunlop et al. 2019; Ghadiri Khanaposhtani et al. 2018; Giamellaro 2014; Harris and Bilton 2019

Feedback on assessments from pilots

6 OTTER pilot assessments

In this section, country specific assessment norms are highlighted. Aspects of the OTTER assessment strategy that pilot countries found useful are reported, mentioning some adaptations where relevant.

6.1 Country-specific assessment practices

6.1.1 Frequency and type

OTTER pilot countries have various exam periods throughout education levels. These are usually summative tests but may include some classroom-based assessments. To capture the extent to which summative examinations are prevalent in each country, the number of state exams, general format of the assessment (exam/other), and compulsory subjects are included below.

Finland has only one high stake exam period, the Matriculation Exam. These exams occur in upper secondary school (high school). The Matriculation Examination consists of a minimum of five tests; one of them, the test in mother tongue and literature, is compulsory for all candidates. The candidate then has to complete four other tests from three different groups: mathematics; second national language; foreign language; humanities and natural sciences. At least one of the tests must be at an advanced level of the subject. Test items may be traditional essay questions, multiple-choice questions, drawing assignments, data analyses, and combinations of these. Questions may include text, pictures, videos, audio recordings, maps, animations and statistics (Matriculation Examination Board 2024).

In Ireland, at primary level students complete standardized tests in mathematics and English in 2nd class (8 years), 4th class (10 years) and in 6th class (12 years). These tests are graded by teachers and reported to the Department of Education. Junior Cycle exams (Department of Education , 2023) take place in Lower Secondary Level (3rd year; 15/16yrs) and are heavily weighted terminal exams with some classroom-based assessments. Students are required to study Irish, English, Maths and History along with optional subjects. Leaving Certificate exams (NCCA, 2024) take place in Upper Secondary Level (6th class; 17/18yrs) predominantly as terminal examinations with some additional assessment methods including oral and aural examinations, practical examinations and assessment of practical coursework at the end of the two-year programme of study. STEM subjects are not deemed compulsory at second level by the Department of Education, but some schools may require students to study science at lower secondary level. take place in Lower Secondary Level (3rd year; 15/16yrs) and are heavily weighted terminal exams with some classroom-based assessments.

Students are required to study Irish, English, Mathematics and History along with optional subjects. Leaving Certificate exams (NCCA, 2024) take place in Upper Secondary Level (6th class; 17/18yrs) predominantly as terminal examinations with some additional assessment methods including oral and aural examinations, practical examinations and assessment of practical coursework at the end of a two-year programme of study. STEM subjects are not deemed compulsory at second level by the Department of Education, but some schools may require students to study science at lower secondary level.

In Hungary there is one high stake exam, the secondary school graduation, matriculation exam (12th class, 18/19 years old students), available in Mathematics, Biology, Chemistry, Physics, Environmental Protection and Water Management, Natural Science, at ordinary and higher level. Students must take a minimum of five exams as per the following: three are compulsory (Mathematics, Literature and History), one chosen foreign language and one freely chosen subject (this is where students usually choose a science subject as universities may require as part of their application process). The results of the exams are optionally included in the points one can get during the university application process. These are typically summative tests, but as a new element in the graduation system, from 2024 onwards, students can prepare a portfolio or project work from science subjects as an alternative to the tests.

In Spain there is one high stake exam which is the university entrance exam. Each autonomous community refers to this type of exam differently. The tests in Catalonia (Canal Universitats, 2024) are available in Biology, Chemistry, Earth and Environmental Sciences, Physics, and Mathematics.

6.1.2 Role of the teacher in pilot countries

In Spain, Ireland and Hungary, assessments are graded and mostly summative. In Ireland assessments are designed and graded by the Irish State Examinations Commission, not the subject teachers.

In Finland, assessment focuses on the pupil's learning, working skills and behaviour. Assessment is mainly formative in nature. Teacher feedback that promotes learning is qualitative and descriptive in nature; it is about interaction that analyses and resolves critical areas of learning. Feedback provided by the teacher intends to make the learning process visible and promote learning. A cornerstone of the Finnish education system is that there is no teacher inspection. Teachers are not supervised or monitored in any way. On the contrary, teachers are considered as experts in their own profession who can assess, reflect, and redirect their own work.

6.2 Reports of assessment practice that worked well in the OTTER pilot countries.

Student portfolios and assessment of the student artefacts was a commonly used practice in all countries, especially in lower grades. Collection of student work throughout the OTTER lab to represent student learning was widely used.

Student self-assessment tools (exit tickets) proved to be useful in Ireland, and with some adjustments also proved successful in Finland and Hungary. Irish schools indicated that a peer-assessment tool, similar to the exit tickets brings an interesting insight into the overall OTTER assessment practice. In Ireland, teachers reported that the IMPACT Student Worksheet was good for assessing ICT Literacy (Appendix D), a knowledge base that they admitted they do not normally assess. In Finland this type of assessment is very unlikely to be used, as digital skills are assessed in other ways.

Teacher self-assessment, similar, to the Teacher Rubric in the Make an Impact Assessment Example (Appendix D) functioned well in Ireland; it is a common classroom practice in Finland, but it is not assessed as such; while in Hungary it was considered rather unusual in its form and depended rather on the teachers' personality and affinity to self-reflection.

Teacher rubrics for the DISCOVER and IMPACT phases were not used in Finland, Spain, or Hungary, but they were considered a nice self-reflective tool for teachers to think about wider objectives in Ireland.

6.3 Needs expressed by OTTER pilot countries

Some pilot teachers reported that they would appreciate forms of summative assessment to be included in the OTTER labs. Due to the flexible nature of the OTTER methodology and the diversity of the countries participating in the OTTER project, summative assessment of this nature were not planned or provided.

Planning documents for the ORIENTATE phase proved to be complicated in most of the countries. Planning documents were not used in Spain or Finland. Teachers in Hungary expressed that they would require further training and explanation from the OTTER Team when completing planning documents. In Ireland, it proved to work well after teachers received some support.

Reviewing feedback country by country, some slight differences were evident. In Spain, it was expressed that because the teachers themselves decided the content to teach during the OTTER Lab, the assessment of the given topic would have fallen on them to assess student learning in terms

of content. These teachers suggested that some guidance on proposed assessment methods would have been helpful.

Teachers in Finland highlighted the need for inclusion of students' summative self-assessment and peer assessment so that it was a more comprehensive reflection over a longer period of time and therefore would encourage teachers and students to pay attention to the bigger learning experience. These teachers emphasised that learning takes time and that this should be taken into account when considering assessment frequency and methods.

Similarly, in Hungary, teachers reported the difficulty in analysing student development in such a brief period. They proposed that sufficient time for implementation was missing. They also reported that there was a lack of guidelines on how to compose (summative) tests that build on OTTER pillars. Teachers also reported difficulties in understanding some of the project language used. They highlighted that a simple translation does not capture the meaning of some terms and contents used.

Ireland indicated the absence of connection with continuous assessments or class tests that are already in place in schools. Game based, summative assessments could also enhance further student (and teacher) engagement and involvement throughout the OTTER labs.

6.4 OTTER Lab adaptation opportunities in pilot countries

In Finland, teachers created their own assessment tools, for example, student learning diaries. The intention of the learning diary was to highlight that the OTTER Lab and all associated activities are one process. They reported that the learning diary made it possible for students to gather their learnings and get a better picture of the whole process. They suggested that it would be useful to create a path during which the students consistently engaged in self-assessment and reflected on what they have learnt. Another adaptation they suggested was that that students give feedback to each other as part of a peer assessment within the OTTER labs. Students could also assess how they are working in groups and as a group and give each other constructive feedback on their joint work and achievements. They suggested that positive pedagogy should be the basis of the OTTER Lab process and assessment. Continuous positive and motivational feedback from the teacher and the peer students has the capacity to empower and encourage students. Lastly from Finland, it was perceived, that the OTTER methodology already emphasises teachers' ownership in the planning and implementation process of OTTER Labs. They commented that the OTTER methodology doesn't dictate exactly what teachers should do but facilitates teachers designing their own resources.

Spanish teachers suggested that providing specific rubrics for each step of the OTTER Lab, in addition to a final rubric may assist teachers. They proposed that one rubric should be of a generic type,

designed to assess the concepts that teachers have decided to include in the OTTER Lab. They indicated that they would value a rubric that would help teachers prepare some type of exam/test to assess the acquired content via a grade. They suggested a second rubric could be more specific, aimed at evaluating the incorporation of skills and abilities.

In Ireland, teachers proposed that similar to the mapping of the Learning Objectives to the curriculum; it would be beneficial to map the Learning Objectives to the exams and assessments already in place so that extra assessment would not as necessary.

In Hungary, teachers suggested that more extended teacher training (probably worthing recognized credits for the pedagogues) would lead to a more in depth understanding of the OTTER methodology and address any assessment issues.

otter

Proposed evaluation for EOC assessment methods

7 Proposed EOC assessment methods

7.1 Summary of overall assessment findings

To enrich our perspective, key findings from sections 2 and 3 are included with the intention of highlighting particular aspects of key importance for the proposed assessment of EOC.

- Assessment plays an influential role in the teaching and learning process, impacting on what is learned, how learning takes place and what is valued.
- Assessment should be planned in alignment with learning objectives and pedagogical approaches to support curriculum coherency.
- The underlying principles of UDL and assessment for learning should be considered when designing assessment approaches to accommodate diverse learning needs of students.
- The analysis of students' work/drawings, observations, and video recording and analysis offer rich insights into student understanding, however most assessment in the literature is based on assessment for research purposes.
- The literature highlights the importance of selecting appropriate assessment tools that not only measure learning outcomes effectively but also capture the diverse aspects of students' experiences in EOC settings.
- Assessment norms in pilot countries mirror the type of assessment approaches the teachers wanted to implement in EOC OTTER labs therefore EOC assessment needs to incorporate flexible tools for assessment.
- Collection of student work throughout the OTTER lab to represent students development of learning knowledge, skills, attitudes, values and ethics facilitate good assessment practices.
- There is a need to emphasise teachers' ownership in the planning and implementation of assessment when conducting EOC.
- More teacher professional development is required to facilitate teachers further enhancing their own assessment literacy so that they can develop EOC assessment tools.

7.2 Overall EOC assessment recommendations

The following actionable suggestions are based on the findings from section 2 and 3 of this deliverable.

• Flexible tools for assessment

Assessment within OTTER should be adaptable so that assessment within EOC considers national specificities and allow room for decisions about assessment types. It should enable flexibility, freedom and choice for teachers, through providing support to enable adaptation of templates to national contexts, and students through the implementation of Universal Design for Learning (UDL) principles. Student portfolios, artefacts, peer and self-assessment, and student learning diaries were just some of the approaches participating teachers found helpful in the OTTER lab.

• Recognising international best practice

While appreciating assessment norms is important, further developing understandings of international best practice regarding assessment approaches is also important. This could include assessment for learning practices, as well as UDL. Supporting teachers (students and parents) understanding of the different forms and purposes of assessment, and how they can engage with this within the existing national context, is an important step in this regard.

• Develop teacher agency and assessment literacy

While providing sample resources are important, teachers should use these resources with the intention of designing their own assessment resources and make key decisions regarding assessment in their classrooms. Teachers understanding of assessment, their assessment literature and the language they have (and understanding of same) to talk about assessment is important in order to support this.

• Formative and summative assessment

The OTTER lab assessment placed a strong emphasis on formative assessment, including assessment for learning. The reality is that many education systems place a strong emphasis on summative assessments. Examples and focus on both forms of assessment, formative and summative, during the OTTER labs may enable teachers to embed EOC experiences within existing systems.

7.3 Revised approach to OTTER

Feedback from pilot countries in relation to assessment methods and future adaptations of the OTTER labs highlight the need to include reflect and assessment cycles within the OTTER Lab so that a collection of assessment checkpoints are embedded throughout the OTTER lab. This has led to a revised approach to encapsulate assessment consistently throughout the cycle (Figure 1).

Figure 1: Revised approach to OTTER including reflect and assessment throughout

7.4 Student Focused EOC Assessment Rubric

In D4.3, an assessment rubric presented information regarding the assessment approach used in the research component of OTTER. The assessment approach was designed to overlay usual assessment practices while conducting EOC therefore this rubric can act as a means of student assessment. Table 3 proposes the types of assessment tools that can be used to assess student's development of skills and competencies. This is not an exhaustive list of assessment types; it is just

a selection of tools an educator may use. As stated earlier in 1.3.6, when educators are designing tools to assess skills, it is important that an educator defines this skill relevant to the student's education level.

The Prepare step is omitted as this section relates mainly to teacher activities in designing the OTTER Lab activities and setting objectives. To assess student skills related to Personal and Social Responsibility, Citizenship and Life and Career during the Orientate phase, a teacher may utilize a pre-survey (See Appendix A) or student artefacts (Appendix B). To assess student skills related to Creativity and Innovation, Communication and Collaboration during the Discover phase, a teacher may utilize a Teacher Rubric including related criteria to evaluate student work/behavior (See Appendix C). To assess ICT Literacy at the Make an Impact phase, a teacher may use a Student Worksheet (See Appendix D). At the final Reflect phase, a teacher may use a Reflect Student Exit Ticket to assess student metacognition and/or knowledge (See Appendix E).

Cycle	Skill/Competency	Sample Instrument used in OTTER Labs	
	Personal and Social Responsibility	Pre-survey	
Orientate	Citizenship	Student artefacts (e.g., data collection,	
	Life and Career	diagrams/drawings, worksheets etc.,)	
	Creativity and Innovation	Teacher Rubric	
	Communication		
	Collaboration		
Discover	Knowledge of (and attitudes		
	towards) environmental issues,	Student artefacts (e.g., data collection, diagrams/drawings, worksheets etc.,)	
	reduced plastic use, and		
	sustainability		
	Learning to Learn	Discover Exit Ticket	
	ICT Literacy	Student Worksheet	
Make an Impact	Knowledge of (and attitudes towards) environmental issues, reduced plastic use, and sustainability.	Student artefacts (e.g., data collection, diagrams/drawings, worksheets etc.,	
	Critical Thinking, Problem Solving, Decision Making		
	Information Literacy	Taaabar Dubria	
	Scientific Literacy		
Final Reflect	Personal and Social Responsibility		
	Citizenship	Post-Survey	
	Life and Career		
	Learning to learn, metacognition		
	Knowledge of (and attitudes		
	towards) environmental issues,	Reflect Exit Ticket	
	reduced plastic use, and		
	sustainability		

7.5 EOC Principles and Suggested Principles in Practice

In this deliverable we have summarized the main assessment practices in education, highlighted key literature regarding assessment of EOC practices and outlined some feedback regarding assessment approaches used in the OTTER labs. Utilizing this information, Education Outside the Classroom Principles and related Principles in Practice are presented in Table 4. An educator may use these principles to indicate where they intend to evidence each EOC Principle within their own practice.

Table 4: Principles and of Suggested Principles in Practice

EOC Principle	Principles in Practice	Evidence of EOC Principle
1. Assessment coherence	 Focus on age appropriate content in line with curricula. Use assessment scaffolding to break tasks down into smaller steps that that focus on component skills. Design assessment sequencing to developmentally build towards larger assignments. 	[Complete with your evidence of practice]
 Assess range of learning objectives including knowledge, skills, attitudes, values and ethics 	 Use a variety of assessment tools that have the capacity to measure a range of learning objectives such as student reflections, model making, investigation design, and debate 	[Complete with your evidence of practice]
 Philosophical and pedagogical approaches related to assessment 	 Universal Design for Learning Assessment that allows flexibility, transparency and choice. Allow students to choose how they complete as assessment, for example, using a concept map, a podcast or an essay written response. 	[Complete with your evidence of practice]

4	. Assessment is	Plan to include reflection and	[Complete with your evidence of practice]
	embedded in the cycle	assessment checkpoints into the	
	of EOC	cycle of the entire EOC cycle rather	
		than just at the start and at the end.	
		For example, using concept	
		cartoons to explore student views	
		during the activity, peer quizzes	
		where student design and answer	
		questions, diagram drawing, graph	
		interpretations, walking debates etc.	
5	. Assessment for	Use criterion-based assessment to	[Complete with your evidence of practice]
	learning and the	facilitate students assessing their	
	importance of reflection	own work and the work of their	
	in terms of peer and	peers. Align with summative	
	self-assessment.	assessment approaches where	
		relevant.	
		Create opportunities for student	
		reflection so that they can evaluate	
		and assess their own learning in the	
		shorterm (exit tickets) or over an	
		extended period (learning diaries).	

<u>etter</u>

6.	Consider pre-existing	•	Consider if learning objectives are	[Complete with your evidence of practice]
	assessment structures		assessed through any classroom	
			assessments	
		•	Explore how summative	
			assessments can be utilised to	
			explore and support learning in EOC	
7.	Developing	•	Access for teachers to engage with	[Complete with your evidence of practice]
	assessment literacy		professional development and high-	
	(teacher and student).		quality resources to provide	
			assessment information.	
		•	Access for students to develop an	
			awareness of the value of engaging	
			with a range of assessment types	
			and the importance of developing	
			new metacognitive strategies.	

8 Reference List

- Affeldt, F., Weitz, K., Siol, A., Markic, S. and Eilks, I. (2015). A Non-Formal Student Laboratory as a Place for Innovation in Education for Sustainability for All Students. *Education Sciences*, 5(3), 238–254.
- Archer, E. (2017). The assessment purpose triangle: Balancing the purposes of educational assessment. In *Frontiers in Education* (Vol. 2, p. 41). Frontiers Media SA.
- Ariosto, A., Ferrarello, D., Mammana, M. F. and Taranto, E. (2021). Math City Map: Provide And Share Outdoor Modelling Tasks. An Experience With Children. AAPP| Physical, Mathematical & Natural Sciences/Atti Della Accademia Peloritana Dei Pericolanti-Classe Di Scienze Fisiche, Matematiche e Naturali, 99, p1-21.
- Au, W. (2011). Teaching under the new Taylorism: High-stakes testing and the standardization of the 21st century curriculum. *Journal of Curriculum Studies*, *43*(1), 25-45.
- Balta, J.Y., Supple, B. and O'Keeffe, G.W. (2021). The universal design for learning framework in anatomical sciences education. *Anatomical Sciences Education*. 14(1), 71-78.
- Bhattacharya, D., Carroll Steward, K. and Forbes, C. T. (2021). Climate education in secondary science: Comparison of model-based and non-model-based investigations of Earth's climate. *International Journal of Science Education*, 43(13), 2226–2249.
- Black, P., Harrison, C., Lee, C., Marshall, B. and Wiliam, D., (2004). Working inside the black box: Assessment for learning in the classroom. Phi delta kappan, 86(1), 8-21.
- Canal Universitats. (2024). PAU and PAP Access Tests Retrieved from https://universitats.gencat.cat/es/proves-acces-PAU-PAP/preparat-PAU/models-examenanys-anteriors/index.html
- Carrington, S., Saggers, B., Webster, A., Harper-Hill, K. and Nickerson, J. (2020). What Universal Design for Learning principles, guidelines, and checkpoints are evident in educators' descriptions of their practice when supporting students on the autism spectrum? *International Journal of Educational Research*, *10*2, 101583.
- CAST (2018). Universal Design for Learning Guidelines version 2.2. Retrieved from http://udlguidelines.cast.org
- Department of Education (2023). Junior Cycle Retrieved from https://www.gov.ie/en/publication/9be14-junior-cycle/
- Dunlop, L., Clarke, L. and McKelvey-Martin, V. (2019). Free-choice learning in school science: A model for collaboration between formal and informal science educators. International Journal of Science Education, Part B, 9(1), 13–28.

⊕tter

- Elwood, J., (2006). Formative assessment: Possibilities, boundaries and limitations. *Assessment in Education: Principles, Policy and Practice, 13*(2), 215-232.
- Ghadiri Khanaposhtani, M., Liu, C. J., Gottesman, B. L., Shepardson, D., and Pijanowski, B. (2018).
 Evidence that an informal environmental summer camp can contribute to the construction of the conceptual understanding and situational interest of STEM in middle-school youth.
 International Journal of Science Education, Part B, 8(3), 227–249.
- Giamellaro, M. (2014). Primary Contextualization of Science Learning through Immersion in Content-Rich Settings. *International Journal of Science Education*, 36(17), 2848–2871.
- Gleeson, J. and Donnabháin, D. Ó. (2009). Strategic planning and accountability in Irish education. *Irish Educational Studies*, *28*(1), 27-46.
- González-Salamanca, J. C., Agudelo, O. L. and Salinas, J. (2020). Key competences, education for sustainable development and strategies for the development of 21st century skills. A systematic literature review. *Sustainability*, *12*(24), 10366.
- Hargreaves, A. (1989). Curriculum and Assessment Reform. Milton Keynes, UK: Open University Press.
- Harris, R. and Bilton, H. (2019). Learning about the past: Exploring the opportunities and challenges of using an outdoor learning approach. *Cambridge Journal of Education*, 49(1), 69–91.
- Hopman, M. and Knijn, T. (2015). The 'turn to parenting': paradigm shift or work in progress?. *International Journal of Child Care and Education Policy*, *9*, 1-17.
- Hymel, S. and Katz, J. (2019). Designing classrooms for diversity: Fostering social inclusion. *Educational Psychologist*, 54(4), 331-339.
- Jin, H., Mikeska, J.N., Hokayem, H. and Mavronikolas, E., (2019). Toward coherence in curriculum, instruction, and assessment: A review of learning progression literature. *Science Education*, *103*(5), 1206-1234.
- Kärkkäinen, S., Keinonen, T., Kukkonen, J., Juntunen, S. and Ratinen, I. (2017). The effects of socioscientific issue based inquiry learning on pupils' representations of landscape. *Environmental Education Research*, 23(8), 1072–1087.
- Kermish-Allen, R., Peterman, K. and Bevc, C. (2019). The utility of citizen science projects in K-5 schools: Measures of community engagement and student impacts. *Cultural Studies of Science Education*, 14(3), 627–641.
- Lange, T., and Meaney, T. (2012). The tail wagging the dog?: the affect of national testing on teachers' agency. In matematikdidaktiska forskningsseminariet, Umeå (2012) (pp. 131-140). Svensk förening för MatematikDidaktisk Forskning-SMDF.
- Matriculation Examination Board. Assessment and Examination (2024). Retrieved from <u>https://www.ylioppilastutkinto.fi/en/assessment-and-certificates/assessment-examination</u>
- Meier, D., (2004). Many children left behind: How the No Child Left Behind Act is damaging our children and our schools. Beacon Press.

Newmann, F., Smith, B., Allensworth, E., and Bryk, A. (2001). School instructional program coherence: Benefits and challenges. Chicago, IL: Consortium on Chicago School Research.

NCCA. (2024). Senior Cycle Retrieved from https://ncca.ie/en/senior-cycle/

- Newton, P. E. (2007). Clarifying the purposes of educational assessment. Assessment in Education: *Principles, Policy & Practice*14, 149–170.
- Ninomiya, S. (2016). The possibilities and limitations of assessment for learning: Exploring the theory of formative assessment and the notion of "closing the learning gap". *Educational studies in Japan*, *10*, 79-91.
- McCormack, O., Gleeson, J. and O'Donoghue, T. (2020). An analysis of national newspaper coverage relating to the Leaving Certificate programme in two newspapers in the Republic of Ireland, *Irish Educational Studies*, 39(1), 1-18.
- Puttick, G. and Tucker-Raymond, E. (2018). Building systems from scratch: An exploratory study of students learning about climate change. *Journal of Science Education and Technology*, 27(4), 306–321.
- Schildkamp, K., and Kuiper, W. (2010). Data-informed curriculum reform: which data, what purposes, and promoting and hindering factors. Teach. Teach. Educ. 26, 482–496.
- Shepard, L. A. 2004. Curricular Coherence in Assessment Design. *Teachers College Record*, *106*(14), 239-249.
- Shiel, G. and Murchan, D. (2017). Understanding and applying assessment in education. Understanding and Applying Assessment in Education, 1-248.
- Voogt, J., and Roblin, N. P. (2010). 21st century skills. Discussienota. Zoetermeer: The Netherlands: Kennisnet, 23(03), 2000.
- Ying, Z., Yuhui, Y., Huifang, C., Mo, C., & Jianping, Z. (2019). The Development and Performance Evaluation of Digital Museums towards Second Classroom of Primary and Secondary School-Taking Zhejiang Education Technology Digital Museum as an Example. *International Journal* of Emerging Technologies in Learning, 14(2), 69–84.

Appendices

9 Appendix A

Orientate Assessment Example: Student Pre/post Survey (9-18yrs)

1 (strongly agree) 2 (agree) 3 (neutral) 4 (disagree) 5 (strongly disagree)

1. I have the right to give my opinion
2. I respect that people can express different cultures, religions, lifestyles, and opinions
3. I can establish a good relationship with people with personalities or interests different
from my own
I am friendly and kind with my classmates
5. I can try to influence issues I care about together with other people.
6. I can participate in developing issues that are important to me.
8. I would dislike working in Science after I leave school
9. A job as a scientist would be interesting
10. I would dislike becoming a scientist because it needs too much education
11. A career in science would be dull and boring
12. I would like to teach science when I leave school
13. When I leave school, I would like to work with people who make discoveries in science
14. I can learn many things from other people
15. I can teach something to other people
16. I make an effort to, as much as possible, fulfill the promises that I make
17. I admit my mistakes and apologize
18. I know that government decisions can affect me in different ways
19. I think that everyone is responsible for taking care of each other.
20. I think that everyone has to consider the affects of our actions on other people.
21. I understand that my choices and actions have an impact on and other people.

Student Pre/post Survey (6-8yrs)

1 (strongly agree) 2 (agree) 3 (neutral) 4 (disagree) 5 (strongly disagree)

1. I know that my classmates can come from different places in the world
2. I can be friends with someone who is different to me
3. I am friendly and kind with my classmates
4. I would like to be a scientist when I grow up
5. I can learn many things from my classmates
6. When I am wrong, I say so

10 Appendix B

Orientate Assessment Example: Sample Orientate Student Artefacts

4. Switching on a lamp What happens or changes?

Where did what happens or changes come from or how did it come about?

How or why did it or could it stop?

Energy family members involved:

etter

11 Appendix C

Discover Sample Assessments

Discover Teacher Rubric

Student Names/Codes: Age Group: Gender(s):

,	After the EOC activity, students were able to:	To a large extent	To some extent	l o little or no extent
1.	be polite and kind to teammates			
2.	acknowledge and respect other perspectives			
3.	make sure all team members' ideas are equally valued			
4.	offer assistance to others in their work when needed			
5.	Improve their own work when given feedback			
6.	follow rules for team decision-making			
7.	organize information well			
8.	complete tasks without having to be reminded			
9.	present all information clearly, concisely, and logically			
10.	understand how knowledge or insights might transfer to other situations or contexts			
11.	find sources of information and inspiration when others do not			
12.	help the team solve problems and manage conflicts			
13.	adapt a communication style appropriate for the purpose, task, or audience			
14.	elaborate and improve on ideas			

Discover Student Exit Ticket

Age 6-8yrs

	്ര	₽ P
I look forward to science lessons in the classroom		
I do not like science lessons		
Science lessons outside are fun		
I dislike science lessons that are outdoors		
School should have more outdoor science lessons each		
week		

Age 9-18yrs

Name/Code:	Age:	Gender:	ப	Ð
Science is one of the most inter	esting schoo	l subjects		
Science lessons are a waste of t	time			
I really enjoy going to science le	essons			
The material covered in science	lessons is u	ninteresting		
I look forward to science lesson	s in the class	sroom		
I would enjoy school more if the	ere were no s	cience lessons		
Science lessons outside are fun	1			
I dislike science lessons that are	e outdoors			
School should have more outdo	or science le	essons each week		
Outdoor science lessons bore m	ne			
Have you anything else to add a [Write your comment in this space]	bout science]	e outside the classroom?		

Discover Student Artefacts

Student Name	Carvin	Dura	Watthew	dirabal	Mogett	Rachel	Katik	lames	Mahadi	Munaf	Darragh	Deterro	Average
pH		7	7 7	7 7	7.6	7.6	7	7.5	5 7	7.8	7.6	7.7	7.32
Ammonia (ppm)		0	0 0.5	5 0	0	0	0	0.5	0.5	0.5	0.5	0	0.21
Nitrate (ppm)		0	0 10	10	10	10	0	C	0 0	0	0	0	3.33
Nitrite (ppm)		0	0 0	0 0	0	0	0	C	0 0	0	0	0	0
KH (ppm)	10	0 10	0 100	100	100	100	100	100	100	100	100	100	100
GH (ppm)	18	0 18	0 180	180	180	180	180	180	180	180	180	180	180
DO Day 1 (ppm)	11	2 11	2 11.2	10.8	10.8	10.8	12.1	12.1	12.1	11.6	11.6	11.6	11.43
DO Day 5 (ppm)	10.8	8 10.8	8 10.88	9.8	9.8	9.8	11.6	11.6	5 11.6	9.7	9.7	9.7	10.50
BOD (ppm)	0.3	2 0.3	2 0.32	2 1	1	1	0.5	0.5	0.5	1.9	1.9	1.9	0.93
Temperature (degrees Celsius)	1	.0 1	0 10	11	11	11	10	10	10	10	10	10	9.565

12 Appendix D

Make an Impact Assessment Examples

Make an Impact – Student Worksheet

Student Name/code:	Age:	Gender:
Give an example of a time when:	Describe your example	
 You used a computer/tablet/phone 	What did you use the device to do?	
You used some type of digital tool to measure/collect or analyse something	What did you use the device to do?	
 You had to find information online that you did not already know 	What was the information about?	
 You created something yourself online/using a digital tool e.g., office tools, surveys, polls, tagging on social media, blogs, websites, games, apps 	What did you create?	

Make an Impact: Student Artefacts

Make an Impact: Teacher Rubric

Scientific	Students and Market	Depth of Knowledge				
Literacy	Students are able to:	Low	Medium	High		
	Use scientifically informed language					
	Explain why the issue/concept is important					
Explaining	Predict what is expected to happen as a result of the Youth Initiative					
Scientifically	Predict why they think something will happen					
	Make links with how the concept relates to society					
	Ask appropriate questions related to the issue					
	Design, plan and carry out Youth Initiative					
Evaluating and designing	Consider how to ensure the information used is reliable					
scientific enquiry	Explain how the approach is fair or ways in which better results could be achieved					
	Sort, group, graph data/evidence					
Interpreting	Look for relationships/patterns in the data/evidence					
data and evidence	Offer explanations from the data/evidence					
scientifically	Communicate findings through appropriate media					
Any Other Comments about student learning during the Youth Initiative?						

13 Appendix E

Reflect Assessment Example: Student Exit Ticket

Student Name/code:

Age:

Gender:

Was there anything you did not like about learning outside the classroom?

Describe 2 things you learned when you did science outside the classroom

Describe 1 thing you would change about learning outside the classroom

Contact

www.otter-project.eu

O @otter_euproject

@OTTER_EU

