The Art of Science Learning

D4.1
Research report: Methodological aspects of science education assessment

María Heras, Isabel Ruiz-Mallén, Karla Berrens, Louis Lemkow

Universitat Autònoma de Barcelona
Universitat Oberta de Catalunya

May 2016

European Commission

Horizon 2020
European Union funding
for Research & Innovation

This project has received funding from the European's Union H2020 research and innovation programme under grant agreement No. 665826

To be quoted as:

Heras, M., Ruiz-Mallén, I., Berrens, K., Lemkow, L. (2016). Research report: Methodological aspects of science education assessment. Deliverable 4.1, PERFORM project. Barcelona.

Peer-review: Eric Jensen, Oriol Marimón, Marina Di Masso

University of Warwick
The Big Van Theory
Universitat Oberta de Catalunya

PERFORM

Participatory Engagement with Scientific and Technological Research through Performance

CONTENTS

FIGURES

TABLES

IMAGES

SUMMARY

This deliverable corresponds to Task 4.1 'Development of an innovative and participatory impact assessment research methodology', aiming to identify those individual, contextual and methodological factors contributing to or detracting from the impacts of the participatory learning process in which secondary school students will be engaged during the PERFORM project. It reports on expert-based and participatory indicators identified through a systematic literature review on assessment frameworks and methodologies used in science learning and engagement activities and a set of exploratory workshops with students conducted in selected schools, respectively. As a result, we identify a total of **93 assessment indicators related to the Responsible Research and Innovation (RRI) values, transversal competences and experiential and cognitive aspects of science learning and engagement**. This deliverable concludes by providing criteria and indicators to evaluate PERFORM participatory science education methods based on performing arts, as well as identifying specific methodological aspects and challenges to be addressed in each specific case study. Findings from this report will be published in international peer-reviewed, open access journals by the authors.

1. INTRODUCTION

PERFORM aims to generate suitable science education methods based on performing arts to foster secondary-school students' motivation and engagement in science, technology, engineering and mathematics (STEM). Such innovative methods will be developed through a **participatory educational process with, for and by students** in four secondary schools in France, Spain and the United Kingdom (UK) in 2017 and 2018. The involved students and their teachers will maintain direct interaction and communication with early career researchers to **promote students' learning and reflection about STEM** concepts, scientists' practice and the impacts and applications of science in their daily lives. Such process –referred to as **PERSEIA,** will also foster students' acquisition of the **values embedded in the Responsible Research and Innovation (RRI) approach** (EC 2012), and the transversal competences required to make informed decisions for full civic participation in a knowledge-based society. Moreover, PERFORM research design and implementation follows such RRI approach.

The RRI approach –guiding the Horizon2020 programme, entails a dynamic and iterative process by which all the stakeholders involved in the research and innovation practice become mutually responsive and share responsibility regarding both outcomes and process requirements so as to align research and innovation agendas with societal needs and concerns (RRI-Tools 2015). The openness of the RRI approach draws attention to the need for **developing rigorous monitoring and assessment methods** of RRI impacts (Strand et al. 2015). In this line, the Rome Declaration of RRI (2014) calls for the review and adoption of metrics and narratives for research and innovation, by, among others, the monitoring of social impacts and the provision of guidelines for such an assessment.

Under the lenses of the RRI approach, the development of performance-based science education activities within PERFORM requires accurate and rigorous assessment approaches that monitor and evaluate their effectiveness in students' engagement and learning about STEM. For that purpose, PERFORM will conduct two types of assessments:

i. an assessment of the punctual interactive events in which performances are played in front of students' audiences, and

ii. an assessment focused on the participatory learning process among students, science communicators, researchers and teachers when developing the science education methods based on performing arts.

In both cases, a mix of qualitative and quantitative methods will be applied, including the use of social media analysis. Specifically in the second type of assessment, PERFORM will adopt a **participatory action research** approach.

Participatory action research is especially relevant in the context of RRI assessment, mainly due to two reasons. First, as already highlighted (Strand et al. 2015, p.9), '*the concept of responsibility is easy to endorse and difficult to define*'. To be operative, RRI requires, thus, a conceptual and practical grounding which might take different shapes depending on the context of implementation and the actors involved and participatory action research methods can be supportive in defining such grounding. Second, to strengthen the legitimacy and use of assessment indicators it is necessary that the actors involved –in this case, in the educational process, assume a sense of ownership (ibid). Participatory processes involving all concerned actors provide opportunities for collectively approaching, exploring and discussing such issues. Thus, in the context of PERFORM, action research theory offers a promising theoretical and methodological framework for approaching the assessment of such innovative science education methods. Through their active participation, students will have the opportunity to change perceptions and attitudes towards science and become more engaged as they experience and reflect through the process, potentially adopting new behaviours based on specific values (Webb 1996).

Furthermore, in the context of PERFORM, **transdisciplinarity** is also essential to assess the impact of such science education methods by grounding scientific concepts into societal contexts and processes, going beyond the viewpoints offered by a single discipline and including actors beyond the academia (Pohl 2008). Transdisciplinarity thus nurtures the assessment from different disciplines and methodological approaches, as for instance in the case of arts and science.

To frame the PERFORM impact assessment design in participatory action research and transdisciplinarity approaches, this report explores a set of assessment frameworks and methodologies from different disciplines used to evaluate science learning and engagement experiences with young people (e.g., educational psychology, science communication, sociology, arts, anthropology) through an academic literature review.

Findings also provide a grounding to define expert-based criteria and indicators of the PERFORM impact assessment in a systematic way, as well as the guidelines to design the corresponding data collection instruments. Moreover, this report describes the involvement of students in the PERFORM evaluation process through a series of exploratory workshops conducted in selected schools in the three case-study countries (Spain, France and United Kingdom). In doing that, it identifies participatory indicators of learning and engagement in STEM that consider contextual factors that are perceived as important in each country, thus complementing the identified expert-based indicators. This is expected to contribute not only to the implementation of the PERFORM project, but also to the development of science education assessments aimed at including RRI values and process requirements. This is a relevant contribution, since, as highlighted by the Expert Group on Policy Indicators for RRI (EC, 2015), it is crucial to find specific indicators adapted for each key policy agenda.

In what follows, this report briefly introduces the context of assessment in science education, including the conceptualization of RRI and the main learning variables to measure as part of PERFORM's assessment. It then presents the research strategy developed to conduct the literature review and the exploratory workshops and the methodology for data collection and analysis. The main findings are then reported, highlighting the results on identified expert-based and participatory indicators for assessing PERFORM learning process and outcomes in terms of students' acquisition of RRI values and transversal competences, as well as cognitive and experiential aspects. This report concludes by examining the implications of these findings for designing PERFORM assessment methodologies.

2. SCIENCE EDUCATION ASSESSMENT IN CONTEXT

2.1. The role and value of science education assessment

Assessment is a fundamental topic in science education and pedagogy, permeating science education curriculum, teaching and learning practices and research (Corigan et al. 2013). Following Black and William (2007), we understand science education assessment as the '*generation and interpretation of evidence about the knowledge, skills, and understanding of learners*' related to STEM (ibid, p. 4). This generation and interpretation of educational data requires a systematic, multistep process (NSES 1996), whose characteristics vary according to the purposes of the assessment.

Three broad purposes of assessment are commonly recognised in the science education literature:

i. Assessment to certify learning, referred to as **summative assessment**. Summative assessment serves to inform an overall judgement of learning achievement of a student at a given point (Black 2005). It is associated with grades and marks and to high-stakes assessments –assessments whose outcomes have a substantial impact in the student, teacher or school.

ii. Assessment to support learning, referred to as **formative assessment.** The purpose of formative assessment is, then, the '*generation and use of the information to assist students' learning, rather than simply to record it*' (Black and William 2007, p. 4). Therefore, this form of assessment is essentially interactive and adaptive and emphasises both: assessment *as* learning, by monitoring learning as it occurs and fostering students' responsibility for their own learning and metacognition; and assessment *for* learning, by using inferences from empirical evidence about students' progress to inform and adapt teaching practices (Corigan et al. 2013).

iii. Assessment to provide a measure of accountability, that is, **public accountability**. This type of assessment is related to summative assessments and its purpose is to provide judgements about a given educational system or institution (Black and William 2007).

Depending on **the purpose/s of the assessment**, its development will be more central or peripheral to science instruction. Summative assessments, for instance, tend to provide insights at the end of the activity, topic or school term and can be decontextualized from the teachers' instruction. It is the case for instance of large-scale science education assessments, like the Programme for International Student Assessment (PISA). On the contrary, the development of formative assessments requires a close interaction with classroom practice, representing a central feature of teachers' instruction.

It is important to mention that these types of assessment are not exclusive, and different assessment purposes can coexist and overlap. For instance, a summative assessment can also provide feedback to enhance students' learning during the process. However, these different purposes require different levels of analytical detail as well as a particular set of assessment tools and procedures to be carried out. Moreover, just as assessment purposes can be synergic, they can also manifest tensions between them. For instance, as Black and William (2007) remind, the pressure to raise achievement can lead (and frequently does) teachers and students to focus their attention on those aspects of the curriculum that are tested in the assessment while neglecting others that might be also important in such process.

Furthermore, as Millar (2013) highlights, there is a crucial, yet often neglected, purpose of science education assessment, which relates to the operationalization of learning outcomes. By providing operational definitions of learning goals and related outcomes, science education assessment plays an essential role in clarifying the learning that is intended in a given science lesson or activity. In words of the author:

> *'Any statement of intended learning outcomes of a programme, or course, or module is inevitably ambiguous and open to (often quite wide) variation in interpretation. Assessment instruments and practices are the tools by which this ambiguity is reduced, perhaps even removed. Assessment operationalizes outcomes and hence defines them. Unless we know what we will accept as evidence of the achievement or non-achievement of any given learning objective, we do not really know what that objective is or means. The job of an assessment instrument (a question or task or a set of these) is to generate this evidence. In doing so, it is more than merely a tool for carrying out a task that is already fully and clearly defined. Rather the assessment instrument becomes an operational definition of the objective'*. (ibid, p. 56)

The strategic value of science education assessment is therefore emphasised as it is acknowledged as a critical element not only for identifying and justifying learning, but also for shaping and potentially enhancing the possibilities of the learning process itself. **This strategic value is crucial in the context of RRI**, in which the lack of an authoritative definition or a consensus on how to understand it, emphasises even more the need to have clear understandings of the impacts to measure (Strand et al. 2015). As the Expert Group on Policy Indicators for RRI reminds us, '*it is difficult to specify a precise, valid and robust indicator for something that is imprecise and changing*' (ibid, p.9). Thus, the PERFORM project, and particularly this research report, expects to contribute to such understandings through the identification and operationalization of science education assessment indicators –and therefore, of learning outcomes and process requirements, relevant in the framework of RRI.

Besides the assessment purpose, **the conception of learning** is also an essential element influencing science education assessments. The predominance or emergence of certain learning paradigms in certain periods of time has shaped different models of assessment and their associated methods. Broadly speaking, the philosophical and psychological shift in the 60-70's towards cognitive psychology and constructivism had important implications in the understanding of learning and, consequently, in the development of educational assessments (Klassen 2006).

The empiricist view of learning, which can be traced back to Aristotle and has been of great influence in Western education until the end of the 20th century, assumed knowledge as a copy of reality, developed through our experiences (Read 2006). Under this perspective, knowledge is just a record of what is received through the senses –without further need for processing, and can be transferred intact from the teacher to the students (Klassen 2006). It takes, therefore, an exogenous approach to learning, which is teaching-centred.

Behaviourist views of learning and behavioural psychology were developed along with techniques of psychological measurement, or psychometrics. In the psychometric approach to assessment, inferences about achievement and competence are typically based upon the aggregation of individual response scores, assigned independently by readers with no additional knowledge about the student (Moss 1994). The primary assessment method that resulted from this approach was the *selected-response test*, in which the students selected one of several possible answers to a question (Klassen 2006). This trend was reinforced by the emphasis on public accountability and the concern over the state of science education from the 1950's on, which fostered the development of standardised examinations to monitor the quality of education (ibid). These examinations were mostly based on rankings and test scores –commonly developed as summative, high-stakes assessments, that could guaranty the comparability of results against specific educational standards.

From the 60's on, new developments on learning theories –represented by authors such as Piaget, Kuhn or Vigotsky, encompassed a shift towards constructivist approaches to learning that challenged traditional educational assessments (ibid). From a constructivist point of view, learning is based on the interactions between the individual and the world in which she or he lives, through an active process of sense-making, influenced by already existing learning structures and prior knowledge, as well as the learners' beliefs, expectations, perceptions and motivations (Read 2006).

One of the key implications of this approach for science education assessments is the importance given to context, either *domain-specific* context through the assessment of disciplinary knowledge, or *real-life* context, through the assessment of students' ability to use or apply knowledge in a given context (Baker et al. 1994). These contextual assessments, in opposition to the traditional assessments applied in psychometrics, require students to apply higher-order thinking skills and other specific practical competences. The assessment practices that emerge from this approach can be classified in three broad categories (Klassen 2006): **conceptual maps**, related to the cognitive context, and how concepts are related; **performance assessment**, related to the practical context and assessing the students' practical development of a task; and **students' portfolio**, related to the classroom context and the assessment of representative samples of students' work.

PERFORM's impact assessment feeds from constructivist approaches to learning and emphasises its formative dimension. In recent years, learner-centred and holistic approaches to learning have emphasised the importance of formative science education assessments (Corigan et al. 2013, Fitzgerald and Gunstone 2013, William 2011). The increasing recognition of the importance of assessment to contemporary science education '*has catalysed research, development, and implementation of new methods of data collection along with new ways of judging data quality*' (NSES 1996). This is reflected on the increasing emphasis on assessment *for* learning approaches within science education, in which assessment is perceived as intrinsic for effective science instruction (William 2011). These approaches include assessment practices such as **authentic assessment**, in which students apply scientific knowledge and reasoning to situations similar to those they will encounter outside the classroom and to situations that approximate how scientists do their work (NSES 1996); or **informal formative assessments** (Eisenkraft 2004, Ruiz-Primo 2011), through which everyday learning activities are used as potential assessments that provide evidence of students' learning in different modes (oral, written, graphic and non-verbal evidence). These approaches have the potential of making students' thinking explicit in non-obtrusive ways (Ruiz-Primo 2011).

Furthermore, the expansion of inquiry-based science education paradigms in the last years has placed an important assessment emphasis in collaborative work and transversal competences and skills (Minner et al., 2010). This is further encompassed by the recent emergence of the RRI approach, which, as mentioned above, promotes a vision of **science education embedding social and ethical principles so as to prepare students for active citizenship in democratic societies** (RRI-Tools 2015). The inclusion of RRI values and process requirements in science education calls for innovations in the development of assessment frameworks integrating these aspects.

The current research report frames a rigorous and accurate assessment approach for the PERFORM project through combining a literature review of international peer-reviewed articles on science education assessment and a set of exploratory workshops with selected students in PERFORM case studies. In doing that, this report expects to contribute not only to PERFORM, but also to the development of further assessment frameworks in science education aiming to address RRI values and process requirements and to operationalize the cognitive and experiential aspects of learning in ways able to better capture its complexity and multidimensionality. In what follows, we provide an overview of the learning aspects to be addressed in PERFORM's assessment.

2.2. Assessment in PERFORM: main concepts and definitions

To design an assessment framework, it is necessary to have a precise understanding of the impacts or outcome variables that the assessment indicators are supposed to measure

(Strand et al. 2015). In the case of PERFORM, a set of learning outcomes and process requirements originating from the RRI approach, and thus related to RRI values, transversal competences and both experiential and cognitive aspects of learning, are included in our assessment as outcome variables.

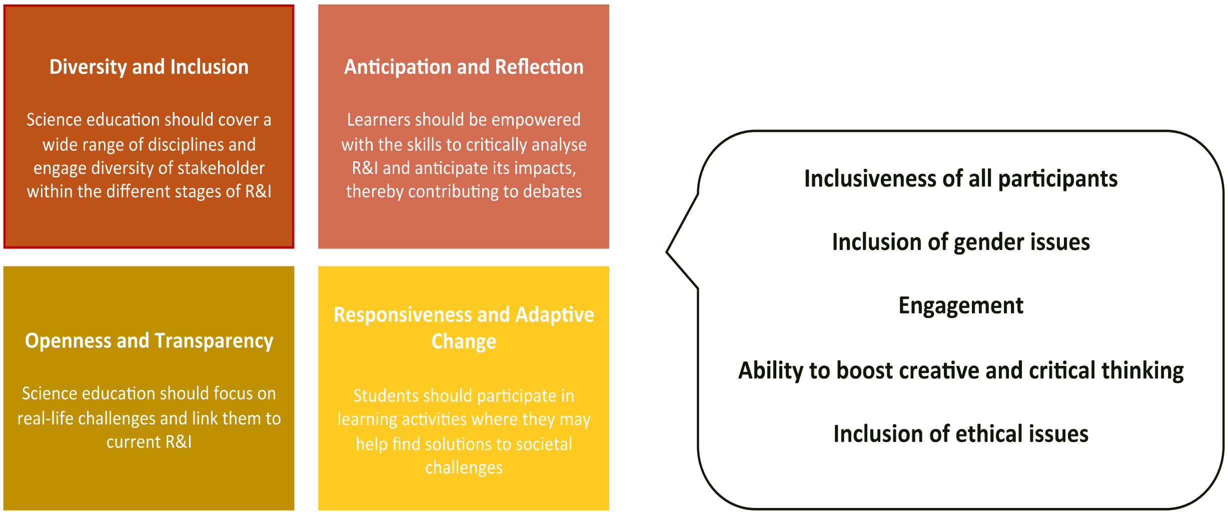
Science education has been identified as a key policy agenda within RRI (Hazelkorn et al. 2015), as part of its aim to achieve socially desirable and socially acceptable ends through an inclusive and deliberative process (Owen et al. 2012). Under the lenses of RRI, science education focuses on preparing students for active citizenship in democratic societies and better equip them with the necessary knowledge, resources and skills to face current complex societal challenges by fostering critical thinking and reflexivity about science and scientific research, and embedding social and ethical principles (EC 2012a, RRI Tools 2015). Along with other educational approaches (see for instance, life-long learning), the RRI approach implies a shift in the focus of science education from learning discrete scientific facts to understanding how to apply science learning to different and new situations, stimulating curiosity, scientific thinking and the understanding of the nature of science (Hazelkorn et al. 2015). Taking into account the different RRI process requirements (i.e., diversity and inclusion, anticipation and reflection, openness and transparency, responsiveness and adaptive change), we have identified five process requirements and learning outcomes of relevance to our assessment (see Figure 1):

i. **Inclusiveness of all participants**: the educational process should be able to reach diverse students' profiles and learning styles and to include different relevant stakeholders.
ii. **Inclusion of gender issues**: the educational process should be sensitive to gender differences and critically approach and manage gender aspects of science and research.
iii. **Engagement**: the ability of the educational process to enhance students engagement with science and scientific research, both in terms of emotional engagement (i.e student's active implication related to intrinsic motivation, affective reasons and/or interest) and cognitive engagement (i.e. sustained, engaged attention during a task or process requiring mental effort).
iv. **Creative and critical thinking**: the ability of the educational process to boost students' ability to question and reframe scientific content, to adopt a systems

thinking perspective, to connect topics with experience and to seek other points of view.

v. **Inclusion of ethical issues**: the educational process should be open, responsive and transparent to participants and reflect ethical aspects of science and research, including values, interests and conflicting perspectives.

Figure 1 RRI process requirements and values to be addressed in PERFORM assessment (based on RRI-Tools 2015)



The shift in the educational focus proposed by the RRI approach emphasizes as well the acquisition of **transversal competences or skills** by learners**.** We adopt the term **transversal competences** following the framework of the European Commission on key competences for lifelong learning (EC 2012b). In the context of education, competences can be understood as '*the ability to successfully meet complex demands in a particular context (…), through the mobilization of knowledge, cognitive and practical skills, as well as social and behaviour components such as attitudes, emotions, and values and motivations*' (Rychen & Salganik 2003, in EC 2012b, p.5). The EC identifies key competences, as those competences particularly necessary for personal fulfilment and development, social inclusion, active citizenship and employment, including those transversal competences that we focus on in our assessment (EC 2006):

i. **Learning to learn** is related to the ability to pursue and organize one's own learning, in accordance with one's own needs, and to the awareness of learning methods and opportunities;

ii. **Social and civic competences** refer to personal, interpersonal and intercultural competences and all forms of behaviour that equip individuals to participate in an effective and constructive way in social and working life;

iii. **Sense of initiative and entrepreneurship** is the ability to turn ideas into action. It involves creativity, innovation and risk-taking, as well as the ability to plan and manage projects in order to achieve objectives. The individual is aware of the context of his/her work and is able to seize opportunities that arise within.

As experiential aspects we include all feelings, emotions and embodied insights generated, expressed or processed through the experience of engaging in science learning. This includes:

i. Attitudes and perceptions both towards science and/or the scientific issues approached and the learning experience.

ii. Experiential aspects related to emotional, body and spatial awareness arising during or as a result of the educational process.

Finally, cognitive aspects in the assessment specifically relate to basic, tacit and conceptual knowledge about science and STEM topics. Such learning outcomes include both the assimilation, acquisition and refinement of explicit scientific facts and concepts and of implicit procedural knowledge. Cognitive aspects are also included in the RRI process requirements and learning outcomes mentioned above, as part of cognitive engagement and critical and creative thinking, and in transversal competences, as part of reflective thinking.

3. RESEARCH STRATEGY TO BUILD PERFORM ASSESSMENT INDICATORS

Our research strategy is developed through an expert-based literature review and several exploratory workshops in PERFORM participating schools. Both have been carried out with the global aim of contributing to the development of task 4.1*: Design of an innovative and participatory impact assessment research methodology tailored to the PERSEIAs participatory processes,* through the proposal and operationalisation of relevant assessment indicators in the context of PERFORM. The next subsections describe their global aim, design and implementation.

3.1. Expert-based literature review

3.1.1. Global aim and focus

As mentioned above, the design of a sound impact assessment strategy for PERFORM needs to take into account how assessment frameworks in the field of science education have addressed learning outcomes and processes requirements related to the RRI approach so far. For that purpose, it was necessary to conduct an expert-based, systematic literature review. This review globally aimed to identify and characterize assessment frameworks used in the context of science learning and engagement with young people. Specifically, it examined the operationalization of: i) RRI values and process requirements, ii) transversal competences, iii) experiential aspects and iv) cognitive aspects. In doing that we identified assessment gaps and challenges relevant to the context of PERFORM and, more broadly, to the development of science education assessments incorporating the RRI dimension (see section 7).

Consequently, this review focused on **assessment frameworks** in the context of **science learning and engagement**, with an emphasis on **RRI values and process requirements**, **transversal competences** and **experiential aspects,** as described in the section above. By **assessment framework** we refer to a set of interlinked criteria, practices and concepts providing a systematic way of data collection, analysis and interpretation to the study of science learning and engagement.

3.1.2 Design and data collection

Taking into account the framework described above, two global questions guided our review:

1) *What assessment frameworks are currently applied in science learning and engagement? To what extend are they transdisciplinary?*
2) *How are RRI, transversal competences and experiential factors included in such frameworks? What are the gaps or methodological challenges related to their inclusion?*

These questions were split in the following specific review questions and sub-questions:

1) *What assessment frameworks can be identified in the selected sample?*
    a. On which disciplines are they based? Are they transdisciplinary?
    b. What is being assessed in these frameworks?
    c. How it is the evaluation conducted?
    d. What are the challenges of each approach for assessing science learning and engagement?
2) *How are transversal competences, RRI and emotional factors included in these frameworks?*
    a. How are these notions operationalised?
    b. What kinds of evaluation indicators are applied for data collection, if any?

Based on these questions, data collection was carried out by three coordinated reviewers, in three consecutive data screenings from December 2015 to March 2016. Previously, the team of researchers agreed on the operationalization of learning criteria related to RRI values and process requirements and transversal competences, based on RRI and science education reports.

The following subsections introduce the keywords, databases and categories of variables that were used for data collection.

Table 1 Steps and calendar of data collection in the systematic literature review

Literature review steps	Schedule
1. Proposal of an operative definition of transversal competences and RRI learning outcomes and process requirements and identification of potential criteria based on literature. 2. Initial screening of materials following the search terms. 3. Second screening: identification of relevant papers within the sample according to search criteria. 4. Distribution of papers among the researchers and test for consistency.	December 2015
5. Data collection and sharing of preliminary analysis.	January – February 2016
6. Third screening according to preliminary results and distribution of papers.	March 2016

Scopus scientific database (https://www.scopus.com/) was used as search engine. We chose Scopus due to its extensive database of peer-reviewed international journals and to ensure the rigour and comparability of the data provided.

Table 2 shows the keywords or search terms that guided the literature review. Keywords were grouped in the following order: (TITLE-ABS-KEY ("science Learning" OR "science engagement") AND ALL (assessment OR evaluation) AND ALL (framework OR approach OR perspective OR method)). Due to scope of the review, we narrowed the search to articles and book chapters as document type, but included all subject areas and all years to present in data range. This search provided a list of **166 scientific papers and book chapters**.

Table 2 Literature review keywords

Context	Object	Descriptor
Science learning Science engagement	Assessment Evaluation	Framework Approach Perspective Method

We carried out a second screening to assess what articles resulting from the scoping fulfilled the inclusion criteria. The three of us read the abstracts of the selected articles and chapters to discard those not relevant to the review. The abstracts were rejected if they fulfilled any of the following criteria:

i. **Focus:** those not directly related to educational activities or projects focused on science learning and engagement (e.g. teaching practices in science education, assessment of curricula in science).
ii. **Context:** those referring to other scientific contexts than formal or informal learning and education with young people (e.g. assessing learning in trainings to teachers).
iii. **Type of article (case study/ review):** those dealing with reviews or secondary data.
iv. **Target group:** those not related to young people (i.e., primary school students to undergraduate students).

As a result, we got a first sample of **47 relevant scientific articles and book chapters**. The excluded articles and the criteria behind their exclusion are listed in Annex 1.

The relevant sample was split in three subsets so as to be reviewed by one reviewer each. To enhance the consistency of data collection and reduce potential bias related to the different reviewers, the three reviewers independently categorized the data of two of the studies prior to data collection. Differences in the data extracted and the interpretation of the variables were then discussed among the reviewers to reach consensus and achieve clarification of the coding criteria and consistency in its application.

Using a **snowball sampling** strategy during the second screening (e.g., other papers cited in the reviewed articles), we identified an additional **set of 38 articles of interest to the review. These were submitted to the same screening process, providing a sample of 20 relevant articles** for a third stage of data collection. Due to the focus of PERFORM on performing arts, we targeted experiences applying any form of artistic practice in science learning and engagement activities, although other topics were also included in this last snowball sampling because of their relevance for STEM education assessment.

Table 3 shows the characterization of the final sample according to document type. Peer-reviewed articles cover a very diverse sample, with 35 different scientific journals, among them: the International Journal of Science Education (n=11), Science Education (n=5), Computers and Education (n=5) and Research in Science Education (n=4). Annex 1 contains the whole list of included journals in the sample.

Table 3 Characterization of the final sample according to document type and journal

Document type	Number of items
Total sample	67
Peer-reviewed articles	63
Book Chapters	4

Table 4 and Table 5 show the different data collection categories and subcategories used in the review to characterise the assessment frameworks and the evaluation of RRI, transversal competences, and cognitive and experiential aspects. Categories of data collection corresponded to: i) variables characterizing the educational activity (such as length, number of participants, topics approached or type of participation); ii) variables characterizing the assessment (such as assessment focus, data collection methods or analysis strategies) and iii) aspects included in the reviewed assessment to address criteria related to the different identified learning outcomes and process requirements. By criteria we mean the conditions that need to be met in order to achieve science learning and engagement. Definitions of each criterion were developed and enriched along the literature review and they are included in section 5.2, in the indicators proposal.

It is important to mention that, although categories of data collection were previously defined, categories related to assessment criteria included an open category for 'other criteria' as well, in order to allow the inclusion of unexpected and emergent criteria. For instance, in the preliminary data collection to test consistency, a new criterion related to basic cognitive aspects of learning emerged. Although this is not the focus of PERFORM, it was included in the review as these aspects are frequently part of science learning and engagement assessments.

Table 4 Categories and subcategories to conduct data collection on educational activities and assessment frameworks

Categories	Subcategories
Basic information	Publication Type; Country of implementation; Objective of the paper Learning approach Context of application; Level of education; Socio-economic status
Educational experience	Topic/s addressed Duration; Number of participants; Specific target group Participatory approach; Type of participation; Manner of interaction
Assessment framework	Focus of the assessment Disciplines applied/ disciplinary background Data collection source
Assessment approach	Type of assessment approach (mix, quantitative, qualitative) Participatory approach
Collection methods	Self-reported scales Close-ended surveys Open-ended surveys Interviews Observation Focus groups Mural evaluation Arts-based methods Others / Key details [Specify, for instance if there's pretest] Takes into account unexpected outcomes? How? Appendix with collection tools
Analysis Strategies	Statistical analysis Type of statistical analysis Qualitative analysis Type of qualitative analysis

Table 5 Categories and subcategories to conduct data collection on assessment criteria and indicators. Based on: RRI-Tools 2015, EC 2012a, EC 2012b and EC 2006.

Categories		Subcategories
Related to	**Process requirements and learning outcomes**	**Assessment criteria**
Responsible research and innovation	Inclusiveness of all participants	Balanced participation
		Fostering dialogue among participants
	Gender	Gender equality in participation
		Approaching critically gender issues
	Engagement	Emotional engagement
		Cognitive engagement
	Ability to boost creative and critical thinking	Questioning and reframing
		Systems thinking
		Connecting topics with experience
		Acknowledging contradictions
		Acknowledging uncertainty
		Seeking other points of view
	Inclusion of ethical issues	Social relevance of topics addressed
		Participants acceptance of process/outcomes
		Connecting scientific topics with values
	Others	[Open category to other criteria]
Transversal competences	Learning to learn	Understanding the value of learning
		Learning autonomy
		Reflective thinking
	Social and civic competences	Communication skills
		Collaborative skills
		Respect for society and environment
		Informed and reasoned decision-making
		Ability to resolve conflicts
	Sense of initiative	Entrepreneurship
		Self-confidence
		Ability to plan & manage projects
	Others	[Open category to other criteria]
Experiential aspects	Feelings and emotions	Enjoyment
		Emotional awareness and reflexivity
		Body awareness
		Empowerment and sense of belonging

	Perceptions and attitudes towards science	Attitudes towards science and the scientific issues approached
		Perceptions of science and the scientific issues approached
	Others	[Open category to other criteria]
Cognitive aspects	Cognitive dimension	Basic knowledge and conceptual change
	Others	[Open category to other criteria]

3.1.4 Data analysis

Data from the reviewed 67 papers and book chapters were analysed with a twofold objective: i) to identify assessment gaps and challenges relevant both for PERFORM and, more broadly, for future assessment methodological developments integrating RRI, and ii) to identify and operationalise relevant assessment criteria and indicators related to RRI values and process requirements, transversal competences, and experiential and cognitive aspects.

Data related to educational activities and assessment frameworks was quantitatively analysed through a basic descriptive analysis to get a broad characterization of the identified categories and subcategories. Data related to data collection subcategories referring to assessment criteria (Table 5) were analysed through a conventional content analysis (Hiesh & Shannon 2005) to identify potential indicators providing specific information of the corresponding criterion. Table 6 includes an example of the content analysis for the gender criteria, within the RRI dimension.

Table 6 Content analysis: example of the identification of GENDER indicators from aspects included in the reviewed assessments

Data collection subcategories	Data from the review	Indicators proposed by reviewers
Gender equality in participation	Number of boys and girls participating in the activity. Distribution of participants by gender.	– Number of students in the activity by gender

	Male and female participation in science activities inside and outside school	– Students' engagement in science in and out school by gender
	Different participation according to gender	– Type of tasks and roles assumed by students in the activity
Approaching critically gender issues	Different affective responses according to gender. Gender differences in response to puppets. Different reactions to the methods according to gender.	– Students' affective responses to the activity by gender
	Differences in gender performance of computer-based games for each grade. Gender equity in science engagement and literacy (compared test scores).	– Students' performance in the activity by gender

3.1.5 Limitations of the expert-based literature review

Results of this literature review show the need to acknowledge the bias produced by the use of Scopus as search engine. As mentioned above, the use of Scopus was prioritised to ensure rigour and comparability of the data set. Consequently, all included articles are written in English and therefore the sample is biased towards Anglophone countries. To deal with this limitation in the context of PERFORM, further exploration could be complemented with insights from French and Spanish literature. Furthermore, we also identified a lack of transparency about the assessment methods reported in some of the articles, which made difficult to conduct the data collection in a systematic way. This lack of transparency is manifested either in terms of missing information about the data collection tools applied and their specific design, and in terms of lack of clarity about the analytical approach used. Therefore, results refer to the information available in the reviewed papers.

3.2. Participatory assessment: incorporating students' views

3.2.1 Global aim and focus

Based on the participatory approach of the PERFORM project, we paid special attention to the inclusion of students' views and opinions about science learning and engagement in the evaluation process by conducting exploratory workshops with students in selected schools in each case study (France, Spain, and the UK). This had a twofold objective: to actively involve the students in the assessment process since the beginning, and to contextualise the PERFORM impact assessment methodology in each educational setting. As a result, we identified both criteria and indicators that students consider important when assessing the impact of science-related activities they experience both inside and outside school.

3.2.2 Design

In December 2015 we designed the workshop protocol. It focused on promoting students' identification of and discussion about key aspects that motivate them to learn and be engaged in science-related activities. The methodological protocol was then reviewed by the coordinators of each case study (Les Atomes Crochus, Science Made Simple and The Big Van Theory) and a researcher from the University of Warwick, and required changes were implemented (see the protocol in Annex 2).

3.2.3 Data collection

An exploratory workshop was implemented in each selected school in each country between March and May 2016, except in the UK where it was not possible to conduct one of the workshops due to the large number of students participating (more than 30) (Table 7). Thus, a total of 161 secondary-school students participated in these 11 workshops: 65 of them in Spain, 57 in France and 39 in the UK. Informed consents to participate in PERFORM research were obtained.

Table 7 Number of girls and boys involved in exploratory workshops, by case study and school

Case study	School	Girls	Boys	Total
Spain	IES Consell de Cent	9	9	18
	Ins Europa	8	8	16
	IES Castellbisbal	6	5	11
	Ins Santa Eulàlia	9	11	20
France	Collège Marie Curie	8	0	8
	Collège la grange aux belles	9	9	18
	Còllege Jean Zay	7	7	14
	Collège Les Toupets	11	6	17
UK	Fairfield High	6	9	15
	Albany Academy	6	6	12
	Derby High	8	4	12

Data were gathered through the post-its students wrote with the aspects they like and do not like when participating in science-related activities, including suggestions to design activities that foster their engagement in STEM. Students' comments and discussion about these topics were also recorded in written notes by the facilitators during the workshops. Workshops were audio-recorded to complement these notes (see data collection table in Annex 3). Facilitators also wrote their perceptions about the mood of the students, their reception of the activity, any unexpected event and any other relevant factors that could affect the implementation of the workshop. Facilitators' comments were thus useful to identify contextual particularities in the development of the activity.

3.2.4 Data analysis

Data were analysed through a conventional content analysis (Hiesh & Shannon 2005). Categories and subcategories referred to 1) criteria and 2) indicators students consider important for a science activity to be motivating, in each case study. Some categories and subcategories were pre-defined according to the criteria and indicators we previously found in our literature review (e.g., '*Student's amusement during the activity*' as an indicator of '*Enjoyment*') while others directly emerged from our collected data (e.g., '*Student's discovery of something not previously known*' as an indicator of '*Enjoyment*'). We also identified those criteria students perceived as the most important for science learning in each case study through the scores they assigned in the barometer exercise.

Image 1 A moment of the exploratory workshop in Institut Europa secondary school in Barcelona.

4. WHAT TYPES OF SCIENCE EDUCATION ACTIVITIES ARE BEING ASSESSED?

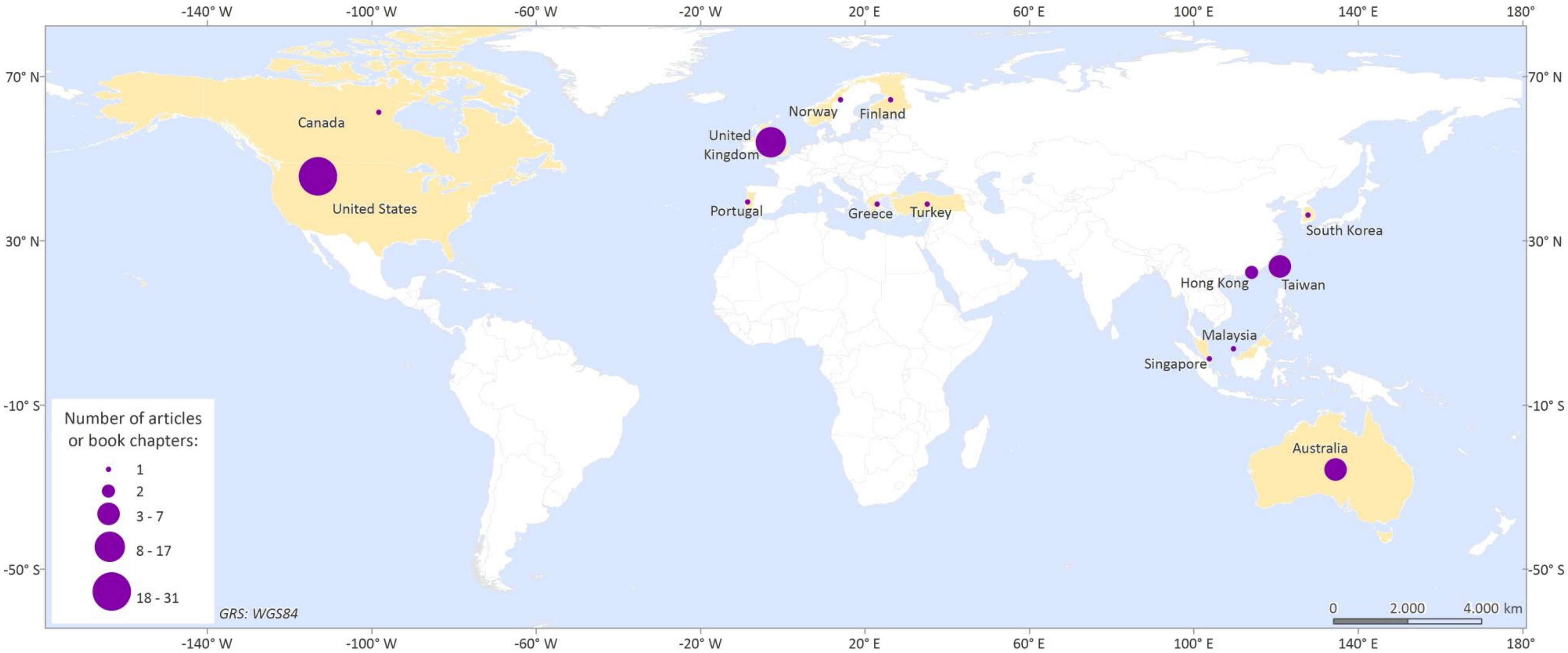
Findings reveal that most of the 67 assessed science education activities identified in the literature review are conducted in formal education settings, most of them focusing on chemistry, physics and life sciences, whereas only a few are about mathematics and engineering, and more than a half include participatory methods and/or approaches.

4.1 Educational contexts and learning approaches

This section provides a global picture of these science learning and engagement activities and their contexts of implementation (see Figure 2).

- Most of the reviewed educational experiences take place in formal education settings (n=57, out of 67).
- Almost half of these experiences are developed in secondary schools (n=31), followed by primary schools (n=17) and universities (n=12).
- Around half of the experiences (n=35) do not explicit a given learning approach.
- Among those that do (n=32), the most applied learning approach is e-learning (electronic learning, n=10), understood as learning that occurs through virtual platforms by integrating information and communication technologies (ICT) tools both in the teaching and learning practice.
- A number of experiences apply as well a diversity of student-centred learning approaches, such as inquiry-based (n=5) and project-based learning (n=3), learning-by-doing (n= 3), collaborative learning (n=3) or situated learning (n=3). In these approaches students' questions, ideas and social interactions are at the centre of the learning experience.
- Most of the activities are located in USA (n=30), UK (n=17) and to a lesser extent, Australia (n=7) and Taiwan (n=4).

Figure 2 Map of countries covered in the literature review sample and number of papers and/or book chapters included for that country



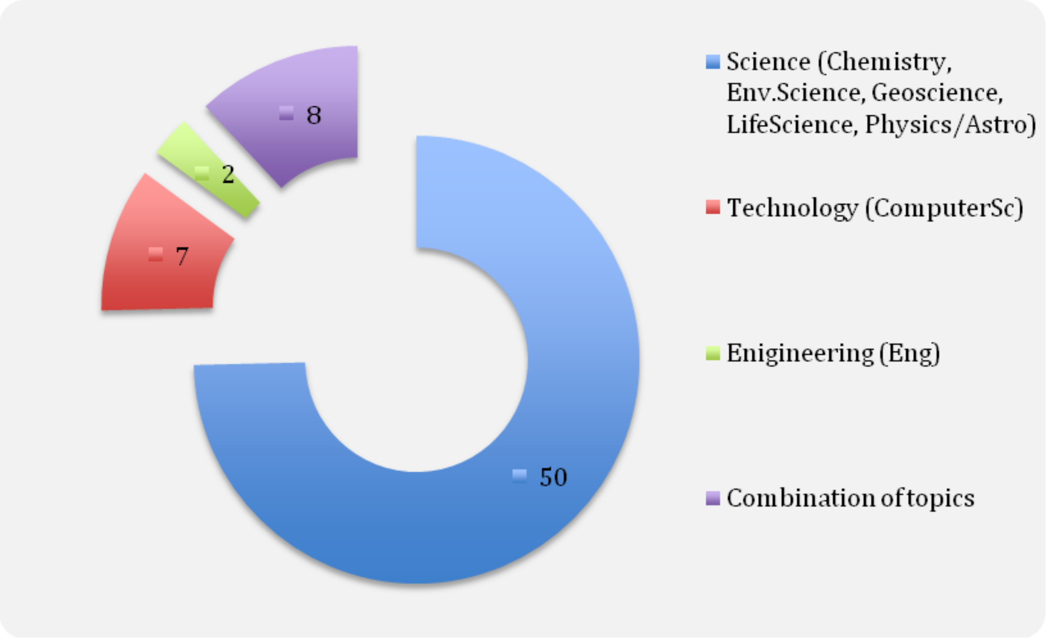
Source: own elaboration with cartographic data of GADM repository (http://www.gadm.org/).

4.2. STEM topics approached and target groups

The scientific topics approached during the activity were also reviewed along with the potential focus of such activities on specific target groups of participants. The following points summarise the main findings in this regard.

- Taking as reference the STEM label, most of the topics approached can be included within the broad category of 'Science', including: chemistry, physics, life sciences, geosciences and environmental sciences (n=50).
- 'Technology' is the topic in seven of the educational activities, 'Engineering' in two, and no activity is exclusively focused on 'Mathematics' (see Figure 3).

Figure 3 Topics approached in the reviewed science education activities (n=67)



- The majority of the experiences (n=62) are not addressed to a specific target group. Only five are focused on specific groups, such as students with difficulties or with high-level of school achievement and students with diverse ethnolinguistic backgrounds.

- An important number of articles (n= 42) do not specify the socio-economic background of their participants. Among those that specify it (n=25), 15 of them are addressed to students from mixed socio-economic backgrounds, and 10 to students from a low socio-economic background.

4.3 Duration and participation

The following points summarise the main findings related to the duration of the activities and the types of participation that were fostered through them.

- Half of the activities (n=34) are developed in several sessions along several weeks, while 22 activities consist of one single session.
- Forty-two of such experiences are designed as participatory activities.
- The main types of participatory practices are related to hands-on activities (participation in labs and experiments, n=30), group discussions (n=28, artistic creation (e.g., drama, drawings, songs, n=15), and information and communication technologies (ICT, e.g. online environments, computer-based games, n=12).
- Less often, participatory activities also include visits and interactions outdoors (n=5) and the active involvement of students in formative and peer assessments (n=2).
- In two cases, interestingly, participants are paid for participating in the activity or in the assessment. Payments are made either to the teachers and mentors through a modest stipend, or to randomly selected students, by taking part in a contest if they participate in the evaluation.

5. WHAT ASSESSMENT FRAMEWORKS CAN BE IDENTIFIED?

Overall, the assessment frameworks used to evaluate the reviewed science education activities focus on students' performance in science through looking at cognitive aspects of their scientific learning. Characteristics of both educational experience and methods applied are analysed in a less extent. Most of such frameworks rely on one discipline but their methodological approaches are built upon different disciplines.

This section provides a description of the assessment frameworks used to evaluate the science learning and engagement activities described above. According to our review questions, these assessment frameworks are globally characterised in terms of their learning outcomes and process requirements addressed, the disciplines applied and their potential transdisciplinarity, as well as the approaches and methods developed.

5.1 Assessment focus and disciplines

The focus of the assessments is diverse in terms of targeted learning outcomes and processes requirements (see Table 8).

Regarding the assessment of learning outcomes, although more than half of the sample (n=39) focuses on two or more of them, the acquisition of conceptual and tacit scientific knowledge is the learning outcome most frequently assessed (n= 44). This is mainly done through the assessment of conceptual understandings and change (including knowledge gains and achievements) and scientific literacy or tacit knowledge (i.e., procedural knowledge gained through experience). General perceptions, values and attitudes towards science or the scientific issues approached –including engagement, motivations or interest, are also included in 19 assessments. Thirteen of the assessments also address students' skills and competences (e.g., cognitive, social, personal) and the extent to which the educational activity enhances them. Only one assessment considers the long-term impacts of learning, in this case, through analysing the persistence and kinds of students' memories of the scientific educational experience.

There are also a number of assessment experiences (n=29) that include procedural aspects related to the science education activity and the learning process generated. In these cases the scientific approach is assessed in terms of its capacity to enhance science teaching and learning and in terms of the participants' reactions to and perceptions of the methods proposed (mainly the students, but also the teachers). Other assessed procedural aspects are the perceptions and attitudes of the students towards their learning experience (n=7), in terms of how this learning is constructed and perceived (i.e., self-awareness of the learning process) and teachers and/or educators reactions to the teaching strategy and their interactions with the students (n=3).

Table 8 Learning outcomes and procedural aspects included in the assessment and number of experiences that include them

Learning outcomes	Number of experiences
Conceptual and tacit scientific knowledge	44
General attitudes, values and perceptions towards science	19
Students' skills and competences	13
Long-term impacts of the education activity	1
Procedural aspects	
The educational activity and the methods or platforms applied	29
Students' perceptions and attitudes towards their learning experience	7
Teachers reactions and interactions with students	3

Regarding the disciplinary and theoretical backgrounds through which these assessment approaches have been generated, as expected, most of the assessments rely on theoretical frameworks coming from science education and educational research (n=55). Theories of motivation and engagement applied to science learning represent an important conceptual background in the assessments. These are also characterized with the contribution of theories and conceptual frameworks from educational, social and cognitive psychology (n=11). To a lesser extent, other disciplines, such as communication and semiotics (n=5) and philosophy, anthropology and sociology (each n=2) provide an analytical base for approaching participants' discourse and their learning experiences.

Interestingly, transdisciplinarity in the assessment frameworks is found mostly in the methods applied, feeding from diverse and different disciplines. Along with conventional data collection methods used in science education, such as self-reported scales, surveys or observation, data collection methods based on information and communication technologies (such as computer-based games and virtual platforms) and on artistic tools (mainly drawings and video) are included in a relevant number of experiences (n=29).

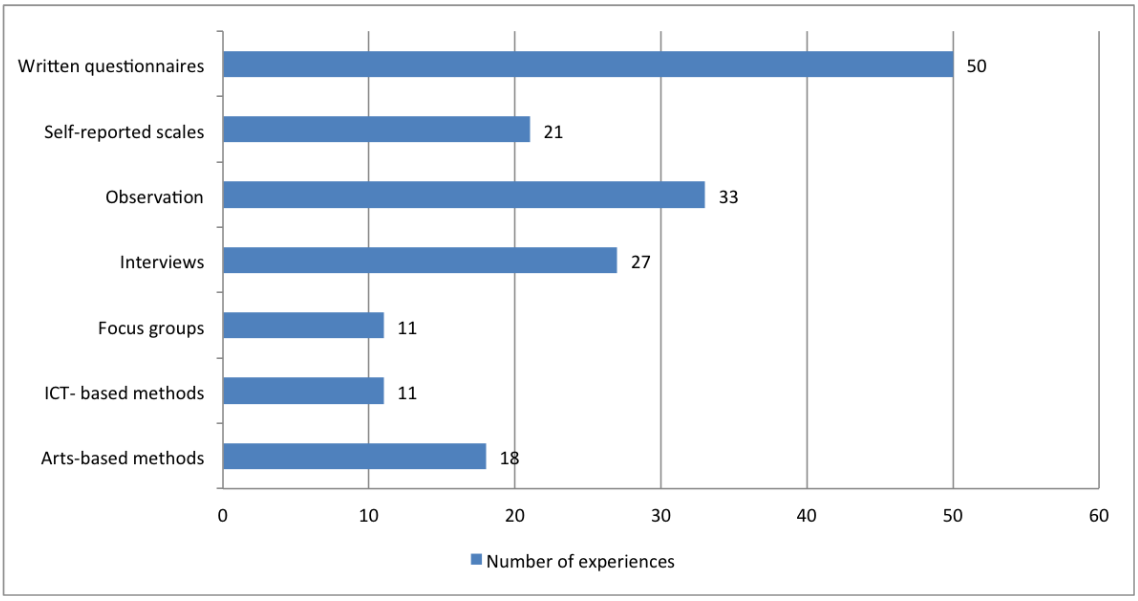
5.2. Assessment design and methods

In most of the cases (n=42), students are the main source of data collection. Only a third of these experiences (n=20) include also teachers or educators while collecting data. Furthermore, assessments including other actors beyond students and teachers are marginal among the sample (n=4). These actors are students' parents, other school staff (e.g., school principal, administrative staff) and undergraduate students participating in the activity as mentors.

Regarding the assessment design, assessments prior to and after the implementation of the activity are used at least in one third of the sample (n=24), in order to track changes in students' cognitive and experiential aspects, skills and values related to their involvement in the educational intervention. Some of the experiences also develop pilot assessments to test and refine the data collection methods, but these are not frequently reported. Similarly, very few experiences use case control groups to compare their data.

Regarding data collection methods, there is a clear predominance of the use of written questionnaires (n= 50). Most of these questionnaires either used only close-ended questions (e.g. multiple choice; n= 19) or combined a majority of closed-ended questions with some open-ended items as well (n= 16). 15 questionnaires used only open-ended questions, of qualitative nature.. Observation (both structured and unstructured); interviews; and self-reported scales were also commonly applied (see Figure 4 for more details). As mentioned above, it is also remarkable the presence of ICT and arts-based methods not only in the development of the educational activity, but also in the assessment. ICT tools are applied in 11 assessments to create different kinds of virtual forums to foster online discussions (n=4), to monitor students' knowledge through interactive computer simulations and games (n=4), and to support conventional methods like surveys or interviews through both online surveys (n=2) and videos (n=1). Arts-based methods are applied in the sample almost as often as self-reported scales (n=18 and n=21, respectively). They are developed in the form of video making (about the topics approached or their learning experiences, n=7), drawings (of scientists, science topics and science class; n= 5), dramatizations (of scientific models and topics, n=4) and learning stories (storytelling based on their learning, n=2).

Figure 4 Number of science education activities including each identified data collection method in the assessments.



Despite the balance between mixed (n=29), quantitative (n=24) and qualitative approaches (n=14), assessment approaches seem to provide more weight to quantitative data. This is suggested by the lower number of assessments taking into account unexpected outcomes (n= 24) –being unexpected outcomes associated to the use of exploratory methods and inductive analysis approaches (see Limitations below).

Finally, due to the relevance of participatory approaches in PERFORM we report findings related to the participatory nature of the assessments analysed. We consider an assessment as participatory when it includes the students either in: i) the design of the assessment approach, ii) its implementation, and/or iii) its analysis. Under this perspective, most of the assessment approaches in the sample are non-participatory (n=53). This implies that students are a source of data, rather than an active agent in the evaluation. Among those approaches actively including participants in the assessment (n=14), participation is frequently developed as peer-review by fellow classmates or older students (in the role of mentors), and in terms of diagnostic and formative evaluation during the process, carried out by students or teachers (for instance, through reflective diaries).

6. EXPERT-BASED CRITERIA AND INDICATORS IN SCIENCE EDUCATION ASSESSMENT

A set of **86 assessment indicators** emerge from our literature review. These relate to the RRI values, transversal competences and experiential and cognitive aspects of science learning and engagement identified in the reviewed science education experiences.

Due to the focus of the PERFORM research and the emergence of the RRI framework, a strong emphasis has been given to the identification of indicators related to **RRI learning outcomes and process requirements, with more than 40 indicators identified.** In doing this, we expect not only to provide a basis for a sound assessment framework in the context of PERFORM, but also to contribute to current research on RRI in science education.

The indicators have been organised according to 32 assessment criteria - corresponding to 11 learning outcomes and process requirements, previously identified in the first stage of the literature review (see section 3.1). The next subsections characterise such assessment criteria identified in the sample and propose the corresponding list of indicators that also emerges from the review.

6.1 Assessment criteria found in the literature review

As mentioned in section 5, students' **cognitive aspects** are, by far, the most assessed. Fifty-six out of the 67 reviewed assessments included cognitive criteria, mostly in terms of the improvement of student's acquisition of basic knowledge about the scientific topic(s) addressed in the activity. Due to the focus of PERFORM, we did not disaggregate cognitive criteria included in the cognitive dimension. Therefore the high frequency of the cognitive criteria (n=58 out of 67) cannot be compared to the frequency of other dimensions' criteria. However, this result does inform of the relevance of the cognitive dimension in assessments, since cognitive aspects are also included in a high number in several criteria identified for the RRI dimension (e.g., 'Cognitive engagement', 'Questioning and reframing', 'Systems thinking') and in the transversal competences dimension (e.g. 'Reflective thinking’).

After the cognitive dimension, the **experiential** one emerges as the second more frequently assessed in the reviewed literature. Although it is included in the same number of assessments as the RRI dimension (n=47) and very close to transversal competences (included in 46 assessments), the experiential dimension contains two of the most cited criteria. These are ‘Attitudes towards science and/or the scientific issues approached’ (n=39) and ‘Perceptions towards science and/or the scientific issues approached’ (n=31). However, with the exception of ‘Enjoyment’ (n=27), criteria related to emotions and feelings associated to the experience –such as emotional awareness and reflexivity, embodiment or empowerment– were not frequently included in the assessments. For instance, the most cited criterion in this group is ‘Emotional awareness’ (included in 11 assessments), while the less present, ‘Sense of belonging’, was included in three assessments.

The dimension of **transversal competences** seems to be as well unbalanced in the assessments. Although criteria related to the transversal competences identified are included in 46 of the assessments, this is mostly done through the inclusion of aspects related to ‘Reflective thinking’ (n=36) –the second most cited criterion and very related as well to the cognitive dimension. A second group of criteria, related to ‘Communication skills’ (n=23), ‘Collaboration skills’ (n=20) and ‘Self-confidence’ (n=21) is also included, but to a much lesser extent. The rest of transversal competences related criteria are included in less than 16 assessments. Among the less included criteria, we found ‘Ability to resolve conflicts’ (n=1), ‘Respect for society and environment’ (n=1), ‘Informed and reasoned decision-making’ (n=9) and ‘Understanding the value of learning’ (n=11).

Finally, the broad list of criteria identified for **RRI** is represented in the assessments mostly through the criteria of ‘Cognitive engagement’ (n=31), related to the attention required during a task or process requiring mental effort. The criteria ‘Questioning and reframing’, also related to cognitive aspects but implying higher order thinking, is the next criteria most included (n=26). There is as well a group of less frequent, but still considered, criteria that include ‘Fostering dialogue among participants’ (n=21), ‘Systems thinking’ and ‘Connecting topics with experience’ (both n=18) and ‘Emotional engagement’ (n=17).

Interestingly, gender and ethical issues –two of the five core aspects of RRI- stand out for their absence in most of the assessments. Criteria related to ethical aspects are especially absent in regards to ‘Connecting scientific topics with values’ (n=2), the ‘Social relevance of topics addressed’ (n=4) and ‘Participants acceptance of process/outcomes’ (n=5). Similarly only 5 assessments within the whole sample address critically gender aspects, beyond the usual identification of participant numbers according to gender. This critical approach to gender in the reviewed assessments has been carried out mostly by addressing gender differences in performance and/or outputs and by looking at different affective responses according to gender.

6.2. Assessment indicators proposed

The following tables provide an overview of the different assessment criteria considered in each learning dimension, their operative definition and the indicators proposed. Due to the overlapping nature of the different learning dimensions addressed, many of the indicators can be applied to several criteria.

6.2.1 RRI values

LEARNING OUTCOME AND/OR PROCESS REQUIREMENT: Inclusiveness of all participants		
Criteria	**Operative definition**	**Indicators**
Balanced participation	Inclusiveness and involvement of all participants, making sure that each one has the opportunity to contribute to the process in an active way.	1. Combination of learning pedagogies and resources in the activity to reach all students in the activity 2. Specific support to students with special needs during the activity 3. Inclusion of other participants in the activity (beyond students) and their expertise 4. Student's sharing of tasks and roles in processes and outputs during the activity Also includes indicators 5, 7 and 9
Criteria	**Operative definition**	**Indicators**
Fostering dialogue among participants	Capacity of the process to build learning upon participants' mutual exchange of ideas and opinions so as to integrate different perspectives and work together.	5. Type of dialectic interactions among students in collective creation and group work, if any 6. Student's use of interactive ICT tools in the activity 7. Characteristics of dialogue between students and teachers in the activity Also includes indicator 3

LEARNING OUTCOME AND/OR PROCESS REQUIREMENT: Gender		
Criteria	**Operative definition**	**Indicators**
Gender equality in participation	Participation differences according to gender	8. Students' engagement in science in and out school by gender 9. Number of students in the activity by gender 10. Type of tasks and roles assumed by students in the activity

Criteria	Operative definition	Indicators
Approaching critically gender issues	Acknowledging and critically reflect about gender differences and their causes and implications	11. Students' affective responses to the activity by gender 12. Students' performance in the activity by gender Also includes indicators 1, 8 and 10

LEARNING OUTCOME AND/OR PROCESS REQUIREMENT: Engagement		
Criteria	**Operative definition**	**Indicators**
Emotional engagement	Active implication in the activity or project, related to intrinsic motivation, affective reasons and/or interest	13. Student's willingness to get involved and participate in the activity 14. Student's feelings when experiencing the activity, if any 15. Student's ability to use the body to express and communicate scientific ideas and concepts 16. Student's involvement of emotions in the process of meaning making 17. Student's further interaction and initiatives related to the activity once it is over Also includes indicator 37
Criteria	**Operative definition**	**Indicators**
Cognitive engagement	Sustained, engaged attention during a task or process requiring mental effort	18. Student's degree of involvement in reasoning and argumentation in the activity 19. Student's ability to develop ideas and engage in higher order thinking 20. Student's ability to ask questions, discuss and develop conclusions and/or solutions 21. Time spent by the student in doing the task during the activity 22. Student's willingness to continue working in the activity out of class Also includes indicators 5, 13, 17, 23, 37, 46

LEARNING OUTCOME AND/OR PROCESS REQUIREMENT: Critical and creative thinking		
Criteria	**Operative definition**	**Indicators**
Questioning and reframing	Promotion of understanding through questions that allow students complex thinking and the possibility to see the issues approached in new or different ways.	23. Student's reframing and/or comprehension of scientific concepts based on rationality 24. Student's ability to formulate and test hypotheses and/or research questions Also includes indicators 19, 20, 25
Criteria	**Operative definition**	**Indicators**
Systems thinking	Holistic approach to analysis that considers the interactions between the constituents of a system	25. Student's ability to relate ideas from multiple topics in multiple contexts 26. Student's awareness of issues of scale when approaching scientific topics 27. Student's ability to identify relations and interactions between different elements of a system 28. Use of learning techniques to represent and/or discuss about the whole system in the activity Also includes indicator 33
Criteria	**Operative definition**	**Indicators**
Connecting topics with experience	Contextualisation of the issues approached within their broader societal context and connection with participants' experience	29. Contextualization of scientific topics within societal challenges in the activity 30. Use of students' previous experiences and knowledge as a basis for learning in the activity 31. Facilitation of students' learning through direct, active involvement during the activity 32. Student's ability to apply science concepts to different tasks and/or contexts
Criteria	**Operative definition**	**Indicators**
Seeking other points of view	Consideration of different perspectives and points of view in participants' discourse	33. Student's ability to consider different perspectives and points of view

LEARNING OUTCOME AND/OR PROCESS REQUIREMENT: Ethical aspects		
Criteria Understanding of the nature of science (NOS)	**Operative definition** Key principles and ideas, which provide a description of science as a way of knowing, and the characteristics of scientific knowledge.	**Indicators** 34. Student's awareness of science contradictions, uncertainty, failure and/or risk 35. Student's awareness of power relations in science Also includes indicator 39
Criteria Social relevance of topics addressed	**Operative definition** Degree to which the scientific issues approached are connected to relevant broader social contexts and challenges	**Indicators** 36. Give students the possibility to make learning choices in the scientific activity Also includes indicator 29
Criteria Participants acceptance of process/outcomes	**Operative definition** Degree to which participants accept and feel ownership of the different learning outcomes and processes involved in the activity	**Indicators** 37. Student's creation of own outcomes in the activity Also includes indicator 36
Criteria Connecting scientific topics with values	**Operative definition** Identification and exploration of the diverse values and normative aspects behind scientific practice and knowledge	**Indicators** 38. Inclusion of scientists' personal stories in the activity 39. Show contrasting perspectives regarding the role of science with and for society 40. Student's reflection about ethical behaviour in research

6.2.2 Transversal competences

Transversal competences		
LEARNING OUTCOME AND/OR PROCESS REQUIREMENT: Learning to learn		
Criteria	**Operative definition**	**Indicators**
Understanding the value of learning	Awareness of one's learning process based on the experience and values developed through engagement with science education practices	41. Student's awareness of the professional value of learning science 42. Student's satisfaction to be able to learn science 43. Student's awareness of the value of experiencing science in a given learning environment
Criteria	**Operative definition**	**Indicators**
Learning autonomy	Ability to pursue and persist in science learning, to organise one's own learning, including through effective management of time and information, both individually and in groups	44. Student's ability to organise their own learning by setting own goals in the process 45. Student's ability to use equipment, technology and/or tools to perform the activity 46. Student's ability to persist in a scientific task despite of failure and difficulty 47. Student's ability to autonomously search for relevant and rigorous information Also includes indicators 19, 20, 22
Criteria	**Operative definition**	**Indicators**
Reflective thinking	Ability to gain, process and assimilate new scientific learning and related life experiences through reasoned thinking and/or discussion, in order to use and apply them in a variety of contexts	48. Student's reflection on her/his own learning during the activity 49. Student's ability to recognize relevant information and use it effectively in the activity 50. Student's assessment and reflection about peers' performance in the activity Also includes indicators 18, 24, 25, 56

LEARNING OUTCOME AND/OR PROCESS REQUIREMENT: Social and Civic competences		
Criteria **Communication skills**	**Operative definition** Ability to communicate ideas about science effectively by using verbal, visual and written tools as well as body language	**Indicators** 51. Student's ability to elaborate and share ideas verbally and written during the activity 52. Student's ability to organise and make meaning from visual information in the activity Also includes indicator 15
Criteria Collaborative skills	**Operative definition** Behaviours that help two or more people work together in the science learning process	**Indicators** 53. Student's willingness to ask for help and/or to help others to perform the activity 54. Student's respect towards others' ideas when doing the activity Also includes indicators 4, 50, 57
Criteria Respect for society and environment	**Operative definition** Behaviours that favour acceptance and respect for others, as well as environmental awareness	**Indicators** 55. Student's willingness to assume a responsible position to socially relevant issues addressed in the activity
Criteria Informed and reasoned decision-making	**Operative definition** Ability to analyse, evaluate, and make sound and informed decisions when transferring scientific knowledge into action	**Indicators** 56. Student's ability to contrast different evidence to provide explanations Also includes indicators 18, 20, 55
Criteria Ability to resolve conflicts	**Operative definition** Ability to approach conflict in a constructive manner through managing the process instead of negate it	**Indicators** 57. Student's ability to contribute to the activity through managing difficulties within the group

LEARNING OUTCOME AND/OR PROCESS REQUIREMENT: Sense of initiative		
Criteria Entrepreneurship	**Operative definition** Ability to turn ideas into action when learning science, including innovation and risk-taking	**Indicators** 58. Student's belief in her/his own ability to perform a scientific activity 59. Student's leadership and/or responsibility in the performance of a group activity 60. Student's performance self-assessment during the activity and of its outcomes Also includes indicator 44
Criteria Self-confidence and esteem	**Operative definition** Perceived capability to effectively accomplish a certain level of performance in science learning, including self-esteem	**Indicators** 61. Student's belief in her/his own ability to do well in a scientific domain 62. Student's belief in her/his own verbal ability to discuss about science Also includes indicators 58, 75
Criteria Ability to plan and manage projects	**Operative definition** Ability to plan and manage science projects in order to achieve objectives	**Indicators** 63. Student's ability to plan and/or perform a scientific task and/or project Also includes indicators 58, 60

6.2.3 Cognitive dimension of learning

LEARNING OUTCOME AND/OR PROCESS REQUIREMENT: Cognitive aspects		
Criteria Cognitive aspects	**Operative definition** Acquisition and assimilation of basic, tacit and conceptual knowledge about science and related topics	64. Student's acquisition of basic knowledge about scientific topic(s) addressed in the activity 65. Student's acquisition of tacit knowledge (procedural information) from experiencing the activity Also includes indicators 18, 23 and 49

6.2.4. Experiential aspects of learning

LEARNING OUTCOME AND/OR PROCESS REQUIREMENT: Feelings and emotions		
Criteria	**Operative definition**	**Indicators**
Enjoyment	Feelings of pleasure caused by doing or experiencing something the person likes	66. Student's interest in science and learning science 67. Excitement caused by science and learning science 68. Student's amusement during the activity
Criteria	**Operative definition**	**Indicators**
Emotional awareness and reflexivity	Student's capacity to identify or express emotions associated to the topics addressed and to reflect upon and through their emotional responses	69. Students' expression and/or embodiment of emotions related to the topic of the activity 70. Student's ability to reflect upon and through her/his emotional responses and make consistent behavioural choices in the activity. Also includes indicators 14, 68
Criteria	**Operative definition**	**Indicators**
Body and spatial awareness	Body movement and expressiveness, sensual awareness, relation with the physical space	71. Student's use of the body to convey meanings and kinaesthetic understandings 72. Student's awareness of the influence of the physical space in their learning and engagement in the activity Also includes indicator 15
Criteria	**Operative definition**	**Indicators**
Empowerment and sense of belonging	Sense of ability to do things and feeling of acceptance as part or member within a group or environment	73. Student's sense of belonging to the community when doing the scientific activity 74. Appropriate design of the activity to address students' resources and competences 75. Student's feeling recognized by other participants beyond their classmates Also includes indicators 58, 61, 62, 81

LEARNING OUTCOME AND/OR PROCESS REQUIREMENT: Perceptions and attitudes		
Criteria	**Operative definition**	**Indicators**
Perceptions of science and the scientific issues approached	Participants' organization and interpretation of science concepts, ideas and topics	76. Student's perceptions of the social value of science 77. Student's perceptions of scientists, scientific careers and/or jobs 78. Student's perceptions of the specific topics approached in the activity 79. Student's perceptions of the way science is taught at schools 80. Student's perceptions of the pedagogic approach and methods used in the activity 81. Student's perceptions of the group in the activity, including sense of belonging Also includes indicators 34, 35, 41, 66, 68, 85, 86
Criteria	**Operative definition**	**Indicators**
Attitudes towards science and the scientific issues approached	Participants' predisposition or tendency to respond positively or negatively towards science concepts, ideas and topics	82. Student's curiosity and interest towards science 83. Student's identification with scientific skills and attributes 84. Student's interest in scientific careers and/or jobs 85. Student's attitudes towards the topics approached in the activity 86. Student's attitudes towards the pedagogic approach and methods used in the activity Also includes indicators 13, 21, 41, 42, 55, 66, 67, 68, 76

7. PARTICIPATORY CRITERIA AND INDICATORS IN SCIENCE EDUCATION ASSESSMENT

A total of **15 indicators corresponding to 11 criteria emerge** from the exploratory workshops with students conducted in our three case studies (Table 9). In general, identified participatory indicators are more related to RRI process requirements and learning outcomes (e.g., inclusiveness, engagement, ethical issues) as well as to experiential aspects (e.g., emotions and feelings) than to cognitive aspects and transversal skills.

Only six of these indicators and criteria are mentioned in the three case studies whereas five are mentioned in two of them and four in one case study, which suggests the appropriateness of using participatory approaches for designing science education assessments in different cultural backgrounds. For instance, only the British students highlight that having the possibility of making learning choices during the science education activity is relevant for their engagement and motivation for science.

Seven indicators emerging from these workshops are different from those previously identified through the systematic literature review. Such new participatory indicators mostly correspond to 'Emotional engagement', 'Cognitive engagement' and 'Enjoyment' criteria, as for instance the use of arts-related methods in the activity and students experiencing surprise or discovering something new while doing the activity. Others relate to 'Balanced participation', 'Fostering dialogue' and 'Body and spatial awareness', such as the inclusion of physical activity or activities outdoor and/or outside school. Another new indicator is the contextualisation of scientific topics through experiencing daily life activities, which corresponds to two RRI criteria: 'Connecting topics with experience' and 'Social relevance of the topics assessed'.

Finally, a new criteria different from those used in the literature review, and the corresponding indicator, emerge from a workshop conducted with French students who highlight the value of doing something useful for science, which we classify into the new criteria of 'Scientific relevance'. Specifically, these French students discussed about the need of perceiving a real own contribution to science to be motivated when doing a science-related activity.

Overall, 'Enjoyment' (e.g., amusement and/or having fun) and 'Connecting topics with experience' (e.g., facilitate students' direct, active involvement) are the aspects scored as the most important for science learning and engagement by the students in the three countries. In this sense, doing experiments, games and hands-on activities is perceived by French, Spanish and British students as essential to be motivated and interested in science-related activities. Moreover, and according to the students in Spain, the teacher plays a key role in fostering such engagement. Most of them complained about the traditional and one-way communication methods and tools their teachers use to teach them about science.

Spanish and French students specifically value the opportunity to do science-related activities outdoors and/or outside the classroom, which relates to the criteria of 'Body and spatial awareness'. Moreover, kinaesthetic aspects are important among the Spanish and British students who perceive physical activity and interaction as key for being engaged in science learning. Finally, and interestingly, in the context of PERFORM, to mix scientific content and artistic methods (e.g., painting) when doing a science-related activity is the most valued aspect for being motivated in one of the French schools.

Table 9 Indicators and criteria identified through the exploratory workshops (indicators and/or criteria not previously identified in the literature review are indicated with and asterisk)

RRI values **Indicator**	**Criteria**	**Quotations from post-its and/or discussion**	**Case studies**
Combination of learning pedagogies and resources to reach all students in the activity	Balanced participation	*'Make sure that everyone participates in the scientific activity'.* *'Make all students participate in science classes and not only those who always talk'.* *'Doing things with your friends - not being split; everyone involved'.*	Spain France UK
Use of interactive ICT tools in the activity	Fostering dialogue	*'Use more videos and videogames in science classes, for instance from Youtube'.* *'To use virtual reality to be able to feel with every sense'.* *'Watching little clips'.*	Spain France UK
Student's experience surprise doing the activity *	Emotional engagement / Enjoyment / Perception of science	*'Teachers must motivate us to learn science by doing something surprising, like experiments, something unexpected, to get our attention'.* *'Do something magical.*	Spain France

		We like to do exciting experiments, flashy things like explosions'.	UK
Use of arts-related methods in the activity *	Emotional engagement / Enjoyment	*'Listen music when learning about science because it motivate us, and it is relaxing too'.* *'Dancing and painting'.*	Spain France
Facilitation of students' learning through direct, active involvement during the activity	Connecting topics with experience	*'Less theory and less taking notes, more debates and exchange of ideas'.* *'Conduct our own experiments; to build something new, like robots'.* *'Do experiments because they are interesting'.*	Spain France UK
Contextualisation of scientific topics through experiencing daily life activities *	Connecting topics with experience / Social relevance of the topics assessed	*'Do something that matter for our generation, something related to our environment'.* *'Do workshops about food and nutrition, to learn how to fix a bike or how to check DNA evidence like in TV'.*	Spain France
Contextualisation of scientific topics within societal challenges in the activity	Connecting topics with experience / Social relevance of the topics assessed	*'Focus on scientific topics that are important for our generation, like climate change and energy'.*	Spain
Give students the possibility to make learning choices in the scientific activity	Social relevance of the topics assessed	*'Give us a choice of what to do'.*	UK
Student's perception of contributing to science through the activity *	Scientific relevance*/Perception of science	*'To do something useful for science, do things'.*	France
Transversal competences Indicator	**Criteria**	**Quotations from post-its and/or discussion**	**Case studies**
Student's ability to elaborate and share ideas verbally and written during the activity	Communication skills	*'Other people share the same interest as you and you can talk to them and you can learn stuff'.*	UK

Cognitive aspects Indicator	Criteria	Quotations from post-its and/or discussion	Case studies
Student's acquisition of basic knowledge about scientific topic(s) addressed in the activity	Cognitive aspects	*'Play educative games to learn something'.* *'Learn about electricity, plants, underwater staff and other topics'.*	France UK
Experiential aspects **Indicator**	**Criteria**	**Quotations from post-its and/or discussion**	**Case studies**
Student's amusement during the activity	Enjoyment	*'Have fun when learning science'.* *'Laughing'.* *'Shocking – weird – funny - wow!'*	Spain France UK
Student discovery of something not previously known *	Enjoyment	*'Be able to discover something new that we didn't know before, something interesting for us'.* *'We like discovering, to learn new things'.* *'Find new things, new facts, more interesting'.*	Spain France UK
Inclusion of physical activity in the activity *	Body and spatial awareness	*'We must have physically active science classes, we do not want to sit and observe all day'.* *'It has to be practical to get us moving so we don't get bored'.*	Spain UK
Inclusion of activities outdoors and/or outside the school *	Body and spatial awareness	*'To do activities outside, like visiting science museums and labs, going to science festivals'.* *'To go out of school to visit exhibitions'.*	Spain France

At the moment, these results are being returned to the participant students in each of the secondary schools, in order to get their feedback and highlight the relevance of their contribution to the assessment development (see Image 2).

Image 2 A moment during the return of results in the IES Consell de Cent secondary school in Barcelona.

8. GAPS AND CHALLENGES IN SCIENCE EDUCATION ASSESSMENT TO ADDRESS WITHIN PERFORM RESEARCH

Based on the findings from both literature review and exploratory workshops, we identify different methodological limitations and weaknesses related to the analysed assessment frameworks and the criteria and indicators explored. Altogether they provide insights about gaps and challenges in the current practice of science education assessment with relevant implications for PERFORM's methodological development.

8.1 Weaknesses of reviewed assessment frameworks

Methodological limitations in the reviewed assessment frameworks are mostly related to the challenging nature of both science learning and engagement, and to the specific design and implementation of the assessment tools in each setting.

The multidimensionality and complexity of science learning and engagement represent an important challenge for monitoring and assessment, as they complicate the provision of accurate explanations and conclusions about the assessed factors and their interrelations –e.g., the direction of effects, causal inference, variance, etc. This challenge is stressed in some of the reviewed assessments through the use of narrow characterizations of learning, the lack of control groups or the interaction with other ordinary scientific activities, which make difficult to disentangle effects from other co-founding variables. The identified lack of qualitative data in some of the assessments or, eventually, of an in-depth exploration of these data, hinders the examination and understanding of such complexity, representing a methodological gap in some studies.

The design of the assessment tools and processes often involves another set of methodological limitations. Identified weaknesses in this regard are mostly related to the generalisation of results by using small and/or not randomly-selected samples in inferential analyses, as well as the self-selection of participants, which might condition their performance in the activity, and the existence of ceiling effects, most commonly associated to questionnaires. Furthermore, the use of partial assessments leading to incomplete evidence –e.g., assessments that do not cover the whole activity, do not include relevant involved actors as data sources, or use only cross-sectional data-, undermines as well the contribution of these studies to the field of science education assessment.

8.2. Limitations of identified assessment criteria and indicators in the literature review

The reviewed assessment criteria and indicators show that RRI values and process requirements are still a gap in current science education assessments. More specifically, and as seen in section 6, while indicators addressing the cognitive aspects of learning are widely implemented, indicators related to gender and ethical aspects are mostly absent. This is a relevant insight in the context of RRI, which emphasises a diversity of learning outcomes and process dimensions that seem to be currently omitted in science education assessment when working with criteria and indicators, beyond its cognitive dimensions.

In this regard, we can distinguish **different levels of assessment**, **in terms of aspects addressed and depth of the assessment**. **A first basic level** addresses **basic cognitive and experiential aspects**, through the assessment of knowledge gains and general attitudes and perceptions towards the scientific topics approached. It addresses learning mostly in terms of the acquisition of new knowledge and how it might change students' perceptions and attitudes towards such issues. This level is covered by most of the experiences and can be approached through quantitative data collection tools, such as close-ended questionnaires. **A second level** expands learning outcomes to transversal **skills and competences acquired** as a result of the educational activity. This can be approached also through performance-based assessments mainly focused on students' inquiry and communication skills. **A third level** of assessment further addresses the **metacognitive, normative and emotional dimensions of learning**. Such an assessment implies a leap forward to include critical thinking and reflexivity, values and emotions and has been rarely found in our review. Indeed, approaching these learning aspects requires mix-methods approaches, in-depth analyses and, sometimes, broad time frameworks which make this level of assessment challenging. Consequently, this type of assessment is limited not only by methodological challenges but also by the nature of the educational activities developed (i.e. goals, processes fostered, time frameworks). Addressing these different levels or scopes of assessment is a challenging but necessary task in the approach to complex learning experiences such as those to be developed in PERFORM.

Moreover, findings from the exploratory workshops conducted with students in selected schools have allowed us to grasp some missing aspects in and/or specific formulations of the expert-based indicators and to identify assessment aspects that are relevant from their perspectives. Among them, experiential aspects, in terms of their engagement and their emotional experience during the activity are the most highlighted. Interestingly, ethical aspects of the RRI approach are also perceived as important by the students, such as having the possibility of connecting STEM topics with their daily life activities and participating in learning activities of scientific and social relevance.

9. IMPLICATIONS FOR PERFORM ASSESSMENT

Through the systematic literature review and explorative workshops, we have detected relevant gaps and challenges to be addressed when designing and implementing PERFORM's assessment. We have also identified a set of expert-based and participatory indicators in the field of STEM education to be considered when evaluating the effectiveness of PERFORM science education methods based on performing arts in fostering secondary school students' learning and engagement in STEM. Furthermore, by using a participatory approach, we have been able to address the specificity of each learning environment and to identify relevant indicators to assess students' learning and engagement in each setting in the context of the PERFORM project. In what follows, we describe the methodological implications of these findings for PERFORM assessment.

9.1. Assessment frameworks

A first insight of the current review is the reinforcement of **assessment transdisciplinarity as a key requirement for PERFORM** to address the complexity and multidimensionality of both science learning and engagement, including RRI values.

The development of a transdisciplinary assessment framework in PERFORM, mostly based on educational psychology, science communication, sociology and performance-based approaches, will not only ground theoretical approaches to the analysis of students' learning and engagement, but also the development of inclusive and innovative data collection tools. The integration of ICT-tools and arts-based methods reflected in the reviewed assessment frameworks suggest a promising contribution of these fields to methodological transdisciplinarity. Science education assessments using both arts and new technologies seem to encompass the development of more student-centred and collaborative pedagogies (Rooney-Varga 2014, Fitzgerald 2013). Indeed, these methods generally emphasise hands-on evaluation approaches in which the students actively employ their creativity, knowledge and competences in the assessment, individually or in groups. Due to such applied and creative nature, these methods are generally seen as less-intrusive assessment approaches and formative in nature (Gold 2015, McGregor 2014, Varelas 2010, Braund 1999), and thus are of interest for PERFORM to assess the participatory educational process in which students will be involved.

A second implication is that **assessment in PERFORM is conceived both as assessment *for* and assessment *as* learning** (Corigan et al. 2013).

The assessment *for* learning approach implies that assessment is understood as a process carefully integrated into the science educational activities –the PERSEIAs, as a reflective and self-reflexive dimension inherent to the students' learning process that can contribute to such learning as it happens. We reject, therefore, the perspective of the assessment as an external activity of summative nature, independent of the educational process developed. The adoption of formative approaches is further relevant in the context of performance-based science education, since assessments in this field are commonly summative and at the end of the process (Odegaard 2003).

The assessment *as* learning approach implies that monitoring and evaluation are also understood as self-reflexive, iterative research processes aiming to contribute to the improvement of the PERSEIAs proposed. Such contributions will take place both during the PERSEIAS development and their pilot testing. This will be possible thanks to the collection of data at different stages of the participatory educational process (prior to, during and after); but also, to the inclusion of self-diagnostic assessment methods that will allow students and researchers to collectively reflect on PERSEIAs learning outcomes and processes. Such iterative process will also allow us to critically reflect on the methodological implementation of the assessment and to adapt it and improve it if needed according to gathered evidence and feedback received. Furthermore, the development of PERSEIAs in selected schools in two different stages along two consecutive school years will facilitate such approach. Through the assessment of the participatory educational processes carried out during the first year in selected schools, we will identify and address potential shortcomings in the methodology that will be reviewed and improved for the second year.

A third methodological implication derived from the review findings is the **opportunity for broadening and enriching data collection sources by including more actors beyond students in the PERFORM assessment**. PERFORM aims to enhance the robustness of the assessment by integrating different perspectives and sensibilities through the inclusion of the participating secondary school students, but also of their teachers, the facilitators of PERSEIAs and early career researchers contributing to the generation of PERSEIAs. The involvement of these actors in the PERFORM assessment will be key not only to enrich the collection of data but also to ensure that it can be formatively integrated into the learning process. The guiding role of teachers and early career researchers during the process will provide a constant conceptual support and assessment of contents. Furthermore, their involvement in the performance-based participatory process will be a key element to address and assess the public engagement required for RRI.

This is further connected to the need of enhancing the active participation of students and other actors in PERFORM assessments. In this regard, and as seen in the review, participatory assessment approaches are still marginal in science education experiences with young people. Derived of such finding, and as part of **PERFORM's commitment** with RRI processes and participation, the forth and last implication for PERFORM assessment consists on paying special attention to the **inclusion of the students in the whole assessment process, from design to analysis** (see Figure 5). Although each case study will tailor their assessment strategy to the specific implementation context, the objective will be to reach the highest participation possible.

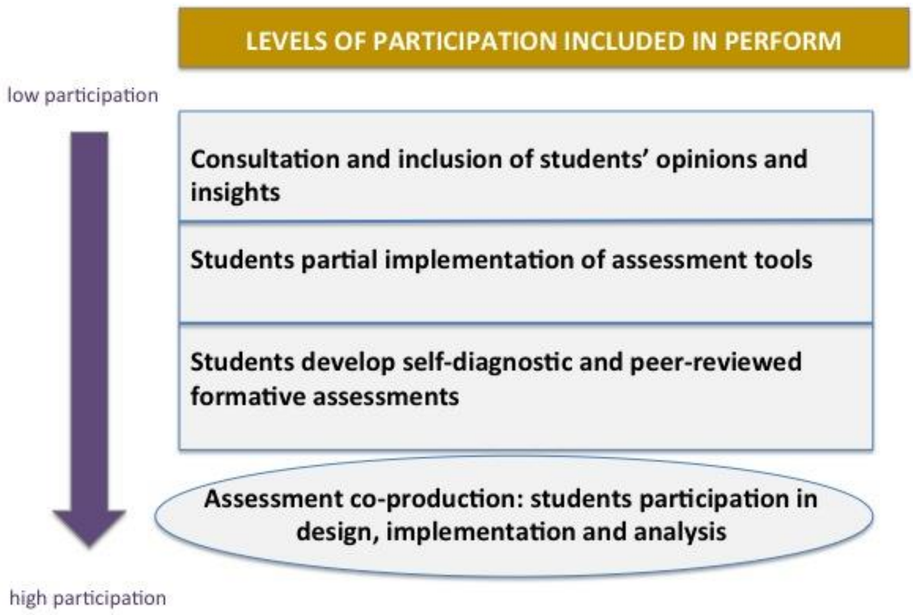
At this stage of the project (Month 7), participant students have been already included in the assessment design, through the explorative workshops and the identification and validation of criteria and indicators relevant to them. This is a rather **basic level of participation**, represented by the implementation of methods to gather participants' opinions and insights about topics of their own interest to be included in the assessment design; such as exploratory workshops or focused discussions. **A second level of participation** is represented by students' implementation of assessment tools designed by others. This is the case for instance of peer-to-peer interviews in which students interview their mates, potentially gathering more honest responses from them, but without further participation in the analysis. **A third level of participation** is represented by the incorporation of self-reflection and peer-review tools in the PERSEIAS to contribute both to the assessment and to students' learning process. When incorporating self-reflective practice, participants have the opportunity to assess their own learning evolution and guide their improvements, through self-diagnostic tools, such as the TWLH[1] and PMI[2] charts,

---

[1] The acronym corresponds to the following dimensions of the chart: what you **t**hink you know, what you **w**ant to know, what you have **l**earned and **h**ow you have learned it.

diagrammatical representations and class discussions (see Fitzgerald 2013). Similarly, the incorporation of peer-review assessments (e.g. assessing other classmates oral presentations, videos, etc.) allows participants to assess other mates' learning or outputs, under their own terms and criteria. The active participation of students in the assessment has the potential to engage them more deeply in the learning process, and empower them by fostering responsibility for their learning and developing learning skills, such as reflexivity about learning and learning autonomy (Kollar and Fischer 2010). Obviously, each level can be more or less participatory depending on the methods applied and the degree of responsibility and decision-making given to the students.

Figure 5 A ladder of assessment participation: identified levels of participation in the reviewed assessments



Finally, **a further level of participation** would result from the combination of all the participation levels and methods mentioned, involving the students in the whole assessment process. Such an assessment co-production would engage students from design

---

[2] The acronym corresponds to the following dimensions of the chart: positives, minuses and interesting things of their own learning performance.

(e.g. intervention in choosing evaluation focus, criteria and/or questions) to implementation and analysis (either direct or through the return and validation of results). **PERFORM project expects to engage students along all these levels of participation,** combining participatory assessment methods so as to engage students in the whole assessment process to the greatest extent possible in each context.

9.2. Assessment criteria and indicators

In terms of criteria and indicators, the **generalised neglect of gender and ethical aspects in the assessment frameworks reviewed is the most relevant insight from the literature review and deserves special attention in the context of PERFORM**.

Both gender and ethical issues are core aspects of the RRI approach. Gender issues are specifically relevant in scientific practice and science education since girls' and boys' attitudes and perceptions are differently influenced by different reference models and negative stereotypes. Interestingly, the limited role of women in science is among the negative stereotypes that prevail across gender in many developed countries (Ruiz-Mallén and Escalas 2012). On the other hand, a focus on the understanding of the nature of science, including the normative aspects, power relations and tensions that intervene and coexist in the construction and sharing of scientific knowledge is crucial to foster critical scientific literacy, active citizenship and participation in science (Yoon 2014, Priest 2013, Klosterman 2010).

If gender and ethical aspects are core issues in science education, then finding ways of critically integrating these aspects in the assessment seems crucial to approach scientific learning and engagement in our methodological development. Sex and gender analysis (Schiebinger and Klinge 2010) will be included in PERFORM's assessment, along with other methodologies, so as to provide further critical rigour in the approach to learning outcomes and process requirements. This emphasis will permeate all stages of research, from design to implementation and analysis. Similarly, indicators assessing students' awareness of science contradictions and uncertainty and the existence of power relations in science, among others, will be carefully included to assess ethical aspects in STEM education.

Based on the participatory approach of the PERFORM assessment design through the explorative workshops, **PERFORM addresses targeted students' views and opinions about science learning and engagement in the evaluation process by including a set**

of indicators not previously identified in the literature review. Participatory indicators relevant for the involved students include experiencing surprise, discovering something new and doing some physical activity when participating in a STEM education activity, as well as the inclusion of art-related methods, doing activities outside the school, and the contextualisation of STEM topics within students' daily life. Furthermore, among these indicators, PERFORM's assessment will emphasise those perceived as relevant by the students but neglected in the literature review, such as indicators corresponding to the ethical aspects of RRI 'Connecting scientific topics with values' and the 'Social relevance of topics addressed'.

At this initial stage of the project, the different implications mentioned above determine an initial prioritisation of the assessment criteria and indicators to be used in the PERFORM assessment from the extensive list proposed in section 5. Table 10 contains identified indicators and their correspondent criteria we suggest to assess PERFORM's participatory educational processes. These indicators consider both the conceptual framework of PERFORM as well as the potentials and limitations of the implementation contexts (i.e., secondary schools in each case study) and time frameworks (i.e. implementation along two different school years).

The identification of RRI assessment criteria and indicators specifically related to science education and engagement with young people moves forward the state of the art in the development of assessment frameworks within the context of RRI and science education as it addresses a key element already highlighted in previous reports on the topic (Strand et al. 2015): the necessary identification of assessment indicators tailored to each policy agenda and RRI context. Furthermore, such a focus on science education and engagement, together with the inclusion of participatory methods, contributes to the development of indicators tailored to specific contexts and stakeholder needs, enhancing their applicability, stakeholder support and usefulness. In this regard, through the identification of RRI-related science education indicators, this research report provides an open and dynamic assessment framework addressing the different RRI learning outcomes and process requirements. We hope that this research effort can guide further methodological developments and implementations in the context of science education and engagement.

Table 10 Summary of indicators and criteria identified in the literature review by learning outcomes and process requirements

Responsible Research and Innovation criteria: BP=Balanced participation; FD=Fostering dialogue among participants; GE=Gender equality in participation; AC=Approaching critically gender issues; EM=Emotional engagement; CE=Cognitive engagement; QR=Questioning and reframing; ST=Systems thinking; CT=Connecting topics with experience; SO=Seeking other points of view; NS=Understanding of the nature of science; SR=Social relevance of topics addressed; PA= Participants acceptance of process/outcomes; CV=Connecting scientific topics with values. *Transversal competences criteria*: VL=Understanding the value of learning; LA=Learning autonomy; RT=Reflective thinking; CM=Communication skills; CL=Collaborative skills; RS=Respect for society and environment; IR: Informed and reasoned decision-making; AR=Ability to resolve conflicts; EN=Entrepreneurship; SC=Self-confidence and esteem; AP=Ability to plan and manage projects; *Cognitive criteria:* CA=Cognitive aspects; *Experiential aspects criteria*: EJ=Enjoyment; EA=Emotional awareness and reflexivity; BA=Body and spatial awareness; ET=Empowerment; PS=Perceptions of science and the scientific issues approached; AS=Attitudes towards science and the scientific issues approached.
* *Participatory indicators not identified in the literature review*

	Responsible Research & Innovation values														Transversal competences											Cognitive	Experiential aspects					
Learning outcome and/or process requirement	Inclusiveness		Gender		Engagement		Critical & Creative thinking				Ethical aspects				Learning to learn			Social and Civic competences					Sense of initiative				Feelings & Emotions				Perceptions & attitudes	
Criteria / **Indicators**	BP	FD	GE	AC	EM	CE	QR	ST	CT	SO	NS	SR	PA	CV	VL	LA	RT	CM	CL	RS	IR	AR	EN	SC	AP	CA	EJ	EA	BA	ET	PS	AS
Combination of learning pedagogies and resources	■			■																												
Inclusion of other participants	■	■																														
Student's sharing of tasks and roles	■																		■													
Type of dialectic interactions	■	■				■																										
Use of interactive ICT tools		■																														
Type of dialogue between students and teacher	■	■																														
Students' engagement in science by gender			■	■																												
Number of students by gender	■		■																													
Type of tasks and roles assumed by gender			■	■																												
Affective responses by gender				■																												
Performance by gender				■																						■						
Willingness to get involved and participate					■	■																										■
Feelings when experiencing the activity					■																							■				
Ability to use the body to express and communicate					■													■											■			

Indicators \ Criteria	BP	FD	GE	AC	EM	CE	QR	ST	CT	SO	NS	SR	PA	CV	VL	LA	RT	CM	CL	RS	IR	AR	EN	SC	AP	CA	EJ	EA	BA	ET	PS	AS
Further interaction and initiatives					■	■																										
*Experience surprise						■																					■				■	
*Use arts-related methods						■																					■					
Involvement in reasoning and argumentation						■											■				■											
Ability to develop ideas, higher order thinking						■	■									■																
Ability to ask questions, discuss, make conclusions						■	■									■					■											
Time spent in doing the task						■																										■
Willingness to continue working out of class						■										■																
Reframing/comprehension of scientific concepts						■	■																			■						
Ability to relate ideas in multiple contexts							■	■									■															
Ability to identify relations and interactions								■									■															
Use of learning techniques to discuss about the system								■																								
Contextualisation of topics within societal challenges									■			■																				
*Contextualisation of topics within daily life									■			■																				
Use of students' previous experiences, knowledge									■																							
Facilitation of learning through direct involvement									■																							
Ability to consider different perspectives								■		■																						
Awareness of science contradictions											■																				■	
Awareness of power relations in science											■																				■	
Possibility to make learning choices												■	■																			
Creation of own outcomes					■	■							■																			

Criteria / Indicators	BP	FD	GE	AC	EM	CE	QR	ST	CT	SO	NS	SR	PA	CV	VL	LA	RT	CM	CL	RS	IR	AR	EN	SC	AP	CA	EJ	EA	BA	ET	PS	AS
Inclusion of scientists' personal stories																																
Show contrasting perspectives about science																																
Reflection about ethical behaviour in research																																
Awareness of professional value of learning science																																
Satisfaction to be able to learn science																																
Awareness of science experiential value																																
Ability to use equipment, technology, tools																																
Ability to autonomously search for information																																
Reflection on one's learning																																
Ability to recognise and use relevant information																																
Assessment/reflection about peers' performance																																
Ability to elaborate/share ideas verbally and written																																
Ability to organise/make meaning from visual info.																																
Willingness to ask for help and to help others																																
Respect towards others' ideas																																
Willingness to assume a responsible position																																
Ability to contrast different evidence																																
Managing difficulties within the group																																
Belief in own ability to perform a scientific activity																																
Leadership and/or responsibility																																

Criteria / Indicators	BP	FD	GE	AC	EM	CE	QR	ST	CT	SO	NS	SR	PA	CV	VL	LA	RT	CM	CL	RS	IR	AR	EN	SC	AP	CA	EJ	EA	BA	ET	PS	AS
Performance self-assessment																							■		■							
Belief in own ability to do well in a scientific domain																								■						■		
Belief in own verbal ability to discuss																								■						■		
Acquisition of basic knowledge																										■						
Interest in science and learning science																											■				■	■
Excitement caused by science/learning science																											■					■
Amusement during the activity																											■	■			■	■
*Discover something unknown																											■					
Expression or embodiment of emotions																												■				
Ability to reflect through/ upon emotional responses																												■				
Use of the body to convey meanings																													■			
Awareness of influence of physical space																													■			
*Physical activity																													■			
*Activities outdoor/outside school																													■			
Design of the activity to address competences																														■		
Feeling recognized by other participants																								■						■		
Perceptions of the social value of science																															■	■
Perception of scientists, scientific careers/jobs																															■	
Perceptions of the topics approached																															■	
Perceptions of the way science is taught																															■	
Perceptions of the pedagogic approach																															■	

Criteria / Indicators	BP	FD	GE	AC	EM	CE	QR	ST	CT	SO	NS	SR	PA	CV	VL	LA	RT	CM	CL	RS	IR	AR	EN	SC	AP	CA	EJ	EA	BA	ET	PS	AS
Perceptions of the group (sense of belonging)																																
Curiosity and interest towards science																																
Identification with scientific skills/attributes																																
Interest in scientific careers and jobs																																
Attitudes towards the topics approached																																
Attitudes towards the pedagogic approach																																
Perception of contributing to science																																

10. REFERENCES

Baker, E., O'Neil, H. F., Jr.,&Linn, R. L. 1994. Policy and validity prospects for performance-based assessment. Journal for the Education of the Gifted, 17(4), 332–353.

Black, P. 2005. Learners, learning and assessment. Editors Patricia Murphy.

Braund, M. 1999. Electric drama to improve understanding in science. *School Science Review*, *81*, 35-42.

Eisenkraft, A. 2004. How do we know what they know. FASS Meeting

ERI-NET 2013. Transversal Competences in Education Policy and Practice. Regional Synthesis Report. UNESCO, Bangkok.

European Commission 2006. Recommendation 2006/962/EC of the European Parliament and of the Council of 18 December 2006 on key competences for lifelong learning, Official Journal L 394 of 30.12.2006. Available at: http://eurlex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2006:394:0010:0018:EN:PDF

European Comission 2012a. Responsible Research and Innovation: Europe's ability to respond to societal challenges. Available at: http://ec.europa.eu/research/science-society/document_library/pdf_06/responsible-research-and-innovation-leaflet_en.pdf

European Commission 2012b. Assessment of Key Competences. Literature review: Glossary and examples. Education and Training 2020 Work programme. Brusselles.

Fitzgerald, A., & Gunstone, R. 2013. Embedding Assessment Within Primary School Science: A Case Study. In *Valuing Assessment in Science Education: Pedagogy, Curriculum, Policy* (pp. 307-324). Springer: Netherlands.

Gold, A. U., Oonk, D. J., Smith, L., Boykoff, M. T., Osnes, B., & Sullivan, S. B. 2015. Lens on Climate Change: Making Climate Meaningful Through Student-Produced Videos. *Journal of Geography*, *114*(6), 235-246.

Hazelkorn, E., Ryan, C., Beernaert, Y., Constantinou, C. P., Deca, L., Grangeat, M., Karikorpi, M., Lazoudis, A., Pintó Casulleras, R. & Welzel-Breuer, M. (2015). Science Education for Responsible Citizenship. Report to the European Commission of the Expert Group on Science Education. European Commission: Luxembourg.

Hsieh, H. F., & Shannon, S. E. (2005). Three approaches to qualitative content analysis. *Qualitative health research*, 15(9), 1277-1288

Klassen, S. 2006. Contextual assessment in science education: Background, issues, and policy. *Science Education*, 90(5), 820-851.

Kollar, I. & Fischer, F. 2010. Peer assessment as collaborative learning: A cognitive perspective. *Learning and Instruction, 20,* 344-348. doi:10.1016/j.learninstruc.2009.08.005

McGregor, D. 2014. Chronicling innovative learning in primary classrooms: conceptualizing a theatrical pedagogy to successfully engage young children learning science. *Pedagogies: An International Journal*, *9*(3), 216-232.

Millar, R. 2013. Improving science education: Why assessment matters. In *Valuing assessment in science education: Pedagogy, curriculum, policy* (pp. 55-68). Springer Netherlands.

Minner, D. D., Levy, A. J., Century, J. 2010. Inquiry-based science instruction-what is it and does it matter? Results from a research synthesis years 1984 to 2002. *Journal of Research in Science Teaching*, 47(4), 474–496.

National Research Council (Editors). 1996. National Science Education Standards. Washington, DC: National Academy Press. Available at: http://www.nap.edu/read/4962/chapter/1

Odegaard, M. (2003). Dramatic science - a critical review of drama in science education, *Studies in Science Education*, (39), 75-102.

Owen et al. 2012. Responsible Research and innovation: From science in society to science for society, with society. *Science and Public Policy* 39: 751-760.

Pohl C. 2008 From science to policy through transdisciplinary research, Environmental science & policy, 11(1), 46-53.

Read, J. R. (2006). Learning Theories. Available at: http://acell.chem.usyd.edu.au/Learning-Theories.cfm

Rooney-Varga, J. N., Brisk, A. A., Adams, E., Shuldman, M., & Rath, K. 2014. Student media production to meet challenges in climate change science education. *Journal of Geoscience Education*, *62*(4), 598-608.

Ruiz-Mallén, I., Escalas, M.T. 2012. Scientists Seen by Children: A Case Study in Catalonia, Spain. *Science Communication* 34(4): 520-545.

Ruiz-Primo, M. A. 2011. Informal formative assessment: The role of instructional dialogues in assessing students' learning. *Studies in Educational Evaluation*, *37*(1), 15-24.

RRI-Tools 2015. RRI Tools: towards RRI in action. Available at: http://www.rri-tools.eu/documents/10184/104615/RRI+Tools+Policy+Brief+%28EN%29.pdf/82ffca72-df32-4f0b-955e-484c6514044c

Schiebinger, L., & Klinge, I. (Eds.) 2010. *Gendered Innovations: Mainstreaming Sex and Gender Analysis into Basic and Applied Research*. Brussels: European Commission.

Sipos, Y., Battisti, B., & Grimm, K. 2008. Achieving transformative sustainability learning: engaging head, hands and heart. *International Journal of Sustainability in Higher Education*, *9*(1), 68-86.

Strand, R. Spaapen, J., Bauer, M. W., Hogan, E., Revuelta, G., & Stagl, S. 2015. Indicators for promoting and monitoring responsible research and innovation: report from the expert group on policy indicators for responsible research and innovation. European Commission: Luxembourg

Varelas, M., Pappas, C. C., Tucker-Raymond, E., Kane, J., Hankes, J., Ortiz, I., & Keblawe-Shamah, N. 2010. Drama activities as ideational resources for primary-grade children in urban science classrooms. *Journal of Research in Science Teaching*, *47*(3), 302-325.

Webb, G. 1996. Becoming critical of action research for development, in O. Zuber-Skerritt (ed.) New Directions in Action Research, London: Falmer Press.

Wiliam, D. 2011. What is assessment for learning? *Studies in Educational Evaluation*, *37*(1), 3-14.

ANNEX 1: Literature review sample

Table A.1. 1 List of included articles and book chapters in the literature review. In grey: articles reviewed in the third screening through snowball sampling.

Paper ID	Bibliographic information			
	Title	**Journal Name**	**Year**	**1st Author**
1	Motivation, Learning, and Transformative Experience: A Study of Deep Engagement in Science	Science Education	2009	Pugh, Kevin
2	Science Engagement and Literacy: A retrospective analysis for students in Canada and Australia	Research in Science Education	2014	Woods-McConney, Amanda
3	Children's Motivation Toward Science Across Contexts, Manner of Interaction, and Topic	Science Education	2013	BATHGATE, Megan
4	Serious games analytics to measure implicit science learning	Serious Games Analytics (Book)	2015	Rowe, Elisabeth
5	Learning Biology Through Innovative Curricula: A Comparison of Game- and Nongame-Based Approaches	Science Education	2015	Sadler, T.D.
6	Impact of Project-Based Curriculum Materials on Student Learning in Science: Results of a Randomized Controlled Trial	Journal of Research in Science Teaching	2015	Harris, Christopher
7	The learning benefits of being willing and able to engage in scientific argumentation	International Journal of Science Education	2015	Bathgate, Megan
8	Lens on Climate Change: Making Climate Meaningful Through Student-Produced Videos	Journal of Geography	2015	Gold, Anne
9	Evaluation of public engagement activities to promote science in a zoo environment.	PloS one	2014	Whitehouse, Jamie
10	Using educational data mining to assess students' skills at designing and conducting experiments within a complex systems microworld	Thinking Skills and Creativity	2015	Gobert, Janice D.
11	Cognitive diagnostic like approaches using neural-network analysis of serious educational videogames	Computers & Education	2014	Lamb, Richard
12	Bioinformatics projects supporting life-sciences learning in high schools.	PLoS computational biology	2014	Marques, Isabel
13	Chronicling innovative learning in primary classrooms: Conceptualizing a theatrical pedagogy to successfully engage young children learning science	Pedagogies: An International Journal	2014	McGregor, Debra
14	Evaluating children's conservation biology learning at the zoo	Conservation Biology	2014	Jensen, Eric

15	Embedding assessment within primary school science: A case study	Valuing Assessment in Science Education: Pedagogy, Curriculum, Policy (Book)	2013	Fitzgerald, Angela
16	Learning progressions as a guide for developing meaningful science learning: A new framework for old ideas	Educación Química	2013	Stevens, Shawn
17	Promoting Students' Interest and Motivation Towards Science Learning: The Role of Personal Needs and Motivation Orientations	Research in Science Education	2013	Loukomies, Anni
18	Science Teaching and Learning Activities and Students' Engagement in Science	International Journal of Science Education	2013	Hampden-Thompson, Gillian
19	Impacts of a STSE high school biology course on the scientific literacy of Hong Kong students	Asia-Pacific Forum on Science Learning and Teaching	2013	Lau, Kwok-chi
20	Exploring Newtonian mechanics in a conceptually-integrated digital game: Comparison of learning and affective outcomes for students in Taiwan and the United States	Computers and Education	2011	Clark, Douglas B.
21	Fat dogs and coughing horses: K-12 programming for veterinary workforce development	Journal of veterinary medical education	2013	San Miguel, Sandra
22	Collaborative Action Research on Technology Integration for Science Learning	Journal of Science Education and Technology	2012	Wang, Chien-hsing
23	Exploring middle school students' conceptions of the relationship between genetic inheritance and cell division	Science Education	2012	Williams, Michelle
24	Korean Students' Perceptions of Scientific Practices and Understanding of Nature of Science	International Journal of Science Education	2014	Yoon, Sae Yeol
25	Student Tools Supported by Collaboratively Authored Tasks: The Case of Work Learning Unit	Journal of Interactive Learning Research	2006	Akpinar, Yavuz; Bal, Volkan
26	Conceptual continuity and the science of baseball: using informal science literacy to promote students' science learning	Cultural Studies of Science Education	2009	Brown, Bryan A; Kloser, Matt
27	Designing collaborative knowledge building environments accessible to all learners: Impacts and design challenges	Computers and Education	2008	Hyo-Jeong, So
28	Scientific Caricatures in the Earth Science Classroom: An Alternative Assessment for Meaningful Science Learning	Science and Education	2008	M.Clary, Renee
29	Assessing Dialogic Argumentation in Online Environments to Relate Structure, Grounds, and Conceptual Quality	InterScience	2007	Clark, Douglas B

30	Contextualising Instruction: Leveraging Students' Prior Knowledge and Experiences to Foster Understanding of Middle School Science	Journal of research in Science Teaching	2008	Rivet, Ann E.
31	Ninth-Grade Student Engagement in Teacher-Centered and Student-Centered Technology-Enhanced Learning Environments	InterScience	2007	Wu, Hsin-Kai
32	An Evaluation of a Nutrition WebQuest: The Malaysian Experience	Eurasia Journal of Mathematics, Science and Technology Education	2008	Wui, Lee Sheh
33	Factors that influence pupil engagement with science simulations: the role of distraction, vividness, logic, instruction and prior knowledge	Chemistry Education and Research Practice	2007	Rodrigues, Susan
34	Design and Reflection Help Students Develop Scientific Abilities: Learning in Introductory Physics Laboratories	Journal of Learning Sciences	2010	Etkina, Eugenia
35	Preferred - Actual learning environment "spaces" and earth science outcomes in Taiwan	Science Education	2006	Chang, Chun-Yen
36	Teachers' collaborative task authoring to help students learn a science unit",2006,"Educational Technology and Society	Educational Technology and Society	2006	Akpinar, Yavuz
37	Science learning through Scouting: an understudied context for informal science education	International Journal of Science Education	2005	Jarman, Ruth
38	A trial of the Five Es: A referent model for constructivist teaching and learning	Research in Science Education	2003	Boddy, Naomi
39	Mapping students' thinking patterns by the use of the knowledge space theory	International Journal of Science Education	1997	Taagepera, Mare
40	Partners in reform: "What's culture got to do with it?"	Urban Anthropology	1997	Kozaitis, Kathryn A.
41	Infusing creativity into Eastern classrooms: Evaluations from student perspectives	Thinking Skills and Creativity	2011	Cheng, Vivian M.Y.
42	Produsage in a/synchronous learner-led e-learning	New Review of Hypermedia and Multimedia	2011	Kazmer, Michele M.
43	Developing a geographic visualization tool to support earth science learning	Cartography and Geographic Information Science	2000	Harrower, Mark
44	The development of an open-ended drawing tool: an alternative diagnostic tool for assessing students' understanding of the particulate nature of matter	Chemistry Education Research and Practice	2011	Nyachwaya, James M.

45	How do clinical clerkship students experience simulator-based teaching? A qualitative analysis	Simulation in Healthcare	2006	Takayesu, James K.
46	An evaluation of multimodal interactions with technology while learning science concepts	British Journal of Educational Technology	2011	Anastopoulou, Stamatina
47	Students in the Director's Seat: Teaching and Learning with Student-generated Video	World Conference on Educational Multimedia, Hypermedia and Telecommunications (Conference Proceeding)	2005	Kearney, Matthew
48	Evidence of Public Engagement with Science: Visitor Learning at a Zoo-Housed Primate Research Centre	PloS one	2012	Waller, Bridget M.
49	The long-term impact of interactive exhibits	International Journal of Science Education	1991	Stevenson, John
50	A learner's tactic: How secondary students' anthropomorphic language may support learning of abstract science concepts	Electronic Journal of Science Education	2011	Dorion, Kirk
51	Dramatising Science Learning: Findings from a pilot study to re-invigorate elementary science pedagogy for five- to seven-year olds	International Journal of Science Education	2012	McGregor, Debra
52	Drama activities as ideational resources for primary-grade children in urban science classrooms	Journal of Research in Science Teaching	2010	Varelas, Maria
53	Science through Drama: A multiple case exploration of the characteristics of drama activities used in secondary science lessons	International Journal of Science Education	2009	Dorion, Kirk
54	Investigating the impact of video games on high school students' engagement and learning about genetics	Computers & Education	2009	Annetta, Leonard A.
55	Multi-level assessment of scientific content knowledge gains associated with socioscientific issues-based instruction	International Journal of Science Education	2010	Klosterman, Michelle L.
56	Student Media Production to Meet Challenges in Climate Change Science Education	Journal of Geoscience Education	2014	Rooney-Varga, Juliette N.
57	Podcasts in Support of Experiential Field Learning	Journal of Geography in Higher Education	2010	Jarvis, Claire
58	Information retention from PowerPoint and traditional lectures	Computers and Education	2009	Savoy, April

59	Framing Interactions to Foster Generative Learning: a Situative Explanation of Transfer in a Community of Learners Classroom	The Journal of the Learning Sciences	2006	Engle, Randi A.
60	What happens when students do simulation-role-play in science?	Research in Science Education	1997	Aubusson, Peter
61	Role play as analogical modelling in science	Metaphor and analogy in science education (Book)	2006	Aubusson, Peter
62	Teaching Science Through Drama: An Empirical Investigation	Research in Science & Technology Education	1984	Metcalfe, Robert
63	Puppets promoting engagement and talk in science	International Journal of Science Education	2008	Simon, Shirley
64	Establishing basic ecological understanding in younger pupils: a pilot evaluation of a strategy based on drama/role play	International Journal of Science Education	1998	BAILEY, S
65	Can Untraditional Learning Methods Used in Physics Help Girls to be More Interested and Achieve more in this Subject?	Research in Science Education in Europe	1999	Tveita, Johaness
66	Electric drama to improve understanding in science	School Science Review	1999	Braund, Martin
67	Using dramatizations to present science concepts. Activating Students' Knowledge and Interest in Science	Journal of College Science Teaching	2000	Palmer, David H.

Table A1.2. List of excluded articles and book chapters and reason for exclusion

EXCLUDED ARTICLES		
Article Reference		Reason/s
1	Onan, A. 2015 A fuzzy-rough nearest neighbor classifier combined with consistency-based subset evaluation and instance selection for automated diagnosis of breast cancer. Expert Systems with Applications, 42(20):6844-6852	Out of focus and context: Evaluation of a breast-cancer diagnosis model
2	Giesbrecht T.a , Schenk B. , Schwabe G 2015 Empowering front office employees with counseling affordances. Transforming Government: People, Process and Policy, 9(4):517-544	Out of focus and context: Face-to-face citizen service encounter in public administrations
3	Forbes C.T. , Sabel J.L. , Biggers M. 2015 Elementary teachers' use of formative assessment to support students' learning about interactions between the hydrosphere and geosphere. Journal of Geoscience Education, 63(3):210-221	Out of focus: Elementary teachers' use of formative assessment to support students' learning about interactions between the hydrosphere and geosphere
4	Sabel J.L. , Forbes C.T., Zangori L. 2015 Promoting prospective elementary teachers' learning to use formative assessment for life science Instruction. Journal of Science Teacher EducatioN, 26(4:)419-445	Out of focus: pre-service teachers' content knowledge and ability to engage in formative assessment practices for science.
5	Forbes C.a , Sabel J.b , Zangori L. 2015 Integrating life science content & instructional methods in elementary teacher education. American Biology Teacher,77(9):651-657	Out of focus: Elementary teacher education
6	Campos-Sánchez A.a b , López-Núñez J.A.b , Carriel V.a , Martín-Piedra M.-A.a , Sola T.b , Alaminos M. 2014 Motivational component profiles in university students learning histology: A comparative study between genders and different health science curricula. BMC Medical Education, 14(1): 46	Out of focus: Motivational component profiles in university students learning histology
7	Hartley S., Millar K.M. 2014 The challenges of consulting the public on science policy: Examining the development of European risk assessment policy for genetically modified animals. Review of Policy Research, 31 (6):481-502	Out of focus: public engagement in science policy making
8	Csaki C.a , Fitzgerald C.b , O'Raghallaigh P.c , Adam F. 2014 Towards the institutionalisation of parliamentary technology assessment: The case for Ireland. Transforming Government: People, Process and Policy, 8(3): 315-334	Out of focus: implementation of a formal parliamentary technology assessment (PTA) capability
9	Köksal M.S.a , Sormunen K. 2014 Advanced science students' understanding on nature of science in Turkey. Asia-Pacific Forum on Science Learning and Teaching, 15(1), Article 9	Out of focus: Study of students' understanding of science
10	Sarkar, M., & Corrigan, D. 2014. BANGLADESHI SCIENCE TEACHERS'PERSPECTIVES OF SCIENTIFIC LITERACY AND TEACHING PRACTICES. International Journal of Science and Mathematics Education, 12(5), 1117-1141.	Out of focus: Study of teachers' perceptions
11	Tan, A. L., & Leong, W. F. (2014). Mapping Curriculum Innovation in STEM Schools to Assessment Requirements: Tensions and Dilemmas. Theory Into Practice, 53(1), 11-17.	Out of focus: Focused on curriculum innovation

12	Täht, K., Must, O., Peets, K., & Kattel, R. (2014). Learning motivation from a cross-cultural perspective: a moving target?. Educational Research and Evaluation, 20(4), 255-274.	Out of focus: Focused on PISA
13	Pride, L. D. (2014). Using learning stories to capture “Gifted” and “Hard Worker” mindsets within a NYC specialized high school for the sciences. Theory into Practice, 53(1), 41-47.	Analysis of narratives around learning in STEM specialised schools
14	Sporea, A., & Sporea, D. (2014). Romanian teachers perception on inquiry-based teaching. Romanian Reports in Physics, 66(4), 1253-1268.	Out of focus: Analysis of teacher's perceptions at kindergarden level
15	Buldu, N., Buldu, M., & Buldu, M. (2014). A Quality Snapshot of Science Teaching in Turkish K-3rd Grade Programs. Egitim ve Bilim, 39(174).	Out of focus: Focused on science teaching
16	Hsieh, T. C., Lee, M. C., & Su, C. Y. (2013). Designing and implementing a personalized remedial learning system for enhancing the programming learning. Educational Technology & Society, 16(4), 32-46.	Out of focus: Focus too specific on a learning system for programming
17	Čagran, B., & Grmek, M. I. (2013). Critical Self-Evaluation: An Attribute of Systemic Behavior: Authors of Natural Science Learning Materials as Evaluators. Systemic Practice and Action Research, 26(6), 537-547.	Out of focus: Focused on self-measurement of authors of science learning materials
18	Fleer, M., & Quiñones, G. (2013). An assessment perezhivanie: building an assessment pedagogy for, with and of early childhood science learning. In Valuing assessment in science education: Pedagogy, curriculum, policy (pp. 231-247). Springer Netherlands.	Review article
19	Fensham, P. J. (2013). International assessments of science learning: Their positive and negative contributions to science education. In Valuing assessment in science education: Pedagogy, curriculum, policy (pp. 11-31). Springer Netherlands.	Review chapter
20	Askew, M. (2013). Issues in Teaching for and Assessment of Creativity in Mathematics and Science. In Valuing assessment in science education: Pedagogy, curriculum, policy (pp. 169-182). Springer Netherlands.	Review chapter
21	Fensham, P. J., & Rennie, L. J. (2013). Towards an authentically assessed science curriculum. In Valuing assessment in science education: Pedagogy, curriculum, policy (pp. 69-100). Springer Netherlands.	Review chapter
22	Taylor, M. (2013). (Re) presenting disaster vulnerability in New Zealand school geography. New Zealand Geographer, 69(2), 158-166.	Out of focus and target group: Focus on teaching approaches to vulnerability
23	Maida, C. A. (2012). Fundamentals: Building Communities of Practice in Comparative Effectiveness Research. In Comparative Effectiveness and Efficacy Research and Analysis for Practice (CEERAP) (pp. 3-21). Springer Berlin Heidelberg.	Out of focus: Communities of practice in research and learning in health care
24	Elson, S. L., Hiatt, R. A., Anton-Culver, H., Howell, L. P., Naeim, A., Parker, B. A., ... & Hajopoulos, K. (2013). The Athena Breast Health Network: developing a rapid learning system in breast cancer prevention, screening, treatment, and care. Breast cancer research and treatment, 140(2), 417-425.	Out of focus: learning networks in research and communication of breast cancer

25	Murphy, C., Lundy, L., Emerson, L., & Kerr, K. (2013). Children's perceptions of primary science assessment in England and Wales. British Educational Research Journal, 39(3), 585-606.	Out of focus: perceptions of assessment
26	Nashon, S. M., & Anderson, D. (2013). Interpreting student views of learning experiences in a contextualized science discourse in Kenya. Journal of Research in Science Teaching, 50(4), 381-407.	Out of focus: students' perceptions of Kenya's learning system
27	LEE, M. H., LIN, T. J., & TSAI, C. C. (2013). Proving or improving science learning? Understanding high school students' conceptions of science assessment in Taiwan. Science Education, 97(2), 244-270.	Out of focus: perceptions of assessment
28	Annetta, L. A., Frazier, W. M., Folta, E., Holmes, S., Lamb, R., & Cheng, M. T. (2013). Science teacher efficacy and extrinsic factors toward professional development using video games in a design-based research model: The next generation of STEM learning. Journal of Science Education and Technology, 22(1), 47-61.	Out of focus and target group: focused on teachers
29	Milutinović, M., Labus, A., Stojiljković, V., Bogdanović, Z., & Despotović-Zrakić, M. (2015). Designing a mobile language learning system based on lightweight learning objects. Multimedia Tools and Applications, 74(3), 903-935.	Out of focus: Focused on language learning and mobile apps
30	Buxton, C. A., Allexsaht-Snider, M., Suriel, R., Kayumova, S., Choi, Y. J., Bouton, B., & Baker, M. (2013). Using educative assessments to support science teaching for middle school English-language learners. Journal of Science Teacher Education, 24(2), 347-366.	Out of focus and target group: Focused on teachers
31	Bell, P., Tzou, C., Bricker, L., & Baines, A. D. (2013). Learning in diversities of structures of social practice: Accounting for how, why and where people learn science. Human Development, 55(5-6), 269-284.	Review, Focus on how learning occurs according to factors of difference among people
32	Lay, Y.F., Khoo, C.H. (2012) Relationships between actual and preferred Science learning environment at tertiary level and attitudes towards science among pre-service Science teachers, Pertanika Journal of Social Science and Humanities, 20(4): 1117-1142	Out of focus: Focused in pre-service Science teachers attitudes towards science
33	Park, S., & Chen, Y. C. (2012). Mapping out the integration of the components of pedagogical content knowledge (PCK): Examples from high school biology classrooms. Journal of Research in Science Teaching, 49(7), 922-941.	Out of focus: Focused on teachers and pedagogical approach (not learning or engagement)
34	Nelson, M. M., & Davis, E. A. (2012). Preservice Elementary Teachers' Evaluations of Elementary Students' Scientific Models: An aspect of pedagogical content knowledge for scientific modeling. International Journal of Science Education, 34(12), 1931-1959.	Out of focus and target group: Focused on pre-service teachers
35	Pinto, M. (2012, May). Information literacy perceptions and behaviour among history students. In J. Broady-Preston, & L. Tedd (Eds.), Aslib Proceedings (Vol. 64, No. 3, pp. 304-327). Emerald Group Publishing Limited.	Focused on Spanish history students' subjective perception of their information literacy

36	Lin, T.-C., Hsu, Y.-S., Lin, S.-S., Changlai, M.-L., Yang, K.-Y., Lai, T.-L., (2012) A review of empirical evidence on scaffolding for science educatio, International Journal of Science and Mathematics Education, 10 (2):437-455	Review of papers, focused on scaffolding
37	Tsai, M. J., Hsu, C. Y., & Tsai, C. C. (2012). Investigation of high school students’ online science information searching performance: the role of implicit and explicit strategies. Journal of Science Education and Technology, 21(2), 246-254.	Out of focus: examination of students' online searching strategies
38	Smith-Jackson, T., Evia, C., Tabor, L., & Benson, K. (2012). Design of an inclusive science learning system for Appalachian children. Theoretical Issues in Ergonomics Science, 13(1), 18-32.	Out of focus and target group: Focused on teachers and parents; identifying requirements for the design
39	Van Est, R. (2011). The broad challenge of public engagement in science. Science and engineering ethics, 17(4), 639-648.	Out of focus: Conceptual paper on the integration of stakeholders
40	BULUNUZ, M. (2014). The Role of Playful Science in Developing Positive Attitudes toward Teaching Science in a Science Teacher Preparation Program. Eurasian Journal of Educational Research, (58).	Out of focus and target group: pre-service teachers' attitudes toward teaching science through play.
41	Jensen, E. (2015). Highlighting the value of impact evaluation: enhancing informal science learning and public engagement theory and practice. JCOM: Journal of Science Communication, 14(3).	Review article; Answer to another paper
42	Fan, L., (2014) Methods for improving the professional level of students majoring in information and computer science. World Transactions on Engineering and Technology Education, 12: 122-126	Out of focus: focused on teachers
43	Tekkumru-Kisa, M., Stein, M. K., & Schunn, C. (2015). A framework for analyzing cognitive demand and content-practices integration: Task analysis guide in science. Journal of Research in Science Teaching, 52(5), 659-685.	Generic, framework proposal for science education
44	Greenfield, D. B. (2015). Assessment in Early Childhood Science Education. In Research in Early Childhood Science Education (pp. 353-380). Springer Netherlands.	Review article
45	Lu, Y. L., Lian, I. B., & Lien, C. J. (2015). The application of the analytic hierarchy process for evaluating creative products in science class and its modification for educational evaluation. International Journal of Science and Mathematics Education, 13(2), 413-435.	Out of scope: empirical case about a specific evaluation method of technological products. Focused on teachers.
46	Schultz-Jones, B. A., & Ledbetter, C. E. (2013). Evaluating students’ perceptions of library and science inquiry: Validation of two new learning environment questionnaires. Learning Environments Research, 16(3), 329-348.	Out of focus: reviews an assessment tool and not students learning
47	Neumann, K., Viering, T., Boone, W. J., & Fischer, H. E. (2013). Towards a learning progression of energy. Journal of Research in Science Teaching, 50(2), 162-188.	Out of focus: reviews an assessment tool and not students learning
48	Scanlon, E. (2012). Open educational resources in support of science learning: tools for inquiry and observation. Distance Education, 33(2), 221-236.	The paper describes two open resource tools but does not include any information on the evaluation

49	Mohan, B., & Slater, T. (2006). Examining the theory/practice relation in a high school science register: A functional linguistic perspective. Journal of English for Academic Purposes, 5(4), 302-316.	Out of the scope: the focus is on teaching practices and the analysis of the linguistic discourse
50	Lustigová, Z., Lustig, F., Mechlová, E., & Malčík, M. (2009). A new e-learning strategy for cognition of the real world in teaching and learning Science. The New Educational Review, 17(1), 305-317.	No evaluation/methods section, only the activity
51	Olckers, L., Gibbs, T. J., & Duncan, M. (2007). Developing health science students into integrated health professionals: a practical tool for learning. BMC medical education, 7(1), 45.	No evaluation/methods section, only the activity and focused on health care professionals
52	Shen, L., & Bai, L. (2006). MutualBoost learning for selecting Gabor features for face recognition. Pattern Recognition Letters, 27(15), 1758-1767.	Not about science learning/engagement
53	Klassen, S. 2006. Contextual assessment in science education: Background, issues, and policy. Science Education, 90 (5): 820-851	It is a review
54	Harris, T.R., Brophy, S.P., ,2005. Challenge-based instruction in biomedical engineering: A scalable method to increase the efficiency and effectiveness of teaching and learning in biomedical engineering. Medical Engineering and Physics, 2 (7): 617-624	It is a review
55	Pringle, R.M., Martin, S.C. 2005. The potential impacts of upcoming high-stakes testing on the teaching of science in elementary classrooms. Research in Science Education, 35(2-3): 347- 361	Not about science learning/engagement
56	Webb, M.E. 2005. Affordances of ICT in science learning: Implications for an integrated pedagogy. International Journal of Science Education, 27 (6): 705-735	It is a review
57	Xu, L. 2004. Advances on BYY harmony learning: Information theoretic perspective, generalized projection geometry, and independent factor autodetermination. IEEE Transactions on Neural Networks, 15 (4): 885-902	Not about science learning/engagement
58	Morejon, R.A., Principe, J.C. 2004. Advanced search algorithms for information-theoretic learning with kernel-based estimators. IEEE Transactions on Neural Networks, 15(4): 874-884	Not about science learning/engagement
59	Honkela, A., Valpola, H.2004. Variational learning and bits-back coding: An information-theoretic view to Bayesian learning. IEEE Transactions on Neural Networks, 15(4): 800-810	Not about science learning/engagement
60	Choi, S., Lee, T. W.2004. A negentropy minimization approach to adaptive equalization for digital communication systems. IEEE Transactions on Neural Networks, 15(4): 928-936	Not about science learning/engagement
61	Wang, S., Schuurmans, D., Peng, F., Zhao, Y. 2004. Learning mixture models with the regularized latent maximum entropy principle. IEEE Transactions on Neural Networks, 15(4): 903-916	Not about science learning/engagement

62	Cruces-Alvarez, S.A., Cichocki, A., Amari, S.-I. 2004. From blind signal extraction to blind instantaneous signal separation: Criteria, algorithms, and stability. IEEE Transactions on Neural Networks, 15(4): 859-873	Not about science learning/engagement
63	Sánchez-Montañés, M.A., Corbacho, F.J. 2004.A new information processing measure for adaptive complex systems. IEEE Transactions Neural Networks,15 (4): 917-927	Not about science learning/engagement
64	Rutkowski, L. 2004. Adaptive probabilistic neural networks for pattern classification in time-varying environment. IEEE Transactions on Neural Networks, 15 (4): 811-827	Not about science learning/engagement
65	Schraudolph, N.N.2004. Gradient-based manipulation of nonparametric entropy estimates. IEEE Transactions on Neural Networks, 15 (4): 828-837	Not about science learning/engagement
66	Iwata, K., Ikeda, K., Sakai, H. 2004. A new criterion using information gain for action selection strategy in reinforcement learning. IEEE Transactions on Neural Networks, 15 (4): 792-799	Not about science learning/engagement
67	Sindhwani, V., Rakshit, S., Deodhare, D., Erdogmus, D., Principe, J.C., Niyogi, P. 2004. Feature selection in MLPs and SVMs based on maximum output information. IEEE Transactions on Neural Networks, 15 (4): 937-948	Not about science learning/engagement
68	Yamada, S. 2004. Recognizing environments from action sequences using self-organizing maps. Applied Soft Computing Journal, 4 (1): 35-47	Not about science learning/engagement
69	Yore, L.D., Bisanz, G.L., Hand, B.M. 2003. Examining the literacy component of science literacy: 25 years of language arts and science research. International Journal of Science Education, 25 (6): 689-725	It is a review
70	Gribble, S.J., Rennie, L.J., Tyson, L., Milne, C., Speering, W. 2000. Negotiating values for the science curriculum: The need for dialogue and compromise. Research in Science Education, 30 (2): 199-211	Not about science learning/engagement
71	Jerant, A.F. 1998.Training residents in medical informatics. Family Medicine, 31 (7): 465-472	Not about science learning/engagement
72	Fernández, H., Asensio, M.1998. Concept mapping as a research tool: Knowledge assessment in social science domain. International Journal of Continuing Engineering Education and Life-Long Learning, 8 (12): 109-123	It is a review
73	Norris, T.E., Coombs, J.B., Carline, J. 1996. An educational needs assessment of rural family physicians. Journal of the American Board of Family Practice, 9 (2): 86-93	Not about science learning/engagement
74	Reiner, M. 1995. Evaluation of a computer integration strategy in a science teacher's professional development program. Studies in Educational Evaluation, 21 (4): 457-473	Not about science learning/engagement, it's about teachers' training
75	Bacchus, C.M., Quinton, C., O'Rourke, K., Detsky, A.S. 1994. A randomized crossover trial of quick medical reference (QMR) as a teaching tool for medical interns. Journal of General Internal Medicine, 9 (11): 616-621	Not about science learning/engagement, it's about medical professionals' training

76	Lewis, M. 1993. Assessing decision heuristics using machine learning. Decision Support Systems, 10 (2): 199-212	Not about science learning/engagement
77	Barto, A.G., Sutton, R.S., Anderson, C.W. 1983. Neuronlike Adaptive Elements That Can Solve Difficult Learning Control Problems. IEEE Transactions on Systems, Man and Cybernetics, SMC-13 (5): 834-846	Not about science learning/engagement
78	Milford, T., Jagger, S., Yore, L., Anderson, J. 2010. National Influences on Science Education Reform in Canada. Canadian Journal of Science, Mathematics and Technology Education, 10 (4): 370-381	Out of focus
79	De Winter, J. 2011.'I no longer dread teaching physics, I now enjoy it!' Participant reflections from the SASP physics course. Physics Education, 46 (2): 159-166	Out of focus: focuses on teachers
80	Su, C.Y., Chiut, C.H., Wang, T.I. 2010. The development of SCORM-conformant learning content based on the learning cycle using participatory design. Journal of computer assisted learning, 26: 392-406	Out of focus
81	Chang, C-Y., Lee, G.2010. A Major E-Learning Project to Renovate Science Leaning Environment in Taiwan. TOJET: The Turkish online journal of Education and Technology, 99(1): 7-12	Theoretical article
82	Lomas, D. 2007. Cognitive Artifacts: An Art-Science Engagement. Conference: Proceedings of the 6th Conference on Creativity & Cognition, Washington, DC, USA, June 13-15	Not about science learning/engagement

ANNEX 2: List of included journals in the literature review

Journal name	Number of articles
Pedagogies: An International Journal	1
Asia-Pacific Forum on Science Learning and Teaching	1
British Journal of Educational Technology	1
Cartography and Geographic Information Science	1
Chemistry Education and Research Practice	2
Computers & Education	5
Conservation Biology	1
Cultural Studies of Science Education	1
Educational Technology and Society	1
Electronic Journal of Science Education	1
Eurasia Journal of Mathematics, Science and Technology Education	1
International Journal of Science Education	7
International Journal of Science Education	4
InterScience	2
Journal of College Science Teaching	1
Journal of Geography	1
Journal of Geography in Higher Education	1
Journal of Geoscience Education	1
Journal of Interactive Learning Research	1
Journal of Learning Sciences	2
Journal of Research in Science Teaching	3
Journal of Science Education and Technology	1
Journal Of Veterinary Medical Education	1
New Review of Hypermedia and Multimedia	1
PLoS computational biology	1
PloS one	2
Research in Science & Technology Education	1
Research in Science Education	4
Research in Science Education in Europe	1
School Science Review	1
Science and Education	1
Science Education	5
Simulation in Healthcare	1
Thinking Skills and Creativity	2
Urban Anthropology	1

ANNEX 3: Methodological protocol of the exploratory workshops to identify participatory indicators

Duration	Minimum 30 minutes
Number of students	Maximum 20
Number or facilitators	Minimum 2 (one person facilitates and the other collects data)
Implementation	In each school, during WP2 explorative workshop with students, preferably at the end of the first session

Objectives of the activity:

- To involve the students in the design of the assessment process
- To include criteria and indicators that they consider important in the assessment of the impact of the project

Focus of the activity:

Key aspects that motivate participants to get engaged and to actively participate in science-related activities.

Materials required:

- Roll paper
- Post-its of two different colours (one for each question, see below)
- Colour tape (for the barometer)
- Video camera or audio recorder
- Power point presentation with the two questions (see below).

Activity description:

- **5' Introduction to the activity:** Introduction to the activity: why are we doing this?

  Context guidelines:

As you know, PERFORM is a project that wants to change the way science is taught, learnt and communicated. We want to make science activities more interesting and motivating. Because of that, we are interested in looking at what happens through this project and in the opinions and feelings of the people participating on it; that is, you! For that reason, we will develop a research all through the process, that allow us identify the things that work and communicate them to other people, and also change those aspects that need to be improved. A team from

the university in Barcelona will be in charge of it and we will help them.
Participation is very important in PERFORM and also in this research. We want this project to be everyone's project. That's why we would like you to be part of the research design too, and give your opinion since the very beginning. We will now develop an activity so we can explore together what is important for you.

Split the group into 5 subgroups of 3 people each (could be through a small quick game or guideline).

- **15' Exploration of questions:** Explain the questions showing the power point:

1st question: *When you are participating in a science-related activity, what are the things you like about it, if any?* (5-10 minutes)
2nd question: *If you were to design a scientific activity for your classmates, how would you do it to make sure to engage them?* (5-10 minutes)
Explain the activity: *In groups, discuss around these two questions and write down in post-it's the aspects you identify as important (one post-it for each idea). You have around 5-10 minutes to discuss each question*.

- **15' Sharing of the key aspects identified by the students and 'group barometer':**

Ask each group for sharing their conclusions by posting their post-its in a collective mural (i.e., on the roll paper) and briefly sharing the conversation they had on each question. Collect notes on each group comments (see attached data collection table).
Read to the students all the post-its while organizing them in clusters or dimensions according to their meanings (e.g., enjoyment, knowledge acquisition, etc.). Participants will have the chance to add whatever they find is missing. Collect notes on the resultant dimensions.
Additionally, each dimension will be explored in terms of support or importance given by participants, through the technique of the barometer. Draw a line in the floor with colour tape representing a degree with three marks: 'very important', 'important', and 'not important'. Ask students to place themselves along the line, according to the importance they give to each dimension in the context of science learning. Collect notes on the resultant scores for each dimension.

Data collection table:

Activity	Students’ responses	Facilitators’ observations
Post-its of the 1rst questions (all groups)	*Please write down here the content of each post-it related to the 1rst question for all groups*	
Comments to the 1rst question	*Include here students’ comments related to their answers to the 1rst question when posting the post-its*	
Post-its of the 2^{nd} questions (all groups)	*Write down here the content of each post-it related to the 2^{nd} question for all groups*	
Comments to the 2nd question	*Include here students’ comments related to their answers to the 2nd question when posting the post-its*	
Resultant dimensions	*Include here the content of post-its related to each resultant dimensions* Dimension 1: Post-its:	
	Dimension 2: Post-its:	
	Dimension 3: Post-its:	
	Dimension *n*: Post-its:	
Comments to the dimensions	*Include here students’ comments related to the resultant dimensions*	
Barometer	*Include here the scores assigned to each dimension* Score dimension 1: Score dimension 2: Score dimension 3: ...	
Other relevant impressions or considerations about the development of the activity *(e.g. mood of the group and reception of the activity, contextual particularities, any unexpected event)*:		