Equipping the Next Generation for Active Engagement in Science

DELIVERABLE D8.16: Evaluation Plan- Indicators, Impact, and Influence

Project Acronym: ENGAGE

Project Name: Equipping the Next Generation for Active Engagement in Science

Call: SCIENCE IN SOCIETY [2013.3.2.2.1-1]

Project Type: Coordination and support actions

Grant Agreement No.: 612269

Project Start Date: 1st January 2014

Project Duration: 36-Months

Due date of Deliverable: Month 12

Actual Submission Date: Month 12

Task Leader: Dury Bayram Jacobs- TUDelft

Report Author(s): Dury Bayram Jacobs (TUD)

Report Collaborator(s): Pat Morton (SHU), Yael Shwartz (WZ), Fotini Chaimala (FOR), Alexandra Okada (OU), Mauro Vannucci (ELS), Maarten van der Sanden (TUD), Marc de Vries (TUD)

Dissemination Level: Public

THE ENGAGE CONSORTIUM

Centre for Science Education – Sheffield Hallam University (Coordinator)	UK
Knowledge Media Institute – The Open University	UK
Institute of Applied and Computational Mathematics, Foundation for Research and Technology	Greece
Innovation in Learning Institute	Germany
eXact learning Solutions	Italy
Traces	France
Valahia University Targoviste	Romania
Weizmann Institute	Israel
Universitat de Barcelona	Spain
Vestfold University College	Norway
Delft University of Technology, Science Education and Communication Department	Netherlands
School of High Pedagogy of Freiburg	Switzerland
Lithuanian University of Educational Sciences	Lithuania
Department of Science Education, University of Nicosia	Cyprus

DOCUMENT HISTORY

Version	Date	Comment	Modifications made by
1	14.07.2014	First Draft	Dury Bayram Jacobs (TUD)
2	11.08.2014	Second Draft	Dury Bayram Jacobs (TUD)
3	19.08.2014	Review	Pat Morton (SHU)
4	15.08.2014	Review	Maarten van der Sanden (TUD) Marc de Vries (TUD)
5	24.09.2014	Presentation of the evaluation report to the partners at Paris meeting and feedback from the partners	Dury Bayram Jacobs (TUD) and partners at Paris meeting
6	24.09.2014	Review	Project partners Dury Bayram Jacobs (TUD)
7	06.10.2014	Integration of review comments and suggestions	Dury Bayram Jacobs (TUD)
8	24.11.2014	Review and feedbacks	Pat Morton (SHU)
	25.11.2014	Review and writing the sub-chapter about Online Courses	Alexandra Okada (OU)
9	26.11.2014	Review and writing the sub-chapter about OER	Yael Shwartz (WZ)
10	26.11.2014	Review and writing the sub-chapter about Knowledge Hub	Mauro Vannucci (ELS)
11	28.11.2014	Review and writing the sub-chapter about face to face workshops	Fotini Chaimala (FOR)
12	11.12.2014	Review	Bernie Stiell
13	16.12.2014	Review	Maarten van der Sanden (TUD)
14	18.12.2014	Final review and submission	Pat Morton (SHU)

Content

Tables

Figures

1. EXECUTIVE SUMMARY

This deliverable describes the strategies of the Engage evaluation plan. The aim of this evaluation plan is to guide the partners of Engage project in the formative and summative evaluation processes, and provide the other stakeholders (EU 7th framework programme, scientists, teachers, science education institutions, etc.) with a practical guide to project evaluation.

Preparation of the evaluation plan, which is a roadmap for Engage project, has been coordinated by TUDelft (partner 11). While preparing the evaluation plan:

- the aims of the Engage project
- the expected outcomes & products
- the timetable of the project activities

were taken into consideration.

All the activities, outcomes & products and impacts on stakeholders of the Engage project will be evaluated by internal and external evaluation. The internal evaluation is led by TUDelft while the external evaluation is led by Centre for Education and Inclusion Research (CEIR) of Sheffield Hallam University. Both evaluators collaborate with each other in order to provide highly qualified and useful evaluation of the Engage project.

As the evaluation plan is a living document, it has been developing continuously according to feedback from the partners, changes in the timetable of the project activities and the needs of the stakeholders.

2. INTRODUCTION

Program Description: ENGAGE

ENGAGE aims to support the EU's ambition by shifting the practice of science teachers towards RRI-based science on a massive scale of 11,750 teachers, and over 2 million students, across 11 different countries. Yet the shift towards RRI-based or 'humanistic' science teaching as it is sometimes called (Aikenhead, 2006), or incorporating inquiry-based pedagogy, is known to be very challenging. Aikenhead underscores the lack of success of professional development (CPD) programmes to transform teachers with a list of reasons for failure – it takes up a whole page. The notion that any event or short-term set of workshops produces lasting change has been discredited. Real transformation such as the one required to move to RRI-based teaching is a long term and a complex process. Our programme synthesizes contemporary models of CPD with our curriculum development expertise into a programme which support teachers through this process of transformation.

ENGAGE synthesizes contemporary models of professional learning and curriculum development. Going beyond training events, its three-stage path will propel teachers in their own inquiry to become comfortable and even expert in RRI teaching. The path requires a first stage, *Adopt,* which achieves take-up on a massive scale. We use a proven approach to provide an easy entry into inquiry-based teaching. Our Open Educational recourses (OER) will combine science-in-the-news contexts with strategies from informal learning to get students talking. An online community of practice supports teacher reflection, while online courses and workshops add coaching and feedback. At the next stage, *Adapt*, the teachers use an expert's toolkit of examples, explanations, anecdotes and activities to help students learn effectively, and finally, in the third step, *Transform,* open-ended Projects put teachers and students into partnership with practicing scientists, to learn about RRI directly.

Our partners bring extensive track records in teacher development and curriculum design. Building on best practice from previous projects, we intend to influence 11,750 science teachers across Europe, in the future we will extend this to pre-service teachers and their trainers.

Users: Project consortium, teacher trainers, teachers, science education institutions

Main Target groups of the project:

- 11,750 teachers used materials
- 2 million students (11-16 years old) reached
- in 11 countries

The main activities and stages of the Engage project are summarized in the figure 3, below.

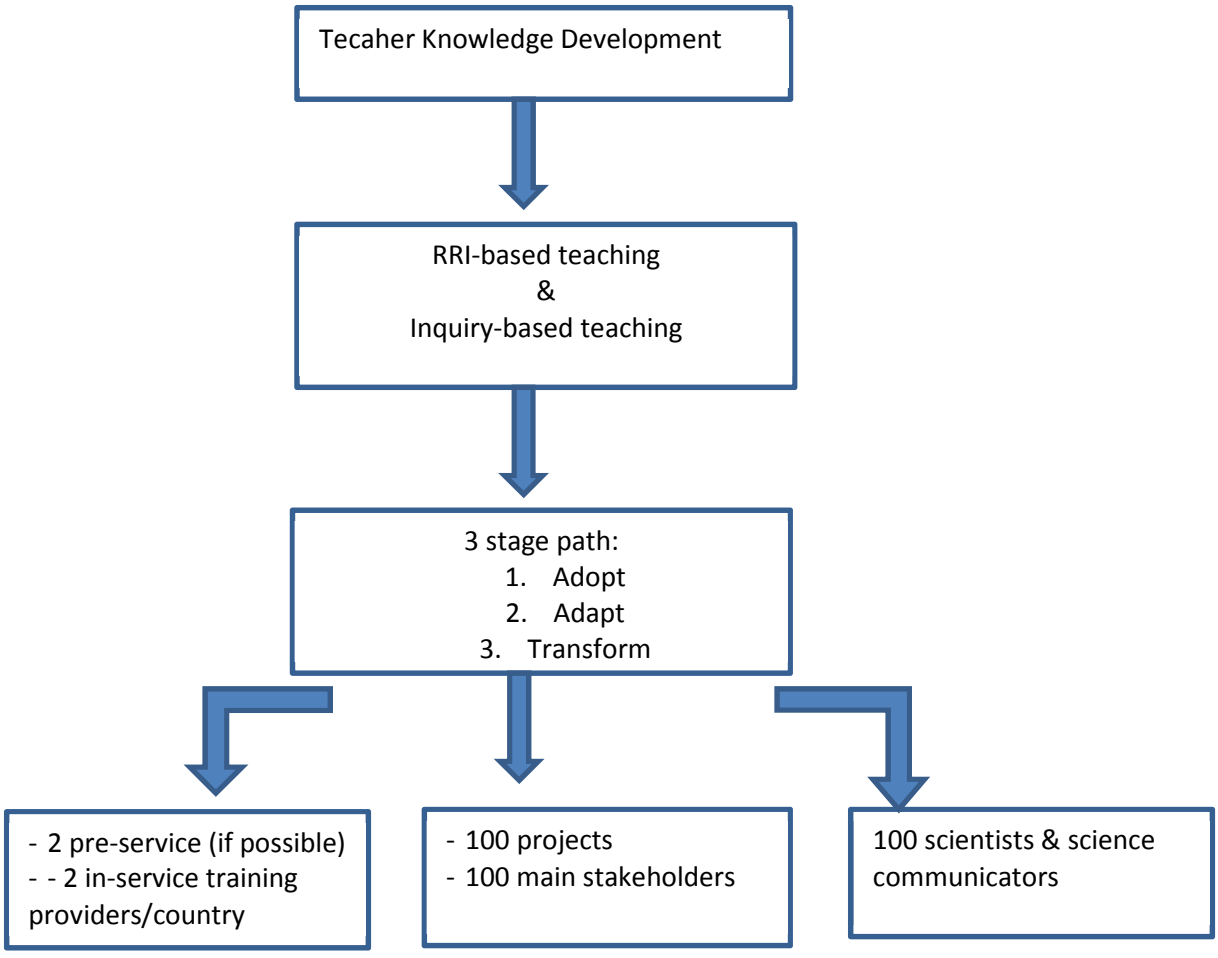


Figure 1. ENGAGE Framework

What is Evaluation?

Evaluation is defined as a systematic data collection process about an activity or an action in order to determine its worth or merit (WAPC, 2003). Patton (1987) adds that the aim of evaluation is also to make judgments about the project, improve effectiveness of the project and inform about future planning. Evaluation gives useful information about the process of project or action and helps to improve it. In general, project evaluation answers two questions to lead two purposes i) does it work? (project's merit) ii) do we need it? (project's value) (NOAA, 2009).

Formative evaluation as defined by Scriven (1991) is an evaluation which is conducted (may be more than once) with the aim of improving the programme, project or product. He also mentions that it is better to do formative evaluation by external and internal evaluator. The aim of the formative evaluation according to Worthen, Sanders, and Fitzpatrick (1997) is improving the programme by providing information to the programme team.

Stake (1976, 4, 6) draws attention to the difference between formative and summative evaluation by his soup analogy: "When the cook tastes the soup, that's formative; when the guests taste the soup, that's summative."

A comprehensive evaluation of the ENGAGE programme is to be conducted by cooperation of internal (TUDelft) and external evaluation (CEIR) teams which are summarized in Table 1. The evaluation design for ENGAGE involves a mix of qualitative and quantitative methods to describe and evaluate the operation of the ENGAGE programme and to assess its impact on participants and other stakeholders.

Table 1. Internal and External Evaluation

Evaluator	purpose	audience	evaluation focus
Internal (TUDelft)	to help the consortium to learn/feedback about the programme	The project team	Implementation and impacts on target groups, with other consortium members
External (CEIR)	Independent, unbiased assessment	The European Commission and other stakeholders	Progress towards objectives determined by the evaluator with the consortium in months 0-4

In the Engage project, evaluation is directed at different purposes:

- Planning evaluation before the programme will give our consortium feedback on our strategies and objectives of the evaluation and how likely they are to be carried out, within each country's context.
- The initial framework will consider assumptions about target population (sampling), interventions, instrument development (surveys, questionnaires) and project outcomes.
- Formative evaluation will provide feedback to the consortium on how we are progressing towards our objectives, so we can improve our delivery. This will be based on the targets we have identified for each step and strategy of the programme.
- Summative evaluation will assess the extent to which we have achieved our objectives and expected impact on stakeholders, and whether the project has justified its costs.

3. EVALUATION PLAN

The evaluation plan which is defined as a roadmap answers the questions ‘what’, ‘how’ and ‘why it matters’. It is expected from an effective evaluation plan that it is a living document and reflecting the changes in the programme or project (DEEP, 2011).

Therefore, the evaluation plan of the ENGAGE project is a dynamic roadmap which reflects the changes in the project. Timing of interventions in the evaluation plan will need to be adjustable to fit with any required changes in the delivery rollout of the other actions in the project. Since the project has many actions, some of them are going parallel, some of them are consecutive, and some others have different timetables, the evaluation plan also needs to fit and reflects them. So, the evaluation plan is not and cannot be straightforward. In this way, it will be possible to have an effective formative evaluation process and to reach intended aims.

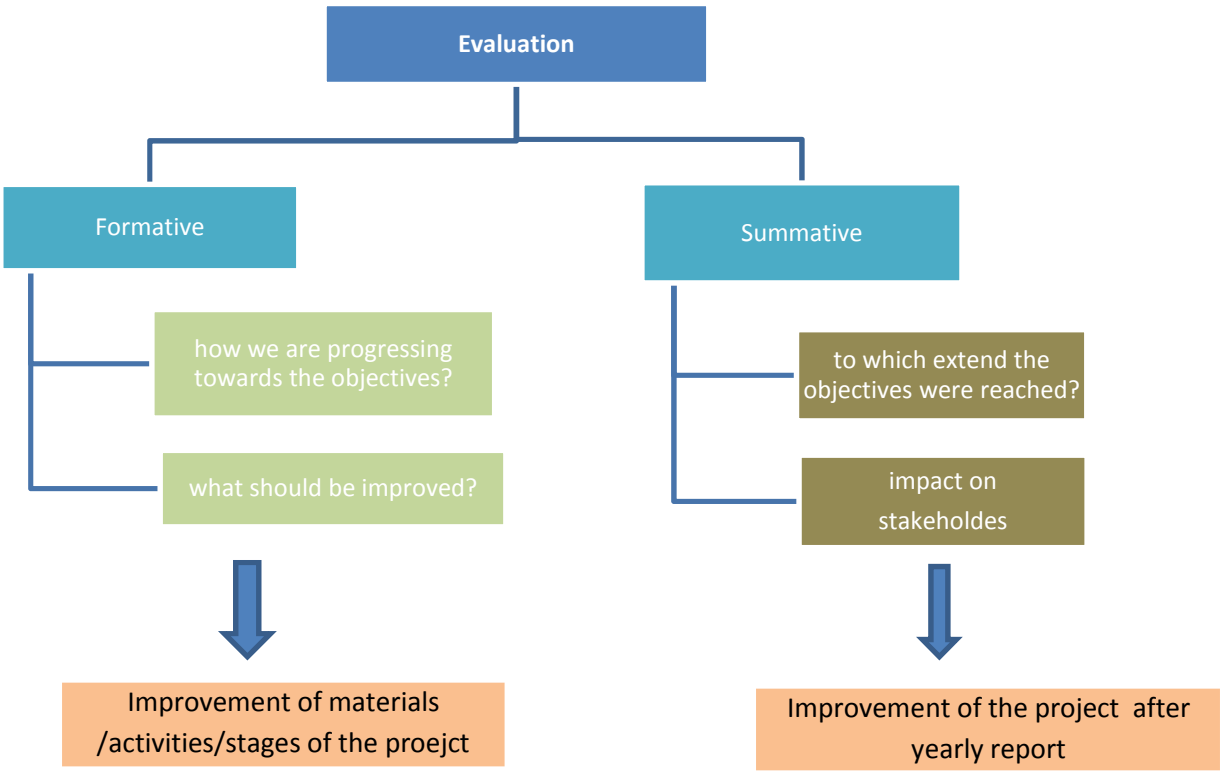


Figure 2. Evaluation framework of ENGAGE project

Purpose of Evaluation (Why are we doing evaluation?)

- to undertake internal evaluation activities of the implementation of the 'Adopt', 'Adapt' and 'Transform' stages
- to report on the participation and impact on teachers and students with respect to project targets
- to report on stakeholder engagement and influence

In addition to these general purposes, the formative evaluation aims:

- to find out how we are progressing towards the objectives of the project
- to find out what should be improved
- to feed the further development of the materials, activities and CPD programme

Adjustments in Evaluation Plan

The WP was originally to be managed by Prof. Dr. Patricia Ossewijer from the Biotechnology and Society department at the Delft University of Technology. Due to changes within the team, she asked colleagues within Science Education and Communication Department at TUDelft to adopt the WP, now led by Maarten van der Sanden from Science Education and Communication Department. The new team had to search for a post-doc researcher. In the meanwhile Delft wrote an outline of the planning they would start in July 2014, to follow on from the pilot plan produced by FAU.

On 1st July post-doc (Dr. Dury Jacobs) started at TUDelft to work on WP8. Starting from this date, TUDelft designed an outline for the evaluation plan. The evaluation plan was presented and discussed during the Paris meeting. During this meeting it was possible to contact and discuss the evaluation plan for effective delivery and frameworks with the external evaluator.

Because of the change of the department of TUDelft as a partner of the project, the project partners made a rapid start for evaluation of the pilot. Therefore FAU (Germany), as leader of the testing and classroom observations task, took over the evaluation task for the pilot at short notice. The evaluation concept for the pilot developed by FAU was developed in the period from May to June in order to assist the testing sites in the UK, Israel, Norway and Cyprus during the critical period of testing.

The main evaluation questions for the ENGAGE testing activities are:

- How easy it is to adopt ENGAGE materials from an organizational point of view?
- How do teachers and students rate the usefulness and quality of the ENGAGE materials?
- How do ENGAGE materials meet teaching and curriculum needs in the different countries?
- How innovative are the ENGAGE materials and the pedagogical approach of the materials?
- How do the ENGAGE materials support the reflective thinking on complex research topics?

The pilot testing took place in Israel, Cyprus and Norway focused on classroom activities, evaluated through observations and interviews. All countries have ethical procedures in place. In contrast, the UK partner focused on a complementary wide dissemination approach. The results of this testing were given in annex 3.

Methods, Approaches and Instruments

There are various methods and data collection instruments that can be used in the evaluation. For example; structured interviews, semi-structured interviews, questionnaires, evaluation stories, photographic history, participant observation, etc.

(WAPC, 2003). The following table summarizes general appropriate data collection instruments.

Table 2. Data collection instruments

Data collection instruments	Knowledge	Skills	Attitude	Behaviour
Interview	X		X	(X)
Focus group	(X)		X	
Questionnaire and survey	X	X	X	(X)
Observation		X		X
Literature review	X	X	X	X
Test	X	X		
Concept maps	X		(X)	
Document or product review	X	X	(X)	
Case study	X	X	X	X

(x): maybe but not always appropriate to evaluate the indicated type of learning.

Source: NOAA, 2009.

Before choosing what methods to use in order to collect data, it is important to know the advantages and limitations of each method. Some of these methods are summarized in Annex 1.

Validity and Reliability of Evaluation Instruments

Validity: the extent to which the evaluation instrument measures what it aims to measure.

Content validity: A way to provide content validity is to ask a panel of experts if your questions adequately sample the content you wish to convey.

Reliability: The extent to which the evaluation instrument yields consistent responses each time it is administered.

Usability: Being reasonable with the length and scope of evaluation.

Field or pilot testing (for validity and reliability): Trying out the instruments under conditions similar to those you expect to have for the evaluation, therefore conduct the field test on a representative sample of people who are likely to participate in the actual project. Field tests answer questions such as: Is the questionnaire providing the information needed to answer my evaluation questions? Are multiple interviewers collecting information in the same way? How consistent is the information obtained with the questionnaire? How accurate is the information obtained with the observation rubric?

In the ENGAGE project, we will develop evaluation tools in the process which is described below.

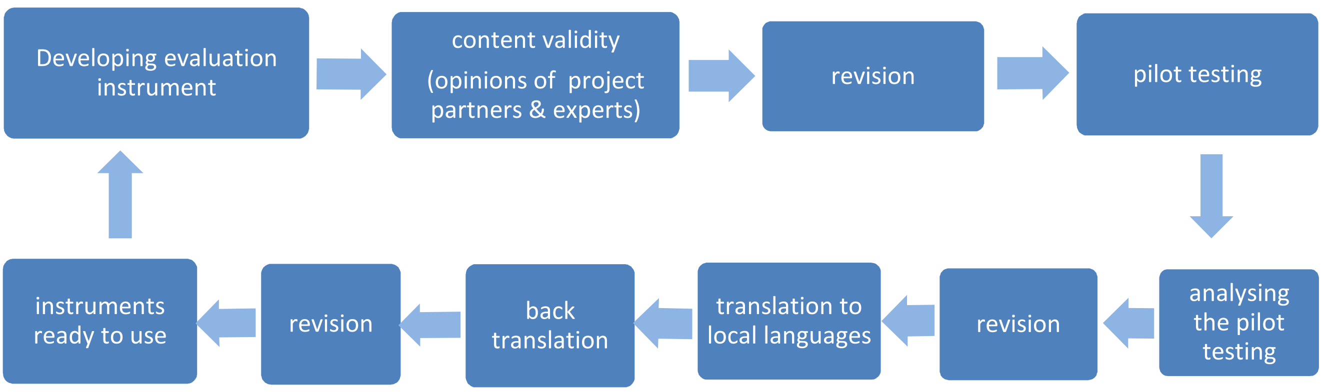


Figure 3. Process of developing an evaluation instrument

As described in figure 3, after an evaluation instrument is developed, the opinions of experts will be asked to revise and improve the instrument. Then there will be a pilot testing stage on teachers to find out whether the instrument provides the needed information or not. Another reason for pilot testing is to try out the online survey system and to perform analysis with the data collected. In this way, it will be possible to develop an evaluation instrument with high validity and reliability.

Evaluation Focus

This provides the opportunity to document how the evaluation focus will be narrowed and the rationale for the prioritization process. Given that there are never enough resources or time to answer every evaluation question, it is critical to work collaboratively to prioritize the evaluation based on a shared understanding of the theory of change identified in the logic model, the stage of development of the program, the intended uses of the evaluation, as well as feasibility issues. This section should delineate the criteria for evaluation prioritization and include a discussion of feasibility and efficiency.

By focusing on the purposes of evaluation, the evaluation focus will be;

- the scope and quality of implementation of the 'Adopt', 'Adapt' and 'Transform' stages
- impact on teachers & students
- stakeholder engagement and influence

The formative evaluation will focus on:

- activities & materials
- courses (f2f workshops & MOOC' s)
- the impact on teachers & students (T8.2)
- the influence on the stakeholders (T8.3)

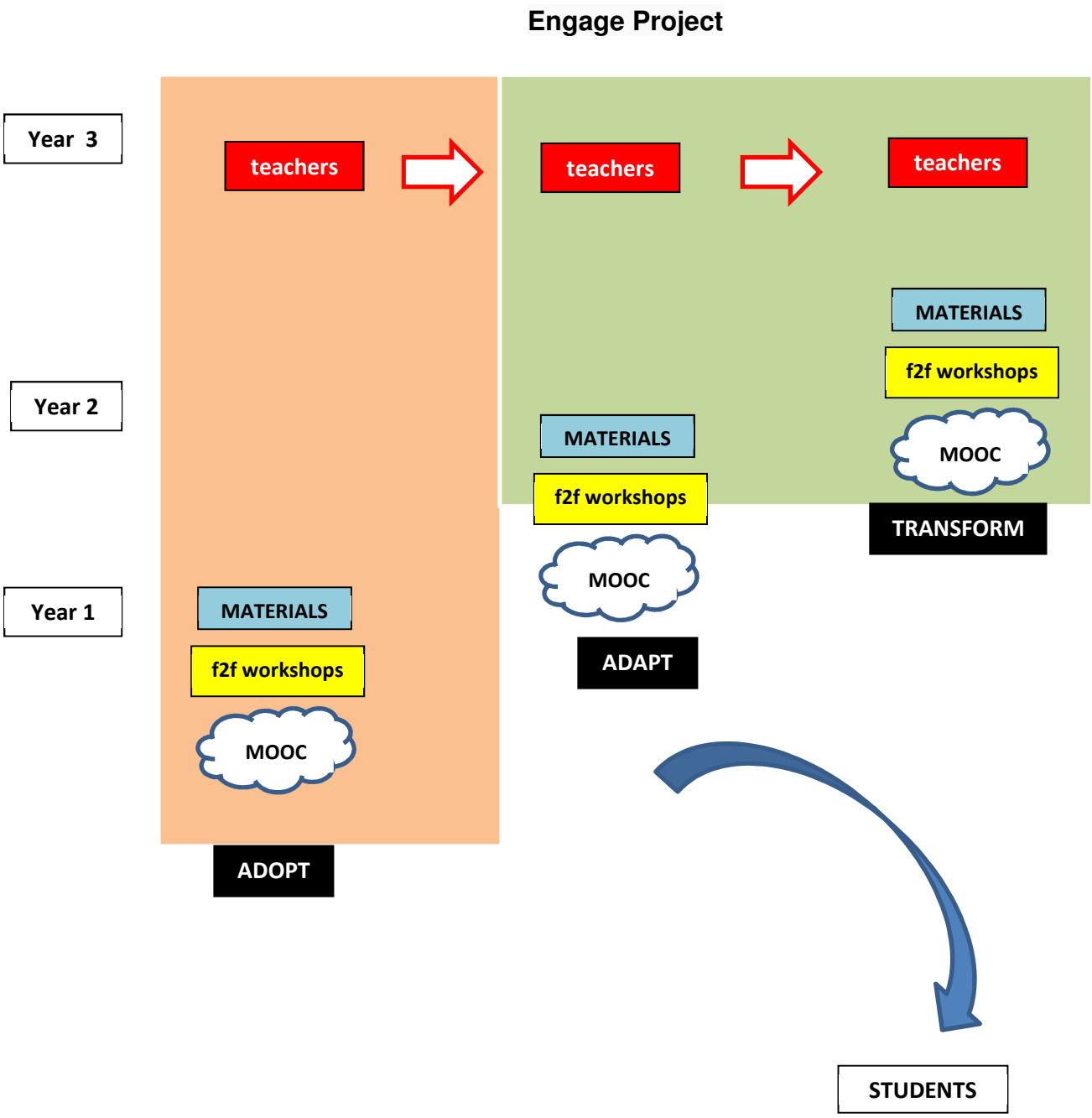
The summative evaluation will focus on the above mentioned points for formative evaluation and:

- the objectives of the project
- the impact on partners (T8.5)
- experiences in different cultures/countries

In the formative evaluation process, the focus mainly will be the open educational materials and the activities (e.g. topicals which are materials for the adopt stage and f2f workshops & MOOC's) since the evaluation of them will result in revising and improving the activities and materials and therefore the impact of the programme. Moreover we will focus also on the impact on teachers to gather information about the impact of our programme on teachers and in order to advise as to whether we are on programme to meet our project targets and if not how to change our interventions.

While internal evaluation will focus on further development of activities, materials and CPD Stages (adopt, adapt and transform) external evaluation will focus on the results of summative evaluation to assess the impact of the project beyond the project partners and to report whether the project reached the aims or not. In this respect, the external evaluation will use the results of the internal evaluation.

Figure 4 summarizes the different dimensions that will be evaluated in the Engage project. During the 3 stages (adopt, adapt and transform) there will be educational materials, face to face workshops and MOOC's which target to have impact first on teachers and then on students. Therefore, the target of evaluation is mainly teachers but in the end of the project we will evaluate the impact of the project on the students, too.


Engage Project
Year 3
Year 2
Year 1
teachers
teachers
teachers
MATERIALS
f2f workshops
MOOC
TRANSFORM
MATERIALS
f2f workshops
MOOC
ADAPT
MATERIALS
f2f workshops
MOOC
ADOPT
STUDENTS

4. FORMATIVE EVALUATION

The aim of formative evaluation is to gain an understanding what the target groups know and don't know to make responsive changes in the programme (Boston, 2002). This type of evaluation focuses on what services are provided and to whom and how. By providing feedback, the aim is to strengthen the programme (NOAA, 2009). According to the feedback necessary changes can be made and in that way the effect of the programme can be ensured.

Questions for formative evaluation can address the following aspects (NOAA, 2009):

i. Are the project resources (time, money, expertise) being utilized as planned? (Resources)
ii. Is the project being implemented as planned? Are the intended activities, products, or services being provided?
iii. What promotional strategies worked or failed? (Activities)
iv. Is the project reaching the target audience? Which targeted audiences participated in the activities?
v. How extensively is the audience engaged in project activities? (Participation)
vi. What are the participants' reactions to the project activities? Do participants perceive immediate benefits from their participation in the activities? (Reactions)

While planning the evaluation, the timeline of the project which is given below will be taken into consideration (DoW, p.18-23):

1. WP4- Adopt commences month 7 (July 2014) to 36 (December 2016)
2. WP5-Adapt commences month 12 (December 2014) to 36 (December 2016)
3. WP6- Transform commences month 16 (April 2015) to 36 (December 2016)

The Engage project is supported by the European Commission under FP7 SIS 612269

Impact Evaluation

The Engage project aims to have impact on mainly three stakeholder groups:

- Teachers
- Students
- The other stakeholders (scientists, countries other than partner countries)

In order to have impact on the above mentioned stakeholder groups, the ENGAGE project has three strategies (materials (Open Educational Resources) strategy, community strategy and courses strategy). Firstly, these strategies and the evaluation approaches of these strategies will be explained and then the evaluation of the impact on teachers, students and other stakeholders.

1. Materials Strategy and Evaluation of the Engage Materials (OER)

Aims of Open Educational Resources (OER):

The aim of developing OER is to provide teachers at the adopt and adapt stages with teaching materials that reflects the project's objectives, framework and strategies. These materials will be downloaded and enacted by teachers, but moreover the framework that is used to develop them will be used in the professional development programmes of the adapt and transform stages to guide teachers in targeting their teaching toward RRI teaching.

What's the purpose?

The prototype Engage materials were designed following a framework that guides all our materials design and production, to help us define our 'RRI curriculum' i.e. what we're teaching, to define the knowledge about the nature of science and scientific practices i.e. inquiry skills, argumentation, etc. that we want to promote through our curriculum. The framework defines top level aims, using the 'understanding by design' methodology.

For more information about OER framework and development process see deliverable 1.2.

Evaluation of the Educational Materials:

Evaluation tool: Online Questionnaire- Q1
Target: Teachers
When: February 2015 (for adopt stage teachers)
Report: # of downloads
Analyzing: descriptive statistics and graphs + content analysis

For each topical the number of downloads will be reported. Teachers' opinions about the topicals will be collected by an online questionnaire. Questions include whether the teachers liked the topical and the activity, would want to use it again, would recommend it to colleagues and the aspects they think should be improved.

The questionnaire was developed in August 2014 and revised according to opinions of the experts from TUDelft and project partners. The expert opinions were collected in three stages and after the each stage the questionnaire was revised and improved. The revisions were done until there was no further feedback.

Currently, the questionnaire is in the translation and back translation process. The partners are going to translate and back translate the questionnaire before the due date 19 December 2014 by using the guide for the translation and back translation (see annex 3). Then, there will be revisions of the translated and back translated questionnaires and after that it will be ready for pilot testing which is planned for January 2015. We have ordered professional online survey software (Survey Gizmo) which will be used in all evaluation processes. Therefore whilst performing the pilot testing, we may test the software, too. The partners will provide the contact information of the teachers from their local network and online survey will be sent by using MailChimp. The questionnaire for the evaluation of the materials (Q1) can be found in Annex 2.

2. Community Strategy

2.1. Knowledge Hub

The second strategy of the Engage project is the Community Strategy. The Knowledge Hub is being developed to reach the aims of this strategy.

The aim of creating knowledge hub is:

- to provide a collection of tools for efficient and effective communication, sharing of ideas, expert articles, media, and delivery of materials
- to create a platform acting as an access point for tools, materials and user-generated contents
- to integrate the platform in the web environment of the existing portals and platforms used by teachers
- to provide information, guidance and support to the use of tools and sharing of experiences

Knowledge Hub is an integration of different platforms to support the teachers to find, download, re-elaborate materials and share ideas. The teacher must be registered in the main site *http://engagingscience.eu* in order to access to all the functionality. The platform is customized for different languages so that each country can have materials and a community of teachers sharing the same interests and educational plans. The Knowledge Hub foresees also the possibility to deliver MOOC courses for teachers in order to covers the three areas of teacher knowledge: pedagogies, how to teach by building on students' ideas, and learning about the nature of science.

Evaluation of the Knowledge Hub:

Knowledge Hub will be evaluated with learning analytics. Learning analytics is an emerging tool which uses analytic tools with the aim of contributing the improvement of learning and education processes (Elias, 2011).

This task will be led by VUT aims to provide learning analytic tools for collecting data about students' opinions, values and knowledge before and after using Engage curriculum resources, for formative and summative evaluation. This will be done by embedding quizzes and short surveys within the activities of the curriculum resources.

3. Courses Strategy

The third strategy is "Courses Strategy". ENGAGE uses a blend of online learning modules, called Massive Online Open Courses (MOOC) and face to face events (Workshops). Both are prototyped in the 'preparation phase' (WP1) and then rolled out in the deployment phase.

3.1. Face to Face Workshops

In addition our OER will include a prototype course for teaching professional development programs. The prototype course will include both a face-to-face and a MOOC version.

F2F workshops will support teachers by:

- Providing simulated practice for teachers, in a supported atmosphere of experimentation learning directly from 'experts'.
- Focusing on successful use of curriculum materials with feedback, which will be given by the course facilitators .
- Practicing using the strategies through simple tasks based on teachers' needs, interests and efficient time.

The features for the teachers' CPD are:

1. Create short interactive online modules,
2. Include input from professionals and expert RRI teachers
3. Provide an RRI toolkit which Includes RRI effective teaching strategies (debate, argumentation, card-games, how to deal with diverse opinions, etc.
4. Make use of our video library

5. Establish and support an online community of practice

For each stage some RRI teaching tools will be developed. For example, the teaching tools for the adopt stage are: using dilemmas and small-group-discussions. These tools will be the focus of the program, and will be exemplified by our materials. For the adapt stage the tools of argumentation and Case Method Tool, the gradual release tool. The question patterns tool will be developed.

Workshops focus on addressing discrete goals, such as learning to use a particular strategy. The goals will be based on the need for teachers and their schools to get maximum practical value from time out of school if they are to get permission to attend.

Timetable:

Partners have flexibility on scheduling the Workshop, depending on the ease of teachers coming out of school:

- Option a) 0.5-1full day interactive workshop format
- Option b) twilight sessions

There will be 1 Workshop per year (Adopt, Adapt and Transform), for 3 years, in each country delivering the programme, making 33 in total. Each partner will team up with a science/discovery centre partner to host the event. The events will be held in the first term of each academic year, to get new teachers engaged, with plenty of time to progress towards further stages.

Period suggested for partners: First year implementation of the ADOPT stage concern the period July 2014-June 2015. We will utilize the first 6 months for June 2014-January 2015 for developing the courses and piloting features of the courses by specific partners. Dissemination of the programme among partners, localization and translation will be held during February-March 2015. The actual programmes for teachers will be held during March-June 2015. Second iteration of the Adopt stage will start at September 2015.

11 Countries: UK, Greece, Germany, France, Romania, Israel, Spain, Norway, Switzerland, Lithuania, Cyprus.

Evaluation of f2f workshops:

In order to find out the impact of f2f workshops it is necessary to know the starting levels of teachers. Since the workshop will be one-day workshop, it is not appropriate to use pre- and post- test. Instead by including questions about the starting knowledge of teachers in the questionnaire which will be conducted at the end of the workshop, we can have idea about the initial levels of teachers. We will use a questionnaire to gather the opinions and to learn about the previous experiences of teachers and the questionnaire will be conducted to entire population. The suggested evaluation instruments are; questionnaire, interview and focus groups. Depending on the size of the group, partners will organize interviews with few teachers (3 teachers), and use questionnaire for the rest. If the group is big, then partners can also organize focus groups discussions in addition to the interviews.

TUDelft will provide a guide to the partners about using these 3 evaluation instruments, how to use them and when to use them. The questionnaire will be prepared by TUDelft with the help of the related WP leader FOR and other members of this WP. The opinions of experts and partners will be asked to improve and revise the questionnaire. Then, the questionnaire will be translated and back translated to/from local languages and it will be conducted at the end of the workshop. The interview will be a semi-structured interview. Semi-structured interviews are very common in small-scale researches. In this way, it possible to gather the data which we need and also give freedom to respondents to talk more and collect more in-depth data (Drever, 19965). The guide for semi-structured interview and interview questions will be prepared by TUDelft and the questions will be translated into local languages by partners. The project partners will make appointments for interviews. The partners will decide according to their size of participants to use or nit to use focus group discussions. It is suggested to use it when the group is big and it is

difficult to organize one to one interviews. If the group is not too big, still partners may decide to use focus group discussion to gather more data about ideas and feelings of teachers. If we can use all these three evaluation instruments we may collect more rich data and have better insight about the impact of the workshop and it is also useful for triangulation purposes.

3.2. Massive Open On-line Courses (MOOC)

Introduction:

ENGAGE Massive Open On-line Courses (MOOC) focus on addressing discrete goals, such as learning to use a particular strategy. Course participants will use a wide range of media and interactive online tools to engage with other colleagues and learn alongside them. These tools include video lectures, videoclips, online discussion boards, blogs, wikis and social networking sites such as Twitter.

The aim of Online Course is to extend the workshop experience as well as provide an introduction for those who could not participated in the workshop. ENGAGE online courses will support teachers by:

- Providing simulated practice for teachers, in a supported atmosphere of experimentation
  learning directly from 'experts'.
- Focusing on successful use of curriculum materials with feedback, which will be given by the course facilitators .
- Practicing using the strategies through simple tasks based on teachers' needs, interests and efficient time.

Content is in development. Its design considers the following components:

- TEACHING TOOLS teaching strategies illustrated with resources (Materials and Video Library)

- ACTIVITIES based on the overall outcomes for Adopt, which includes teachers‘ tasks.
- ASSESSMENT procedures for assessing teachers' learning

Timetable:

1. January 2015: Completion based on WP1 CPD framework and WP2 EDX specifications. WP4 will lead the dissemination and WP2 will open the registration
2. February/March 2015: Pilots and final improvements
3. April 2015: Online Courses will be ready to start based on pilot-partners country calendar

   The foreseen dates of the starting of MOOC’s are:

   - Adopt April 2015
   - Adapt October 2015
   - Transform March 2016

Evaluation of MOOC’s:

The evaluation of MOOC’s will be done with an online questionnaire. We will use pre-and post-test design to find out the impact of the MOOC’ s. In order to find out the impact of MOOC’ s, we need to know the starting levels of teachers. Therefore, the pre-and post-test design is more appropriate. For this purpose, TUDelft will prepare the questionnaire together with OU. Then, the opinions of experts and project partners will be gathered to revise and improve the questionnaire and also for content validity. The next step will be the translation and back translation of the questionnaire to/from partner languages. TUDelft will provide a guide for this purpose to the partners.

4. Evaluation of the Impact on Teachers & Students

The evaluation of impact will focus on three stakeholders:

- teachers
- students
- the other stakeholders

The Impact on teachers:

For the impact evaluation it is important to keep in mind the objectives of the project and indicators. The objectives that are related to teachers are summarized below in Table 3.

Table 3. Objectives of Engage project (teachers)

What to evaluate	Target group	Indicators (measures of success)	How to evaluate
Purpose (project goals) More **science teachers:** 2. are able to use RRI techniques with the support of our exemplar materials **(adopt)** 3. have begun a transition with significant change in either their beliefs, knowledge or classroom practice **(adapt)** 4. have made substantial changes to their beliefs, knowledge and classroom performance **(transform)**	Science teachers	2. At least 25% of teachers using the 'adopt' ENGAGE programme have achieved what we have defined as level 3 in the 'RRI integration' model i.e. that their integration of our 'Topicals' goes beyond motivational and casual use, and has a purposeful intention in their curriculum. 3. At least 25% of teachers using the 'adapt' ENGAGE programme have made a significant positive shift in at least 2 of the 5 'dimensions of RRI teaching' model. 4. At least 25% of teachers using the 'transform' ENGAGE programme have made a significant positive shift in at least 4 of the 5 'dimensions of RRI teaching' model	Questionnaire and/or semi-structured interviews

Developing teachers' practice involves changing their beliefs, knowledge and pedagogies. ENGAGE builds on existing models from the RRI literature to describe and quantify this impact.

In order to find out the impact of Engage programme on teachers we will gather the opinions of teachers about our three strategies (materials, community and courses strategies). Our instruments are online survey (mostly consisting of closed, pre-structured questions which offer respondents fixed alternative responses), semi-structured interviews and focus groups. We will repeat the evaluation procedure for three stages of CPD: adopt, adapt and transform.

The evaluation timetable is given in tables 4, 5, 6 and 7 below for the stages Adopt, Adapt and Transform respectively. These tables give an overview for 3 years and includes the evaluation instruments, the target and the planned time to use them. Following this table, the approaches and models which we are going to use for impact evaluation are summarized.

Evaluation Timetable:

Table 4. Evaluation timetable-ADOPT

Evaluation focus	Target	Measure of success	Evaluation Tool	Year 1	Year 2	Year 3
OER	Teachers	At least 25% of teachers achieved level 3 in RRI int. model	Online Questionnaire-Q1 (entire population) interview (small sample in each country) (if applicable to partners)	**Q1(v.1). Ready:** Nov. 2014 **Trans. & back trans. :** Dec. 2014 **Pilot:** Jan. 2015 **Survey online:** Feb. / March 2015 **Data anal.** : Apr.2015 **Report** : May 2015	**Q1(v.2):** Feb. 2015 **Trans. & back trans. :** March 2015 **Pilot:** Apr. 2015 **Survey online:** May 2015 **Data anal.** : June 2015 **Report** : July 2015	**Q1(v.3):** July 2016 **Trans. & back trans. :** Aug. 2016 **Pilot:** Sep. 2016 **Survey online:** Oct. 2016 **Data anal.** : Dec. 2016 **Report** :Jan. 2017
f2f workshops	Teachers		Questionnaire-Q2 (entire population) Interview (min. 3 teachers) Focus groups (if applicable)	**Q2 (v.1) Ready:** Dec.2014 /Jan. 2015 **Trans. & back trans. :** Jan. 2015 **Pilot :** Jan. 2015 **Interview quest.:** Jan. 2015 **Translation:** Jan. 2015 **Guide for focus groups**: Feb. '15 **Data analy.**: Apr. 2015 **Reporting:** May 2015		
MOOC's	Teachers		Pre-and post- test design - Online questionnaire (Q3)	**Q3 (v.1)Ready:** Dec. 2014 **Translation:** Jan.. 2015 **Pilot:** Jan. 2015 **Online:** Feb. 2015	**Q3 (v.2)Ready:** May 2015 **Translation:** June 2015 **Pilot:** July/Aug. 2015 **Online:** Sep. 2015	**Q3 (v.3)Ready:** Oct. 2015 **Translation:** Nov. 2015 **Pilot:** Dec. 2015 **Online:** Jan. 2016
Change from adopt to adapt	Teachers	RRI teaching model	25% achieve significant shift	Feb. 2015	May 2015	Feb. 2016
Impact	Teachers	5 dimension model	Questionnaire / Observation	Feb. 2015	May 2015	Feb. 2016
Impact	Teachers	Report	Global Report of Adopt stage	May/June 2015	Oct. 2015	Jan. 2014

Table 5. Evaluation Timetable-ADAPT

Evaluation focus	Target	Measure of success	Evaluation Tool	Year 2	Year 3
OER	Teachers	At least 25% of teachers make significant positive shift in at least 2 of the "5 dimensions of RRI teaching model"	Q1 (entire population) interview (small sample in each country) (if applicable to partners)	**Q1. Ready:** March 2015 **Trans. to partner lang.:** Apr. 2015 **Pilot testing:** Apr. 2015 **Survey is online:** May 2015 **Data anal.** : June 2015 **Report** : July 2015	**Q1 Ready:** July 2016 **Trans. to partner lang.:** Aug. 2016 **Pilot testing:** Sep. 2016 **Survey is online:** Oct. 2016 **Data anal.** : Nov. 2016 **Report** :Dec.2016
f2f workshops	Teachers		Q2 (entire population) Interview Focus groups	**Quest. ready:** Apr. 2015 **Translation:** May 2015 **Q2 Ready to use:** June 2015 **Interview quest.:** May 2015 **Translation:** May/June. 2015 **Guide focus groups**: May. '15 **Data analy.**: July/Aug. 2015 **Reporting:** Sep. 2015	
MOOC's	Teachers		Pre-and post- test design Online questionnaire	**Quest. Ready:** July 2015 **Translation:** Aug 2015 **Pilot:** Sep. 2015 **Online:** Oct. 2015	**Quest. Ready:** .Nov. 2015 **Translation:** Dec. 2015 **Pilot:** Jan. 2016 **Online:**Feb. 2016
transition	teachers		25% achieve significant shift in 2 of 5 RRI dimensions of RRI teaching model	June 2015	
Impact	Teachers	5 dimension model	Questionnaire / Observation	June 2015	
Impact	Teachers	Report	Global Report of Adapt	Nov. 2015	Dec.2016

Table 6. Evaluation Timetable-TRANSFORM

Evl. focus	Target	Measure of success	Evaluation Tool	Year 3
OER	Teachers	At least 25% made a significant positive shift in 4 of 5 dimensions of RRI teaching model	Online Questionnaire-Q1 (whole population) interview (small sample in each country) (if applicable to partners)	**Instr. Ready:** July 2016 **Trans. to partner lang.:** Aug. 2016 **Pilot testing:** Sep. 2016 **Survey is online:** Oct. 2016 **Data anal.** : Nov. 2016 **Report** : Dec. 2016
f2f workshops	Teachers		Q2 (entire population) Interview Focus groups	**Quest. ready:** Dec. 2015 **Translation:** Jan. 2016 **Q2 Ready to use:** Feb. 2016 **Interview quest.:** Dec. 2015 **Translation:** Jan. 2016 **Guide focus groups**: Jan. 2016 **Data analy.**: April. 2016 **Reporting:** May. 2016
MOOC's	Teachers		Pre-and post- test design Online questionnaire	**Quest. Ready:** .Dec. 2015 **Translation:** Jan. 2015 **Pilot:** Feb. 2016 **Online:** March 2016
Substantial Change via Transform	Teachers		RRI teaching model	Jan. 2015
Impact	Teachers	5 dimension model	Questionnaire / Observation	Jan. 2015
Impact	Teachers	Report	Global Report of Transform	Dec. 2016
Impact	Students	**Attitude indicators:** feel more confident about the process of reaching an informed viewpoint **Knowledge and skill indicators:** be aware of the wider context of RRI issues, critically analyse sources of evidence, evaluate a technology, use ethical reasoning, and argue with reasoning. **Behaviour indicators**: read, watch or discuss something related to science and technology outside the classroom	Online Survey	**Survey (v.1):** .Jan. 2016 **Trans.&back trans. :** Feb. 2016 **Pilot:** March 2016 **Online:** April 2016

Table 7. Evaluation Timetable-Other Stakeholders

Evaluation focus	Target	Measure of success	Evaluation Tool	When
Impact	Students	Opinions of students	Online Questionnaire (whole population)	**Instr. Ready:** July 2016 **Trans. to partner lang.:** Aug. 2016 **Pilot testing:** Sep. 2016 **Survey is online:** Oct. 2016 **Data anal.** : Nov. 2016 **Report**: Dec. 2016
The influence	Partners	Opinions of project partners	Interview & Roundtable discussions	At partner meetings
The influence	scientists	Opinions of scientists	Online survey & interviews	**EXTERNAL EVALUATION**
The influence	Countries other than partner countries	The impact of ENGAGE extends beyond targeted countries, and after the project ends		**EXTERNAL EVALUATION**

RRI integration model for Adopt step impact:

We will use part of the model of Solomon & Aikenhead (1994), to describe the degree to which science and society content is integrated with traditional science content.

Level 1) Little RRI content – for motivational purposes (at this category little learning about RRI content would be expected; motivation by RRI content holds little promise for meeting RRI objectives).
Level 2) Casual infusion of more RRI content but without coherent purpose
Level 3) A purposeful infusion of RRI content giving even more time to RRI

We consider that fully reaching the 'Adopt' step is equivalent to teachers at Level 2 in this model, because this first step is just about usage.

Dimensions of change model for Adapt/Transform steps impact:

Reference:

Solomon, J. & Aikenhead, G. (Eds.) (1994). STS education: International perspectives on reform. New York: Teachers College Press.

The impact needs to address three aspects of how teachers operate: their beliefs, their knowledge and their skills (Bell and Gilbert, 2004). Of these, beliefs are probably the most important: “The cornerstone for successful change … must be that new knowledge can only be successfully coupled to new actions in the classroom if a teacher’s underlying beliefs are supportive of the desired change. Beliefs are the key to change” (Gilbert, 2010).

To evaluate all three areas is a complex undertaking but can be simplified using a model devised by Ratcliffe & Grace (2003; 159). They researched teaching about the nature of science, its practices and processes, and identified 5 dimensions of practice (table 8), which consist of two extremes.

Table 8. The five dimensions of practice seen in teaching the nature of science

NOVICE RRI TEACHER	DIMENSION	*EXPERT RRI TEACHER*
	1. Teachers' knowledge of RRI	
Anxious about understanding	<---------->	Confident that they have a sufficient understanding
	2. Teachers' conception of their own role	
Dispenser of knowledge	<---------->	Facilitator of learning (interpreted for ENGAGE as building on students' existing ideas)
	3. Teachers' use of RRI pedagogies like discourse	
Closed and authoritative	<---------->	Open and dialogic
	4. Teachers' conception of learning goals	
Limited to knowledge gains	<---------->	Includes the development of reasoning skills
	5. The nature of classroom activities	
Student activities are contrived and inauthentic	<---------->	Activities are authentic and owned by students

Source: Ratcliffe & Grace (2003; 159), Bartholomew et al (2004; 664)

Their model fits with ENGAGE for several reasons:

1. It specifically includes the 3 areas of knowledge we are attempting to foster: pedagogies for RRI-teaching (dimension 3), building on students' ideas (dimension 2), RRI content (dimension 1).
2. It is consistent with the pedagogy of IBSE, aimed at developing a skillset as much as content (dimension 4)
3. It recognises the importance of 'authentic materials' like the ENGAGE Topicals, sequences and projects, as characteristic of RRI-based teaching (dimension 4)

ii. For our purposes, we have slightly amended dimension 2 to make more explicit the knowledge of students' existing ideas - which has been shown to be critical factor in successful CPD approaches (see Guskey & Huberman 1995), in student learning.

5 dimensions to the left hand side describes less successful RRI teaching, and the right hand side more successful RRI teaching.

Following Ratcliffe & Grace (2003), we will use this 5 dimension model as a measure of 'RRI teaching'. Evaluation will turn each dimension into 5 steps with statements to determine a teacher's position, for example:

Dimension1: Teachers' knowledge and understanding of nature and science
It's too much to expect that teachers gain detailed knowledge of all the aspects of socio-scientific issues. / Teachers need a realisation of the extent of their understanding of socio-scientific issues, and appropriate pedagogical knowledge and skills

Dimension 2: Teachers' conceptions of their own role
Instruction must provide experiences and information from which learners can build new knowledge/ Teacher should provide opportunities for students to learn key concepts and discover the tools that they need for learning.

Dimension 3: Teachers' use of discourse
The teacher is a transmitter of knowledge and should provide necessary information to students / I believe that students learn more, remember it longer, and apply it to new situations better if they learn through experience, rather than through being told/ the teacher is a transmitter of knowledge.

Dimension 4: Teachers' conceptions of learning goals
It is important to teach students all the content of science curriculum / It is important to develop students' abilities such as critical thinking, analytic thinking and asking questions.

Dimension 5: the nature of classroom activities

Teacher should prepare the learning activities or should control the activities which are prepared by students and make students to use them / Students should choose/prepare activities according to their ideas.

Evaluation of The Impact on Students

We have defined a set of assessable outcomes on students from two directions:

1. Our RRI curriculum, the set of knowledge and skill areas relating to the nature of science and its impact on society, as taught through ENGAGE Materials.
2. What students need to believe, think, and do: "How do you make a decision about the use of science and technology in the context of conflicting expert advice?" (Aikenhead, 2006, p98).

There are three areas of impact we expect on the 88,000 students whose teachers have reached the Transform stage (see project Impact on students, p.72, DoW):

Attitudes:

1. Students will feel more confident about exploring an issue about new science and technology, explaining it to others, and in persuading others of their point of view (Reiser & Berland, 2008).
2. Students will adopt an unbiased attitude towards emerging technologies where knowledge is uncertain, emotion responses with rational reflection

Knowledge and skills:

3. Evidence: know how to ask pertinent questions, to obtain evidence, to use it as the basis for decision making, to understand the limitations of scientific evidence, to have criteria for critically judging the credibility of scientists and media sources
4. Technology: understand concepts of probability and risk, know how to identify benefits and drawbacks and weigh these up

5. Values: understand the social processes engaged in by scientists, to understand the institutional aspects of science, to identify the value positions held by scientists Argumentation: be able to marshal data and reasoning in support of a claim, being able to anticipate and deal with rebuttals.

Behaviour:

- Students will be more likely to read, watch or discuss an issue related to RRI outside the classroom.

For each of these knowledge and skills objectives we will use the progression of outcomes set out in the US Next Generation Curriculum Framework (NRC, 2012), which allows us to quantify to what extent students have made progress.

The overall objective of the project is to equip future students for future RRI participation. Even though the interventions are on teachers, and the student impact is largely indirect, we believe it is critical to define this impact. The literature is clear that the most effective CPD is that which is directly targeted at improving student learning (Guskey, 2000).

We will evaluate the impact on the students on the last year of the project. Since, we aim to have impact on students through their teachers; some time is needed to see this impact. The valuation of student impact is necessary if ENGAGE is to leave a legacy. In order to evaluate the impact on students we will use online survey by using our professional online survey software (Survey Gizmo).

In the process of evaluation of impact on students we will take into account the project objectives and indicators that are related to students. These objectives and indicators are summarized below in table 9.

Table 9. Objectives and indicators (Students)

What to evaluate	Target group	Indicators (measures of success)	How to evaluate
Overall objective: 1. Students acquire the knowledge and skills, attitudes and behaviours that allow them to engage effectively with RRI-based science as future citizens and in their own lives.	students	As a result of the ENGAGE programme, students of teachers using the 'transform' programme will on average be more likely to show a positive change in one indicator from Knowledge/Skill, Attitudes and Behaviour **Attitude indicators:** feel more confident about the process of reaching an informed viewpoint **1.1 Knowledge and skill indicators:** be aware of the wider context of RRI issues, critically analyse sources of evidence, evaluate a technology, use ethical reasoning, and argue with reasoning. Behaviour indicators read, watch or discuss something related to science and technology outside the classroom	Questionnaire • Self-evaluation of the students against mentioned indicators (on the 2nd column) • Teachers evaluate their students against the mentioned indicators

5. The Influence on the other stakeholders (T8.3)

This will form part of the external evaluation plan which is outlined in further detail in D8.17, but summarised below.

Evaluation tool: Online surveys and selective interviews with a sample of key stakeholders
Target: stakeholders (pre-service and in-service teachers, scientists, countries other than partner countries of the project)

In order to find out the influence on the other stakeholders (pre-service and in-service teachers, scientists, countries other than partner countries of the project) we will develop a range of appropriate methods. These are likely to include online surveys of

scientists in Year 3 (transform phase) and selective interviews/Flash meetings with other key stakeholders. The indicators for the evaluation of the Engage project on the stakeholders are given in table 10 below.

Table 10. Stakeholders and indicators for evaluation

What to evaluate	Target group	Indicators (measures of success)	How to evaluate	Remarks
'Main groups' of other **stakeholders:** 5. As a result of exposure to the project, more pre- and in-service trainers include RRI practice 6. Scientists are more confident and capable to engage with teachers and students 7. The impact of ENGAGE extends beyond targeted countries, and after the project ends	Stakeholders (pre-service and in-service teachers, scientists, countries other than partner countries of the project)	5. Most pre-service teachers and training providers surveyed believe that ENGAGE Materials have helped them towards 'RRI teaching' 6. At least 50% of scientists acting as 'RRI experts' in our Projects feel more confident about interacting with teachers and students in discussion of socio-scientific issues. 7. Teachers in the Netherlands, Italy and Portugal – outside the Programme – who use Materials rate them as 'very useful' in their teaching.	online questionnaire small number of follow-up telephone/Flash meeting interviews	Since the potential number of stakeholders in this group could be very high this can be a short online questionnaire.

During Year 2, all partners will be asked to identify a sample of pre- and in-service teachers and scientists in their respective countries. Contact lists of stakeholders' names and email addresses will then be used as the basis for designing a sampling frame that will ensure an appropriate representation of stakeholder characteristics. Depending on the numbers and nature of these, the samples could be purposive or randomised. Online surveys will be designed and translated in time for distribution to the sample of stakeholders during the transform phase in Year 3. Questions will focus on their awareness, use/engagement and views on the RRI materials.

Online respondents would be asked to indicate whether they would be willing to take part in a follow up semi-interview that would be conducted by telephone or via Flash meeting. Depending on the depth and richness of evidence required, arrangements would be made to gather this additional information from a smaller sample of each stakeholder group in each country (including, as appropriate, stakeholders in the Netherlands, Cyprus and Italy). The aim of this would be to explore in further detail, the extent to which the Engage programme and materials have made a lasting and sustainable change to the ongoing development of teacher training and science education in each of the partner countries and beyond.

6. The impact on partners

Evaluation tool: semi-structured face to face interviews at project meetings, roundtable discussions, skype or flash meeting interviews
Target of questionnaire: Partners

At the end of each year, the external evaluator will be responsible for gathering feedback from partners, about the progress of the WPs they have been involved in, their activities, outcomes, achievements and views of the communication and management of Engage. In addition, they will be asked to identify key issues requiring action as the programme progresses – overall and with regard to the adopt, adapt and transform processes. Further questions on impact will be developed for Years 2 and 3, as the measurable outputs of the programme become evident Annual reports and internal documents with analysis, suggestion for actions, to be used for continuous improvement of the professional learning programmes.

In the report, we can also include information about expected outcomes of the projects and how far could we realized them. Table 11 summarizes expected outcomes and key activities.

Table 7. Project outcomes and key activities

Results (project outcomes)	Check Points
i) Large number of students are exposed to RRI-teaching	a. 2 million students aged 11-16, across partner countries who are reached at least once through the Programme, through their teacher's use of ENGAGE Materials
ii) Large number of science teachers have taken the 'adopt' step	ii) 11750 teachers have used our topical Materials, and used the online 'just in time' content, in our Community. 645 teachers have attended ENGAGE Workshops
iii) A significant proportion has taken the 'adapt' step	iii) 3020 teachers have used our teaching sequences 570 teachers have enrolled in an online Courses
iv) A proportion of these go further to engage in creative, mentoring and training practices transform sage	(iv) 480 teachers have used a Project, and partnered with an RRI scientist/communicator 150 teachers have taken a mentoring role in our community 200 teachers have inducted colleagues using our RRI training pack.
(v) ENGAGE materials, videos of RRI pedagogies, and courses are integrated in pre- and in-service training.	(v) 2 pre-service and 2 in-service training providers have integrated elements of ENGAGE, within each country.
(vi) The scientific community is influenced to gain a greater understanding of engaging with teachers and students	(vi) 100 scientists and science communicators take part in school partnerships in our Projects
vii) Project outputs are showcased at a celebration event, captured and published on Knowledge Hub and teachers continue to adopt the Programme after the project ends.	(vii) 100 projects are showcased, to 100 main stakeholders, at the high profile RRI festival, in each country The rate of new teacher adoption of Materials at the end of Year 3, matches previous years. 1 training provider per country has taken up either our online courses or Workshops, so that they continue. Policy report is circulated to at least 1000 key recipients throughout EU.
Key activities (Strategies) a) Materials Strategy, producing curriculum materials for classroom experimentation for Adopt, Adapt and Transform teachers	(a) 20 Topicals produced over 2 years: 10 by month 18, 20 by month 30) over 2 years (see Materials). 20 Sequences, produced over 2 years: 10 by month 18, 20 by month 30)over 2 years 20 'Projects' produced, and experts sourced over 1.5 years: 10 by month 24, 20 by month 30)
b) Community Strategy: establish and support an online community of practice (CoP)	(b) 20 sets of 'just in time' online content produced, 10 by month 18, 20 by month 30) over 2 years (see Community) 100 'expert RRI teachers recruited as mentors who facilitate discussions
c) Courses Strategy: to provide Workshops and online modules.	(c) 33 ENGAGE Workshops, in 11 countries (1 per year) 154 online courses (module 1: 2 x per year, over 3 years, module 2 and 3,, 2x per year, over 2 years,, in 11 countries
(d) Dissemination Strategy: promoting ENGAGE to main groups of stakeholders, to teachers beyond partner countries, and to ensure the project legacy.	(d) 11 'RRI festivals,' high profile events to promote ENGAGE to new teachers. Localised dissemination plans for each country, with targets for each main stakeholder group. Viable exploitation plan, to show how all elements of ENGAGE can be embedded within other repositories and networks, or accessed by for continued hosting and maintenance of ENGAGE Knowledge Hub. Publication of short policy focused report, circulated amongst key governmental and other stakeholders.

5. REFERENCES

Aikenhead, G. S. (2006). Science education for everyday life: Evidence-based practice. New York: Teachers College Press.

Bartholomew, H., Osborne, J. & Ratcliffe, M. (2004).Teaching students 'ideas-about-Science': Five dimensions of effective practice. Science Education, 88(2), 655-682.

Boston, C. (2002). The Concept of Formative Assessment. ERIC Digest.

Developing an Effective Evaluation Plan (DEEP). Atlanta, Georgia: Centers for Disease Control and Prevention, National Center for Chronic Disease Prevention and Health Promotion, Office on Smoking and Health; Division of Nutrition, Physical Activity, and Obesity, 2011.

Drever, E. (1995). Using Semi-Structured Interviews in Small-Scale Research. A Teacher's Guide. Scottish Council for Research in Education, Edinburgh.

Elias, T. (2011). Learning Analytics: Definitions, Processes and Potential. Retrieved December 2 from http://learninganalytics.net/LearningAnalyticsDefinitionsProcessesPotential.pdf

National Oceanic and Atmospheric Administration U.S. Department of Commerce (NOAA) (2009). Designing Education Projects a Comprehensive approach to Needs Assessment, Project Planning and Implementation, and Evaluation. 2nd edition. BridgeWater Education Consulting, LLC.

Ratcliffe, M., & Grace, M. (2003). *Science Education for Citizenship. Teaching Socio-Scientific Issues.* Maidenhead: Open University Press.

Scriven, M. (1991). Evaluation thesaurus (4th ed.). Newbury Park, CA: Sage.

Solomon, J. & Aikenhead, G. (Eds.) (1994). STS education: International perspectives on reform. New York: Teachers College Press.

Stake, R. E. (1976). Evaluating Educational Programmes: The Need and the Response: A Collection of Resource Materials. OECD. ISBN 9264115358.

Western Australian Planning Commission (WAPC) (2003). Coastal Planning and Management Manuel. Retrieved from http://www.planning.wa.gov.au/dop_pub_pdf/5_project_evaluation.pdf on 23.06.2014.

Worthen, B.R., Sanders, J.R., Fitzpatrick, J.L. (1997). Program Evaluation: Alternative Approaches and Practical Guidelines. (2nd Ed). White Plains, NY: Addison Wesley Longman.

ANNEX 1: Evaluation Methods

- **Questionnaires**

Questionnaires can be conducted in person, by telephone, or by mail. They are used to quickly obtain

information from a wide variety of people. Questionnaires are typically inexpensive, can be completed anonymously, and are easy to compare and analyse. It is possible to involve many people, although it may only appeal to a certain section of the community, and responses from certain sections of the community may be limited.

Useful further reading can be found in De Vaus, D. (1995) Surveys in social research.

- **Semi-structured interviews – In person**

Semi-structured interviews are informally guided. Some questions are predetermined and new questions are developed from the discussion. Questions are mainly open, providing an opportunity for the respondent to provide an opinion.

Semi-structured interviews are used to understand an interviewee's experiences and impressions. Open question example: How effective do you believe the group has been in addressing the original objectives of the project?

Patton (1990) Qualitative methods in Evaluation and Research provides further reading on semi-structured interviews.

- **Participant observation**

Data is collected by listening, watching, and documenting what is seen and heard. Through asking questions, and by noting comments, behaviours and reactions, useful information is provided to the evaluation process. The participant observation method gathers accurate information about how a group and project operates in the field.

ANNEX 2: Questionnaires

1) Questionnaire for the evaluation of the materials and the activities – Teachers' opinions (Q1)
2) Questionnaire for the evaluation of f2f workshops (Q2)
3) Questionnaire for the evaluation of MOOC' s – (Q 3)

Q1. Questionnaire for the evaluation of the Engage materials strategy – Teachers' opinions

Target of questionnaire: Teachers

Aim: To gather the opinions of teachers about the materials (topicals) of ENGAGE project.

Analysis: descriptive statistics and graphs + content analysis

About You:

1. Name of school: ……
2. What is your gender: ☐ female ☐ male
3. What is your age?
   ☐ Under 25 ☐ 25- 29 ☐ 30- 39 ☐ 40 -49 ☐ 50 -59 ☐ 60+
4. What is the highest level of formal education that you have completed?
   ☐ Post-secondary non-tertiary education ☐ Short-cycle tertiary education (at least 2 years)
   ☐ Bachelor's or equivalent level ☐ Master's or equivalent level ☐ Doctoral or equivalent level
5. What subject is your highest qualification?
   ☐ Physics ☐ Mathematics ☐ Chemistry ☐ Biology ☐ Earth Science ☐ General ☐ Other
6. How long have you been working as a teacher?
   ☐ This is my first year ☐ 1-2 years ☐ 3-5 years ☐ 6-10 years ☐ 11-15 years ☐ 16-20 years ☐ more than 20 years
7. Do you have responsibilities other than teaching?
   ☐ School head ☐ Vice-school head ☐ Head of department ☐ Head teacher
   ☐ Curriculum coordinator ☐ Coordinator of a specific grade level
8. How long have you been teaching science?
   ☐ This is my first year ☐ 1-2 years ☐ 3-5 years ☐ 6-10 years ☐ 11-15 years ☐ 16-20 years ☐ more than 20 years

About Engage Materials:

1. Engage Materials used at……..
   - ☐ Primary school
   - ☐ General secondary school
   - ☐ Vocational or technical secondary school

2. Have you used inquiry-based teaching before? □ Yes □ No
   If "yes", How? ……

3. Which material(s) did you use? (you can check more than one choice)
   - ☐ Sinking island
   - ☐ Ban the beds
   - ☐ Take the test?
   - ☐ Attack the giant viruses
   - ☐ Ban Cola?
   - ☐ Car wars
   - ☐ Grow your own body
   - ☐ Three parents
   - ☐ What does the fox say?

4. How many students totally were involved in the activity/activities;
   ☐ less than 30 ☐ 30 to 44 ☐ 45 to 54 ☐ 55 to 64 ☐ more than 65

5. What is the average age of the students who took part;

Sinking island	☐ 11	☐ 12	☐ 13	☐ 14	☐ 15	☐ 16
Ban the beds	☐ 11	☐ 12	☐ 13	☐ 14	☐ 15	☐ 16
Take the test?	☐ 11	☐ 12	☐ 13	☐ 14	☐ 15	☐ 16
Attack the giant viruses	☐ 11	☐ 12	☐ 13	☐ 14	☐ 15	☐ 16
Ban Cola?	☐ 11	☐ 12	☐ 13	☐ 14	☐ 15	☐ 16
Car wars	☐ 11	☐ 12	☐ 13	☐ 14	☐ 15	☐ 16
Grow your own body	☐ 11	☐ 12	☐ 13	☐ 14	☐ 15	☐ 16
Three parents	☐ 11	☐ 12	☐ 13	☐ 14	☐ 15	☐ 16
Making Decisions	☐ 11	☐ 12	☐ 13	☐ 14	☐ 15	☐ 16
What does the fox say?	☐ 11	☐ 12	☐ 13	☐ 14	☐ 15	☐ 16
Eat insects	☐ 11	☐ 12	☐ 13	☐ 14	☐ 15	☐ 16
Solar Roadways	☐ 11	☐ 12	☐ 13	☐ 14	☐ 15	☐ 16
Ebola	☐ 11	☐ 12	☐ 13	☐ 14	☐ 15	☐ 16

6. To what extent was the teacher guide of the material useful to you?

Sinking island	☐ 1	☐ 2	☐ 3	☐ 4	☐ 5	☐
Ban the beds	☐ 11	☐ 12	☐ 13	☐ 14	☐ 15	☐ 16
Take the test?	☐ 11	☐ 12	☐ 13	☐ 14	☐ 15	☐ 16
Attack the giant viruses	☐ 11	☐ 12	☐ 13	☐ 14	☐ 15	☐ 16
Ban Cola?	☐ 11	☐ 12	☐ 13	☐ 14	☐ 15	☐ 16
Car wars	☐ 11	☐ 12	☐ 13	☐ 14	☐ 15	☐ 16
Grow your own body	☐ 11	☐ 12	☐ 13	☐ 14	☐ 15	☐ 16
Three parents	☐ 11	☐ 12	☐ 13	☐ 14	☐ 15	☐ 16
Making Decisions	☐ 11	☐ 12	☐ 13	☐ 14	☐ 15	☐ 16
What does the fox say?	☐ 11	☐ 12	☐ 13	☐ 14	☐ 15	☐ 16
Eat insects	☐ 11	☐ 12	☐ 13	☐ 14	☐ 15	☐ 16
Solar Roadways	☐ 11	☐ 12	☐ 13	☐ 14	☐ 15	☐ 16
Ebola	☐ 11	☐ 12	☐ 13	☐ 14	☐ 15	☐ 16

7. To what extent did the materials help you in the teaching process?

Sinking island	☐ 1 (not at all)	☐ 2	☐ 3	☐ 4	☐ 5 (great extent)	☐ does not apply
Ban the beds	☐ 1 (not at all)	☐ 2	☐ 3	☐ 4	☐ 5 (great extent)	☐ does not apply
Take the test?	☐ 1 (not at all)	☐ 2	☐ 3	☐ 4	☐ 5 (great extent)	☐ does not apply
Attack the giant viruses	☐ 1 (not at all)	☐ 2	☐ 3	☐ 4	☐ 5 (great extent)	☐ does not apply
Ban Cola?	☐ 1 (not at all)	☐ 2	☐ 3	☐ 4	☐ 5 (great extent)	☐ does not apply
Car wars	☐ 1 (not at all)	☐ 2	☐ 3	☐ 4	☐ 5 (great extent)	☐ does not apply
Grow your own body	☐ 1 (not at all)	☐ 2	☐ 3	☐ 4	☐ 5 (great extent)	☐ does not apply
Three parents	☐ 1 (not at all)	☐ 2	☐ 3	☐ 4	☐ 5 (great extent)	☐ does not apply
Making Decisions	☐ 1 (not at all)	☐ 2	☐ 3	☐ 4	☐ 5 (great extent)	☐ does not apply
What does the fox say?	☐ 1 (not at all)	☐ 2	☐ 3	☐ 4	☐ 5 (great extent)	☐ does not apply
Eat insects	☐ 1 (not at all)	☐ 2	☐ 3	☐ 4	☐ 5 (great extent)	☐ does not apply
Solar Roadways	☐ 1 (not at all)	☐ 2	☐ 3	☐ 4	☐ 5 (great extent)	☐ does not apply
Ebola	☐ 1 (not at all)	☐ 2	☐ 3	☐ 4	☐ 5 (great extent)	☐ does not apply

8. Why did you choose to use this/these material(s)? (you can click more than one choice)

□ The case was interesting for the students

□ It was fun to discuss the subject

□ It was related to school curriculum

□ It was easy to use in the class

□ Other……

9. What did you <u>not</u> like about this/these material(s)? (you can click more than one choice)

□ It was boring

□ It was too easy

□ It was too difficult

□ It was not relevant to school curriculum

□ It required prior knowledge which students do not have yet

□ It required long preparation time

□ Nothing, I liked all of it

□ Other……

10. Would you use these material(s) again?

□ Yes □ No □ Don' t know

Would you like to tell, why?.....

11. How interesting are the contents of the materials to students in your opinion?

Sinking island	☐ 1 (not at all)	☐ 2	☐ 3	☐ 4	☐ 5 (great extent)	☐ does not apply
Ban the beds	☐ 1 (not at all)	☐ 2	☐ 3	☐ 4	☐ 5 (great extent)	☐ does not apply
Take the test?	☐ 1 (not at all)	☐ 2	☐ 3	☐ 4	☐ 5 (great extent)	☐ does not apply
Attack the giant viruses	☐ 1 (not at all)	☐ 2	☐ 3	☐ 4	☐ 5 (great extent)	☐ does not apply
Ban Cola?	☐ 1 (not at all)	☐ 2	☐ 3	☐ 4	☐ 5 (great extent)	☐ does not apply
Car wars	☐ 1 (not at all)	☐ 2	☐ 3	☐ 4	☐ 5 (great extent)	☐ does not apply
Grow your own body	☐ 1 (not at all)	☐ 2	☐ 3	☐ 4	☐ 5 (great extent)	☐ does not apply
Three parents	☐ 1 (not at all)	☐ 2	☐ 3	☐ 4	☐ 5 (great extent)	☐ does not apply
Making Decisions	☐ 1 (not at all)	☐ 2	☐ 3	☐ 4	☐ 5 (great extent)	☐ does not apply
What does the fox say?	☐ 1 (not at all)	☐ 2	☐ 3	☐ 4	☐ 5 (great extent)	☐ does not apply
Eat insects	☐ 1 (not at all)	☐ 2	☐ 3	☐ 4	☐ 5 (great extent)	☐ does not apply
Solar Roadways	☐ 1 (not at all)	☐ 2	☐ 3	☐ 4	☐ 5 (great extent)	☐ does not apply
Ebola	☐ 1 (not at all)	☐ 2	☐ 3	☐ 4	☐ 5 (great extent)	☐ does not apply

12. Was the activity at the right level of difficulty for most of the students?

Sinking island	☐ 1 (too easy)	☐ 2 (about right)	☐ 3 (adequate)	☐ 4 (too hard)	☐ 5 does not apply	☐ 1 (too easy)
Ban the beds	☐ 1 (too easy)	☐ 2 (about right)	☐ 3 (adequate)	☐ 4 (too hard)	☐ 5 does not apply	☐ 1 (too easy)
Take the test?	☐ 1 (too easy)	☐ 2 (about right)	☐ 3 (adequate)	☐ 4 (too hard)	☐ 5 does not apply	☐ 1 (too easy)
Attack the giant viruses	☐ 1 (too easy)	☐ 2 (about right)	☐ 3 (adequate)	☐ 4 (too hard)	☐ 5 does not apply	☐ 1 (too easy)
Ban Cola?	☐ 1 (too easy)	☐ 2 (about right)	☐ 3 (adequate)	☐ 4 (too hard)	☐ 5 does not apply	☐ 1 (too easy)
Car wars	☐ 1 (too easy)	☐ 2 (about right)	☐ 3 (adequate)	☐ 4 (too hard)	☐ 5 does not apply	☐ 1 (too easy)
Grow your own body	☐ 1 (too easy)	☐ 2 (about right)	☐ 3 (adequate)	☐ 4 (too hard)	☐ 5 does not apply	☐ 1 (too easy)
Three parents	☐ 1 (too easy)	☐ 2 (about right)	☐ 3 (adequate)	☐ 4 (too hard)	☐ 5 does not apply	☐ 1 (too easy)
Making Decisions	☐ 1 (too easy)	☐ 2 (about right)	☐ 3 (adequate)	☐ 4 (too hard)	☐ 5 does not apply	☐ 1 (too easy)
What does the fox say?	☐ 1 (too easy)	☐ 2 (about right)	☐ 3 (adequate)	☐ 4 (too hard)	☐ 5 does not apply	☐ 1 (too easy)
Eat insects	☐ 1 (too easy)	☐ 2 (about right)	☐ 3 (adequate)	☐ 4 (too hard)	☐ 5 does not apply	☐ 1 (too easy)
Solar Roadways	☐ 1 (too easy)	☐ 2 (about right)	☐ 3 (adequate)	☐ 4 (too hard)	☐ 5 does not apply	☐ 1 (too easy)
Ebola	☐ 1 (too easy)	☐ 2 (about right)	☐ 3 (adequate)	☐ 4 (too hard)	☐ 5 does not apply	☐ 1 (too easy)

13. What do you think of the design of the materials?

☐ 1 (very poor) ☐ 2 ☐ 3 ☐ 4 ☐ 5 (very good)

14. By using these materials, to what extent were the following skills of your students improved?

Be able to question and evaluate the evidence for a scientific claim

□ 1 (not at all)	□ 2	□ 3	□ 4	□ 5 (great extent)	□ does not apply

Be able to analyse issues, apply knowledge, come to reasoned opinions, express these clearly, and consider possible actions

□ 1 (not at all)	□ 2	□ 3	□ 4	□ 5 (great extent)	□ does not apply

Critically evaluate the strength of the evidence for a claim about emerging science/technology, from a media report

□ 1 (not at all)	□ 2	□ 3	□ 4	□ 5 (great extent)	□ does not apply

Come to an informed opinion on a life, community or society decision, taking into account scientific and other perspectives

□ 1 (not at all)	□ 2	□ 3	□ 4	□ 5 (great extent)	□ does not apply

Argue for his/her opinion on a socio-scientific issue

□ 1 (not at all)	□ 2	□ 3	□ 4	□ 5 (great extent)	□ does not apply

Evaluate possible solutions to science/technology problems, developing criteria, applying knowledge, and using data

□ 1 (not at all)	□ 2	□ 3	□ 4	□ 5 (great extent)	□ does not apply

15. To what extent did the materials enrich the way you teach science?

□ 1 (not at all)	□ 2	□ 3	□ 4	□ 5 (great extent)	□ does not apply

16. Do you use discussion as a part of your teaching method?

□ Yes □ No

If so, how do you use? ……

17. Does your teaching methods allow students to arguing, analysing, questioning and communicating?

□ Yes □ No

If yes, How? Could you please explain? .....

18. Will you recommend these materials to your colleagues?

□ Yes □ No □ Don' t know

19. Do you have any suggestions/proposals/comments? ……

Would you like to receive e-mails about new materials, trainings, courses, etc…

Your e-mail: ……

This is the end of the questionnaire…

Thank you very much for your contribution…

ANNEX 3: Back Translation Guide

A back translation is when a translated document is translated (back) into the original language. The idea is that the author can verify whether the translation covers all aspects of the original.

How to Perform Back Translation

1. Original questionnaire is in English
2. Translate the original questionnaire into local language (translators should be native speakers of the target language)
3. If several translators are involved, one of them evaluates the translations and reconciles' these into one version.
4. One or more back translators (a different person than the one in step 2) translate the reconciled version back into English without having access to the original questionnaire.
5. If several back translators are involved, one of them evaluates the back translations and reconciles these into one version.
6. A group of bilingual reviewers evaluate the back translation and compare it with the original questionnaire.
7. The reviewer needs to be aware that the two versions are not going to be the same; if they were, that could be a sign that the back translator had access to the original. Any discrepancies between the original version and the reconciled back translation can best be resolved by having the author discuss these, with those who have performed the reconciliation of the translations and of the back translations.
8. The translation can then be accepted for further testing.
9. Send it to TUDelft (Dury)

P.S.

The back translator must be aware that the job concerns a back translation. This means he will stick to the source text a bit more than normal. It is especially important that small errors or weak sentences in the text are not 'ironed' away, which he otherwise would do. The main point of a back translation is obviously to find shortcomings in the original translation, and any errors must be marked and commented.

Reference: Andriesen, S. (2008). Benefiting from Back Translations. http://www.medilingua.com/pdf/BackTranslationsICTSummer%202008.pdf

ANNEX 4: Classroom Testing and Observation Report

WP1 CLASSROMM TESTING AND OBSERVATION REPORT (T1.4)

The aim of this task is to organize the first testing of the ENGAGE approach and materials. The testing is taking place in the period May to August in four countries, namely UK, Israel, Norway and Cyprus. The main objectives are a) to test the first developed materials under real teaching situation in different European countries b) get feedback on the current developments and collect ideas for further developments and start disseminating the ENGAGE Project amongst different teaching communities all-over Europe.

In the first six months of ENGAGE the testing approach has been outlined and discussed with all involved partners. The basic approach and aims of the testing have been defined collaboratively and a testing roadmap was agreed. Along with the organisation of the testing a formative evaluation approach has been discussed, developed and approved in order to collect as much feedback as possible for the further development of ENGAGE. Four evaluation instruments have been developed and partially translated in the testing for collecting the data, more specifically these are: Observation guidelines and results form, interview guidelines and questions, teachers' questionnaire and students' questionnaire. For supporting the evaluation results analysis, an evaluation matrix has been developed for a structured and uniform analysis of the evaluation results in the four countries.

Description of the Evaluation Instruments: As described above four evaluation instruments have been implemented: observation guidelines, semi-structured interview guidelines and a questionnaire for teachers, and students' questionnaire.

The Teachers questionnaire is divided in three main parts:

1. General information on the activity, materials used student's age and number, school or organisation performing the activity, etc.
2. The second part focuses on the pedagogical added-value of the performed activity (interesting, engaging, importance, acceptance of students, easiness to use, etc.), as well as challenges raised during teaching and learning.
3. The third part of the questionnaire focuses on suggestions for improvements and further ideas of interesting topics.

The Teachers interview guidelines: provided a set of questions (semi-structured) to be able to make comments on:

- Contents of interview (a) transliteration and b) report based on the structure of the profile protocol)
- Added value for teaching
- The view of the teacher
- The view of the students
- Suggestions for improving existing tools
- Suggestion for future developments

The Student's questionnaire focuses on questions (likert-scaled and open) if they liked the lesson, why and what they liked and found it engaging and want to make similar activities more frequently.

The Observation guidelines are divided in four parts:

1. General information on the activity, materials used student's age and number, school or organisation performing the activity, teaching subjects, teaching methods.
2. Databased observations taken throughout the activity about how the activity was set up, circumstances and reactions of teacher and students, discussion topics, etc.
3. Subjective impressions and questions of the observer: interpretation of the databased observations
4. Summary analysis of the observation in form of comments by the observer and 8 likert-scaled questions.

The project will rely on much quantitative data early in the roll-out, in terms of numbers of downloads, as an indicator of how well the Adopt stage is going.Therefore it makes sense to collect this data now, and identify any problems with our assumptions early. The UK partner established a large mailing list, and have started weekly mailings. Two materials are being published every week, to give multiple opportunities for emails. In July, when teachers have had more time to start using the Materials, they will be encouraged to submit reviews, or online feedback, to give their views on the Material quality. The testing and evaluation process is taking place in-line with the objectives outlined in the DoW. At current no problems or deviations from the DoW are expected.

The main aim of the early piloting is to undertake necessary corrective measures and improve the development of further materials. Therefore six classroom activities (described above) were used for piloting the initial project developments and focus on three key questions aimed at investigating both the usefulness and quality of the material. Additionally the piloting aims at generating new ideas for improving the activities and developing new ideas or to make it easy to use:

1. Do teachers and students like the ENGAGE material and what do they like?
2. What could be improved?
3. What could be other interesting topics?

Materials and Procedure of the Pilot Study

In this very early phase of the ENGAGE project six classroom activities have been developed and introduced for the pilot in the four participating countries (Norway, UK, Cyprus and Israel). These first classroom activities focus on in general interesting and challenging topics: "Ban Coke", the implications of coke and sugar consume; "Grow your own body", growing cells, tissues and organs; "Three Parents", reproduction in humans; "Car Wars", effects of spoiling resources and environmental pollution; "What does the fox say?", using technology for understanding animals emotions; "Attack of the giant viruses", learn how to interrogate sources to separate science fact from fiction. "Ban the beds" and "Take the test"

The learning materials used are based on the concepts and experiences of the upd8 project in the UK (www.upd8.org.uk), which have been produced and used in schools for more than 10 years, and which had a very high reputation among science teachers. In each classroom activity we characterized the scientific topic and the RRI topic. For example: for the activity: "grow your own body" the scientific topic is organs

(biology) and the RRI topic is argumentation; for the topic "What does the fox says?" the scientific topic is waves (physics) and the RRI topic is evidence.

The Engage materials and evaluation instruments were first produced in English and afterwards localized to the language and piloting demands of the local partners. Changes in the items with respect to national specifics were done in a very careful way, especially taking into consideration that transnational comparison should still be possible.

The evaluation took place in school classes in four different countries: Israel, Cyprus, Norway and United Kingdom. Engage project members of Israel, Cyprus and Norway contacted personally teachers in secondary schools, informed them in detail about the Engage Project, ENGAGE Material and the use of the first classroom piloting and asked teachers to participate and test the material in their classes. The teachers that were selected to run the activities have a solid background in teaching socio scientific issues, which have a strong linkage to RRI agenda. In the UK similar materials have been used for more than 10 years. This project partner performed the piloting mainly through the ENGAGE website, to focus the effort on an email marketing campaign to reach large numbers of science teachers, and encourage them to sign up and download the pilot Materials from the website. The approach was chosen for two reasons: 1) ENGAGE Materials are based on similar Materials from science upd8 which have been produced and used in schools for more than 10 years, and which had a very high reputation among science teachers. No further data needed to confirm this. In the UK, the challenge is to recruit science teachers in large numbers to use the new materials. 2) A key part of the ENGAGE model is to recruit very large numbers of teachers in to the Adopt stage. Thus it is necessary to test at the pilot stage the ability to use email marketing to reach teachers, and whether the website is sufficiently easy to use to enable large numbers of teachers to access the Materials.

In all four countries teachers could choose which or how many of the six different materials dealing with RRI topics they want to try. They then run the lessons with their pupils according the instructions of the materials. Due to the very different situations and school systems in the four countries each partner could decide which evaluation instrument would be most feasible to record the results of the testing.

Sample

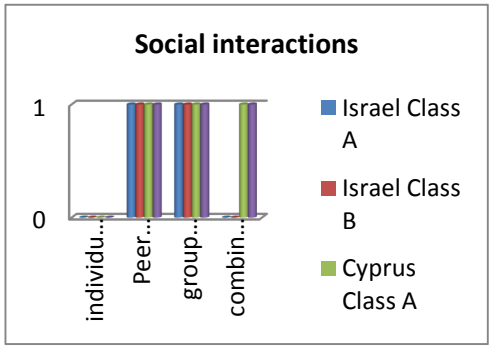
These first test results of the Engage activity materials base on piloting in Israel, Cyprus, Norway and UK.

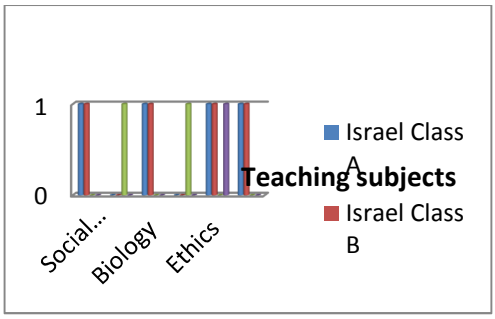
In Israel the activities took place end of May, more precisely in May 28 and May 30 2014. Two teachers in two different towns (Tel Aviv, Arad) tested the Engage material “Ban Coke” in their classes (A and B) with 35 pupils (age 14 to 15 years, 8th and 9th grade classes). Evaluation took place performing teacher’s interview and by using students’ questionnaire.

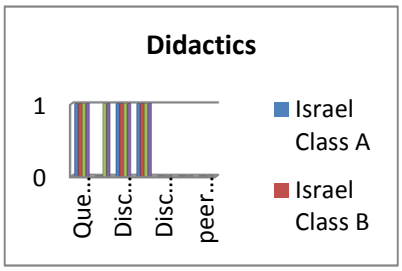
For Cyprus the activities took place on July 14th and 18th 2014 involving two classes and a total of 41 students (age 13 to 14 years, grades?). Two teachers of one school (Nicosia) used and tested the Engage material, one class (Class A) tested the RRI Material “Car Wars” and one class (Class B) the RRI material “Three Parents”. The evaluation instruments adopted for this testing are: interview with the teachers before and after, and observation.

In Norway thee activities took place on June 6th 2014. One teachers of one school (Tonsberg?) used and tested the Engage material “Ban Coke” in two of her classes (A and B) with 50 pupils (age 15, 9th grade classes). The evaluation instrument adopted for the piloting in Norway were the observation guidelines.
For United Kingdom piloting is taking place mainly through the ENGAGE website, the effort is focusing on an email marketing campaign to reach large numbers of science teachers, and encourage them to sign up and download the pilot Materials from the website. The testing results from UK are also reported below.

Findings and analysis from the classroom testing:
Answers of teachers (questionnaire and interviews):
Three of four teachers (two each from Israel and Cyprus) would like to use also other ENGAGE materials; one teacher was not yet sure because he/she felt not competent enough. Also three teachers found the topic interesting, one moderately interesting. Three teachers answered that the topic is important; one thought it is very important. Teachers stated that they found it easy to use the material and that most of their students could complete the task and were engaged. But all teachers suggested to improve the material with more detailed instruction and especially more information about the topic, including references. Further suggestions or new ideas were to include role play or group debates in higher classes. Although being positive, some doubts were brought up during the interviews in terms of the nature of students and difficulties of the teachers to manage this kind of activities (see Annex A: Report from Israel).
The following graphs describe teaching subjects and didactics:







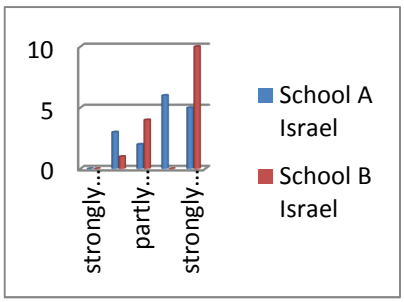
Answers of students (questionnaire)

In Cyprus the questionnaire was not applicable but students were asked in both classes to indicate whether the activity was interesting and engaging or not. Overall students of both classes found the activities interesting and engaging because these were different from their normal lessons.

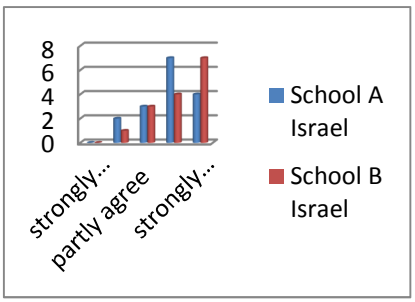
In Israel 31 of 35 students answered the questionnaire. Both classes from Israel expressed interest and engagement in the activity. The engagement and interest were manifested in the likert-scale and open questions (regarding their affective and RRI related perception). Especially positively students mentioned the chance to express their opinion and conduct a discourse. But both classes expressed problems with understanding the basic scientific concept in the activity("Ban Coke"). As well as teachers also students asked for more information and evidence.

The following graphs reflect the answers of students regarding a selection of closed questions:

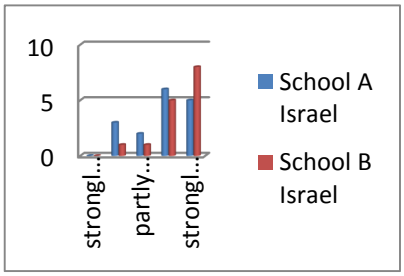
Graph 1: I found the activity very engaging



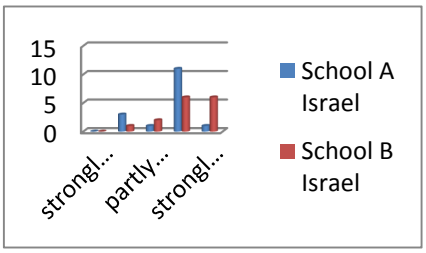
Graph 2: I would like to do this type of activities frequently



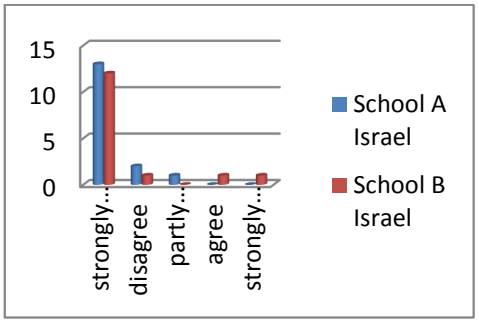
Graph 3: I thought the topic was very interesting



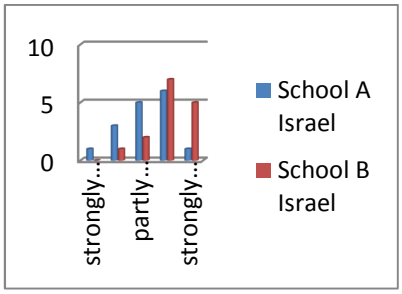
Graph 4: I would imagine that most students would like to do this activity



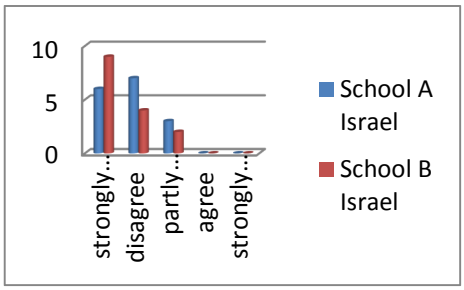
Graph 5: I didn't know what to do



Graph 6: I didn't know what to do



Graph 7: I didn't know what to do



Observations (observation grid)

Observations from Norway

Summary analysis from the observer: according the impressions and notes taken during the activity to answer eight questions within a likert scale (strong agreement, agreement, partly agreement, rather disagreement, disagreement).

The activity objectives were met: partly agreement

The answers the teacher gave to students' questions were clear: agreement

The materials provided illustrative examples. No answer

The activity was well facilitated: agreement

The activity support students learning the basic concepts of the topic: rather disagreement

The activity was an effective way for students to learn on the topic: rather disagreement

Students were actively engaged in the activity: partly agreement

The activity overall was effective: partly agreement

Observations:

Comment from the observer: The activity was carried out in two classes with slight alterations (Class A was divided in groups of 3-4 pupils, in class B students were divided mainly into pairs).

Both the evidence cards and the argument cards are small. In class A the students had some difficulties sharing and reading the cards in the groups (3-4 students had to share one set of cards). In particular the graphs on the cards D) and 1) were difficult to read. In the next session (class B) the students were divided into pairs, which seemed to make the sharing of cards and tasks easier

Overall the students seemed to have limited previous knowledge about the claims that were raised (that sugar acts like a drug, or that sugar may be the prime cause of increasing obesity).

The number of evidence cards is rather high (8). Most groups seemed to sort the evidence cards quickly after a rather superficial discussion. There were very little further discussions of the arguments after the initial sorting was completed. The groups seemed somewhat confused about what to do next (in particular in class A several students complained that the activity was boring).

The time set for the 1st plenary (5 min) also left the teacher only very short time to discuss the different evidence cards more thoroughly. In addition, the only categories for sorting are "strong vs weak evidence *that increasing sugar consumption is causing increase in obesity*" even though some of the cards must rather be considered as evidence AGAINST this claim (e.g. card H).

In class A, the teacher asked the groups to tell which cards they assessed as strong/weak evidence in their own chosen order. With this approach a few cards were left out and not assessed at all. In class B the teacher systematically went through the cards from card A to card H and made a table showing the responses from the different groups. This approach gave more structure to the session. On the other hand a closer discussion of the evidence was not possible in either class within the time available. The class did not reach a clear conclusion from the evidence cards before the argument cards were distributed.

In general, the groups were able to sort the argument cards rather quickly as arguments for or against a ban on sugary drinks. There was some further discussion of the arguments after the sorting (the students generally highlighting arguments that supported their own view). Some students also wanted to raise other arguments that were not covered by the cards, and were somewhat frustrated that the activity did not open for this.

In both classes the teacher finished by going through the argument cards 1 through 8 with the students. This resulted in another table where the cards were (correctly) sorted as potential arguments in favour of a ban on soft drinks for under-18s, or arguments against a ban. Again, this sorting of arguments (5 for, 3 against a ban) did not lead to a clear conclusion. A few students seemed puzzled by the lack of a clear outcome from the activity, whereas several others seemed largely indifferent.

Subsequent comments from the teacher

The teacher asked the students about their comments and views on the activity in the following science lesson. She reports the following main points from the discussions:

Some students were antagonized by the key question (should we ban soft drinks for under-18s) since they consider this an unacceptable limitation of their freedom to choose. This affected their motivation for the activity in a negative direction

There were too many cards (given the time constraint for the activity)

Some students found the cards difficult to read (in particular students of non-Norwegian speaking background)

A number of students were puzzled by the lack of right or wrong answers

Several students found the activity excessively "theoretical" (compared to other science activities that incorporate practical/hands on work)

Summary analysis from observants

The activity was carried out following the given instructions and schedule quite closely. However, based on our own observations and subsequent comments from students/teacher we would rate the activity as only partially successful.

Obviously one can´t draw very strong conclusions from this very limited pilot, however we would like to suggest the following points for further development:

Since (at least in Norway) few students have little previous knowledge about the issue more time should be used to introduce the activity

The cards should be made larger and more easily read

The number of evidence cards is quite high, and time limited. Either the time frame should be extended, and/or different groups could be given a smaller number of cards and told to examine the claims more closely. The students could also be instructed specifically to find arguments that support one particular opinion and make the best possible case for this view (e.g. in a role play, court case etc.)

Many students came out of the activity without knowing if they had succeeded in attaining the goals of the activity. The learning objectives should be made more concrete/accessible to students

Observations from Cyprus

Data-based observations:

Class A ("Car Wars"): The teacher changed the lesson considerably. He added an introduction in which he asked students to brainstorm about different forms of energy the know, and then he structured the entire lesson as an argumentation debate in which students were working in groups (using the information provided by the lesson plan and information from the internet) to choose which energy they prefer and why. The teacher did not use the game activity because he did not find it fitting his learning style and the information on how to use the game was not clear.

Class B ("Three Parents"): The teacher changed the lesson because she was worrying that the students were not able to understand the information (more background information should be provided). She started with a discussion on what genetics are and what are SSI. The issue of cloning came in the discussion and the students were asking questions about this issue. Then the teacher presented the scenario as in the activity. Then the students were assigned in groups, and groups were divided to have two different roles, those of parents and those of doctors. Then a debate followed, with comments which indicated that the students lacked the knowledge to understand how the fertilization with three parents would happen.

Interpretations/Questions/Comments

Class A ("Car Wars"): The students were working mostly in their groups and they were free to search information on-line using the computers in the classroom. The teacher also provided a word document with some relevant links in Greek, and some videos with information but most of the students chose to search for their own information. The final discussion mostly emphasized on that some of the solutions were not yet available in Cyprus.

Class B ("Three Parents"): The discussion at the end mostly focused on questions about the process of the fertilization, and how DNA is "transferred" to the baby. The kind of questions showed that the students did not have the background knowledge necessary to help them understand this topic.

Students were asked in both classes to indicate whether the activity was interesting and engaging or not. Overall for both lessons they found the activities interesting and engaging because these were different from their normal lessons.

Findings and analysis from the UK web-based testing

The UK context in terms of testing Materials in classrooms is a little different to other pilot countries The Engage Materials for Adopt, which our team has been responsible for developing, are based around an approach that we previously developed for the UK - called 'science upd8'. Upd8 had a huge user base, and a large amount of user feedback and positive testimonials. Therefore we did not feel it was productive to focus our efforts on duplicating this feedback from classroom testing.

Instead we determined that the important contribution we could make here was to test the Engage marketing and dissemination. Our targets are to reach very large numbers of science teachers in each country with Materials. This has to rely on effective strategies for web-based publication and communication. So this has been our approach, with a number of questions we investigated:

Email marketing: Could we get large numbers to read our newsletter, and click on the website links?

Web publishing: could we create a website which attracts large numbers of teachers to sign up, download and use materials?

Teacher feedback: We sought user feedback in the form of online reviews of the Materials on the Engage website.

Email marketing

The main objective was to find the best strategies for getting teachers to open and read our newsletter, click on the website link, and then to sign up on the Engage website to download Materials.

We first assembled a large, high quality database of science teacher contacts. This was collated from our previous projects, and critically we had these professionally verified, so that we could start email marketing without having huge 'bounce rates', which might cause our emails to be rejected by ISPs as spam.

Our approach, following that of upd8, was to release the 8 pilot Materials two at a time, on a weekly basis. This meant we had 4 opportunities, one a week, to email teachers informing them of new content. This was designed to create a positive expectancy among teachers, and form a habit of going to the Engage site for teaching contemporary science issues

Teachers have little time to read email, so it is vital to test factors which influence the rate of 'opens' and 'clicks' - which determine how many visit at the website. We used the standard marketing method of successive A/B testing to discover the most effective: subject of email, format of the message, day of the week, and time of day. We used the online service Mailchimp to organise, monitor and analyse our 'Campaigns'. Attractive HTML emails were sent, showcasing the latest published Materials, and emphasising the benefits to teachers, importantly, using their language, not ours.

The conclusions were:

Subject of email: containing words like 'try' and 'new', 'exciting',

Format of message: visual, highlighting what's new on the site, and how it relates to the curriculum:

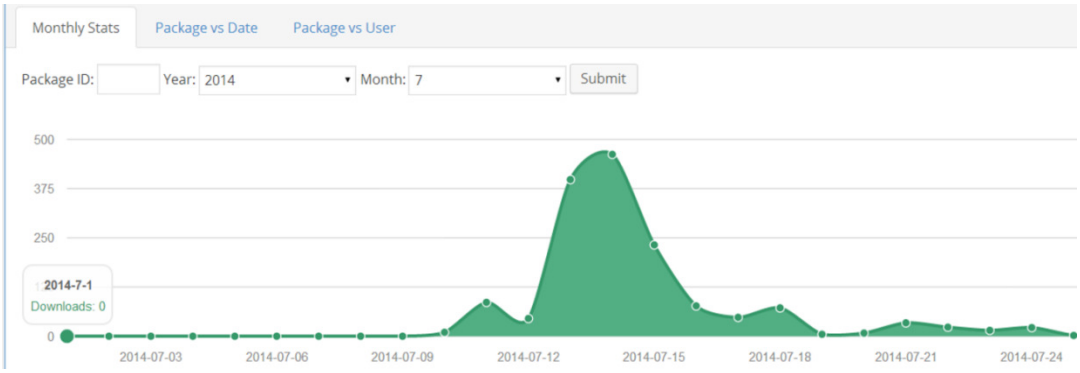
Day of week: Sunday or Wednesday

Time of day: Sunday 3.30-4.30pm, or weekday 8.30-9.30am, or 5.30-6.30pm

The result of successive rounds of A/B testing is that we managed to achieve a high 26.7% 'open rate' for the emails.

The positive impact of effective email marketing is underlined by the chart below, which shows the spike of visits and downloads to the Engage website, in the hours after the emails were sent (the first spike represents the A/B testing, sent to a small sample).

Peak after mailing

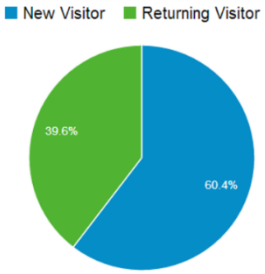


Web publishing

The website immediately began attracting large numbers of teachers, after its launch in early June. As the table shows, there were over 3,500 unique visitors per month, who found the site interesting enough to stay quite a long time exploring.

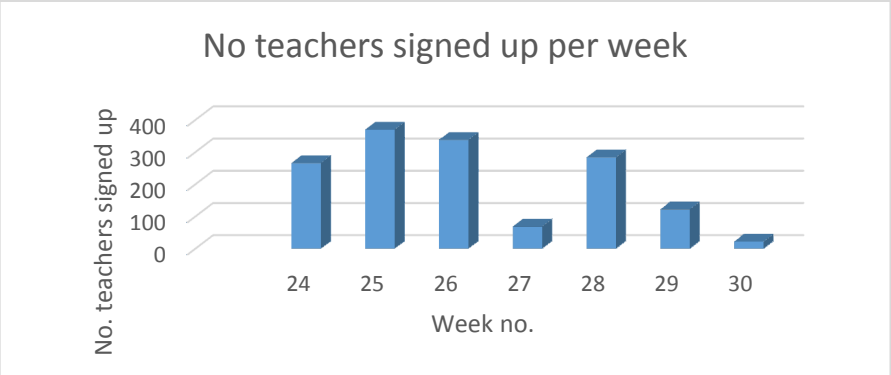
Month	Unique visitors	Page views	Ave. session duration
June	3613	19,456	3 min 24 s
July - until 18th (end of school)	2320	13,396	3 min 55 s

What is also very encouraging for the success of Adopt is that we had some success in establishing the beginnings of a habit for teachers to visit the site regularly. By July, almost 40% of visitors were returning to the site to look at new Materials.



We helped to design and iterate the Engage website, to make sure that it encouraged teachers to sign up, and made it easy to do so. Teachers could only download once they had registered, and links and explanations were put in several places, with reminders when teachers re-visited.

In just 7 weeks, 1500 teachers signed up. The chart shows that there were large numbers of signups in each of the 5 weeks where we sent an email newsletter (24,25, 26, 28, 29).



Teachers downloaded 8,500 Materials files in this period, which is clear evidence that they found the Materials attractive. The earliest published Materials were downloaded the most, and they were all similarly popular - with around 400 downloads each.

Ban Cola	428
Attack of the Giant Viruses	419
Three Parents	398
What does the fox say?	395

What was also encouraging was that teachers did download the 'teachers guide' as well as the Presentation materials, in most cases.

Teacher feedback

From the online reviews, we know that many teachers liked the Materials and found them useful. We received 13 comments and all of them were positive about the Materials.

We encouraged teachers to give comments, not only on the site, but also with an email directed at those teachers who we knew from the website sites that had already downloaded and were regularly reading the email newsletters. Here are some examples of the positive feedback, with just the occasional constructive criticism.

Take the Test

"This resource engaged a class of Year 8 boys – it was well presented, easy to navigate around and some of the slides were useful as worksheets. Setting the science in context helped them to understand the importance of pedigree diagrams and has given them an excellent platform for GCSE Science when they study this for their exams. The issues/dilemmas of taking a test, the ignorance of some and possible prejudice of others gave the series of lessons an extra dimension for the boys to hook their knowledge and understanding of genetic inheritance onto."

"I wished I had found this website earlier."

Three parents

"Thanks for preparing these fantastic resources, this one was particularly topical having just had a large article in local newspaper- students enjoyed task and really got them thinking — looking forward to more of the same"

"Fantastic work -many thanks."

Ban the beds

"I asked my year 8s to vote on which "engage" lesson they wanted to learn more about and they voted for this lesson. Despite being a Friday a week before the end of term, they were all really interested in the topic and were looking at the evidence in a critical way. I can really recommend this lesson; it's easy to adapt, add differentiation and make it your own."

Attack of the giant viruses

"I downloaded this activity this morning and used it this afternoon! It worked really well to capture the attention of my rather passive lower set Yr10. The post-it note activity, identifying level of concern, worked very well. Overall an enjoyable lesson and I will definitely use it again."

Ban Cola

"This lesson worked well but needs more content added to make up a full hour lesson.
Beware of the video starter. I had tried this out at home with no problems but when I tried it at school I found the site was blocked."

Conclusions and Discussion

Testing of Activities

4 x Ban Coke (2 in Israel, 2 in Norway)

1 x Three parents (Cyprus)

1x Car wars 1 Class (Cyprus)

In Israel and Cyprus students and teachers liked the activity and would like to use further Engage Materials.

In Norway the activity was only partly successful.

Results in the UK regarding the investigation of effective strategies for web-based publication and communication are very promising: high % of email readers, successful Web publishing (registrations, downloads), positive feedback from teachers.

Suggestions for improvements:

All teachers suggested to improve the material

more detailed instruction

especially more information about the topic, including references.

More time for the activity

As well as teachers also students asked for more information and evidence (Ban Coke).

Suggestions or new ideas

include role play or group debates in higher classes.

Annex A: Detailed report from Israel

Student's questionnaire:

The work was divided according to the questionnaire structure. So the first part of the analysis will be concerning the first 7 Likert questions.

The questions are sorted to 3 categories: engagement, content and interest.

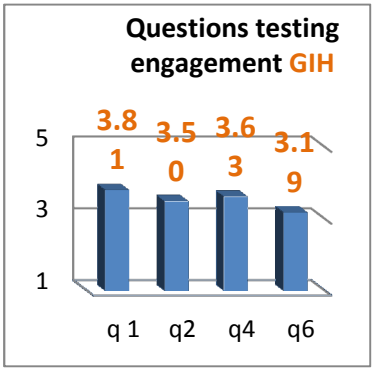
Questions testing student's engagement

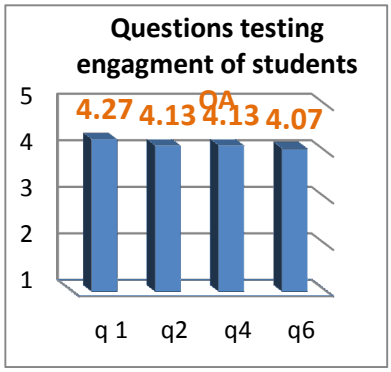
I would like to do this type of activities frequently

I found the activity very engaging

I found the various parts of the activity engaging

I would imagine that most students would like to do this activity

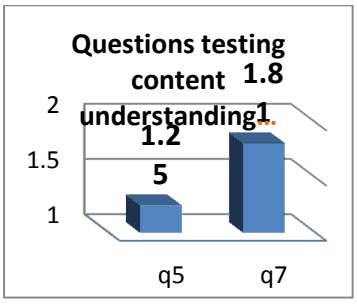


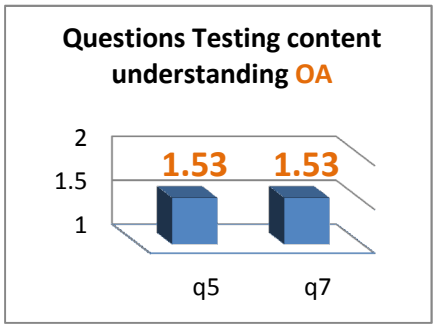


Questions testing content understanding

I didn't know what to do

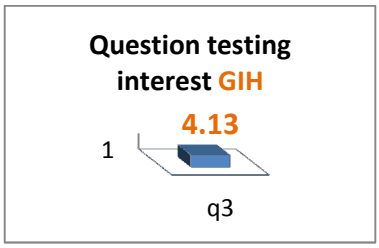
I needed to learn a lot of things before I could perform this activity

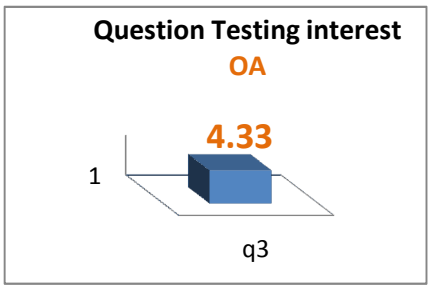




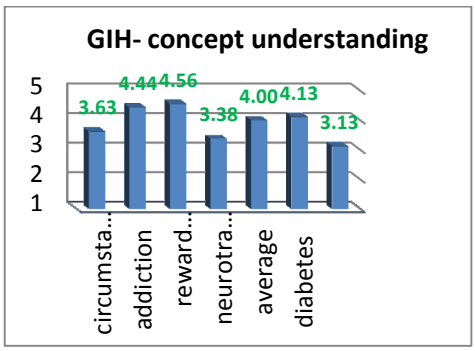
Questions testing interest

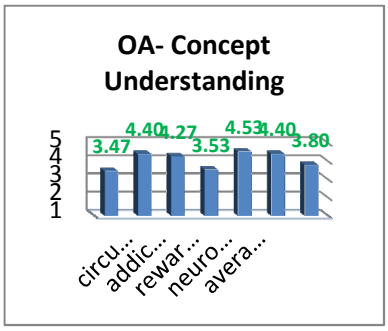
I thought the topic was very interesting





Understanding concepts regarding the activity Ban Coke





Open Questions in the questionnaire

Questions that regard to affective

I liked the activity because…..

I didn't like the activity because……

The activity is similar to what we do in classroom because….

Questions that regard to RRI

Why there is still a debate between scientists about the question whether sugar causes obesity?

Are you in favor of ban soft drinks sale to under 18? which evidence and arguments did you use?

GIH Group

Student no.	RRI	Affective
1	"The mechanism about the connection between sugar and obesity is not fully understood that's why scientists still argue" I am against the ban there is no obvious prof that soft drinks cause you damage. "Teenagers have the right to choose what to drink"	"It was more interesting than a regular lesson" "The subject was pretty boring"
2		"I liked the subject because it was very relevant for people my age"
3	"There is no mechanism that can explain clearly that sugar causes obesity"	The activity does not look like anything we do in the class, it is much" more

	"I am in favor a ban of soft drinks because there is an increase in the number of obese teenagers and I believe the ban can make these phenomena stop"	interesting"
4	There is no mechanism…. "I am in favor of a ban of soft drinks because there is an increase in the number of obese teenagers and I think it will help to prevent this phenomenon"	"I liked the activity very much, it is much more interesting than sitting and listen to the teacher" "The activities helped me to understand the material" "The activity doesn't look like anything we do in the class it is much more interesting"
5	"Because there are people that have a diet without sugar and they are still obese" No I am against the ban because each person should make is own choices"	I liked the activity because I got more understanding about the connection between obesity and sugar uptake. "In this kind of activity we can learn much more from conversation and not parroting"
6		
7	Ban soft drinks is wrong, other methods can be used to raise awareness ; advertisements , commercials and more	Because it was related to us, the students who consume soft drinks in everyday life. There is a lot discussion and work within groups
8	There are not enough evidence that link clearly between consumption of sugar and obesity. "I think that consumption of soft drinks should be an individual decision. Let alone that there is no solid proof that sugar consumption is dangerous for maintaining one's health"	"I didn't like…because the activity wasn't related directly to what we have learned lately on chemistry" The activity didn't relate directly to chemistry but to the society in Israel and legalization
9	It wasn't proofed scientifically that sugar is related to obesity. I am against, because I eat many snacks and drink soft drinks every day and still I am not obese on the contrary.	I liked …. Because I discovered how much sugar I drink each day.
10	Don't know I am against because soft drinks are tasty	I wanted to know more about the subject.
11		
12	There is no chemical and physical explanation that prove the link between sugar consumption and obesity. I am against because as long that there is no scientific proof for the link between sugar consumption and obesity there is no right to ban sugar.	I liked the lesson because it was chemistry lesson and it ended up without learning chemistry.
13	There is a debate because there is no biological evidence yet	I liked…because we didn't learn boring school stuff but just general education.
14		

15	I am against because soft drinks are as part of everyday life	I liked the activity because it took into consideration our opinions
16	"I do not think soft drinks are that bad, besides ban them will make youngsters to wanting and consuming it even more"	"I liked the activity because it took into consideration our opinions, so we were engaged"

Summary of the data GHI

It appears according to the results that the majority of students found the activity engaging and interesting. According to the students most of the content was clear. The concepts that were less clear to the students were Neurotransmitters, infectious disease and circumstantial link. The first two concepts are not that crucial to understand the activity, however it may seem that the meaning of circumstantial link wasn't explicitly explained.

RRI

We can infer from the results that there were some levels of understanding the disagreement of scientists among them on the link between obesity and sugar consumption.

The lowest level just didn't relate to the question.

The medium level related vaguely to the lack of evidence without mentioning what the evidence are actually lacking.

The highest level related to the lack of knowledge of biological and chemical mechanism.

The majority of the students (14 out of 16) were against the ban and here is a summary of their arguments

If there is not enough scientific evidence, there is no moral ground to ban soft drinks to youngsters.

Teenagers are individuals, and therefore they have the right to make their own decisions.

There are other methods to deal with high consumption of sugar, like making the public aware of its risk.

Egocentric arguments

Soft drinks and sweets have no effect on my weight.

Soft drinks and sweets are tasty.

Affective

IN general it can be asserted that most of the students liked the activity. Here is a summary of their reasons:

Because it was interesting

Because we (the students) weren't passive and had the chance to express our opinion.

Because it made me understand the amount of sugar I consume.

Because it was not related to the chemistry curriculum.

OA Group

Student number	RRI	Affective
1	There are not enough evidence. I don't agree because it's not normal	It was interesting and fun.
2	Because many experiments are conducted and each time you get different results. I am against the ban because it makes no difference what age you are it will have the same effect.	The activity was fun and elaborated the general knowledge.
3	I am in favor of the ban. Because soft drinks cause damage to the body and adolescence is a critical period in which the body develop.	The lesson wasn't set up as regular lesson. It was like a big inquiry question.
4	There is a debate among scientist because there are some who say there is a link and some say there isn't. I am against the ban; let everybody do what's best for them.	The activity was interesting.
5	Because each scientist lays on his research and pre knowledge. I am in favor of the ban to 10 years old and under and not 18 and under because sugar helps keeps you concentrated.	It was interesting. The students were involved and practiced in the class.
6	Because obesity can be caused by consumption of products without sugar. I am against because each person is responsible for his own deeds. There is not enough evidence for a damage caused by consumption of soft drinks.	There is no need in notebooks and I could express myself in a better way so I really liked it.
7	They don't have enough evidence. I am against the ban because soft drinks are armful full adults as well as for minors.	It was fun to measure the sugar. It was connected to what we learn at class to carbohydrates.
8	Don't know I am in favor of the ban because kids have no control if their consumption.	The activity was very interesting.
9	Because there are experiments that show sugar is not the main cause for obesity. I am in favor of the ban because that's how teenagers will understand soft drinks aren't healthy and could keep a healthier lifestyle.	Because I understood what I am consuming and from now on I will try to consume as less as possible and keep a better lifestyle. Really liked it.
10	Because they are not sure sugar is the main cause for obesity. I am in favor of the ban but only until the age of 14, 18 is too steep.	I liked it because it was updated and interesting subject. I liked that in the class we raise different ideas.

11	There is no serious proof for what happens, why and because of what. I am against because each one and his parents should decide whether to consume or not.	It was interesting, I learned things I didn't know and have been introduced to a different point of view.
12	There is no proof and no way to explain the link between obesity and sugar consumption. I am against because each one responsible of his own deeds, and who doesn't want allowed not doing so.	It was interesting and not a regular subject.
13	Because it wasn't proofed scientifically. I am in favor of the ban because soft drinks are bad to one's health, and only after the age of 18 you take full responsibility for your actions.	It was interesting. I liked that there is a discourse and participation of students in the class.
14	Don't know I don't agree because everyone can do whatever he likes.	The activity was boring. There was no home work.
15	I am in favor of the ban because it is bad.	It wasn't interesting.
16		

Summary of the data OA

It appears that most of the students showed interest, understanding and engagement in the activity. The content concepts that were less clear to the students are: Circumstantial link, neurotransmitters and infectious disease (3.47, 3.53, 3.8 respectively).

In general we can assert that the answers to the different questions were to say the least; unarticulated, including laconic, superficial answers which don't leave much room for interpretation. So despite the encouraging results of the Likert scale analysis it seems there is a gap between what the students understood and their engagement in the activity.

RRI

Some of the answers (3 out of 15) for the reason of debate between scientist on the link between sugar and obesity are simply circular arguments or repetition of the question in different words.

"There is a debate among scientist because there are some who say there is a link and some say there isn't"

Other answers just say there are not enough evidence without elaborating what kind of evidence exist or miss in order to complete the picture. Some are more simplistic than that and claim for no serious proof or no scientific proof. This kind of answers exemplify that the students didn't fully understand the meaning of scientific consensus and scientific evidence.

Some of the students were against the ban of coke for the next reasons:

The effect of soft drinks is the same on adults as on youngsters so if adults are allowed to consume soft drinks so do youngsters.

Everyone is responsible of his own deeds, so there is no need to ban let the people decide.

Some were in favor of the ban for the next reasons:

The ban should be for a population of 14 or10 year olds and under, 18 is to steep.
Soft drinks are bad for one's health and adolescence is a critical page in persons live so preventing steps should be taken to maintain a healthy lifestyle.

Affective

There is a correlation between the results in the Likert scale and the open questions. The main reasons students mentioned for liking the activity were:
It was interesting
It elaborated their knowledge
The activity allowed them to conduct a discourse.

Summary

In comparison between the two groups there were few resembles and number of major differences. Just as a reminder there is a one year difference between the classes, GIH students are at 10^{th} grade and OA are at 9^{th} grade.
Both groups expressed interest and engagement in the activity. In both cases the engagement and interest were manifested in the Likert and open questions. In both classes mainly in the GIH group, students expressed their positive feedback, regarding to the chance to express their opinion and conduct a discourse. It is very encouraging because it seems as if the opportunity given to the students to conduct a discourse and exchange opinions is strongly embedded within the activity itself.
It seems both classes didn't have special problems with understanding basic scientific concept in the activity. However three concepts were less understood than others (Circumstantial link, neurotransmitters and infectious disease). Circumstantial link should be explained more explicitly in order to emphasis the reasons for debate between scientists. The other concepts are expected not to be understood because they require a background on biology studies which the students presumably were lacking.

Analysis of the RRI ideas the students captured, reveals differences between both classes. It appeared in the two questions regarding two the RRI ideas in the activity ban coke. In general GIH class students related to the chemical and physical mechanism scientists are researching, and to the difficulty to get to a consensus on the issue. In comparison OA class mentioned that there is not enough evidence, or there was no scientific proof, which is not true, plenty of evidence was presented to them. While some of GIH students mentioned the need for a consensus among scientists OA students seemed to use overused slogans.
Very similar results were revealed in the second open RRI questions. While GIH answers were diverse and gave different arguments in favor or against ban of soft drinks, OA class answers were superficial and some from an egocentric point of view.
These differences in conceptualizing RRI ideas brought up in the activity may rise from the difference in age and ability in expression in writing. Another reason can be two different teachers that conducted the activity a little bit different from each other.

Teacher's interview: Conclusions from the interview with the teacher DD

The view of the teacher

The teacher asserted that she really enjoyed the activity and showed positive views regarding the activity.

"it's like a mini research; a student is going to decide about something, not according to a gut feeling but according to different opinions. And I must say that because of the activity I changed my mind about what I was thinking about sugar and obesity. I will be much more reluctant to speak about sugar as a cause of obesity, because the biochemical mechanism is far more complicated than consumption of fats"

The teacher tells that the activity made her reflect on her own teaching and we can see that she even used her background as a chemist to explain her self the debate between scientists.

Will you be happy to teach these kinds of activities?

"Millions of them just hand them to me. You need to pay attentions it correlates with the curriculum, and I could give you the themes, energy in 9th grade, food chemistry to 9th and 11th grade"

It is clear the teacher enjoyed the activity and seas a great value in the activity.

In the next examples we can see that the teacher's point of view and perspective correlate with RRI principles. She brought up the examples spontaneously without even asking her for.

…"they were introduced to methods that I wish they'd use them with almost every issue they are facing in their life"

"You should let them express themselves; this is something that I am very interested in, How you make them (the student) express themselves, more than usually"

"I took a presentation of david kahn and localized it to my needs, it was about the challenge of energy, food and water…"

"And I asked them let's talk for a second why do you think you got candies from the company?"

"Yes, I told them that third of the energy consumptions is due to cattle nurturing, and they attacked me saying: "why all the teachers tell us to stop eating meat" and I told them that I deliver them the information, they have the choice to use it however they'd like.

You should waken them up, shake them a little bit, they do not think…"

Suggestion for future developments

The teacher had many suggestions and tips to conduct the activity in the future. She sometimes used the "us" as referring to the people who should do the refinement. I believe that after this activity she sees herself as part of something that can make change and she is willing to take an action in order to be a part of it.

"In higher grades you can conduct a debate between two groups of students"

"What was missing in this activity is that we gave them a problem and they negotiated it but eventually there is no closure for what can we do in order to solve the problem…
let us put the last slide in the presentation with relevant references. Things I have summarized by heart, they should be visible. They say a good lecture presents at the beginning what is going to be discussed, during the lecture what is being discussed and at the end of it what was discussed. You need to repeat it"

"Both activities are very similar, I would make something different in the second activities, something like a roll play"

"You mean that the student will play the characters on the argument cards?"
"Yes something like that"

Doubts

Although being positive, some doubts were brought up during the conversation. They generally didn't refer to the activity itself but to the nature of students and the difficulty of the teacher to manage this kind of activities.
"I don't know if many teachers are prepared to let this degree of freedom in the classroom"
The teacher was reluctant to the wish of other science teachers to have these kinds of activities.
"I don't know how much it affects them, there are some students it does…
"You know they are junior high students they recently found out that they can speak of their minds, and they don't actually know what to do with it. Like a three year old boy that says no to each thing you offer him even if it's good"
"So why it happens to be, that so many students don't have any opinion, they haven't thought of so many issues. It really concerns me"
"Chemistry, philosophy, phycology and physics they are all intertwined. How could you possibly create a scientist if he doesn't know how to think?"

Annex B: Detailed report from Norway

Classroom Observation Tool – HiVe Pilot

Date: ___6.6.2014__ Topic of the activity: _____Ban Coke_____

Class: _____9th _____ No. of Students: __25 + 25 (Class A, Class B)

Teacher: ____Inger Holhjem_____

School: ____Varden_____Observer: ___Harald Bjar/Elin Aschim__

Time class began: ____12.30___ Time class Ended: ___13.30___ (Class A)

Time class began: ____13.45___ Time class Ended: ___14.45___(Class B)

Learning Objectives of the activity: What are students supposed to learn, understand and discuss about?

In the teacher guide for the activity these are defined as:_Students will

Apply knowledge about food and health

Use evidence to decide whether a correlation is causal

Evaluate and synthesise information to make a decision about a health-related issue

Part II: SUMMARY ANALYSIS

Use the notes taken during the activity to answer the questions below.

1. The activity objectives were met.

☐ 1 strongly agree ☐ 2 ☒ 3 ☐ 4 ☐ 5 strongly disagree

2. The answers the teacher gave to students' questions were clear.

☐ 1 strongly agree ☒ 2 ☐ 3 ☐ 4 ☐ 5 strongly disagree

3. The materials provided illustrative examples.

☐ 1 strongly agree ☐ 2 ☐ 3 ☐ 4 ☐ 5 strongly disagree

4. The activity was well facilitated.

☐ 1 strongly agree ☒ 2 ☐ 3 ☐ 4 ☐ 5 strongly disagree

5. The activity support students learning the basic concepts of the topic.

☐ 1 strongly agree ☐ 2 ☐ 3 ☒ 4 ☐ 5 strongly disagree

6. The activity was an effective way for students to learn on the topic.

☐ 1 strongly agree ☐ 2 ☐ 3 ☒ 4 ☐ 5 strongly disagree

7. Students were actively engaged in the activity.

☐ 1 strongly agree ☐ 2 ☒ 3 ☐ 4 ☐ 5 strongly disagree

8. The activity overall was effective.

☐ 1 strongly agree ☐ 2 ☒ 3 ☐ 4 ☐ 5 strongly disagree

NB The activity was carried out in two classes with slight alterations (improvements) from the first to the second case. The summary covers piloting in both classes.

Notes on environmental factors

The activity was carried out in a normal classroom with about 25 students (9th grade/15 years)

Time: Friday afternoon (from 12.30) on a warm, sunny day with open windows. There were some problems with noise from adjacent rooms and outdoors (lawn mower etc.)

In order to make room for the activity the teacher had made some adjustments to the normal time schedule of the two classes. This caused some initial confusion and a few disturbances through the sessions.

Teacher set-up:

Teacher followed the teacher´s guide for the activity quite closely, i.e.

The two observants from HiVe were presented, followed by a short introduction to the activity.

The class was divided into groups (Class A: groups of 3-4. Class B: students divided mainly into pairs)

The teacher asked a student to demonstrate physically the sugar content in a 500 ml bottle of Fanta (25 sugar cubes/50 gram)

A 3 minute trailer to the documentary “Fed up” (which discusses sugar and obesity) was shown on projector. The trailer had American language, no subtitles and the sound was slightly low, and was probably somewhat difficult to follow for a significant number of students.

The teacher used the slides to introduce the topic for discussion (should sugary drinks be sold to people under 18). The evidence cards were then distributed to the groups for discussion

Observations:

Both the evidence cards and the argument cards are small. In class A the students had some difficulties sharing and reading the cards in the groups (3-4 students had to share one set of cards). In particular the graphs on the cards D) and 1) were difficult to read. In the next session (class B) the students were divided into pairs, which seemed to make the sharing of cards and tasks easier

Overall the students seemed to have limited previous knowledge about the claims that were raised (that sugar acts like a drug, or that sugar may be the prime cause of increasing obesity).

The number of evidence cards is rather high (8). Most groups seemed to sort the evidence cards quickly after a rather superficial discussion. There were very little further discussions of the arguments after the initial

sorting was completed. The groups seemed somewhat confused about what to do next (in particular in class A several students complained that the activity was boring).

The time set for the 1st plenary (5 min) also left the teacher only very short time to discuss the different evidence cards more thoroughly. In addition, the only categories for sorting are "strong vs weak evidence *that increasing sugar consumption is causing increase in obesity*" even though some of the cards must rather be considered as evidence AGAINST this claim (e.g. card H).

In class A, the teacher asked the groups to tell which cards they assessed as strong/weak evidence in their own chosen order. With this approach a few cards were left out and not assessed at all. In class B the teacher systematically went through the cards from card A to card H and made a table showing the responses from the different groups. This approach gave more structure to the session. On the other hand a closer discussion of the evidence was not possible in either class within the time available. The class did not reach a clear conclusion from the evidence cards before the argument cards were distributed.

In general, the groups were able to sort the argument cards rather quickly as arguments for or against a ban on sugary drinks. There was some further discussion of the arguments after the sorting (the students generally highlighting arguments that supported their own view). Some students also wanted to raise other arguments that were not covered by the cards, and were somewhat frustrated that the activity did not open for this.

In both classes the teacher finished by going through the argument cards 1 through 8 with the students. This resulted in another table where the cards were (correctly) sorted as potential arguments in favour of a ban on soft drinks for under-18s, or arguments against a ban. Again, this sorting of arguments (5 for, 3 against a ban) did not lead to a clear conclusion. A few students seemed puzzled by the lack of a clear outcome from the activity, whereas several others seemed largely indifferent.

Subsequent comments from the teacher

The teacher asked the students about their comments and views on the activity in the following science lesson. She reports the following main points from the discussions:

Some students were antagonized by the key question (should we ban soft drinks for under-18s) since they consider this an unacceptable limitation of their freedom to choose. This affected their motivation for the activity in a negative direction

There were too many cards (given the time constraint for the activity)

Some students found the cards difficult to read (in particular students of non-Norwegian speaking background)

A number of students were puzzled by the lack of right or wrong answers

Several students found the activity excessively "theoretical" (compared to other science activities that incorporate practical/hands on work)

The observants did not have the opportunity to observe the teacher/class in a "regular" science lesson. However, the teacher reports that the class discusses scientific questions and matters from the media/in the

news quite frequently. Several students find science interesting, and occasionally bring scientific knowledge and news from programs on TV etc. into the classroom.

Summary analysis from observants

The activity was carried out following the given instructions and schedule quite closely. However, based on our own observations and subsequent comments from students/teacher we would rate the activity as only partially successful.

Obviously one can´t draw very strong conclusions from this very limited pilot, however we would like to suggest the following points for further development:

Since (at least in Norway) few students have little previous knowledge about the issue more time should be used to introduce the activity

The cards should be made larger and more easily read

The number of evidence cards is quite high, and time limited. Either the time frame should be extended, and/or different groups could be given a smaller number of cards and told to examine the claims more closely. The students could also be instructed specifically to find arguments that support one particular opinion and make the best possible case for this view (e.g. in a role play, court case etc.)

Many students came out of the activity without knowing if they had succeeded in attaining the goals of the activity. The learning objectives should be made more concrete/accessible to students

Annex C: Detailed observation report from Cyprus

Date: A: 14.7.2014 / B: 18.07.2014 School: Learning in Science Group Summer School

Class: 13 – 14 years old students No. of Students: Class A: 21, Class B: 20

Time class began: 09:00 a.m. Time class ended: 10:20 a.m. (Class A)

Time class began: 09:00 a.m. Time class ended: 10:20 a.m. (Class B)

Teacher: Marios Papaevripidou (Class A) / Christiana Nicolaou (Class B)

Topic of the activity: A: Car War / B: Three Parents Observer: Maria Evagorou

Data-based observations:

Class A: The teacher changed the lesson considerably. He added an introduction in which he asked students to brainstorm about different forms of energy the know, and then he structured the entire lesson as an argumentation debate in which students were working in groups (using the information provided by the lesson plan and information from the internet) to choose which energy they prefer and why. The teacher did not use the game activity because he did not find it fitting his learning style

Class B: The teacher changed the lesson because she was worrying that the students were not able to understand the information (did not have background knowledge). She started with a discussion on what genetics are and what is SSI. The issue of cloning came in the discussion and the students were asking questions about this issue. Then the teacher presented the scenario as in the activity. Then the students were assigned in groups, and groups were divided to have two different roles, those of parents and those of

doctors. Then a debate followed, with comments which indicated that the students lacked the knowledge to understand how the fertilization with three parents would happen.

Interpretations/Questions/Comments

Class A: The students were working mostly in their groups and they were free to search information on-line using the computers in the classroom. The teacher also provided a word document with some relevant links in Greek, and some videos with information but most of the students chose to search for their own information. The final discussion mostly emphasized on that some of the solutions were not yet available in Cyprus.

Class B: The discussion at the end mostly focused on questions about the process of the fertilization, and how DNA is "transferred" to the baby. The kind of questions showed that the students did not have the background knowledge necessary to help them understand this topic.

Students were asked in both classes to indicate whether the activity was interesting and engaging or not. Overall for both lessons they found the activities interesting and engaging because these were different from their normal lessons.