Report from EAP on assessment materials (Part A)

SAILS

**Strategies for Assessment of
Inquiry Learning in Science**

**Strategies for Assessment of
Inquiry Learning in Science**

D7.1 Report from EAP on assessment materials (Part A)

Authors: Odilla Finlayson, Eilish McLoughlin, Deirdre McCabe and PSC
Project name: Strategies for the Assessment of Inquiry Learning in Science (SAILS)
Project number: 289085
Start date: 01/01/2012
Duration: 48 months
Lead partner for this deliverable: Dublin City University
Project coordinator: Dublin City University
Contact: info@sails-project.eu
Website: www.sails-project.eu

Introduction

This report is in two sections:

- Section 1 is the report by Professor Jens Dolin, Member of the External Advisory Panel, on the assessment strategy and instruments developed to date;
- Section 2 is the response and action plan developed by the PSC to the report.

SECTION 1

Report on the SAILS Project

from Prof. Jens Dolin, member of the SAILS External Advisory Panel, June 2014.

This report is based on reading of the SAILS documents, following the subscription updates through many months and participating in the Szeged meeting May 5th. Before this I have also accompanied Eilish McLoughlin on a trip in Belgium where we presented our two projects for teachers, developers and school leaders, which gave a good overview of the project and how teachers can benefit from it. As External Advisory Panel member, I see myself as a critical friend, meaning that I within a frame of acknowledging the project's big achievements will focus on the weaknesses I have found. And it is important to stress, that the overall impression of SAILS is that it is without doubt significant, well driven and with high impact. The whole project management, logistic and communication seem very efficient and following the updates you get an impression of very active and committed participants. Having said this, and I must stress this context of a solid, well-functioning project, I will concentrate on areas where I see possibilities for strengthening the project.

1.1   The framework

The aim of SAILS is to support teachers in adopting an inquiry approach. And realizing assessment as one of the main drivers in teaching, SAILS will provide teachers with a framework for assessment of inquiry learning in science. This framework is then a central tool for organising teachers' implementation of IBSE sequences, including assessment. The framework and its use are thus very important for the project.

The deliverable D2.2 sets up the framework describing the assessment areas and how to assess them in the context of IBSE. The framework has the following elements: Content (from three disciplines), Cross Disciplinary content, Reasoning, Scientific literacy and application, and Inquiry skills. These elements could all be relevant to assess, but their internal connections and the differences between them, seem not to be defined and described in a much operationalized way. Scientific literacy, for instance, is in the PISA project, which is referred in the framework, defined by the other elements in the framework and the domain of inquiry skills includes reasoning and application, which also are independent domains. The elements are put together in figure 1, adding 'resources for inquiry', which not seems to be defined. These overlappings and conflicts with other, well known, frameworks, might cause problems when teachers have to interpret the different elements and report the outcome.

Another aspect of the framework is that it seems as if its construction has been steered by some well-known topics and items more than by the more generic aspects defining IBSE (like the Linn, Davis & Bell, 2004, definition you are using). You sometimes have the impression that the need for an approach for assessing these topics, that you have refined thorough years of practice, define the framework, more than trying to find a way to assess students going through problem solving and inquiry processes. This might be important when the transferability of the assessment methods is to be established. A description too close to specific topics might undermine the transferability.

1.2   The cases

The project has selected app. 20 inquiry lessons within different topics. These lessons are described according to a given structure (very much based on fig. 1: The involved inquiry skills, suggested assessment procedure etc.) -and then called units. All units are tried out in classrooms and the results are reported according to a given template (milestone 7) – and are then constituting a case. These cases are the foundation for the further work, e.g. evaluating the appropriateness of the chosen assessment for the given topic, and characterising the inquiry skills across the units. So, the quality of cases is a core concern. To be able to conclude across the cases, they must be based on the same understanding of the constituent parts of the case: Competence, inquiry, learning progression, assessment as well as all the elements in fig. 1. And here you should consider whether your definitions and descriptions are precise enough to be operational.

Cases are – as genre – a very broad and open description, while comparison calls for narrowness and closeness, which means detailed descriptions of the concepts and procedures used. This is of cause a balance between a very valid description, broad and close to the concrete practice, and a very reliable description, based on a common format.

1.3   Formative and summative assessment

Another point of attention is your use of assessment as such. You seldom distinguish between formative and summative assessment. I find this a crucial distinction in any assessment. The terms describe the purposes for which the assessment is done, not the task or the method itself! (but students are often aware of what purpose a given task is given for).
It is my firm belief that all assessment methods can be used both formatively and summatively – but with very different results! If the purpose is to show that the assessment used can deliver summative results, you will look at other issues than if the purpose is to enhance the learning of the student.
So, I will strongly recommend that the cases more precisely declare for what purpose the assessment has been done.

1.4   The reliability of the assessment

A special concern here is fact that teachers are not reliable scorers or assessors of student learning without using strategies to reduce the measurement error. So it might be a good idea to introduce and rehearse strategies for enhancing reliability, like

1. Use established 'best practices' for the construction of assessments
2. Use of scoring guidelines (rubrics)
3. Increase teacher understanding of student learning (to inform a consistent interpretation of student responses)
4. Form Communities of Practice building a common understanding of expectations of student work.

A discussion of the collected cases can be seen as a way to enhance reliability.
The whole aspect of comparing and extracting results across the individual cases is based on a reliability of the data you are comparing. Not necessarily in 'inter rater reliability' measures but at least with some considerations on the degree of common standards for assessment.

1.5   Learning progression as the basis for formative assessment

Linked to this problem, is the importance of being aware of the learning progression steering the learning of the given topic. To engage successfully in the practice of formative assessment teachers need

"… a sound model of students' progression in the learning of the subject matter, so that the criteria that guide the formative strategy can be matched to students' trajectories of learning" (Black, P. and Wiliam, D. (1998). Assessment and Classroom Learning. *Assessment in Education: Principles, Policy and Practice, 5, 7-73, p. 37*.)

The big challenge is (1) to find the right level of description - not too detailed (not to give too strong a steering of the teaching) and not too gross-grained (not being able to give specific feedback on student activity) and (2) to have a common approach towards progression.

It might be useful to initiate discussions among the teachers of the learning progressions within the different cases.

I do not see the above considerations as challenging the quality and usefulness of the SAILS project, but rather as areas where small adjustments can give big improvements.

1.6   Cooperation with other FP7 projects

As coordinator of another FP7 IBSE-oriented project, ASSIST-ME (*Assess Inquiry in Science, Technology and Mathematics Education*), it is striking how ASSIST-ME and SAILS are very much alike:

- We share the same premises and use the same vocabulary and refer to the same reports and references
- We use the same definition of IBSE
- We perform a lot of identical work (mapping the field, describing assessment forms, implement new methods etc.)
- We partner with experienced teachers
- We focus on the implementation process of the assessments
- We include stakeholders in the project with the aim to promote change of educational policy
- etc.

This gives many possibilities for cooperation and exchange of experiences - which we should benefit from! So, one strong recommendation is to setting up closer links between the two projects.
This is luckily enough under establishment.

SECTION 2

Response by PSC to Report by EAP member Prof Jens Dolin
July 2014

The PSC welcomes the input of Prof. Jens Dolin and his advice for the project. Prof Dolin was sent the original DOW, and initial deliverables from WP2 and 3. He was also given access to the members' area of the website to see the ongoing work of the SAILS project.

In March, he met with Sally Reynolds (ATiT) and Eilish McLoughlin (DCU) from the SAILS project as they were involved in presenting SAILS and ASSIST-ME projects to a Belgium stakeholders meeting. This gave Prof. Dolin a good overview of the aims and the initial outputs of the project. He then attended the General Assembly (GA) meeting in Szeged in May 2014 where he addressed the group and gave an oral presentation of his advice on the SAILS project. Time was given to discussing his recommendations with the group.

The role that he has described for himself in relation to SAILS is that of a critical-friend and his input is very valuable to clarify in identifying gaps or lack of clarity in the work of the project. His written report was received in June 2014 and the Project Steering Committee (PSC) have considered his report and now make a response to his report.

1. Overview

    We welcome his positive comments that his overall impression of SAILS is that *"it is without doubt significant, well driven and with high impact. The whole project management, logistic and communication seem very efficient and following the updates you get an impression of very active and committed participants"*.

    Action: We will endeavour to maintain the active commitment and co-operation between consortium members in an effort to strengthen the working relationships within the group members and to value the input by each member to the overall outcomes.
    His further comments are given in the context of strengthening the work of the project.

2. The Framework

    The development of the framework has been identified as a critical element within the SAILS project "SAILS will provide teachers with a framework for assessment of inquiry learning in science. This framework is then a central tool for organising teachers' implementation of IBSE sequences, including assessment". A key criticism here is that the elements identified that will be included in the final Framework (as outlined in D2.2) need further definition so that their similarities and differences may be clarified. Also several of these aspects already have associated frameworks (e.g. PISA framework for scientific literacy, 2010) so clarity is required to avoid difficulties in interpreting different elements.

    ACTION: The need for clarity and definition is well made and the consortium is developing a common understanding of these elements – as outlined in D2.3. Further clarification of the elements will be noted in the final framework.

    Another aspect that he drew attention to was the assessment of particular elements within topics versus assessing students going through an inquiry process. This distinction may be

important in terms of transferability of the assessment methods as he states that "a description too close to specific topics might undermine the transferability".

ACTION: The idea of transferability between topics is interesting and it will be considered when the case studies are analysed, to determine if there are indeed transferable assessments or if it is instead focussed on discipline or topic. This analysis will be ongoing within the evaluation as part of WP3.

3. The Cases

Prof. Dolin had a good understanding of the development of the SAILS unit incorporating several case studies. He identifies two main points here. Firstly, as the case studies form a very important part of the unit and are the foundation for the evaluation of the appropriateness of the assessment, there needs to be a common shared understanding of the constituent parts of the case. Secondly, he highlights the difficulties within comparison of case studies, balancing between "very valid description, broad and close to the concrete practice, and a very reliable description, based on a common format".

ACTION: Analysis of the current case studies as part of D3.2 has shown the need for a detailed guide in an effort to generate more consistent and comparable information in the case studies. Also, following the analysis conducted as part of D2.3 and D3.2, the content of the SAILS UNIT (incorporating case studies) will be discussed at the next GA and agreed. Following this, a number of exemplar units will be developed.

4. Formative and Summative Assessment

Prof. Dolin pointed out that within the units and case studies, the use of formative and summative assessment had not been distinguished. In agreement with SAILS, he did highlight that the methods of assessment could be used both formatively and summatively. As the results of assessment could be used to enhance the learning of the student versus generating summative results, his recommendation is that this be clarified.

ACTION: Within the case studies, the purposes of the assessment will be included.

5. The Reliability of assessment

Prof. Dolin noted in his report that "teachers are not reliable scorers or assessors of student learning without using strategies to reduce the measurement error". He suggested that teachers need to be given opportunities to rehearse strategies for enhancing reliability, such as looking at 'best practice', use of rubrics, examine student responses and build Communities of Practice to share common understanding of expectations of student work.

ACTION: This was considered to be a very important part of SAILS and it is addressed within both WP4 and WP5. Within the Teacher Education Programme, teachers must be given time to discuss and identify good elements of student work. This has been incorporated within WP4. Also as the teachers will trial and implement inquiry and assessment practices after attending the Teacher Education Programmes, they will all be part of the electronic Community of Practice (CoP) to share both their experiences and resources within WP5.

Common standards of assessment will be interesting to investigate across all partners – but maybe this is more important to develop within each partner group with their teachers.

6. Learning Progression as the basis of Formative Assessment

Prof. Dolin indicated that teachers needed a good knowledge of learning progression so that they can guide students along this progression. Particular challenges he identified here is "to (a) find the right level of description - not too detailed (not to give too strong a steering of the teaching) and not too gross-grained (not being able to give specific feedback on student activity) and (b) to have a common approach towards progression".

ACTION: Learning progression is considered to be important in teaching; brief discussions of learning progression of skills and competencies are implicit in SAILS Teacher Education Programmes.

7. Co-operation with other FP7 projects
Prof. Dolin highlighted the common elements of SAILS to another FP7 project which he co-ordinates ASSIST-ME. While the objectives of the two projects differ, clear synergies were identified and Prof. Dolin suggested that the two projects might work more closely together.

ACTION: Co-operation and exchange with ASSIST-ME was welcomed. There was already contact between the two projects as they share a common partner. This has materialised in a joint session at Scientix in October 2014 and also a joint session at NARST in 2015.