

[www.sails-project.eu](http://www.sails-project.eu)



## **Report from EAP on assessment materials (Part B)**

### **D7.3 Report from EAP on assessment materials (Part B)**

---

Authors: Bronwen Cowie, Odilla Finlayson, Eilish McLoughlin 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](mailto:info@sails-project.eu)  
Website: [www.sails-project.eu](http://www.sails-project.eu)



The research leading to these results has received funding from the European Union's Seventh Framework Programme for research technological development and demonstration under grant agreement no 289085  
This deliverable contains original unpublished work except where clearly indicated otherwise. Acknowledgement of previously published material and of the work of others has been made through appropriate citation, quotation or both. This document does not represent the opinion of the European Union, and the European Union is not responsible for any use that might be made of its content.

**SAILS External Advisory Panel Member Report**  
**Professor Bronwen Cowie, The University of Waikato**  
**April 2015**

## **Preface**

Thank you for the invitation to be part of an ambitious and important project. It was clear to see that despite its diversity the group is committed to the project and greatly valued the opportunity to discuss and debate the case studies of IBSE from their different contexts.

## **Introduction**

The SAILS project has the objective of developing the teaching, learning and assessment of inquiry based science education (IBSE) for second level students (aged 12-18 years) across Europe. The focus is on the development of resources and of models for teacher education in IBSE, both pre-service and in-service. The proposed resources include unit plans, strategies and frameworks for the assessment of IBSE skills and competences and case studies of the units in action in different country settings. The intention is that through this unified approach a sustainable model for IBSE will be achieved for each of the twelve participating countries. This goal, and the work package sequencing, which progresses through a process of developing a framework for understanding the various aspects of science inquiry, to consideration of how these aspects (skills and attitudes) might be taught and assessed fits with current research that indicates making substantive change requires attention to curriculum, pedagogy and assessment (e.g. Bernstein, 1971, 1990). Unfortunately, assessment is often the last area to receive attention meaning change in assessment tends to lag behind and often impedes changes in curriculum and pedagogy (Bell & Gilbert, 1996). The SAILS project avoids this trap by having an explicit focus on formative assessment as part of curriculum and pedagogical change whilst also paying attention to the demands of summative assessment. In so doing, it raises the possibility that assessment will act as a driver of innovation and learning (Wyatt-Smith, 2014).

In what follows I comment on the structure and focus of the case studies of practice that have been developed by teachers who have taught the SAILS units. This commentary relies on my viewing of a selection of the cases and my participation in the conversations that took place at the SAILS General Assembly meeting in March 2015. My commentary is structured around the themes that emerged. In most instances I provide an overview of the theme and then highlight some of the questions/ issues associated with that theme. I was struck that the matters raised by the cases and conversations scope the issues and challenges being discussed by the assessment research community. The group could usefully consider whether these questions are explicitly included as part of the final version of the materials made available for teacher education/ teacher educators. While I suspect the decision might vary from country to country depending on country policy and practice context there would seem to be some merit in teachers knowing that others are experiencing and struggling to resolve some/ many of the challenges they are facing.

### **The structure of the case studies**

The SAILS partners in each country have put a lot of time and effort into the development of the teacher case studies– they were interesting and engaging. The use of a set sequence of prompts to structure the cases eased the reading demands, although there was sometimes a sense of repetition.

In most instances the case study reports indicated that project partners were not in the classroom gathering data on student and teacher actions and interactions during lessons. This perhaps explains the lack of specific detail on student ideas in most of the cases. This lack is a limitation in terms of providing illustration and evidence for teachers about the nature of possible student responses, and of formative interactions and feedback. Where data does exist it brings the cases to life - if national partners have more examples of student written work, preferably annotated with teacher feedback, I would advise that this be included.

The discussions at this general assembly meeting, which involved partners from at least three different countries discussing case studies of teachers piloting the same unit, were rich and robust. These discussions helped contextualise and highlight the nuances and implications of different teacher actions. This material could usefully be synthesised to provide an overview to each unit and its set of case studies. As I see it, one benefit of the presenting 3-4 cases of implementation will be that teachers will not be led to think there is one right way to introduce and for students to undertake a particular inquiry. Highlighting differences in the cases, along with reasons for these differences, will signal to teachers that they can and should adapt them to meet the needs and interests of their students/ setting.

I did wonder, especially when thinking about teachers for whom SAILS units might be one of the first inquiries they orchestrate, whether additional information could be added to explain how teachers might set the scene to introduce a particular unit. In my experience a successful first step is key to teacher comfort and student engagement/ buy-in when teachers are trialling new kinds of activities.

A suggestion for the unit leaders is to consider how some units/cases might be annotated to highlight variations across and within different pedagogical approaches such as the way a mindmap/ brainstorm was conducted and why and how the information gained was used etc. I observed some groups having discussing on this point and I suggest this aspect be considered in the final presentation of the SAILS Units.

The Units read as more or less open - could this be signalled? In making this suggestion I also recognise that teachers can enact inquiries as more or less open so it is not possible to be prescriptive.

There was some indication in the general assembly discussions that some units more readily lent themselves to high levels of student engagement and discussion. For example, the trials and the wider experience of the Kings College group indicated that the “Floating oranges” unit typically garners a positive response from students and offers teachers a positive experience of what productive inquiry might sound, look and feel like for them and their students. The Floating Orange cases illustrate variations in how it can be approached and evidence of teacher formative practice. Perhaps this unit or others could be identified as being useful introductory ones? More detail of likely student responses and possible teacher feedback could be included in just this (and the one other unit and its cases) as starter units.

## **Themes from the case studies and the general assembly discussions**

### ***The role of group work/ collaboration***

Most of the cases included some use of group work, which is not surprising given the criteria. Groups varied in size from 2 to 4/5 students. Across the cases there was some indication that teachers were more or less experienced and comfortable with using group work. As teachers tend to overlook the fact that two students do constitute a group, it may be worth pointing this out as a legitimate but low risk first step to involving a class in group work given this can pose challenges for teachers and for students.

In many instances case reports indicated that the groups were self-selected by the students, although in some cases teachers gave reasons for allocating students to groups. There is potential include a prompt somewhere in the material to provoke teachers to consider their reasons for the use of group work/ collaboration as part of inquiry and to prompt them to consider how and why they might allocate students to different groups. Reasons could be distilled from the cases - to make the inquiry task manageable through the distribution of roles, to provide access to a range of ideas and safe opportunities to discuss these, because scientists share and debate ideas as part of generating new knowledge etc. Bringing this decision and its pedagogical implications to the fore is likely to have flow on effects due to its influence their expectations and focus for evidence gathering and feedback as part of formative assessment.

### ***The unit for assessment – individual, group, class***

From an assessment point of view the unit of assessment, that is whether the focus is on individual thinking (as in test responses), how an individual uses resources, group problem-solving/ discussion/ inquiry or how the collective thinking of a class evolves is linked to how learning is viewed – individual cognition/ construction or a situated or sociocultural understanding (Gipps, 1999; Cowie, 2005; Elwood, 2006). The units and cases do not focus on this aspect but discussion of this point might be an interesting addition to the SAILS Teacher Education Programme. Then again it might simply complicate matters.

The focus in the cases and the general assembly discussion was more practical than theoretical and tended to revolve more around what would count as adequate and useful information and for what purposes: Was it sufficient to sample a small group of students within a whole class discussion to make judgments about next steps/ formative responses? What constitute an adequate sampling of individual and group responses for formative feedback purposes? Was it /when was it appropriate to focus on assessment group progress? How to track individual progress within a group? There was some evidence that teachers keeping a detailed record of all individuals across all groups was not practical – they were taking all their time recording against criteria and so had no time for teaching/ formative assessment. To some extent this aspect links back to the earlier point about the how and why of group work/ collaboration. This might be an issue that this included as a discussion point, with indications that to some extent the decision relates to purpose. I suggest the teams are targeting some of these issues in their professional development sessions and if not it may be important to discuss this matter with teachers so that they are clear about reasons for various foci and can discuss these with those who might pressing them to be accountable for individual/ class progress.

At the general assembly, some groups discussed how information might be collated over time to allow teachers to report on individual progress and achievement across time and contexts.

This would also be a worthwhile focus for discussion, including for example, how ICTs might support this given the cases read as though students spontaneously used some of these means. Attention to this aspect has theoretical and political implications – attention to context and time aligning with sociocultural views of learning and attention to development over time aligning to an interest in developing the competencies for lifelong learning.

***Which learning outcomes to teach and which to assess?***

An overview of the 20 units and the possible learning outcomes for each was presented in a draft version of Deliverable 3.3. At the general assembly meeting, the small and whole group discussions on the Units raised a number of issues in relation to the nature and number of learning outcomes associated with each unit both in terms of an explicit teaching and an explicit assessment focus. Groups also considered which units might better illuminate and require particular skills.

Obviously, teachers and students cannot reasonably be expected to teach/learn to all the possible learning outcomes that are/ could be encompassed within any one unit. However, given the importance attributed to teachers and students be aware of the intended learning as part of teacher formative and student self-assessment (Black & Wiliam, 1998), some clarification of the unit learning outcomes is warranted. Some thoughts on this follow:

- Could the particular outcomes that are the focus for each case be identified more clearly? Would it be worth including a map of which outcomes each case foregrounds within those that are possible for each unit?
- Would it be possible to identify which units more readily lend themselves to fostering particular outcomes so that teachers could make a more informed selection of units to use based on their goals for their students inquiry skill development? Could the team/ should the team identify some sort of sequence that details how a sample of the units might be worked through to develop increasingly sophisticated understandings of a particular skill? This would take account of the notion that learners need to experience an idea in more than one context (Nuthall suggests three) in order to reasonably master this and to be able to transfer it to other contexts.
- Is there a need to ensure consistency within the language and meaning of the same skill/ skill set across the units and case studies? Would a glossary help address this matter?

***Use of a variety of modes /media for student to demonstrate their learning***

There was evidence across the cases of teachers using a range of sources of information on student ideas/ learning – student questions, brainstorming, their actions and products, writing, drawing, videos, photographs etc. Some cases included examples where one teacher had provided more than one opportunity and mode for students to use - talking, writing, using a cell phone to take video. This is in line with current thinking about how to address equity issues within assessment (Gipps & Murphy, 1995; Stobart, 2008) and could be highlighted. This said, there is quite a strong reliance on written products in part I suspect because teachers generated the cases. I suggest consideration be given to commenting on this reliance in terms of how it articulates what is important in the inquiry process which is, by implication, the final product rather than the process of investigating. Is this what the team intends?

This variety of information sources raises questions that are the subject of debate in the field of assessment such as how teachers might amalgamate and or report on diverse sources of evidence as part of summative assessment. Is it valid/ reliable for learning to be summed up through the accumulation of data across a range of times and settings.

***Grading / the use of rubrics/ use of teacher – student developed criteria***

Many of the cases included mention of the use of rubrics. It was not clear however to what extent this use enabled or constrained teacher formative focus on what students were learning as compared with whether they were learning what the teacher intended/ had planned for. Torrance and Pryor (1995) distinguish between divergent and convergent formative assessment in this regard. Torrance (2007) warns against 'criteria compliance' as a replacement for learning with understanding.

Across the cases there are examples of more and less sophisticated use of rubrics – in a number of cases it would appear that this was the first time teachers had used criteria of quality/ success criteria/ rubrics within their teaching. It was interesting that a number of teachers reported that next time they would share the criteria with their students to support student self-assessment. In other cases, teachers reported incorporating previously established and used criteria and or developing criteria with students as well as active student use of criteria. Teachers might find it helpful to have a guide that highlights which units describe starting and which more sophisticated use of criteria. This would allow them to map a developmental progression for their use of criteria.

***Feedback***

There was very little evidence of feedback as dialogic or of written feedback in the cases I reviewed. This is of concern given that this is the most challenging and most often omitted step in formative assessment. If possible I would suggest asking the participating teachers about this aspect and also including some comment about this aspect in the final version of the Units.

I am not able to make a substantiated comment on the use of peer feedback but this would seem likely to have been part of group work – mentions could be highlighted to signal this as an idea worthy of development by teachers who read the cases.

***Equity issues***

Gender – this is addressed directly through the reporting structure. The most striking example of this was the teacher who took account of his Muslim student's needs.

Cultural – I was surprised that there appeared to be very little on cultural difference in terms of world views but I suspect this is due to the nature of the topics. From my own experience, the egg drop activity would raise issues around the use/ potential waste of food in some cultures. The oil spill activity could also provide a forum for discussion about how different cultures think about their relationship to the land and the broader and deeper implications of pollution beyond the economic to the spiritual. This would certainly be the case in New Zealand.

See: Solano-Flores, G., & Nelson-Barber, S. (2001). On the cultural validity of science assessments. *Journal of Research in Science Teaching*, 38(5), 553-573.

Values – The Woodlice Unit group discussed the ethics of working with woodlice. To what extent does the group view values/ ethics as an aspect of scientific literacy in the sense that scientists operate within in a particular ethical framework or code?

Second language learners - this issue was not really addressed in any of the cases I read. The assumption seemed to be that all students in a class would have the same first language. Is

this the case? I assumed that in most of the participating countries classes would include students for whom the language of instruction was not their first language. Does this mean that this matter is taken for granted and or so effectively addressed it is not an issue worthy of attention across the cases, final materials and or teacher education programmes?

### ***The role of ICT***

Some cases talk about the use of cell phones to take video and photos but on the whole very little mention was made of the role ICTs can play in assessment – documenting evidence of learning, serving as a repository of evidence, supporting more immediate/ detailed peer and teacher feedback.

See: Williams, P. J., & Newhouse, C. P. (Eds.). (2013). *Digital representations of student performance for assessment*. Rotterdam, The Netherlands: Sense Publishers.

### **References**

- Bernstein, B. (1971). On the classification and framing of educational knowledge. In M. F. D. Young (Ed.), *Knowledge and control: New directions for the sociology of education* (pp. 47-69). London: Collier Macmillan.
- Bernstein, B. (1990). *The structuring of pedagogic discourse. Volume IV Class, codes and control*. London: Routledge.
- Cowie, B. (2005). Pupil commentary on assessment for learning. *The Curriculum Journal*, 16(2), 137-151.
- Gipps, C. (1999). Sociocultural aspects to assessment. *Review of Educational Research*, 24, 353-392.
- Hay, P. J., & Penney, D. (2013). *Assessment in physical education: A socio-cultural perspective*. London: Routledge.
- Stobart, G. (2008). *Testing times: The uses and abuses of assessment*. London: Routledge.
- Torrance, H. (2000). Postmodernism and educational assessment. In A. Filer (Ed.), *Assessment: Social practice and social product* (pp. 173-188). London: Routledge Falmer.
- Williams, P. J., & Newhouse, C. P. (Eds.). (2013). *Digital representations of student performance for assessment*. Rotterdam, The Netherlands: Sense Publishers.
- Wyatt-Smith, C., Klenowski, V., & Colbert, P. (2014). *Designing assessment for quality learning*. In C. Wyatt-Smith, V. Klenowski, & P. Colbert (Eds.), *Foundation book in the series: The enabling power of assessment*. Dordrecht, The Netherlands: Springer.

**Response to: SAILS External Advisory Panel Member Report**  
**Professor Bronwen Cowie, The University of Waikato**  
**April 2015**

The SAILS project steering committee welcome and thank Professor Cowie for her input and advice for the SAILS project and acknowledge the work that she has done in terms of scrutinising a range of draft units and case studies that will ultimately inform the wider community on inquiry and its assessment in the classrooms. Her subsequent input to the General Assembly discussions in March 2015 is also gratefully acknowledged.

Professor Cowie's report has raised many issues and it is interesting to note that she comments that 'the matters raised by the cases and conversations scope the issues and challenges being discussed by the assessment research community'. Therefore we acknowledge that within the project, that it will not be possible to develop a fully coherent strategy for assessment of all inquiry skills; however, our aim is to highlight a number of inquiry skills and to give examples, through case studies, of how different strategies can be used in their implementation and assessment. To this end, we have focussed on the trialling and collation of 20 SAILS Units, which each present case studies from at teachers from across at least three of the participating countries.

We welcome the comments supporting this approach, that 'teachers will not be led to think there is one right way to introduce and for students to undertake a particular inquiry..... Highlighting differences in the cases, along with reasons for these differences, will signal to teachers that they can and should adapt them to meet the needs and interests of their students/setting'.

Professor Cowie has raised number of issues which will inform/add to our discussions and further development of the SAILS inquiry and assessment materials.

**Structure of Case studies:**

The acknowledgement of the extensive work carried out by the consortium and teachers in the generation of the units and case studies is welcomed. The value and limitations of the case studies are outlined and the consortium will attempt to record as much information as possible in the case studies, e.g. specific detail on student ideas, teacher first steps, student and teacher actions and interactions during lessons.

As the case studies describe short teaching episodes (up to a few classes), it is not envisioned that this sequence of activities would solely be used for the development of a particular skill. However, this point should be highlighted within the Teacher Education Programmes and particular sequences should be identified and utilised accordingly.

A key component of the final version of each SAILS Unit will be the inclusion of a descriptive synthesis to 'contextualise and highlight the nuances and implications of different teacher actions', as informed by the small group discussions on each Unit and its case studies.

The requirement of clear articulation of each of the skills will be addressed in the forthcoming SAILS Framework for assessment of inquiry skills (Deliverable 2.4).

**Themes from Case Studies and GA Discussion:**

The contributions of Professor Cowie to the discussions on the Case Studies were particularly welcomed and her comments have directly been incorporated into the preparation of the final versions of the Units and case studies. She has highlighted the need for the inclusion of as much detail as possible in the case studies, particularly including the teacher's reasons and rationale for the structure of their lesson e.g. why they used group work, why did they decide on the size of the group, their desired learning outcomes, the modes of assessment used, etc.

Professor Cowie has raised a number of issues that we recognise as key aspects of inquiry and assessment, and believe these should be addressed within the Teacher Education Programmes. We intend to use the collection of Units and case studies that have been developed through the project to address these key ideas in the final SAILS Teacher Education Programmes:

- Role of group work in the inquiry classroom
- Unit of assessment – is this individual, small group or whole class
- Sequencing of activities and development of inquiry skills and their assessment
- Modes of assessment and their use, including ICT
- Development and use of assessment criteria
- Nature and role of feedback.

Finally, recommendations regarding equity issues were highlighted by Professor Cowie and include gender, cultural, values and second language learners. We feel these issues should be highlighted within the Teacher Education Programme by making use of particular examples that have been included in the case studies.