



Report of the Working Group on International Collaboration in the Evaluation of Inquiry-Based Science Education (IBSE) programs

Contents

Executive summary

Introduction

1.0 Overview of the project

- 1.1 Rationale
- 1.2 The meaning of IBSE
- 1.3 The aims of the project

2.0 IBSE in operation: goals and implementations

- 2.1 Learning objectives for IBSE: what should students know?
- 2.2 Indicators and outcomes of IBSE

3.0 Purposes and procedures of program evaluation: general considerations

- 3.1 Purposes of program evaluation
- 3.2 Stages of programme implementation and evaluation
- 3.3 Activities of the evaluator and of the teacher

4.0 Evaluation for the improvement of program implementation

- 4.1 Outline of procedures

5.0 Evaluation of the effects of IBSE

- 5.1 Procedures for developing an evaluation of program effects
- 5.2 Developing tests and tasks to measure student outcomes of IBSE

6.0 Mechanism for the organisation and implementation of the IBSE International Collaboration in Evaluation Process

- 6.1 Proposed mechanism for implementation of the local evaluation
- 6.2 Dissemination of the Working Group Report to regional and international meetings sponsored by Academies

7.0 Recommendations

- 7.1 Recommendations in relation to this report

7.2 Recommendations for additional work by the IAP

Appendices

- A1 Profiles of Working Group members and observers
- A2 List of reviewers of the draft report
- B Selected bibliography
- C1 Countries known to have IBSE program in operation
- C2 Some brief descriptions of some programs
- D Research in conjunction with IBSE
- E Examples of instrument for collecting evaluation data
 - E1 Classroom observation
 - E2 Questionnaires for teachers, administrators and students
 - E3 Interview guidelines

Report of the Working Group on International Collaboration in the Evaluation of Inquiry-Based Science Education (IBSE) programs

Executive Summary

This report has been produced by the International Working Group set up by the InterAcademies Panel (IAP) on International Issues. The Working Group was charged with developing a proposal for collaboration with the evaluation¹ of the implementation of Inquiry-Based Science Education (IBSE) programs² for pre-secondary school³ students in different countries. At least 30 countries, both developing and developed, are known to be implementing some form of IBSE in some of their pre-secondary schools, creating a need for information about the impact of these programs on students and on teachers.

Aims and procedures

The aims of the Working Group are to provide collaboration in relation to one or more of the following

- the collection of evidence about the extent to which what is being implemented in classrooms matches the intentions of the IBSE curriculum, to feed back to the local program developers information as to how implementation could be improved;
- the development of instruments for assessing students' learning in science that are suitable for assessing the outcome of IBSE;
- the design of evaluation projects that enable a comparison to be made between IBSE programs and conventional science education programs, or that enable assessment of students' progress from year to year within an IBSE program;
- the design of research projects that aim to increase understanding of IBSE and which can lead to the refinement of teaching methods, material and professional development for IBSE.

Once established, following the recommendation at a workshop in Stockholm (September, 2005), the working group sought to achieve these aims through two meetings, held in Washington (March 2006) and Paris (June 2006), and through

¹ Evaluation is the term used here for the process of collecting evidence and making judgments about systems, programs, materials, procedures and processes. We use the term 'assessment' to refer to the process of collecting evidence and making judgments relating to outcomes for students' and others, such as achievements and attitudes.

² Programs refers not only to the written and non-written materials and equipment for use in the classroom, but the professional development provided for teachers, for school management, inspectors, etc, and the mechanisms for informal support from collaboration among schools.

³ We use the term 'pre-secondary' to denote classes and schools where the students are mainly taught by a single non-specialist teacher and where the students' range in age from 5 to 12 or 13 years. In some countries these schools are described as 'primary', in others as 'elementary' and sometimes 'middle' schools at the upper part of this age range.

preparing and revising drafts of this report between meetings. Nine international experts, including scientists, science educators and program developers were asked to review an advanced draft of the report and their views have been taken into account in producing the final document.

The meaning of IBSE

The Working Group views IBSE as being not a single pedagogical method, but an approach having key features that can be implemented in various ways. IBSE shares some features with traditional science education but differs from it in many respects that go beyond the manipulation of materials to the key factor of engaging students in identifying relevant evidence, in critical and logical reasoning about it and in reflection on its interpretation. Some key distinguishing characteristics of IBSE are:

- Students are developing concepts that enable them to understand the scientific aspects of the world around them through their own thinking using critical and logical reasoning about evidence that they have gathered. This may involve them in first hand manipulation of objects and materials and observation of events; it may also involve them in using evidence gained from a range of information sources including books, the Internet, teachers and scientists.
- Teachers are leading students to develop the skills of inquiry and the understanding of science concepts through the students' own activity and reasoning. This involves facilitating group work, argumentation, dialogue and debate, as well as providing for direct exploration of and experimentation with materials.

The Working Group has further spelled out these general statements in terms of the inquiry practices of the teachers, the inquiry experiences of the students and the outcomes of students' learning of inquiry processes, science concepts, attitudes and dispositions.

Evaluation roles

Evaluation has several roles in the development and implementation of IBSE programs. The Working Group has focused its work on the roles that evaluation can take once a program has been developed, or has been adopted or adapted from a program developed elsewhere. The important roles of evaluation in the course of program development are not included, as this would extend the project beyond its capacity. It is recommended that a separate project be set up to serve these purposes.

Implementation of an already identified program passes through various stages along a dimension from early implementation, where only a small number of classes and schools may be involved and not all aspects intended may be in operation, to more advanced implementation, where the program has been well established in most respects in a larger number of schools. The processes and role of evaluation vary according to the point reached along this dimension of implementation. At early implementation stages the focus of evaluation is formative, to provide information about how the implementation can be improved. The evaluation is designed to address questions such as: Are the activities that the teacher and students are engaged in consistent with those intended in the program? Is the content which engages the teacher and students

consistent with the intended content? Do the teacher-student interactions and the student-student interactions match those intended? Is the nature of the classroom discourse consistent with that envisaged in the program? At the early stages the focus needs to be on classroom practices rather than on student achievement; the assessment of student outcomes becomes worthwhile at the later stages, when there is evidence that the students have the experiences that can lead to intended learning. Thus, at these later stages, in addition to evidence about how teachers use the IBSE materials and students' experiences, data are collected on students' inquiry skills, knowledge and understanding of science concepts and attitudes.

Evaluation procedures

When the purpose of the evaluation is to provide and feed back information about the early implementation of a program, the data collected will come from close observation of the teacher and students, from interviews with and questionnaires for teachers, administrators and students, and from review of teachers' journals and students' notebooks and artefacts.

At later stages of implementation the evaluation will provide summative information on students' learning of inquiry skills, science concepts and attitudes, in addition to information about classroom processes. The design of the evaluation will depend on the questions to be addressed. In some cases these will be about the extent to which the desired outcomes as specified in the program are realized. In other cases the questions will focus on whether student outcomes are the same or different (better, worse) when the IBSE program is compared with other science education programs. Addressing questions of IBSE compared with other programs requires careful evaluation design, choice of IBSE and of comparison classes and selection and creation of instruments. Where possible, existing assessment instruments ought to be used, supplemented by assessment items that require the specific skills and understanding that are outcomes of IBSE. Examples of evaluation instruments are given in appendix E to this report.

Reporting and feedback procedures

Countries participating in this project will be assisted in producing reports at different levels of detail for different audiences. Where appropriate, local and international evaluators will hold seminars to provide face-to-face feedback to different users of the results. However, the purpose of the international context is to add value to each individual evaluation study by sharing procedures and findings across countries. Thus, in addition to its unique value to the country in which the evaluation was conducted, the findings will contribute to a wider understanding of the effects of IBSE in different circumstances.

It is clearly understood that collection of reports from different countries will not be used to rank or compare countries in terms of the quality of IBSE programs nor of the performance of their students. The development of synoptic reports, bringing together findings from different countries will only be used to extend the vision of each participating country as science education continues to be developed internationally as well as permitting the accumulation of evidence on the effects of IBSE across contexts and countries.

Operation of the international evaluation project

It is proposed that the operation of the international project will be overseen by an International Oversight Committee (IOC) of about four or five members. The IOC will initially draw up a list of international experts in IBSE evaluation and research from which members are recruited for International Evaluation Teams (IET) to support particular evaluation projects. Countries wishing to participate in the international project will approach the IOC and, with its help, produce a proposal indicating the purpose of and the potential design of the evaluation study. The IOC will then identify, from the list of experts, an IET whose members will support the local evaluation team in planning, carrying out and reporting the evaluation. Reports will be published by each country and, in addition, will be available to the IOC for the purpose of producing a synthesis of reports from different countries which will then be available to all.

Recommendations

In relation to the international evaluation project, the Working Group recommends that:

- Countries implementing IBSE programs should evaluate the extent and effect of implementation and set up local evaluation teams for this purpose.
- An International Oversight Committee be set up to manage the provision of support for implementation evaluation to countries who request such support.
- International Evaluation Teams work with local evaluation teams to support the evaluation and enhance local capacity to conduct it.
- The reports of procedures and findings of evaluations conducted in various countries are made available to others and are synthesised by the IOC to add to general understanding of IBSE and its implementation in various contexts.

It is further recommended that the IOC should report to the IAP annually on the International Collaboration in Evaluation Project and, through the IAP, should:

- Develop a handbook to guide evaluation design and procedures and a compendium of pertinent research literature for the evaluation of IBSE.
- Hold an international workshop to discuss the outcomes of the synthesis of reports from different countries at a time when sufficient data have been accrued.

Finally, the Working Group identifies the following work outside this evaluation project that it recommends the IAP to consider in relation to IBSE:

- A project to provide help for countries in the design and development or adaptation of IBSE programs.
- The creation of guidelines aimed at enabling countries to generate for inclusion in their IBSE programs, materials that provide help for teachers in assessing their students and evaluating their own effectiveness in implementing IBSE.
- Establish a program of post graduate fellowships that will support capacity building in countries where there may be a lack of evaluation and research expertise.
- The promotion of research projects aimed at increasing knowledge and understanding of the nature of inquiry learning in science and how it is best implemented in various contexts.

Report of the Working Group on International Collaboration in the Evaluation of Inquiry-Based Science Education (IBSE) programs

Introduction

This report is the product of the International Working Group on Science Education. The Group was set up by the InterAcademies Panel (IAP) on International Issues following the decision of the Academies of Sciences of the World to work in collaboration on the improvement of pre-secondary science education, under the leadership of the Chilean Academy of Science and the co-ordination of Professor Jorge Allende. A workshop on Evaluation of Inquiry-Based Science Education, convened by the IAP Science Evaluation Program and held in Stockholm in September 2005, recommended the establishment of a small multinational Working Group to produce a report building on the discussions of the workshop. At its first meeting in Washington, March 16/17 2006, the Working Group members and observers planned and began work on the preparation of a proposal for ‘an International Platform that could provide assistance in the evaluation of Inquiry-Based Science Education (IBSE) projects and programmes carried out in different countries of the world.’ (Terms of Reference). The Working Group agreed on the form of the report, to include guidelines on evaluation purposes, possible designs and methods, and on the dissemination and implementation of the international evaluation project. Members formed into sub-groups to work on different aspects of the proposed report.

Subsequent to the first meeting the sub-groups continued to work, communicating through e-mail, to create the sections and appendices that are brought together in this report. Comments from all members, following circulation of successive drafts, were used to prepare the report discussed at a meeting in Paris, June 15/16, 2006. Following that meeting, a revised draft was sent for comment to nine reviewers who were international experts including scientists, science educators and program developers. Reviews were received from those are listed in Appendix A2. This draft takes account of the reviewers’ comments and suggestions.

The first of the seven sections of the main body of the report gives an overview of the rationale, aims, purposes and benefits of the international evaluation project. The project is intended to serve different purposes according to the stage(s) of implementation of IBSE in a particular country. It is emphasised that a purpose that the project does not serve is to promote comparison of student achievement among countries, one of several points that distinguishes it from international surveys such as PISA⁴ and TIMSS⁵.

⁴ Programme for International Student Assessment www.pisa.oecd.org

⁵ Trends in International Mathematics and Science Study <http://nces.ed.gov/timss>

The brief discussion in section 1 of the meaning of IBSE, is developed in section 2 of the report into the kinds of practices and student learning outcomes that characterise the classrooms where IBSE is in action. Section 3 discusses some general considerations concerning the purposes of evaluation at different stages of implementation of a program along a dimension from a focus on improving implementation to a focus on the learning outcomes that are achieved once an IBSE programme is being implemented as intended. This section serves as an introduction to sections 4 and 5, which deal respectively with evaluation for improving implementation and evaluation of the effects and outcomes of a well established IBSE program.

Section 6 spells out how the international evaluation project will operate, the help that participant countries will receive, the implications of commitment to evaluation of IBSE, and how information about the project will be disseminated. Finally, a brief concluding section lists recommendations from the Working Group. Some of these recommendations follow from the arguments and evidence in the report; others refer to work that is outside the remit of the Working Group but is considered important for the further development and understanding of IBSE across the world.

Several appendices provide material that illustrates, or expands on, points made in the main sections of the report.

All members and observers of the Working Group have had some part in writing this report. Individual contributions are difficult to identify in many cases as successive drafts of sections and the whole are modified collaboratively. Thus we list all as joint authors of the report. Wynne Harlen has had the main responsibility for editorial drafting of the report.

Authors of the report (please see Appendix A1 for brief profiles)

Members of the Working Group

Julie Campbell
Ernst Hamburger
Wynne Harlen
Glenn Hultman
David Klahr
Jean Matricon
Jayashree Ramadas
Patricia Rowell
Richard Shavelson

Observers

Bruce Alberts
Jorge Allende
Philip Bell
Sally Goetz-Shuler
Pierre Léna
Jean Moon
Yves Quéré
Senta Raizen

1.0 Overview of the project

Inquiry-Based Science Education (IBSE) is being widely advocated by science educators across the world and is being implemented in some schools in both developing and developed countries. A list of those countries known to have programs in operation in pre-secondary schools is given in Appendix C, which also includes brief profiles of a selection of five programs to show something of the range of ways of putting IBSE into practice. While IBSE shares some features with traditional science education, its main distinguishing characteristics are summarised in Box 1.

This overview section seeks to provide the reasons for the development of the evaluation project, how IBSE is understood within the project and what the project's aims include and exclude. In other words, it deals with 'why' the project is proposed, whilst later sections provide guidelines to 'what' and 'how'.

Box 1

Summary of some distinguishing characteristics of inquiry-based science education (for more detail see section 2.2)

- **Students** develop concepts that enable them to understand the scientific aspects of the world around them through their own thinking using critical and logical reasoning about evidence that they have gathered. They will be involved in
 - first hand manipulation of objects and materials and observation of events;
 - using evidence from a range of other sources of information including books, the Internet, teachers or scientists;
 - raising questions for investigation, making predictions, planning and conducting investigations, solving problems, testing ideas, reflecting on new evidence and developing new hypotheses
 - collaborating with others, sharing their ideas, plans and conclusions; advancing their own understanding through dialogue with others.
- **Teachers** lead students to develop the skills of inquiry and the understanding of science concepts through the students' own activity and reasoning. This involves facilitating group work, argumentation, dialogue and debate, as well as providing for direct exploration of and experimentation with materials and access to information sources.

Note on terminology.

We use the term 'pre-secondary' to denote the age range and type of school with which this report is concerned. This is roughly from age five to 12 or 13 years, when students

are largely taught by a single non-specialist teacher. In some countries, schools for students of this age range are called ‘primary’, in others ‘elementary’ and sometimes ‘middle’ schools at the upper end of the given age range.

We are aware that the terms ‘evaluation’ and ‘assessment’ are used in different ways and in some case interchangeably. Here we use ‘evaluation’ to denote the process of collecting evidence and making judgments about programs, systems, materials, procedures and processes. ‘Assessment’ is used here to refer to the process of collecting evidence and making judgments relating to outcomes, such as students’ achievement of particular goals of learning or teachers’ and others’ understanding. The processes of assessment and evaluation are similar but the kind of evidence, the basis of judgement and uses of the data differ.

We refer to ‘programs’ here meaning the whole range of actions, resources and initiatives needed to implement IBSE. These are likely to include not only the provision of relevant written and non-written materials and equipment for use in the classroom, but the professional development provided for teachers, for school management and for inspectors or advisers who influence what happens in classrooms, and the informal support from collaboration among schools in the program. Steps taken to inform parents and the wider public are also relevant, as is ensuring that agencies responsible for the curriculum and examinations are aware of the changes being made and the reasons for them.

1.1 Rationale

1.1.1. Reasons for the adoption of IBSE

Learning science through inquiry promises to improve students’ understanding, participation and enjoyment in relation to scientific activities and contributes to improving general education. Through engaging in the processes of scientific inquiry, students acquire scientific literacy, meaning a general understanding of: the important ideas of science, the nature of scientific investigation and the evaluation and interpretation of evidence. These are outcomes of education that are important for all citizens, not only for those who will pursue science-based occupations in adult life.

The development of scientific literacy has to start in the pre-secondary school grades, for the ideas and frameworks that students need to know are ‘big’⁶ ones that cannot be taught directly but need to be built from the ‘small’ ideas relevant to the objects and events familiar to the students. The role of pre-secondary school science is to help in building understanding by ensuring that the ‘small’ ideas are consistent with evidence and not just the students’ preconceived ideas. Research in students’ own ideas⁷ has shown how important this is.

⁶ ‘Big’ ideas denote the broader, more abstract concepts that are formed from the ‘small’ ideas developed through specific activities. Examples are concepts of systems, change, adaptation.

⁷ See for example:

Black, P. J and Lucas, A. M. (Eds) (1993) *Children’s Informal Ideas in Science*. London: Routledge
Harnqvist, K and Burgen, A. (1997) *Growing up with Science. Developing Early Understanding of Science*. London: Jessica Kingsley and Academia Europaea

There are both theoretical and practical reasons for enthusiasm to embrace IBSE as a way of teaching and learning science. The theoretical arguments relate to what can be achieved in terms of goals of learning in a rapidly changing and increasingly technological society. To prepare for this future, students need: to understand key science concepts, rather than merely knowing certain facts; to recognize the application of these concepts to issues they face in their everyday lives and as citizens of the world; to appreciate the nature, including the strengths and limitations, of science; to know how to learn so that they can use these skills throughout life beyond school. As the OECD points out, students cannot learn in school everything they will need to know in adult life. They must acquire in school the cognitive skills and attitudes for successful learning in future life. ‘Students must become able to organise and regulate their own learning, to learn independently and in groups, and to overcome difficulties in the learning process. This requires them to be aware of their own thinking processes and learning strategies and methods.’⁸

The claim that IBSE can lead to these outcomes is underpinned by modern views of how learning takes place, in which learning is seen as the construction and reconstruction of knowledge by students. Whilst being open to other perspectives on learning, one of the purposes of the project proposed in this document is to provide evidence relating to the claim that learning depends on the active participation of students, leading to the progressive development of understanding, in which inquiry skills play a key role as each new experience is linked to existing ideas. We take up later, in section 2, what this means for the role of teachers and students in the classroom.

In terms of evidence from practice, there is a limited amount of evidence from research and evaluation relating to the processes and outcomes of IBSE and how these compare with those of traditional transmission-based science education. One of the results of this project will be to add to the body of evidence about the outcomes of implementation of IBSE.

1.1.2 IBSE and education in other subjects

Among the distinguishing characteristics of IBSE are its emphasis on rational thinking and the development of ideas by learners through discussion and argument as well as through first-hand investigation of objects and materials. For example, one of the overarching aims of the NTA program in Sweden (largely developed from the ‘Science and Technology for Children’, one of the IBSE programs in the United States) is the encouragement of language development for both students and teachers. Language development, both written and spoken, is also considered a central aim in the French program *La main à la pâte*. Furthermore, through the development of thinking about evidence, the value of IBSE extends beyond language literacy and mathematics to thinking in other subjects, particularly geography, history and social studies. These

Osborne, R. and Freyberg, P. (1998) *Learning in Science. The Implications of Children’s Science*. Auckland: Heinemann

Eaton, S.S. and Viennot, L. (2001) *Reasoning in Physics. The Part of Common Sense*, Amsterdam: Kluwer Academic Publishers

⁸ OECD (1999) *Measuring Student Knowledge and Skills. OECD Programme for International Student Assessment (PISA)* Paris: OECD p. 9.

arguments add weight to the case for implementing IBSE from the earliest years of education, for it contributes to basic literacy and learning in other areas as well as developing understanding of the natural or made world around. Once IBSE is well established as part of the students' experience then the effects may well be detectable in performance across other core subjects of the pre-secondary school curriculum.

1.1.3 Questions raised about IBSE

Any innovation is inevitably faced with questions about whether its claims can be substantiated. 'Does it work?' is a common question. Another question is whether it can be implemented on anything other than a small scale, where teachers and schools can be given support that would not be available on a large scale.

In the case of IBSE, many searching questions are being asked. Although there is no single and fixed form of IBSE, all approaches to it, when first introduced, are likely to require a change in teaching methods and in the relationship between teacher, students and materials, and not just in the materials that are used in the classroom. Without the more fundamental change in teaching, it is all too easy for classroom materials to be adopted without students interacting with them in the way that is intended and which is, indeed, necessary for the intended learning to ensue. Thus before asking about what students are learning from IBSE, there is a prior question:

Is IBSE being implemented as intended?

This question has to be divided into several sub-questions, such as:

Are the principles of IBSE reflected in

- the materials (both those used in active investigation and written) made available to the students?
- the way the students interact with the materials and in what they produce?
- how the students interact with each other?
- the teachers' actions and interactions with students?

Only when it is clear that IBSE is well represented in students' experience, and has been so for a reasonable time, is it appropriate to ask about students' learning outcomes. Otherwise, false empirical conclusions can be drawn leading to inaccurate decisions.

1.1.4 The importance of evaluation

It is important that these questions are asked and addressed, for IBSE requires a considerable investment of resources in teacher education, materials and in the effort that teachers have to make to bring about the change in practice required. The purpose of this project is to help countries to answer questions about the processes and outcomes of IBSE whether the implementation of change is just beginning, is at an advanced stage, or is at some intermediate stage (see section 3.2).

1.2 The meaning of IBSE

1.2.1 Definitions of inquiry

The notion of inquiry as an important element of education is not new. It featured in many commentaries and program in education in the 1960's and in several programs in the 1980's⁹. However, more recent discussion has helped to define it in terms of what it means for teaching and learning. Although it is not confined to science education, the definition from the United States' National Research Council (NRC) is given in that context:

Inquiry is a multifaceted activity that involves: making observations; posing questions; examining books and other sources of information to see what is already known; planning investigations; reviewing what is already known in light of experimental evidence; using tools to gather, analyze, and interpret data; proposing answers, explanations, and predictions; and communicating the results. Inquiry requires identification of assumptions, use of critical and logical thinking, and consideration of alternative explanations¹⁰.

A statement from a United States' National Science Foundation (NSF) publication spells out more about students' experiences:

Inquiry teaching leads students to build their understanding of fundamental scientific ideas through direct experience with materials, by consulting books, other resources, and experts, and through argument and debate among themselves. All this takes place under the leadership of the classroom teacher¹¹.

Other definitions, such as embodied in the 10 principles of inquiry developed in the program in France¹², the nine principles of the Chinese program (see Appendix C2), and the five Es of the Australian program¹³ concur with the views expressed in these quotations. From these and other sources we can identify some of the key characteristics of IBSE, as set out in Section 2.

1.2.2 The role of inquiry in learning

It is important to avoid the impression that in an IBSE program all science activities involve all the characteristics of inquiry. The broad definitions of IBSE indicate a complex set of activities for both students and teachers. These are undoubtedly necessary to achieve certain goals of science education such as the development of 'big'

⁹ Raizen, S. (1991) The reform of science education in the U.S.A. Déjà vu or de novo? *Studies in Science Education*. Vol 19, 1-41

¹⁰ NRC (National Research Council) (1996) *National Science Education Standards*. Washington DC: National Academy Press p23

¹¹ NSF (National Science Foundation) (1997) *The Challenge and Promise of K-8 Science Education Reform* Foundations, 1. Arlington, VA: NSF p7

¹² Charpak, G., Léna, P., Quéré, Y. (2005) *L'Enfant et la science*, Paris: O. Jacob [translated into Spanish, Los Niños y la Ciencia, Siglo XXI, Buenos-Aires, 2006].. See also, *La main à la pâte. La science à l'école primaire*, (G. Charpak dir.) (1996) Paris: Flammarion [translated into Chinese, Portuguese, Vietnamese, Serbo-croatian]

¹³ The Australian *Primary Connections* teaching and learning model: Engage, Explore, Explain, Elaborate, Evaluate www.science.ord.au/primaryconnections

ideas¹⁴, the ability to use such ideas in solving problems, the understanding of how scientific ideas arise from evidence and the dispositions of mental flexibility and respect for evidence. However, there are other things that are more efficiently learned by direct instruction. These include, for example, procedures, conventions, names and basic skills of using equipment, which are best taught as they are needed. Thus it is not expected that all science learning will involve inquiry, but it is equally important to ensure that inquiry is used where it is appropriate. For example, students can be taught directly how to name parts of a plant, but to understand the function of the parts, and therefore the needs of plants, may be best achieved through investigation.

Further, which particular inquiry skills will be involved in any particular instance will depend on what is being investigated. First-hand manipulation of subjects under study is not always possible, as not all objects of study are directly accessible to students. For example, when trying to understand the reason for the apparent movement of the Sun, Moon and stars, these are not susceptible to controlled investigation but they can be studied by careful observation, the use of models and the interpretation of findings in relation to different theories or ideas. It follows that not all inquiries will show all the characteristics of IBSE.

In summary, the appropriateness of using inquiry will depend on the goals of the activity and the inquiry skills used will vary with the subject matter and with the age and experience of the students. However, over a period of time it is expected that students have experience of various inquiries and that, in combination, these activities give students opportunities to use and develop all aspects of inquiry.

1.3 The aims of the project

1.3.1 The range of help offered

This project has been initiated by the IAP with the express purposes of collaboration with countries which have programs for implementing some form of IBSE, or where implementation is being planned, to improve science education at the elementary level through evaluation. The Working Group set up to do this is planning, through this collaboration, to help countries develop human resources and materials to conduct evaluation and/or research projects according to the stage(s) reached in their implementation of IBSE. It is envisaged that the IAP project will provide support in relation to one or more of the following

- gathering evidence of how what is implemented in classrooms matches the intentions of the IBSE curriculum, to feed back to the local program developers information as to how improvements could be made
- developing instruments for assessing students' learning in science that are suitable for assessing the outcome of IBSE
- design of evaluation projects that enable a comparison to be made between IBSE programs and conventional science education programs, or that enable progress of students from year to year within an IBSE program to be assessed

¹⁴ See note 6

- design of research projects that aim to increase understanding of IBSE and which can lead to the refinement of teaching methods, material and professional development for IBSE.

In some countries, where implementation is just beginning and only some aspects may be in place, the first of these is likely to be most relevant. (Evaluation with this purpose of improving the program implementation is known as *formative evaluation* in relation to implementation.) This would involve developing and using means of collecting evidence of learning processes and experiences, such as by observation, study of student notebooks, interviews and self-reporting by students and teachers of their practices and preferences. Countries where IBSE is well established may also need to have evidence of whether the intended learning is indeed taking place and to what extent it is similar or different from learning arising out of ‘traditional’ science education. (Evaluation with the purpose of assessing the products of the program is known as *summative evaluation*.) Whether or not there is demand for comparison with other ways of learning, many countries may wish to be able to report on students’ learning and to follow the progression in learning of students taught through IBSE. Countries may also wish to develop research into aspects of IBSE alongside the evaluation of how well it is being implemented or its impact on learning. Some of the kinds of research questions that arise are summarised in Appendix D.

Specific proposals for delivering this help are outlined in section 6 of this paper.

1.3.2. The unique purposes of the evaluation project

At the same time as outlining the project’s main aims, it is useful to make clear what it does not intend to do. The project is *not* creating a program for the comparison of one country with another nor, within a country, the comparison of schools and teachers with each other.

International programs for assessing student outcomes already exist in the IEA and PISA surveys and it is important to recognise how the international project proposed here is distinct in its aims and procedures. These international surveys, and those within countries, such as the NAEP¹⁵ in the United States or the *Evaluation diagnostic des élèves en sciences en fin de primaire*¹⁶ in France, provide information from a random sample of schools and students about student outcomes at specific grade levels. They yield rich information about achievement across a range of curriculum outcomes, but this means that individual students take a small sample of items and the survey results are not meaningful at the student, class or school levels, but only when aggregated at national, state or selected district levels. Although these surveys can report comparative achievement by student variables such as gender, race/ethnicity and geographical region, these are general statements which do not necessarily apply to each and every school. Moreover, their scale precludes the collection of direct and detailed information about classroom processes, practices and experiences, although reported information

¹⁵ National Assessment of Educational Progress <http://nces.ed.gov/nationsreportcard/>

¹⁶ This evaluation will be applied from 2007 onward. For more information : www.education.gouv.fr/stateval/default.htm

about these is collected through student, teacher and school questionnaires. However, particularly at early stages of IBSE program implementation, when the stress needs to be on experiences and practices rather than on student achievement, it is critical to obtain detailed information and observations made at the classroom and individual student level—information not available from larger-scale testing programs.

For evaluation at the stage of full implementation, when attention is given to student outcomes, the IAP project will make use of information and instruments from international surveys, as from other research, consistent with its quite different aims. The IAP project will not provide ready-made instruments for countries to use. Even though similar guidelines will be provided for all countries, the instruments developed within each country for evaluation of learning processes and outcomes are unlikely to be such as to allow comparisons to be made. The intention is clearly to enable each country to gather information required for evaluation of IBSE and to learn from the experience of other countries as to how to ensure further progress in the development and implementation of their own IBSE programs. This learning will be helped by countries including some common items in their evaluation instruments (see section 5.1.2) but this would be for the purposes of learning about the effectiveness of IBSE in different contexts, not the comparison of countries.

1.3.3 The benefits of participation

The values of participating in this international enterprise are of two main kinds. First are those concerned with the provision of assistance to those conducting the evaluation within a country. These benefits include:

- access to international experts, and eventually to written guidance in the form of a handbook, to support the evaluation and enhance local capacity to conduct it;
- procedures for quality assurance in the way that an evaluation is carried out, lending weight to the findings;
- access to exemplary assessment practices and instruments;
- access to relevant research via the international network of experts.

The second kind of help concerns the development of understanding of IBSE and what is needed to promote it in different systems. These benefits include:

- the provision of guidelines based on international experience as to what are essential aspects of IBSE where training and support should be focused;
- the types of initial teacher training and professional development that have been found necessary and effective in other countries;
- access to reports and results of evaluations carried out in other countries, showing what is possible in varying contexts and conditions;
- some ways in which IBSE can be interpreted and adapted to different cultures and traditions.

This second set of benefits will depend on the involvement in the project of a range of developed and developing countries; the more extensive the participation and thorough the commitment, the greater the value for each country taking part.

2.0 IBSE in operation: goals, and implementation

Evaluation of IBSE implementation and outcomes requires the application of the same rigorous methods that students are expected to learn about science itself. Thus IBSE must be carefully expressed in terms of its ultimate goals for student learning, the characteristics of its implementation, and the nature of the evidence that will be persuasive about its effects.

2.1 Learning objectives for IBSE

IBSE advocates a set of learning objectives that go beyond the traditional goals of science education. These can best be characterized in terms of an emergent, multifaceted view of what science is¹⁷. These facets include the following:

2.1.1 *Knowledge about the natural world*

Making sense of events and phenomena is the fundamental core of both traditional and IBSE approaches to science education. That said, there is little consensus about precisely what to include, what to exclude, and in what sequence to present the vast amount of accumulated knowledge about the natural world that could potentially be part of any science instruction at any grade level. In Section 2.2.2 below, we offer a short list of what this body of knowledge might comprise, but it is only suggestive, and definitely not as a listing of “core knowledge”. Such decisions are at the heart of the many different science curricula that exist between and within different countries around the world, and we do not attempt to create a global curriculum in this document. Similarly, the inclusion of some education in technology may or may not be part of science education, depending on students’ level and countries’ conception of technology. We assume some flexibility on this point, neither explicitly including it, nor excluding it.

2.1.2 *A process of observing, questioning and experimenting*

This facet emphasises the peculiar attitude which does not take natural phenomena for granted, but notices them, pinpoints their occurrence, their regularities, and questions the potential existence of an explanation (or several) for the observed or produced effects. The emergence and formulation of a question is the basic ingredient of a scientific attitude. It is closely connected to the progressive mastery of an appropriate, accurate and rigorous language to formulate the question and the attempted answers, then later to the use of mathematical tools enabling measurement, comparison and symbolic formulation of natural phenomena.

2.1.3 *A process of logical reasoning about evidence*

This facet emphasizes the role of processes of ‘reasoning about evidence’ that can be applied across all science content areas and beyond (see section 1.1.2). These processes

¹⁷ Lehrer, R., & Schauble, L. (2006). Scientific thinking and science literacy: Supporting development in learning contexts. In (Eds) W. Damon, R. Lerner, K. Anne Renninger, and I. E. Sigel *Handbook of Child Psychology, Sixth Edition, Volume Four: Child Psychology in Practice*. Hoboken, NJ: John Wiley & Sons

can be broadly defined to include imagination and intuition, formal logic, ways of finding out, and problem-solving strategies. More specifically, students engaged in scientific inquiry must be able to use strategies for solving ill-defined problems¹⁸, to coordinate theory and evidence from both experimental and observational procedures¹⁹, to distinguish patterns of evidence that do and do not support a definitive conclusion, and to isolate parameters and understand the logic of experimental or observational design. These processes and skills are not tied to any specific content domain and thus reflect some degree of generality and transferability.

2.1.4 *A process of conceptual evolution*

This facet emphasizes the progressive modification of the learner's conceptions, when confronted with new observations or experiments, and the associated reasoning. This process may include, as among scientists, progressive changes in pre-conceptions or emergence of new concepts of an abstract nature. As concepts are developed, they become more interconnected and powerful, leading to theory development. Evidence from new experiments or observations may bring about change through the gradual accretion of new facts and knowledge or by the replacement of a previous idea by a different one. For students as well as for scientists, the existing non-scientific ideas can delay the development of new ideas, unless these earlier ideas are recognized. This perspective places more emphasis on processes of conceptual change than on the mastery of general processes and skills discussed in 2.1.3. It recognises science as a human endeavour to make sense of the world around.

2.1.5 *A process of participation in scientific practices*

The view of science-as-practice is emphasized by anthropologists, ethnographers, social psychologists, and those psychologists and educators who study 'situated cognition'²⁰. New scientific content always emerges in a particular cultural and social context, from which it progressively reaches universality. Science-as-practice suggests that observing and questioning, theory development and reasoning are components of a larger ensemble of activity that includes networks of participants and institutions, specialized ways of talking, writing, arguing, modelling, and depicting scientific data and phenomena.

2.2 Indicators and outcomes of IBSE

18 Klahr, D. & Simon, H. A. (1999) Studies of Scientific Discovery: Complementary Approaches and Convergent Findings. *Psychological Bulletin*, 125 (5), 524-543;

Newell, A., & Simon, H. A. (1972). *Human problem solving*. Englewood Cliffs, NJ: Prentice-Hall.

19 Kuhn, D. (1989). Children and adults as intuitive scientists. *Psychological Review*, 96, 674-689;

Kuhn, D., Amsel, E., & O'Loughlin, M. (1988). *The development of scientific reasoning skills*. Orlando, FL: Academic Press.

20 She, H. C. (2004). Fostering "Radical" conceptual change through Dual Situated Learning Model. *Journal of Research in Science Teaching*, 41(2), 142-164

The ultimate measure of the success of IBSE is how well it advances students toward an understanding of the natural world, the development of skills and attitudes of inquiry and an appreciation of the multifaceted view of science just outlined. Section 5 discusses approaches to how such learning outcomes can be assessed. Other measures, listed below, such as relating to teacher training, materials development, classroom implementations, and student behaviours in class, are all input measures to the ultimate output, which is student learning. However, the basis of any valid evaluation of IBSE outcomes must be evidence that IBSE is indeed in operation.

Thus we begin the description of this evaluation process at the beginning, by listing the types of things that may serve as indicators that IBSE is being implemented by teachers with some fidelity to its basic tenets. The indicators are expressed in terms of practices of the teachers and experiences of the students. We then give some examples of outcomes that these IBSE practices and experiences aim to achieve.

Together, but not individually out of context, the items in these lists identify IBSE and distinguish it from programs that are based on texts and not based on inquiry. Although manipulation of materials and equipment ('hands on') may be the most obvious difference between these inquiry-based and non-inquiry-based programs, the key distinction is in the engagement of students, in IBSE programs, in identifying and finding relevant evidence, using critical and logical reasoning in interpreting it and in developing new ideas ('minds on').

2.2.1 *Indicators of IBSE in operation*

These lists bring together, in very broad strokes, activities of teachers and experiences of students that together describe IBSE in operation. As just noted, no single descriptor, taken out of context of other items in the lists, is uniquely associated with inquiry; rather, the combination of descriptors captures what is meant by inquiry-based teaching and learning.

Inquiry practices of the teacher

Teachers will be

- Providing opportunity for students to encounter materials and phenomena to explore or investigate at first hand
- Arranging for discussion in small groups and in the whole class about procedures that are planned or have been used, to identify alternatives and ways in which the approach to particular investigations might be improved
- Encouraging tolerance, mutual respect and objectivity in class discussion
- Providing access to alternative procedures and ideas through discussion, reference to books, resources such as the Internet and other sources of help
- Setting challenging tasks whilst providing support (scaffolding) so that students can experience operating at a more advanced level
- Teaching the techniques needed for advancing skills, including the safe use of equipment, measuring instruments and conventional symbols
- Encouraging students through comment and questioning to check that their ideas are consistent with the evidence available
- Helping students to record their observations and other information in ways that support systematic working and review

- Encouraging critical reflection on how they have learned and how this can be applied in future learning
- Using questioning to encourage the use of inquiry skills.

Inquiry experiences of the students

Students will be

- Gathering evidence by observing real events or using other sources
- Pursuing questions which they have identified as their own even if introduced by the teacher
- Raising further questions which can lead to investigations
- Making predictions based on what they think or find out
- Talking to each other or to the teacher about what they are observing or investigating
- Expressing themselves using appropriate scientific terms and representations with understanding both in writing and talk
- Suggesting ways of testing their own or others' ideas to see if there is evidence to support these ideas
- Taking part in planning investigations with appropriate controls to answer specific questions
- Using measuring instruments and other equipment appropriately and with confidence
- Attempting to solve problems for themselves
- Using a variety of sources of information for facts that they need for their investigation
- Assessing the validity and usefulness of different ideas in relation to evidence
- Considering ideas other than their own
- Reflecting self-critically about the processes and outcomes of their inquiry.

2.2.2 Outcomes of IBSE

Here we propose a preliminary list of cognitive and attitude objectives indicative of what students who participate in ISBE should know and/or be able to do. These will vary in detail from program to program and there is no intention to produce a definitive list or to impose a set of expected outcomes. The purpose is to exemplify the learning that IBSE is designed to achieve. Again, no one item in these lists, out of context, uniquely defines what we are trying to portray.

Students' use of inquiry processes

Students will be able to:

- Identify questions that can be answered by investigation
- Develop descriptions, hypotheses (explanations) and predictions
- Identify and gather information relevant to an investigation
- Think critically and logically in analyzing and interpreting evidence and in drawing conclusions
- Communicate, report and reflect on procedures and explanations, both in talk and in writing in their science notebooks.

Students' understanding of science concepts

Students will have a non-trivial knowledge base, appropriate to the age and stage of education that includes information about key science concepts. A definitive list of these is not attempted, but for the purposes of illustration, they may include (depending on the country and the age of the students) concepts relating to:

- The characteristics, behaviour and needs of human beings and other living organisms
- Diversity and adaptations of organisms
- Population and ecosystems
- Materials and how they can change or be changed
- Motion and forces
- Energy forms, transfer and conservation
- The Earth, the Solar system and the universe
- Geological change

Students' attitudes and dispositions

Students' will show evidence of:

- Curiosity, interest and enjoyment in their science work
- Respect for evidence and honesty in their investigations
- Willingness to change ideas based on evidence
- Willingness to consider alternative ideas
- Sensitivity towards people and other living things and the environment
- Willingness to review self-critically the process and outcomes of their inquiry
- Recognition of the value and limitations of scientific investigation
- Intent to engage in science learning and application in the future
- Working collaboratively with others.

3.0 Purposes and procedures of program evaluation: general considerations

The characteristics of a fully developed IBSE curriculum, including learning goals and indicators of implementation of materials, teaching activities, student activities, and embedded assessments of student learning in science, are spelled out in the preceding sections. In this section we consider various purposes of program evaluation related to stages of implementation. Then in sections 4 and 5 we summarize evaluation approaches appropriate to evaluation for the purposes of improving implementation and identifying the effects of the program.

3.1 Purposes of program evaluation

In relation to purposes, there are three sequential, though overlapping, stages in program evaluation:

- Evaluation for purposes of developing a program.
- Evaluation to establish the extent to which the program, once developed, has been implemented in a number of settings—including effective adaptation to the local context.
- Evaluation of the effects (impact) of the program—short-term and long-term.

In this evaluation project, we assume that a country or jurisdiction has either chosen to adopt/adapt an existing program such as *La main à la pâte* or NSRC's *Science and Technology for Children* (see Appendix C2) or it has developed its own program(s). Thus the starting point is an already identified program, not the development of a program, and so we do not deal with the first of the above three purposes. However, as stated in section 7.2, the Working Group does recommend to the IAP that help be provided, to countries that request it, in the design of new IBSE programs. Our focus is the second and third purposes, the evaluation of program implementation and evaluation of program effects and impact. Nevertheless, several of the methods for gathering and using relevant information are quite similar across all three purposes for program evaluation.

3.2 Stages of program implementation and evaluation

In order to ascertain a program's effects and impact, there must be reasonable evidence that the program is being implemented appropriately, including necessary adaptation to local contexts that nevertheless preserves the program's essential features and leads to local ownership. Though program implementation is not a smoothly proceeding process from the first introduction to reasonably full implementation, we posit several stages in this process.

Because it is important to adjust evaluation activities to address the purpose of the evaluation we are offering an evaluation framework that identifies a progression of evaluation goals.

- In the introductory stage of a program's implementation, interest is in using the evaluation results for improving program implementation, so many of the data gathering strategies will be near the classroom or classroom-based.
- When the program is more fully realized in all the classrooms in which it has been introduced interest is in the effects of the program, and the means of data gathering are likely to be more distant from actual classroom activities and may include standardized tests or assessments.
- In the intermediate stages of the implementation continuum, it may be appropriate to look for intermediate outcomes, such as changes in teaching practices.

3.2.1 Evaluation at early stages of implementation

In the early stages, the purpose of evaluation is to establish to what extent the selected IBSE program is being implemented as intended (including necessary adaptation). At these stages, evaluation is *formative* in relation to implementation (see section 1.3.1). Evaluation questions should address the presentation and delivery of science education in the classroom, that is, to what extent the learning and teaching implemented in the classroom matches the intent of the IBSE program being introduced. Questions might include the following: Are the activities that the teacher and students are engaged in consistent with those intended in the program? Is the content that engages the teacher and students consistent with the intended content? Do the teacher-student interactions and the student-student interactions match those intended? Is the nature of the classroom discourse consistent with that envisaged in the program?

3.2.2 Evaluation at the stage of more advanced implementation

It would be inappropriate to use evaluation results obtained in the early stages of implementation to draw conclusions about a program's effects and impact. Only when the IBSE program is more fully realized is it appropriate to conduct *summative* evaluation (see section 1.3.1), that is, to ask questions about overall effects; for example, evaluation of student performance in science across all the classrooms using the IBSE program and—if desired—impact compared to alternative or 'traditional' programs. As implementation of the program moves closer to what is intended, the purposes of evaluation change, as do evaluation methods and specific measures, ranging from proximate (close to the classroom) measures in the beginning stages of implementation to more distal (externally designed) measures in the later stages. At the point of a well implemented program, comparative measures or measures over time become of greater concern.

3.2.3 Types of information required for evaluation

While the evaluation questions, the design, methods, and instrumentation should be responsive to the purposes of the evaluation, in general, information about the system supporting the program, about teaching practices and about student outcomes needs to be collected and analyzed. This will comprise:

- Support provided at the system level for sustaining IBSE
- How teachers use the IBSE materials
- Inquiry practices of the teacher
- Inquiry experiences of the students
- Students' use of inquiry processes
- Students' understanding of science knowledge, as shown by application in novel situations
- Students' use of scientific vocabulary and representations in reporting and discussing their inquiries
- Students' attitudes or dispositions towards science and learning science.

However, the focus of data collection will shift as evaluation goals progress from interest in how a program is being implemented to the overall effects or outcomes of a program for students and teachers.

3.3. Activities of the evaluator and of the teacher

The activities required of the evaluator and teacher in relation to the evaluation will vary according to the stage of implementation of the program. At the stage of well advanced program implementation, when the purpose of the evaluation is to ascertain the program's effects, data gathering will include many—if not all—classrooms and teachers implementing the program. Information gathering activities and the interpretation and use of evaluation results move from close proximity to the classroom to large-scale evaluation across many classrooms conducted for summative purposes, with a corresponding change in the activities of teacher and evaluator.

3.3.1 Evaluation activities of the Teacher

The activities of the teacher in respect of the evaluation will take different forms depending on the program and the kind of evidence that needs to be gathered. In the early implementation stages, the teacher's activities in program evaluation include:

- Informal and formal recording of classroom activities.
- Using the information to improve classroom enactment of the IBSE program.
- Assessing students' ideas, skills and attitudes through continuous observation and discussion as part of teaching.
- Using the information about students' learning to help them make progress.
- Ensuring students are aware of their learning goals.
- Helping the external evaluator understand the realities of the classroom.
- Responding to oral and written questions related to what is being implemented and why.
- Active participation in discussions with the evaluator.

Note that most of these activities are part of good teaching in general and are indispensable when a new program is being introduced.

When the IBSE program implementation appears to be more advanced in a number of classrooms, the design of the evaluation, instrumentation used, and analysis and reporting of results become largely the responsibility of an external evaluator, although committees of teachers may well be called upon to advise in this work. In addition to the above activities, classroom teachers may be asked to respond to teacher surveys and questionnaires and perhaps to administer externally designed assessment of students' learning in science and students' attitudes. In some cases the administration of student questionnaires and tests may be carried out by a person from outside the school, to enhance credibility. Teachers should also have timely access to the results of the summative evaluations and recommendations for changes in curriculum and teaching practice drawn from the evaluation findings.

3.3.2 Activities of the Evaluator

In the early implementation stages, the evaluator works with the teachers implementing the IBSE program to develop evaluation procedures for gathering information that the teacher can use to improve the implementation. The evaluator also looks for larger findings across classrooms. Activities carried out jointly with the teachers include:

- Formulating evaluation questions.
- Selecting or developing appropriate instruments for observation and criteria for assessing student progress.
- Selecting or developing written or oral questions to probe teachers' understanding of inquiry and the goals of IBSE.
- Keeping logs or journals of teaching activity.
- Conducting continuous observation of classroom interactions and discourse.
- Analyzing and interpreting the information gathered.
- Making recommendations for improvements in the implementation of the IBSE program based on the evaluation results.
- Holding seminars to stimulate further thinking and understanding of IBSE and its implementation.

The evaluator will also wish to make observations and conduct teacher and student interviews independently. These points are taken further in section 4.

As with evaluation approaches and instruments, the teacher and evaluator activities represent a continuum as the purpose of the evaluation changes from gauging the extent of program implementation to assessing effects of the program on students and teachers. As the purpose of the evaluation shifts to informing on effects of the IBSE program and student outcomes, either in terms of its own goals or in comparison to other programs, the role of the external evaluator assumes much greater importance. In conjunction with the designers of the IBSE program and the teachers who are implementing it, the evaluator:

- Formulates evaluation questions based on the information needs and expectations voiced by policy makers, educators, and other interested parties
- Designs the evaluation, including selecting the methods to be used
- Develops the sampling design, if necessary

- Develops and/or adapts existing instrumentation, including teacher, administrator, and student questionnaires and student tests and scoring guidelines
- Analyzes data
- Reports results in reader-friendly reports
- Collaborates with teachers and program leaders to formulate recommendations for improvement of the program and its implementation.

These points are taken further in section 5.

4.0 Evaluation for the improvement of program implementation

In this section we discuss in outline procedures involved in evaluation at the early stage of implementation of an IBSE program. The purpose of the evaluation is to provide feedback into the program and its implementation. The concern is, therefore, with the extent to which intended classroom processes and experiences are enacted. It is the purpose of this project to offer and provide through collaboration, assistance as required with all activities and procedures outlined here.

4.1 Outline of procedures

In the early stages of program implementation, evaluation should be carried out as close to the classroom as feasible. This also makes possible the timely use of results to improve implementation. The evaluation proceeds through these steps:

- Identification of questions for the evaluation to address
- Identification of data to answer the evaluation questions
- Selection of data gathering methods and instruments, according to resources available
- Collection of data
- Analysis of data
- Feedback at class, school, and program levels as appropriate

These are not distinct, but interconnected, steps. As will be seen, decisions about one step affects how other steps can be accomplished. For example, there is no point in collecting so much data that it can only be handled by computer, if the resources for computer analysis are not available; or extensive video-taping when the time for reviewing is limited. So the processes of thinking through these steps, and making decisions about how to put them into practice is iterative in any evaluation. We now consider how they can be put into practice for formative evaluation of implementation, noting that this project is offering technical expertise as indicated in section 6.1.2.

4.1.1 Identification of evaluation questions

Preliminary collection of basic data about the program, its goals, materials, teacher training, the implementation timescale, etc., is a prerequisite for specific evaluation questions to be identified. It may well be that what is implemented is a selection of aspects of IBSE rather than a 'full blown' program. The evaluation should look not only at how well what is intended is implemented but also at how adequate this is as an IBSE program as envisioned in this report and in the in-country definition of IBSE.

The purpose, then, is to find out what is being implemented, how well and how it can be improved. The questions are likely to be such as:

- What features of IBSE are and are not reflected in the program being implemented?

- What features of IBSE are and are not being practised as intended?
- How well do teachers, school principals, inspectors and others involved in the implementation, understand and believe in the aims of IBSE?
- Do teachers, principals and others believe that these aims can be achieved within IBSE programs?
- To what degree do teachers display the skills and knowledge required to implement an IBSE program?
- To what extent do students' experiences of science reflect the aims of IBSE?
- Can the inquiry-based activities be accessed by students of all abilities and backgrounds?

To help understanding of why certain aspects of IBSE are or are not being implemented in the classroom the evaluation may also address questions such as:

- What professional development is available for teachers, principals and others to learn about the aims and practices of IBSE?
- What training in IBSE have teachers actually experienced?
- How useful did teachers find this training?
- Does the teacher receive on-going help with IBSE implementation from within or outside the school?
- Are all the necessary materials available for the implementation of the IBSE program?
- Do allocations of budget and teacher planning time support IBSE implementation?

Questions should be agreed in discussion with all involved, particularly the program co-ordinator, the teachers, science educators and scientists, so that there are realistic expectations of what can be achieved through the evaluation.

4.1.2 Identification of data to address the questions

It is necessary to be clear about what data are required before data collection methods and instruments can be chosen. There will always be a wide range of data that potentially could be collected relating to each question; the lists in section 2.2 indicate the kind of data to be considered. Making the selection in the case of a particular evaluation involves refining the evaluation question and clarifying what features of IBSE are intended to be implemented. It will also depend on the timescale of the evaluation. For example, for some questions, such as relating to the effectiveness of an on-going teacher development program, evidence may be needed of change over time in classroom practice. Teachers' and/or students' views and practices of IBSE at time 2 would be compared with time 1.

4.1.3 Selection of data gathering methods and instruments

Optimally, the data collection methods will include:

- Observation of teacher and student activities in the classroom, particularly the characteristics of their inquiry activities (Appendix E1)
- Noting the science content of lessons
- Examining teacher and student resources
- Logging classroom discourse
- Examining teacher journals

- Questionnaires for teachers, students and administrators (Appendix E2)
- Noting teacher assessment practices
- Examining student writing/drawing, including student notebooks and laboratory records
- Interviewing teachers and selected students (Appendix E3)
- Noting continuous assessment of student progress through teacher observation and records.

The specific instruments to be used in a particular evaluation have to be tailored to the questions and the context, but it is not necessary to start from scratch. Examples exist in previous work and a key purpose of this international project is to ensure that each evaluation can benefit from evaluations conducted in other places. Appendix E provides some examples of data gathering methods that have been used and can be adapted. (Note that this is just the beginning of a collection of useful examples that the project will make available to participating countries via a website or other means).

4.1.4 Collection of data

Design of the evaluation and details of data collection procedures depend on the local circumstances, such as the number of schools involved, their location and the ease of communication. Clearly the human and other resources available also frame what is possible in a particular case.

If only a small number of schools is involved, then regular visits by an evaluator may be possible. During visits interviews, observations and perusal of documents and student work can be carried out. If resources allow, some activities can be video-taped, either by a local evaluator or by the teacher. On the other hand, if schools are spread across a wide geographical area then questionnaires are likely to be the main feasible method for data collection, combined with a few visits to a sample of schools. When tele-communications are good, interviews can be conducted by telephone. Video-tapes of lessons, where possible, can be sent to the evaluators or reviewed by teachers themselves. To obtain some idea of students' work, teachers can be asked to set specific tasks and to send examples of what they consider to be good, medium and poor responses to the evaluation team or program co-ordinator.

Where resources are tightly limited, priorities have to be carefully considered. There is a strong case, in formative evaluation, for prioritising visits to as many schools as can be afforded to conduct classroom observation, interview teachers and students and collect examples of teachers' journals and students' notebooks. Before or between visits, teachers can be asked to make a record of specific aspects of science lessons under a given set of loose headings.

In the course of studying implementation, a major question concerns the effects of the implementation on the understandings, skills and pedagogical beliefs of teachers. For formative assessment, evidence on these is likely to be best gathered by interview. Some guidelines on developing interview schedules are given in Appendix E3.

4.1.5 Data analysis and interpretation

Both qualitative and quantitative data should be analyzed in a scientifically justifiable manner, such that the analysis reflects the structure of the data collected and directly addresses the questions driving the formative evaluation. Instruments used to collect data should be first examined for their reliability and validity. Analysis of the IBSE implementation outcomes should take into account student, classroom (teacher), schools and community factors (hierarchical organisation of schooling).

The results of the data analyses should be displayed clearly in tables or graphs with discursive summaries. The intent is to display areas of strengths and to pinpoint where improvements need to be made. Information about where improvements are needed serve as basis for discussion among teachers, program leaders and local evaluators about the way to improve the implementation of IBSE.

4.1.6 Feedback at school and program levels

This is the most important part of formative evaluation, for if the information is not used to improve program implementation, then the evaluation is not serving a formative purpose. An important role of the local evaluation team (see 6.1.1) is to ensure that feedback is provided. Feedback should be based on evidence and not be judgemental, but indicate where improvements can be made for the purpose of discussing how to take action. Feedback should be at a number of levels to optimise the value of the evaluation:

- to individual teachers, about what they are doing/not doing in relation to what was intended and what this suggests for changes to aim for
- to schools, about the extent of implementation in and any variation across classes and about what this indicates about the support the school should provide
- to program co-ordinators, about what is being implemented and what support, such as further resources and training appears to be needed.

At all these levels, information about what others have done in similar circumstances can be of great help. Thus case studies of classes, schools and countries at early stages of implementation need to be prepared as part of the evaluation. Where facilities and resources permit, these might take the form of, or include, video-taped material on CDROM.

5. Evaluation of the effects of IBSE on students and teachers

This section considers evaluation relating to questions about whether an IBSE program leads to the outcomes for students and teachers that are expected. It is appropriate to try to answer such questions when an IBSE is well established in a number of a country's schools. We deal first with the steps in conceiving, designing and reporting an evaluation of effects and then with the development of assessment tests and tasks for assessing IBSE outcomes. Although not the only evidence of a program's effectiveness, the assessment of student outcomes is necessarily a prominent part of the data to be collected in a summative evaluation.

5.1 Procedures for developing an evaluation of program effects

When implementation of an IBSE program is well established across a number of schools it is appropriate to ask questions about its effects on students through a summative evaluation. Until well established, as noted earlier, outcomes are not likely to be a valid indication of the effects of IBSE, and so the purpose of evaluation at that stage is to improve implementation. But reaching a stable implementation does not mean uniformity and it is still necessary to gather information about classroom processes, so that different learning outcomes can be related to different learning experiences. Hence, both qualitative and quantitative methods need to be used to establish teaching practice and student outcomes.²¹ Again we note that technical assistance will be available to the local evaluation team through the support proposed in this project (see section 6.1.2).

The steps in a summative evaluation are the same as in formative evaluation (see 4.1) but the questions and consequently the methods will be different. We now look at how the steps are operationalized in a summative evaluation.

5.1.1 Summative evaluation questions

Summative evaluation of an IBSE program can be designed to address different questions. One is to determine whether, and to what extent, the program results in the desired outcomes as specified in the program. Another question asks whether student outcomes are the same or different (better, worse) when the IBSE program is compared to other science education programs.

The motivation for asking these questions is often to justify the decision to introduce an IBSE program. It may be a school principal wanting to justify school policy to parents, governors and local politicians; or those responsible for education policy at national or regional level wishing to convince the electorate to support a change in science education; or the program developer wanting evidence to persuade more teachers and

²¹ For a discussion of selected evaluation techniques, see the *2002 User-Friendly Handbook for Project Evaluation*, NSF 2002.

school principals that the effort required for implementation is worthwhile. The evaluation questions should be negotiated by the evaluators with those who need answers, so that there is agreement about the evidence that the evaluation can and cannot provide and the possible conclusions that could be drawn from it.

5.1.2 Identification of data

The principal focus of attention in summative evaluation is the extent to which students have achieved the learning identified in sections 2.1 and 2.2.2. The particular items selected will depend on the questions being raised in the evaluation. When the evaluation focus involves overt comparison of an IBSE program with another science education program, or with ‘traditional’ practice prior to the introduction of IBSE, several questions arise about the kinds of student outcomes to be assessed; for example, whether the evaluation will address only the common core of desired outcomes, or capture additional outcomes specific to either the IBSE program or an alternative program. We suggest that, in such cases, all the goals of each program are sampled in the assessment.

Even when the question is couched in terms of whether intended goals are achieved, the outcomes are judged in terms of some expectations, based on judgements or on existing or anticipated levels of performance. Thus some comparison is involved in order to interpret the data and so the same questions arise about the range of outcomes included.

In an international evaluation project, there are considerable benefits from including some common measures (marker variables) of students’ outcomes in the evaluation in different countries. The purpose of this strategy would not be to compare countries but the effects of IBSE in different circumstances. The various country studies might be thought of as replicates of one another under vastly different conditions. This would yield valuable information to help understanding of conditions under which implementations work better or do not work so well.

In addition to the measures of student outcomes, the data collected should include process data about classroom experiences and interactions. This need not be wide-ranging, as in formative evaluation, but focus on the ‘key variables’ (indicated in the questions in section 4.1.1).

Teachers are key players in the implementation of IBSE programs. Thus in the course of studying implementation of an IBSE program, it is important to find out what changes in teachers’ practices are occurring as a result of changes in teachers’ understanding of IBSE and in their pedagogical beliefs. Appendix E2 provides some examples of evaluation methods used to assess such teacher outcomes.

5.1.3 Selection of data gathering methods

While the collection of evidence relating to key variables may involve some of the data collection methods listed in 4.1.3, the following are major additions to the list:

- Surveys of teachers and administrators on teaching policies and practices.

- Questionnaires to teachers on understanding of IBSE and what it requires in practice.
- Questionnaires to students on classroom activities.
- Tests and tasks for assessing students' understanding and use of scientific principles and inquiry skills.
- Surveys of students' attitudes.

There are many possibilities for assessing student cognitive and attitude outcomes:

- Standard national tests, already in place.
- Written tests in conventional form (e.g. multiple choice).
- Tests or tasks embedded in classroom materials.
- Practical tests or tasks.
- Extended performance tests or tasks.
- Concept mapping.
- Attitude questionnaires. (Appendix E2)
- Collection of work portfolios.
- Computer-based tests.

The use of existing tests, such as are in operation in many countries at the end of primary school, is clearly the simplest option. It may also be the most convincing where the IBSE program is being compared with a traditional one. It shows the extent to which learning through inquiry transfers to learning in science as conventionally conceived. However, this would not be adequate to answer questions about whether all the goals of learning through IBSE, as identified in section 2, are being achieved. For this, methods are needed that provide opportunities for problem solving, use of evidence, interpretation of data, rational thinking and application of knowledge. While it can be argued that it is reasonable to see if students taught through non-IBSE programs also achieve the same outcomes as those in IBSE programs, there can be ethical objections to facing students with tests of skills and application which they have not been taught.

5.1.4 Collection of data: design

If IBSE programs are to be compared to other programs, how will comparison groups be created? The classical way to address this problem is to assign the potential pool of students (or classrooms or schools, depending on the particular situation) randomly to either the IBSE program or the alternative program(s). Although this may be difficult to carry out in some settings, it is methodologically the most defensible approach, and creative ways have been found to carry out randomized experiments of this type.

An alternative is to select comparison groups from existing classrooms in which IBSE and “traditional” programs are already being taught. However, this introduces a number of difficulties, all related to failing to take into account differences other than curricular ones between the comparison groups. For example, teachers may volunteer their classrooms, which is likely to introduce bias into the evaluation, as volunteers are likely to be different from non-volunteers in their implementation of any program. Also, matching student and teacher variables between the IBSE classrooms and non-IBSE

classrooms is likely to be extremely difficult, when student characteristics (level of previous science performance, family income and status) and teacher characteristics (level of science specialization, amount and kind of professional development, years of experience with program) are likely to be relevant. Complex equating procedures must be used to account for such differences.

How should samples be selected? If the comparison program is widely taught, then it is likely that the evaluation will have to be limited to a sample of such classrooms due to cost and time constraints. The smaller the sample, the greater the errors introduced by using a sample rather than all the classrooms. Non-response or incomplete information provided by some of the sampled classrooms introduces further problems, since these classrooms are likely to be different from the classrooms that respond in full. Errors also are introduced if respondents misunderstand the instruments.

Whatever the methods used, there is likely to be resistance from those using the alternative program(s) who may see little need for the comparative evaluation and may find it burdensome and intrusive.

Any attempt to develop valid conclusions about the comparative efficacies of alternative science education programs, say IBSE versus a “traditional” program, must resolve these difficulties in the evaluation design, methods, and analysis of data. The resulting reports must be clear about how these issues were addressed and with what success.

5.1.5 Data analysis

Both qualitative and quantitative data should be analyzed in a scientifically justifiable manner such that the analyses reflect the structure of the data collected and directly address the questions driving the evaluation. Use of data should be preceded by reliability and validity analysis of the measurements. Estimation of the effects of IBSE on student outcomes should take into account the hierarchical organisation of schooling, where students are found in classrooms, classrooms are found in schools and schools are found in communities.

When data are collected longitudinally, the analysis should take into account the effects of the time difference between repeated measures (that is, some differences may have occurred just because of the passage of time). In estimating IBSE effects, local contexts and pre-existing differences among students, teachers and schools should be taken into account. Finally, where feasible, multiple data sources should be used to triangulate on interpretation of IBSE effects.

Analyses of these kinds also yield information to research questions surrounding IBSE (see Appendix D). It is often beneficial and economical to both to combine the evaluation with appropriate research studies.

5.1.6 Communication and reporting of evaluation findings

There is widespread interest in the outcomes of any educational innovation. Parents, teachers and the general public are thus likely to be as anxious to know about the effects

of IBSE as are those initiating and funding the evaluation. Reporting to these groups is an important role for the local evaluation team. There will be several audiences for the report of findings and while there are existing channels for reporting to the science education and research communities, appropriate means have to be sought for communication to other audiences. In terms of written reports, it is suggested that three types of report are necessary: the full account of methods, analyses and results; a summary of about 20 pages; and a one or two page bulletin of the main findings and implications. The last two types of documents may need to be written in different forms for different target readership, e.g. for teachers, for the general public and for local and national policy-makers. The help of the media should also be enlisted for wider dissemination.

However, the proposed international project is designed to magnify the value of any one evaluation project by placing it in the context of other evaluations of IBSE. In addition to its unique value to the country in which it was conducted, evaluation findings will contribute to a wider understanding of the effects of IBSE in different circumstances. The development of synoptic reports, bringing together findings from different countries will extend the vision of each participating country as science education continues to be developed in response to changing circumstances and understanding of teaching and learning.

5.2 Developing tests and tasks to measure student outcomes of IBSE

In section 5.1.2 we noted that, in designing a summative evaluation, questions have to be answered about what measures of student outcomes are to be used. Good cases can be made both for using existing tests and for using tests designed to assess the particular outcomes of IBSE. This section provides guidelines for developing measures relating to IBSE outcomes, whether used alone or in combination with existing tests that may already be in use.

There are several stages in the development of tests and tasks for assessing student outcomes. Each test item or task (referred to as 'items' for brevity) should be clearly related to specific outcomes and hence created with the outcome(s) as the starting point. The aim in item development is to present the students with a situation that requires them to use the particular process or concept that is indicated in the IBSE outcomes (such as those listed in section 2.2.2). It is likely that there are various types of item that validly meet this requirement, so identifying them is the first stage in the development process. The various stages for each outcome are as follows:

1. Identifying the types of item that will require the students to use the process or apply the conceptual understanding described in the outcome.
2. Selecting content, or subject matter, appropriate to the outcome and to the experience of the students.
3. Drafting items and marking schemes (protocols), followed by review and revision.
4. Trial of revised items and marking schemes with a small sample of students in the target group, followed by revision.
5. Pilot trial of items and marking schemes.

6. Further review of items and marking schemes.
7. Repeat of steps 5 and 6 as necessary (and feasible).

5.2.1 Identifying item types

An item type is described in terms of what is presented to the student and what the student is required to do in response. For example:

Process: Identify information relevant to conducting an investigation

What is presented in the item: A proposition that can be investigated

What is required of the student: To say what would be necessary to conduct the investigation

Or: To select from given alternatives the best way of conducting the investigation and explain the reasons for the choice.

For the understanding of concepts, an example is

Concept area: Energy transformations.

What is presented in the item: Description of a situation involving the concept.

What is required of the student: To make a prediction about what will happen or an explanation of the events described.

Or: To select from given predictions or explanations and give reasons for the choice.

There are several other possibilities for item types for the same processes and concepts. Since these should have a common focus on the process and concepts that are goals of IBSE they should be widely applicable across countries. Thus it should be possible to provide evaluators in different countries with a list of types which they can adapt, if necessary, but do not need to develop *ab initio*.

5.2.2 Selecting content

The content or subject matter of an item needs to be chosen so that it is familiar from everyday life and yet not likely to have been the context for teaching (in which case the item may be answered by memory). For example, in the case of the process cited above, the proposition might be 'larger balls take longer to fall to the ground than smaller balls'. Students are likely to be familiar with balls of different sizes but unlikely to have undertaken this particular investigation. In the case of the concepts relevant to energy transformation, an item might show three pictures of a man pulling a heavy roller up a sloping plank, with the plank at three different angles. Students are asked to say which involves the man in using least energy or whether it is the same in all cases, and to give a reason for their choice.

5.2.3 Drafting items and marking schemes

Drafting the items requires decisions to be made about the form in which they are to be presented. Although computer-based tests are being developed for certain purposes, in many countries they are not yet a viable option for testing a number of students at the same time. Thus we consider here the development of items of more conventional

kinds, where students read instructions and give responses on paper, even if they also undertake some hands-on task.

It is important to minimise the amount to be read, since the reading skills needed may be a hurdle to achievement. On the other hand, in order to present a real context for the item some description is needed. Drawings are helpful for written items, but can sometimes give unintended cues. Decisions also have to be made as to the format: multiple-choice, short-answer or extended answer. Multiple-choice questions often have a considerable reading demand, whilst other forms make a demand on writing and, of course, on others being able to read the students' writing. Field trialling with a small number of students is essential to reduce problems of communicating the item content. However, before that stage is reached, the items should be subject to careful review and critique by a groups of teachers of the age range of students for which the test is intended.

The preparation of marking schemes (scoring guidelines) is a key part of drafting items. In the sorts of items needed to assess IBSE outcomes, there may often be more than one acceptable response and possibly also degrees of 'correctness'. Whether and if so, how, marks are to be allocated to partially correct or incomplete answers is a matter of judgement. In a constructed-response item it is important to signal to the student what kind of answer is required. So, for instance if marks are going to be awarded for the description of an investigation, the question should indicate this by including instructions such as 'make sure you say what things you would use, what you would do with them and how you would decide the answer'.

5.2.4 Trial with a small sample of students

This stage is a trial with a few students in a few classes, not a full-scale pilot of the whole test. The small scale means that it is possible to observe the students answering the questions and then, immediately after they have finished, to talk through their thinking, how they understood what was required and how they decided their answers. This discussion will reveal whether the students' responses were the outcome of the intended thinking or whether they used recall, guesswork or some unintended cues in the way the item was presented. It may be possible to make changes to ensure that an item assesses what is intended or it may have to be abandoned.

5.2.5 Pilot trial of items and marking schemes

A larger scale pilot trial of items put together as tests is necessary to investigate feasibility across a number of classes as well as details of test administration. Are the items suited to the age range of the students? Does the test take the amount of time intended? What problems do student encounter in understanding what to do (particularly if they are used to a different form of test)? What instruction should be given to students? Can the items be reliably marked? Such questions are the ones that need to be answered before the test can be used in a summative evaluation across a number of schools.

5.2.6 Revision of items and marking schemes

If answers to the questions above are less than satisfactory, then further changes may need to be made. If required changes are extensive, a repeat trial ought to be carried out. At all times when changes are made it is essential to check that the test items remain a valid reflection of the outcomes to be assessed. It is all too easy for concerns about reliable marking and other technical aspects to narrow the range of the test and so impoverish the information that can be obtained about what students are learning through IBSE. The final judgement of the relevance of the test items is one to be made on educational grounds, by informed reviewers.

6.0 Mechanism for the organisation and implementation of the IBSE International Assistance Evaluation process

On the basis of the previous sections of this Report, the following mechanism for international assistance in the evaluation of IBSE programs and dissemination of information is proposed. The mechanism will be piloted as a matter of urgency. This will not only help to clarify details of the mechanism but will inform the development of a handbook to assist local evaluation teams.

6.1 Proposed mechanism for implementation of the local evaluation

6.1.1 *Committees/Teams Integral to the Process*

The IAP proposes to set up a committee and teams as follows and as summarised in the organizational chart below.

International Oversight Committee: The IAP and a subset of the Working Group will identify 4-5 persons who are willing to serve on the International Oversight Committee (IOC). A Chair of the IOC will be elected, and membership will be rotated in 2 to 3 year cycles. The Committee will operate *ad honorem*.

The IOC will be the body responsible for the overall running of the international assistance evaluation and play a central role in quality assurance. Its meeting and travel expenses will initially be borne by the IAP and it will report annually to the IAP Science Education Program on the evaluation work.

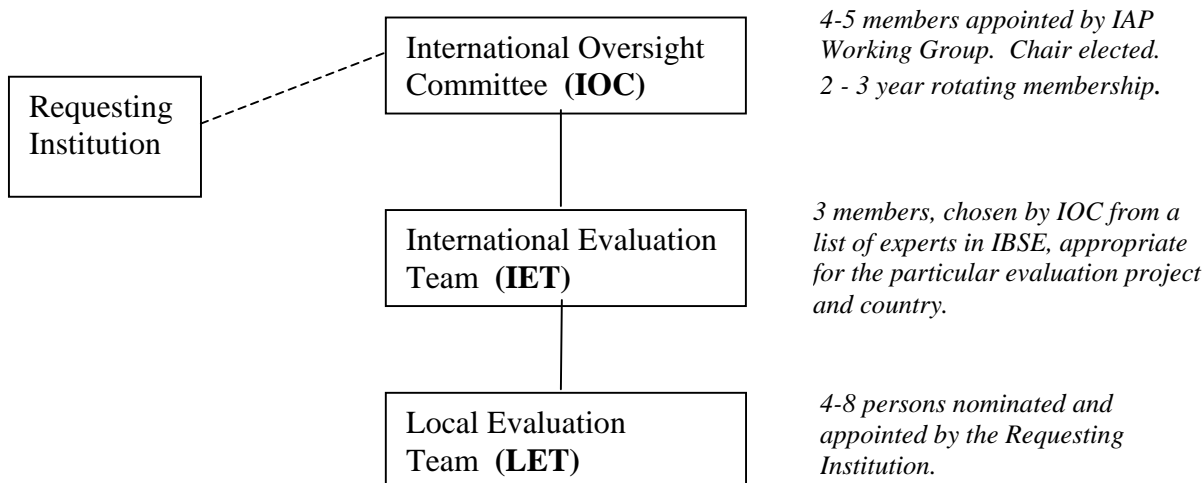
International Evaluation Team(s): The IOC will draw up a list of international experts in IBSE evaluation and research who may be willing to serve on an *ad hoc* International Evaluation Team (IET).

There will be a new IET for each evaluation. Each IET consists of 3 persons appropriate for the country and type of evaluation project.

Local Evaluation Team: The Institution (e.g. National Academy or Ministry of Education) requesting international assistance with evaluation of its IBSE Program will nominate a Local Evaluation Team (LET) that is external to the Institution.

The LET should consist of 4-8 persons with expertise in evaluation and implementation of science education in elementary schools.

Organizational Chart



6.1.2. Process for Development, Review, and Validation of Evaluation Projects

The Institution or other body responsible for the local evaluation project (e.g. National Academy or Ministry of Education) will apply to the IOC for review. Information submitted for the review will be guided by an Application Form provided by the IOC and will include:

- An overview of education in the locality to be evaluated and history of IBSE implementation.
- Details of the project or program being implemented, including grade/age level, and the objectives of the evaluation.
- The proposed design of the evaluation activity (see Sections 3, 4 and 5 and appendix E of this report).
- Its nominated LET (see 6.1.1) and details of how this evaluation team was selected.
- The intended use(s) of evaluation results.

Prior to submission of the application, the institution or body requesting the evaluation (Requesting Institution) should seek advice on the above from the Chairman of the IOC. At this stage the Chairman will also provide information on the criteria to be used in the evaluation process. This information will vary depending on what the Requesting Institution seeks to determine and accomplish through the evaluation process.

Once the application has been reviewed and accepted by the IOC, the IOC will select the IET appropriate for that evaluation project.

The IET will make an initial visit to the country of the Requesting Institution to discuss the development of the evaluation methodologies for the project with the LET. It is envisaged this work will require one week and will result in a report addressed to the Institution by the IET outlining the agreed evaluation methodologies.

The LET will then carry out the evaluation project according to the agreed methodologies. One or more members of the IET may be available to provide advice (either electronically, by phone, or in person) at strategic times during this process.

The IET will make a second visit of one week's duration to review and discuss the results of the evaluation project with the LET. The results and conclusions drawn by the LET will undergo rigorous scrutiny by the IET before international validation is provided by the IOC.

The validated evaluation report will be provided to the Requesting Institution, via the IOC, which will retain both hard and electronic copies.

The Requesting Institution has ownership of the report and thus retains the right to publish separately.

The IOC retains the rights to publish a synthesis of this report in combination with similar ones from other countries. The report will not be used to rank or compare the quality of IBSE programs from different countries or compare individual teachers or students.

6.1.3 Financial Commitments

The IAP or another body that undertakes this task will be responsible for costs associated with the work of the IOC. This will include the salary of a staff member who spends a substantial part of his or her time organizing the activities described here that are overseen by the IOC

The Requesting Institution will be responsible for all costs associated with the local evaluation project, including expenses of the LET.

The Requesting Institution will be mainly responsible for finding the financial resources to cover the costs associated with the two site visits by the IET which would include economy air fares for the IET members and their expenses during the visit (hotel, meals, transportation, etc). Members of the IET may receive an honorarium as agreed by the persons and Institutions involved.

If the Requesting Institution requires one or more members of the IET to make additional trip(s) during the evaluation process to assist the LET, the additional costs must be covered by the Requesting Institution.

6.2 Dissemination of the IAP Working Group Report to regional and international meetings sponsored by Academies

The content of the report will first be released at a meeting hosted by IAP Working Group in Santiago, Chile, September 25-26, 2006. It is expected that representatives of

about 50 of the 93 affiliated National Academies will attend. Particularly targeted will be those countries likely to be interested in accessing the IAP Working Group-generated evaluation process of their IBSE Program. Electronic versions of the report will be available to delegates prior to this meeting.

Information on the report will also be disseminated at the IAP General Assembly meeting in Alexandria, Egypt in December 2006 and at proposed IAP Regional meetings for the Americas, Asia and Oceania after the Santiago meeting.

7.0 Recommendations

7.1 Recommendations in relation to this report

This report has put forward arguments for changes in pre-secondary science education that enable students to learn through inquiry-based activities. The value of making this possible has been expressed in terms of development of scientific literacy and the understanding, skills and attitudes needed by citizens in the rapidly changing world where learning must continue throughout life. The report has spelled out what IBSE means in practice for students, teachers and schools, indicating the considerable change that implementation is likely to require. It is essential that the effort and investment required are justified by evidence of the improvement in student outcomes, teachers' understanding and the quality of education experiences in science that are claimed. The international project proposed here is designed to provide assistance with this endeavour. The Working Group therefore recommends that:

- Countries implementing IBSE programs evaluate the extent and effect of implementation and set up local evaluation teams for this purpose.
- An International Oversight Committee be set up to manage the provision of support for implementation evaluation to countries who request such support.
- International Evaluation Teams work with local evaluation teams to support the evaluation and enhance local capacity to conduct it.
- The reports of procedures and findings of evaluations conducted in various countries are made available to others and are synthesised to add to general understanding of IBSE and its implementation in various contexts.

It is further recommended that the IOC should report annually to the IAP on the International Assistance in Evaluation project and, through the IAP, should:

- Develop a handbook to guide evaluation design and procedures and a compendium of pertinent research literature on IBSE and its evaluation.
- Hold an international workshop to discuss the outcomes of the synthesis of reports from different countries at a time when sufficient data have been accrued.

7.2 Recommendations for additional work by the IAP

The international project which is the subject of this report has been designed specifically to focus on the evaluation of the implementation of already developed IBSE programs. However, it is recognised that many countries would welcome help at earlier stages, in the development of programs. Further, there are specific aspects of programs and of research into their effects where international collaboration and sharing of expertise are of particular value. Thus the Working Group has identified the following work outside this evaluation project that it recommends the IAP to consider in relation to IBSE:

- A project to provide help for countries in the design and development or adaptation of IBSE programs.
- The creation of guidelines for countries aimed at enabling them to generate and incorporate into their IBSE programs, materials that provide help for teachers in assessing their students and evaluating their own effectiveness in implementing IBSE.
- Establish a program of post graduate fellowships that will support capacity building in countries where there may be a lack of evaluation and research expertise.
- The promotion of research projects aimed at increasing knowledge and understanding of the nature of inquiry learning in science and how it is best implemented in various contexts.

Brief biographical sketches of IAP Working Group members and observers

Dr. Bruce Alberts, a respected biochemist with a strong commitment to the improvement of science and mathematics education, has returned to the Department of Biochemistry and Biophysics at the University of California, San Francisco, after serving two six-year terms as the president of the National Academy of Sciences (NAS). During his tenure at the NAS, Alberts was instrumental in developing the landmark National Science Education standards that have been implemented in school systems nationwide. Dr. Alberts is also noted as one of the original authors of *The Molecular Biology of the Cell*, a preeminent textbook in the field now in its fourth edition. For the period 2000 to 2009, he serves as the co-chair of the InterAcademy Council, a new organization in Amsterdam governed by the presidents of 15 national academies of sciences and established to provide scientific advice to the world. Widely recognized for his work in the fields of biochemistry and molecular biology, Dr. Alberts has earned many honors and awards, including 14 honorary degrees and an investiture as Commander of the Order of the British Empire. He currently serves on the advisory boards of more than 10 non-profit institutions. He is an Overseer at Harvard University, a Trustee of the Carnegie Corporation of New York, a Trustee of the Gordon and Betty Moore Foundation, and the president-elect of the American Society of Cell Biology.

Dr. Jorge Allende made pivotal contributions to the understanding of protein synthesis through his characterization of elongation factors and mammalian aminoacyl-transfer RNA synthetases. He was also a pioneer in studying the mechanism of hormonal induction of amphibian oocyte maturation. In more recent years, Dr. Allende has been studying the structure, function and regulation of two ubiquitous protein kinases, CK1 and CK2, which are involved in the phosphorylation of key cellular proteins. In addition to his prominence in scientific research, Dr. Allende has been a leader, both nationally and internationally, in the field of science education and in the establishment of scientific networks in Latin America. He is a member and former President of the Chilean Academy of Sciences, a member of the Latin American Academy of Sciences, Vice President of the Third World Academy of Sciences, and a Foreign Associate of the U.S. National Academy of Sciences and of the U.S. Institute of Medicine. He is also Director of the ECBI project of the Chilean Academy of Sciences and the Ministry of Education and Coordinator of the Science Education Program of the Inter American Network of Academies of Sciences and of the Inter Academy Panel. His honors include the Chilean National Prize for Sciences 1992, Doctor Honoris Causa and the University of Buenos Aires, Argentina, Grand Cross of the Scientific Merit awarded by the President of Brazil.

Dr. Philip Bell is Associate Professor in the College of Education at the University of Washington where he also serves as the Administrative Lead of the Cognitive Studies in Education Graduate Program, and as Chair to the Committee on Faculty Affairs. His research is in the areas of science education, design and study of learning environments and technologies, theories of learning and instruction in science and history, children's epistemologies and images of disciplines, influence of pervasive technologies on development, and design-based research methods in education. Dr. Bell is currently a member of NRC Board on Science Education. He is a member of many educational associations, including the American Educational Research Association (AERA), National Association for Research in Science Teaching (NARST), Society for Social Studies of Science (4S), Association of Internet Researchers (AIR), and the Cognitive Science Society (CSS). He has served as a member of several advisory boards, including the Interim Board of Directors for the International Society of the Learning Sciences (ISLS), and the Center for Information Technology in Science (ITS). Dr. Bell received his Ph.D. in Education in Human Cognition and Development from the University of California, Berkeley.

Professor Julie Campbell is a Senior Principal Research Fellow of the National Health and Medical Research Council of Australia, Director of the Centre for Research in Vascular Biology at the University of Queensland, and Director of the Wesley Research Institute at the Wesley Hospital. Her major research interest over the last 30 years has been the cell biology of vascular smooth muscle in the normal artery wall and in diseased states such as atherosclerosis. Recently she has tissue-engineered vascular, bladder and uterine grafts grown in the peritoneal cavity, from autologous bone marrow-derived cells. In 1995 she

was awarded the Wellcome Australia Medal and in 2000 was elected to the Australian Academy of Science. In 2003 she received a Centenary Medal, and in 2004 was named a 'Queensland Great' by the State Government. In 2006 she was awarded the Order of Australia, Officer (AO) in the General Division by the Australian Government. In 2006 she became Secretary of Education and Public Awareness for the Australian Academy of Science with the brief to oversee inquiry based science education programs for primary school (Primary Connections: Linking Science with Literacy) and for early to middle high school (Science by Doing).

Professor Ernst W. Hamburger was faculty member and sometime head of the Department of Experimental Physics, *Universidade de São Paulo(USP)* from 1960 until retiring 2003. He conducted research in Experimental Nuclear Physics, Nuclear Structure and Reaction Mechanisms. Directed Curriculum Development Project in Physics for middle school, 1969-75; production of films for physics teaching, 1972-76. Set up graduate program on Physics Teaching, collaboration of Physics Institute and Education School in USP, 1973. As Director of *Estação Ciência* (where he still collaborates) a science center of USP, 1994-2003, increased activities of science exhibitions, science fairs, videos, theatre, internet, educational and social projects. Member of *Academia Brasileira de Ciências* since 1967 and Coordinator of Inquiry Based Science Education project in Brazil (for the *Academia*) and specifically in São Paulo, since 2001. Past director of *Sociedade Brasileira para o Progresso da Ciência* and of *Sociedade Brasileira de Física*. Prizes for Popularization of Science: awarded by Brazilian National Research Council, 1994; by Latin American Network for Popularization of Science, 2003; by UNESCO, "Kalinga", 2000. Member *Ordem Nacional do Mérito Científico*, Brazil, 2005.

Dr. Wynne Harlen has held several high ranking position including Sidney Jones Professor of Science Education and Head of the Education Department at the University of Liverpool and Director of the Scottish Council for Research in Education. She is now semi-retired and has an honorary position as Visiting Professor at the University of Bristol and undertakes some consultancies. She was an Osher Fellow at the Exploratorium, USA, in 1995 and a consultant and regular visitor to the Institute for Inquiry giving workshops in primary science and particularly in assessment. Dr. Harlen was also a consultant and co-director of a research project, funded by the NSF, at TERC, Cambridge, MA from 1999 to 2002. She was awarded the OBE by the Queen for services to education in 1991 and was given a special award for distinguished service to science education by the Association for Science Education (ASE) in 2001. She edited *Primary Science Review*, for the ASE from 1999 to 2004 and she serves on the editorial board of three international journals. Her publications include 25 research reports, over 150 journal articles, 26 books of which she is sole or joint author and contributions to 35 books and on science education and on assessment. She graduated from Oxford University an Honours Degree in Physics and obtained her PhD through research in educational evaluation at the University of Bristol.

Dr. Glenn Hultman is Professor in Pedagogic Practices - a multidisciplinary field, within teacher training, which integrates different disciplines in order to give a new perspective on various aspects of the role of schooling and teaching in society - at the Department of Educational Science and head of the Graduate school in Pedagogic practices at Linköping University. His research areas focus on knowledge creation/learning; leadership; teachers work, and the dynamics of change processes, within schools, public organizations and small business companies. Other research areas include evaluation, leadership training, the interplay between research and practice, interactive research and development of scientific methods (i.e. the contextual interview and the observview). His current research projects are knowledge creation and the development of schools - a study of the teacher's interactive working conditions; learning processes among children, young people and adults & the teacher's role; learning and understanding -the form and content of teaching and pupil learning in the natural sciences in the K-12 school (a study of STC in Sweden). He received his Ph.D. in Education from Linköping University

Dr. David Klahr is a Professor of Psychology at Carnegie Mellon University. Throughout his career, Dr. Klahr has focused on the analysis of complex cognitive processes in such diverse areas as voting behavior, college admissions, consumer choice, peer review, problem solving and scientific reasoning. Dr. Klahr's most recent research efforts have focused on the thinking processes that support children's understanding of the fundamental principles underlying scientific thinking. This work includes both basic research with pre-school children and more applied classroom studies of how children learn about experimental science. He has worked in a wide variety of schools in the Pittsburgh region, focusing on children's ability to learn how to design and interpret simple experiments. He recently served on two

National Research Council Committees, one of which produced the publication *Implementing Randomized Field Trials in Education, Strengthening Peer Review in Federal Agencies That Support Education Research, and Advancing Scientific Research in Education*. He received his Ph.D. in Organizations and Social Behavior from Carnegie Mellon University.

Dr. Pierre Léna is Professor Emeritus, Université Denis-Diderot (Paris 7), and currently Education secretary (Délégué à l'éducation et la formation) at the French Académie des sciences. As an astrophysicist, associated for his whole career with the Observatoire de Paris, he has contributed to the development of infrared astronomy by numerous observation programs from ground based, airborne and space telescopes, devoted to the study of the Sun, of interstellar matter and star formation. He has heavily contributed to the design of the European *Very Large Telescope* in Chile, especially for high angular resolution imaging, developing adaptive optics and optical interferometry. He has led the graduate school of astrophysics, a joint program of several Paris Universities, for many years. Dr. Léna has always been deeply involved in education matters, as President of the Société française de Physique (1989), President of the Institut national de recherche pédagogique (1991-1997) and one of the co-founder (with Georges Charpak and Yves Quéré) of *La main à la pâte*, a renovation program for science education at primary school level, which began in France in 1996 and successfully expanded in France, Europe and many countries through various collaborations. He is a member of the Académie des sciences in France, of the Académia Europeae and the Pontifical Academy of sciences.

Dr. Jean Matricon graduated from the Paris University in 1957 in Physics and Chemistry and received his Ph.D. in 1966 in theoretical solid state physics (superconductivity). After a postdoctoral position at Berkeley University, he had a position of Assistant Professor at the Paris University and became in 1970 full professor of Physics at the Université Paris7 Denis-Diderot up to his retirement in 1998. During that time, mainly devoted to teaching physics to students, he did research on different subjects in solid state physics and biophysics. He was also involved in many activities of science communication, like writing books and participating as scientific consultant in science museum exhibitions. Since he has retired, he is very much involved in the process of teaching Science to children, following “La main à la pâte” which is the French version of the “Hands On” method.

Dr. Jean Moon is a Senior Program Officer and Director of the Board on Science Education at the National Academies (USA). In this capacity she is responsible for providing leadership, through the work of the Board, to the National Research Council (NRC) on policy, research, and knowledge issues in science education, both in formal and informal environments, from early childhood through adulthood. She works with other NRC boards and committees in the Center for Education to build studies critical to policy, education, research, and practitioner communities. Prior to her position at the Academies, Dr. Moon was a program officer and education advisor at the ExxonMobil Foundation, overseeing development of the Foundation's pre-college and higher education portfolio. In 2001, Dr. Moon was invited to the University of Uppsala in Sweden as a scholar-in-residence to work on building communication skills in undergraduate science courses. While in Sweden, she worked with the Ministry of Education in Stockholm on matters of assessment and teacher education. A current interest of Dr. Moon's is the intersection of science, science education, and public policy. Dr. Moon received her Ph.D. from the University of Wisconsin – Milwaukee with an emphasis in learning and development as well as education within urban settings.

Ms. Senta Raizen is the Director of the National Center for Improving Science Education at WestEd. She headed a study sponsored by the U.S. Department of Education resulting in a series of reports dealing with science education reform at the secondary and elementary levels. She has led a number of major evaluation efforts, including evaluation of several federally funded sponsored programs that provide pre-service education and professional development for science and mathematics teachers, and has been active in international studies of science education and student achievement in science. Her current work includes development of a new framework and accompanying test specifications to guide the science assessment to be administered by the National Assessment of Educational Progress (NAEP) in 2009 and beyond. She has authored or edited a number of books and numerous articles in science and technology education. Recent books include: *Comprehensive Teacher Induction: Systems for Early Career Learning: Enhancing Program Quality in Science and Mathematics*; *Bold Ventures*, a three-volume series of case studies of mathematics and science innovations in the U.S. She is a fellow of the American Association for the Advancement of Science and a member of the American Educational Research Association.

Dr. Jayashree Ramadas serves on the faculty of the Homi Bhabha Centre for Science Education at the Tata Institute of Fundamental Research in Mumbai, India. She has been associated with summative and formative evaluation of formal and nonformal educational programs, classroom observations and survey research. Her Ph.D. thesis also addressed the phenomenon of students' alternative conceptions, at a time when this was yet to be well-recognized in the literature. Later, she did post-doctoral work with Rosalind Driver at Leeds, Michael Shayer at Chelsea College, and Seymour Papert at the Learning and Epistemology Group, MIT Media Lab. She has carried out research on students' conceptions in several areas including light, motion and Galilean relativity, the human body, plants and living things and the role of experiments in school science. Her current research is on visual thinking in science. She has taught courses at the graduate level on cognitive science, children's drawings and visuospatial thinking. She has worked in educational projects in rural Maharashtra, in schools of the Bombay Municipal Corporation (and as Visiting Scientist) in inner city schools of Leeds, UK, Boston, MA and Newark, NJ, USA. She holds a Master's degree in Physics from the Indian Institute of Technology, Kanpur and a Ph.D. in Science Education from the University of Pune.

Dr. Patricia Rowell is Professor Emeritus, University of Alberta, Canada.. Her research has entailed case studies of school science in many different contexts. She carried out one of the case studies in the 1984 Science Council of Canada national study of science education, and recently completed a follow-up study of elementary science education in the province of Alberta. She has participated in international projects in Botswana, Namibia, South Africa, Australia and China, addressing curriculum development and pedagogy in primary school science. Her current research interests focus on the discursive nature of teaching and learning science in elementary science classrooms and on the discourse of programs for schools in science centers. She has been an active member of the Centre for Mathematics, Science and Technology Education at the University of Alberta since its inception, and has contributed to many of the resources published by the Center for use by elementary teachers. She has coordinated and co-authored a recent series of climate change materials for elementary schools, funded by the federal government. Dr. Rowell received her Ph.D. from the Faculty of Education at the University of Alberta.

Dr. Richard Shavelson is Margaret Jacks Professor of Education and the former I. James Quillen Dean of the School of Education at Stanford University (1995-2000). He is also professor of psychology and Senior Fellow at the Woods Institute for the Environment at Stanford. Considered one of the leading experts on the measurement of human performance, Dr. Shavelson pioneered work in the measurement of human performance in zero and lunar gravities, in military job performance, and in science achievement. His work on generalizability theory, a statistical theory of the dependability of behavioral measurements) is widely used to model human-performance measurements. Dr. Shavelson's current work involves the study of inquiry-based science instruction and its impact on students' knowledge structures and performances. Other work includes a study of accountability and assessment in higher education and new standards for measuring students' science achievement in the National Assessment of Educational Progress. Before joining Stanford, he was dean of the Graduate School of Education and Professor of Applied Statistics at the University of California, Santa Barbara. He is author of *Statistical Reasoning for the Behavioral Sciences*; co-author with Professor Noreen Webb of the book, *Generalizability Theory: A Primer*; and co-editor of the National Research Council's publication, *Scientific Research in Education*. He received his Ph.D. in educational psychology from Stanford University.

Ms. Sally Goetz Shuler is the Executive Director of the National Science Resources Center, an organization of the National Academies and the Smithsonian Institution, with a mission to improve K-16 science learning and teaching in the U.S. and throughout the world. As one of the co-founders of the NSRC two decades ago, her leadership has created an organization committed to establishing effective science programs for all students. She has formed numerous strategic partnerships with national academies, academic institutions, corporations, and museums that are resulting in the development, implementation, and evaluation of research-based products and services for improving science education programs for school districts, states, and countries. In addition to the NSRC, her three decades of national and international experience in K-16 science education have also included 15 years of teaching math and biology and serving as a chair, trustee, or advisor for numerous boards and organizations. These include being a member of the Lemelson Advisory Board of the Smithsonian National Museum of American History; member of the Board of Trustees for the Keystone Center; the Merck Institute for Science Education Advisory Board; Chair of the Science Education Program of the Burroughs Wellcome Fund;

Chair of the Assessment Committee for the National Youth Science Camp; Advisory Board Members for the Math/Science Partnership Comprehensive Projects of Rutgers University and the Boston Science Partnership; Expert Panel for Washington State; a member of the DC Children and Youth Investment Trust Corporation Academic Advisory Panel; member of National Assessment of Educational Progress Science Steering Committee; and membership on the National Advisory Board of the Centers for Ocean Science Education Excellence.

Reviewers of the draft report

Pierre Brochu (Canada)
Rosa Deves (Chile)
Norma Sbarbati de Nudelman (Argentina)
Yasushi Ogura (Japan)
Helen Quinn (USA)
Edith Saltiel (France)
Padma M. Sarangapani (India)
Per-Olof Wickman (Sweden)

Selected Bibliography

The following are widely available published sources relevant to the report and additional to the references given in footnotes in the text.

1. Research and evaluation methods

- Guba, E.G. and Lincoln, Y.S. (1997) *Fourth Generation Evaluation*. Newbury Park: Sage.
 Harlen, W. (1975) *Science 5-13: a Formative Evaluation*. London: Macmillan Education
 Patton, M. Q. (2001) *Qualitative Research and Evaluation Methods*. Sage.
 Patton, M. Q. (1997) *Utilization-Focused Evaluation: The New Century Text*. Beverly Hills; Sage.

2. Students' learning and development

- Anderson, J. R., Greeno, J. G., Reder, L. M. & Simon, H. A. (2000) Perspectives on learning, thinking, and activity. *Educational Researcher*, Vol.29, No. 4, p.11-13
 Bransford, J.D., Brown, A.L. and Cocking, RR. (Eds) (2000) *How People Learn: Brain, Mind Experience and School*. Washington DC: National Academies Press
 Cobb, P. (1994) Constructivism in Mathematics and Science Education. *Educational Researcher*, Vol.23, No. 7, p.4
 Cobb, P. & Bowers, J. (1999) Cognitive and Situated Learning Perspectives in Theory and Practice. . *Educational Researcher*, Vol.28, No. 2, p.4-15
 Driver, R., Asoko, H., Leach, J., Mortimer, E. & Scott, P. (1994) Constructing Scientific Knowledge in the Classroom. . *Educational Researcher*, Vol.23, No. 7, p.5-12
 Nersessian, N. J. (1992) How do scientists think? Capturing the dynamics of conceptual change in science. In R. N. Giere (Ed.) *Cognitive models of science. Minnesota studies in the philosophy of science* (Vol. 15). Minneapolis: University of Minnesota Press. p. 3 - 44.
 Smith, J., di Sessa, A., & Roschelle, J. (1993). Misconceptions reconceived: A constructivist analysis of knowledge in transition. *Journal of the Learning Sciences* , 3, 115-163.

3. Science education practices

- Harlen, W. and Qualter, A. (2004) *The Teaching of Science in Primary Schools.(4th Edition)* London: David Fulton Publishers
 Prophet, R. and Rowell, P.M. (1990). "The Curriculum Observed". Chapter 1 in C.W. Snyder, Jr. and P.T. Ramatsui (Eds.), *Curriculum in the Classroom: Context of Change in Botswana's Junior Secondary School Instructional Program*. Gaborone: MacMillan.
 Rowell, P. M. & Ebbers, M. (2004). Constructing explanations of flight: A study of instructional discourse in primary science. *Language and Education*, 18 (3), 264-280.
 Rowell, P. M. & Ebbers, M. (2004). Constructing school science: Competing discourses in an inquiry-based elementary program. *International Journal of Science Education*, 26(8), 915-934.
 Rowell, P. M. (1997). Learning in school science: The promises and practices of writing. *Studies in Science Education*, 19-56.
 Schoultz, J., Hultman, G. and Lindkvist, M. 2005. Apprentices in context and complex didactical situations. Pupils' learning in science teaching, In *Educational Science in Sweden, 2005 – findings and future research*. Stockholm: Swedish Research Council, Committee for Educational Science, 2005: 13.

4. Student assessment issues

- Harlen, W. (2005) The role of assessment in the implementation of science in the primary school. In (Eds) T. Ellermeijer and P. Kemmers *Science is Primary*. Proceedings of the 2004 European Conference on Primary Science and Technology. Amsterdam: AMSTEL Institute 27 -48
 Harlen, W. (2003) *Enhancing Inquiry through Formative Assessment*. San Francisco: Exploratorium
 Neill, H, Bursch, P., Schaeffer, R., Thall, C., Yohe, M., and Zappardino, P. (nd) *Implementing Performance Assesments*. Cambridge, MA: Fairtest
 Pellegrino, J.W., Chudowsky, N. and Glaser, R. (1999) *Knowing What Students Know: The Science and Design of Educational Assessment*. Washington DC: National Academies Press

5. Recent conferences of relevance to IBSE in pre-secodnary schools were held in the following locations:

- Budapest (1999):** Conférence mondiale de l'Unesco, *La Science pour le XXIème Siècle: un nouvel engagement*, 26 juin – 1er juillet 1999
- Beijing (2000) :** *International conference on primary school science and mathematics education (CCBS/ICSU)*, nov. 2000.
- Monterrey, Mexico (2001, 2003) :**
- *Conferencia internacional sobre investigacion en la ensenanza de la ciencia* (11/13 septembre 2001);
 - *Second Monterrey international conference on K-12 education – social and economic impact of research based science education* (11-13 mai 2003);
- Vatican (2001):** *The Challenges for Science. Education for the Twenty-First Century*. Working Group, 19-21 November 2001, Vatican City, 2002 (published in *Scripta Varia*, vol.104).
- Kuala Lumpur, Malaysia (2004):** *2nd International Conference on Primary and Secondary Schools Science and Mathematics Education*, June 16-18, 2004.
- Rio (2002):** *ICSU conference on science and mathematics education – 21-23 septembre 2002*.
- Amsterdam (2003):** *ESERA Conference, Research and the Quality of Science Education*, Noordwijkerhout, August 19 - 23, 2003.
- Lima (2003):** *Encuentros Científicos de Invierno*, July 1-3, 2003.
- Santiago-du-Chili (2003, 2004) :**
- *Segundo Encuentro Latinoamericano sobre la Enseñanza de las Ciencias en la Educación Básica*, Santiago, 7 & 8 octobre, 2003
 - *International workshop on evaluation of inquiry based science education (IBSE) projects*, Jan 22-23, 2004.
- Bielefeld, Germany (2004):** *Ausserschulische Lernorte*, Feb. 16-18, 2004.
- Erice, Sicily (2004, 2005):**
- *La main à la pâte dans le monde : échange, partage, formation – July 14-18, 2004*
 - *European summer school for primary science trainers – July 9-14, 2005*
- Dakar (2005):** *Symposium international sur l'élaboration d'un programme régional d'enseignement des sciences en Afrique*, 10/11 mars 2005 (IAP/NASAC/ASTS).
- Berlin (2005):** *Science is primary – Internationale Konferenz zum naturwissenschaftlichen Grundschulunterricht – Sept. 27/28, 2005*
- Zlatibor, Serbia (2005):** *South-East European Summer School for Hands-on Primary Science Education*, June 11-16, 2005.
- Corpus Christi, Texas (2005):** *International Conference on Children's Rights and Education for the 21st Century*, Corpus Christi, July 26-30, 2005.
- Edmonton (2005):** IANAS Seminar
- Kampala (2005):** NASAC Seminar
- Stockholm (2005, 2006):**
- *Evaluation of Inquiry-Based Science Education Programs*, InterAcademy Panel (I.A.P.) Workshop, Royal Swedish Academy of Science, Stockholm, Sweden, September 21-23, 2005.
 - *Scienceeduc European Network Final Conference*, Royal Swedish Academy of Science, Stockholm, Sweden, September 21-23, 2005.

C1 Countries known to have IBSE programs in operation at some stages of pre-secondary education

Afghanistan	Estonia	Portugal
Argentina	France	Senegal
Australia	Germany	Serbia
Belgium	Hungary	Slovenia
Brazil	India	Spain
Cambodia	Italy	Sweden,
Canada	Malaysia	UK
Chile	Mexico	USA
China	Netherlands	Venezuela
Colombia	Norway	Vietnam
Egypt	Panama	

Some examples of IBSE programs

France: *La main à la pâte*

La main à la pâte (Lamap) is a program to renew primary school science education (maternelle + 5 years) initiated in France in 1996 by Georges Charpak, supported by the Académie des sciences. LAMAP has progressively transformed science teaching in the highly centralized system of French public schools, by providing a set of pilot centers (2000 classes), an Internet site for teachers (www.lamap.fr), kits and numerous aids for teachers, publishing guides and influencing the new French curriculum (2002). Ten simple principles articulate the role of inquiry in learning. These relate to the need for experiments and observation, the close relation of science acquisitions with language learning, the development in students of tolerance and critical thinking, and the involvement of families, other educational institutions and scientists.

From a few percent of science teaching in French classrooms in 1996, the proportion has raised to 30-40 % ten years later, with the creation of many local resource centers. A systematic evaluation of primary school science education will be conducted in 2007, but partial results have indicated the positive impact on children's learning. The success of these actions, reported by the media, has led to the establishment, in the fall 2006, of a new experimental program of integrated science and technology teaching in the two first years of secondary school. *La main à la pâte* is collaborating, in various ways, with many countries (nearly 30) of various cultures, helping them to develop inquiry based science education (see for example the site in Arabic : www.lamap.bibalex.org/). It leads a European network (*Pollen*) aimed at the creation of 12 "seed cities" in Europe, with schools that exemplify key aspects of learning and teaching science through inquiry (www.pollen-europa.net).

Brazil: ABC na Educação Científica – Mão na Massa

The program *ABC na Educação Científica – Mão na Massa* was started in 2001 as a collaboration between the Académie des Sciences of Paris and the Brazilian Academy ABC. That year was the first of several in which in several a group of nine primary school teachers and educational officers visited schools and centers of the *La main a la pâte* (*Lamap*) project in France for ten days. Over the same time about 15 teacher training programs were set up in 7 states, most being are in Science Centers in Universities, or have University support. In S. Paulo, the materials developed are mostly distributed during teacher training sessions, as printed sheets and experimental prototypes, which the schools can reproduce for class use. Each training session lasts about 4 hours for up to 40 teachers and can be part of a university extension course.

Originally the French texts were adapted; currently mostly new texts are being developed. The work follows, in general, the "ten principles" of *Lamap*, and has been mainly concentrated on the first years of schooling (ages 7 to 11), when each class has a single teacher. The connection between science teaching and literacy is stressed. The program is also applied to education of young adults who are not literate, handicapped children and also in kindergarten.

Internet sites (<http://educar.sc.usp.br/maomassa/> in S. Carlos and http://www.ciencia.usp.br/site%5F2005/mao_na_massa/default.html in S. Paulo - temporary address) contain materials for teachers and reports of application in classroom. A teacher training program via the Internet, after a few days of training in the science center, has been successfully put in place in S. Carlos.

United States: National Science Resources Center Science and Technology for Children (STC) Program

The STC Program™ is a comprehensive, research-based K-through-9 science curriculum program designed to improve the learning and teaching of science with all K-9 students in the United States and throughout the world. One of three current IBSE programs in the USA, STC is developed by the National Science Resources Center (NSRC) and published by Carolina Biological Supply Company. The STC Program™ consists of 32 instructional units designed with K–9 learning progressions in the life, earth, and physical sciences and technology. The STC Program is comprised of two curriculum series: Science and Technology for Children® (STC®), a 24-unit curriculum for students in kindergarten through grade six, and Science and Technology Concepts for Middle Schools™ (STC/MS™), comprised of eight rigorous courses for students in grades six through nine. Both programs were developed with major support from the National Science Foundation and numerous private foundations and corporations.

The STC Program materials for a unit or course include a comprehensive teacher's guide, set of student books, a literacy component, support materials for teachers, and the equipment and materials needed to conduct each of the lessons. Each of the 32 STC units provides students with opportunities to investigate important science topics in depth during an eight- to ten-week period. Through these experiences, students learn age-appropriate concepts, develop problem-solving and critical thinking skills, and acquire scientific attitudes and habits of mind. In conducting investigations, their design allows students to work cooperatively as well as independently to assess their prior knowledge and skills; ask questions; make and test predictions; record, share, and discuss their findings; monitor their learning; and apply the skills and knowledge they have gained to new situations.

Additional information about the STC Program can be found on the NSRC Web site – www.nsrconline.org – or on the Carolina Biological Web site – www.carolina.com.

China: 'Learning by Doing'

"Learning by Doing" (LBD) is a pioneer science education program initiated jointly by the Ministry of Education (MOE) and China Association for Science and Technology (CAST) in 2001, which promotes Hands On Inquiry Based Learning and Teaching on Science and Technology in kindergartens and primary schools (age 5-12), aiming to bring out student-centered instruction learning, foster children's scientific way of thinking and of living, cultivate qualified citizens with higher scientific literacy and facilitate the progress of quality education for all.

1. Orientation to every child and taking into account differences between individual children;
2. Laying the foundation for the child to learn through his/her lifetime, and more importantly, for children to learn how to live;
3. Selecting content closely related to children's life and surrounding environment;
4. Students at the center of active exploration and experience of the process of inquiry;
5. Teachers are supporters and instructors of children's learning;
6. Using encouraging assessment, including formative assessment;
7. Implementing science activities in collaboration with scientists and educators;
8. Mobilizing forces in communities and families to support;
9. Using the Internet to increase exchange and cooperation with domestic and international counterparts.

To support LBD, a research center has been established in Southeast University and a website www.handsbrain.com developed to connect all experiment school. The site provides books to guide teaching, related curricula and teaching materials, teacher's training etc, including on line teacher's training and supporting.

Starting with 44 schools in four major cities in 2001, LBD program now has spread to more than 400 kindergartens and primary schools as cores in 20 cities. The program has supports from the Chinese government and private sectors, and is based on international cooperation with ICSU-CCBS, French LAMAP and IBSE of IAP.

India: *Small Science*

Small Science is an inquiry-based curriculum for Grades 1 through 5 developed by the *Homi Bhabha Centre for Science Education (HBCSE)* of the *Tata Institute of Fundamental Research (TIFR)* in Mumbai, India. Based on research and field-work, the materials are further developed and refined through classroom trials. Sample chapters of the English, Hindi and Marathi versions of *Small Science* can be viewed at <http://www.hbcse.tifr.res.in/smallscience> .

Science in the early grades is seen as providing a rich and meaningful context in which to sharpen one's qualitative and quantitative tools of learning. Evaluation of students' learning in *Small Science* is therefore focused on skills of oral expression, reading, writing and calculating, along with observation, design, drawing and construction.

Over the last three years *Small Science* has served as a resource for National and State education bodies. About 25 schools around the country voluntarily implement it as their regular curriculum, while many others use it as supplementary material. The implementing schools span a range from elite urban to schools for children of tribal communities and migrant workers, and from conventional mainstream to progressive "alternative" schools. The *Small Science* website and mailing list collect feedback and offer support to teachers and parents. There is closer follow-up in two schools within the city of Mumbai. Teacher professional development has been initiated with a few schools in rural Maharashtra State. In 2007, with wider availability of the materials, the program is expected to be scaled up.

Research in Conjunction with IBSE Evaluations

Contributed by Richard Shavelson

The distinction between evaluation and research is a fuzzy one at best. Indeed, a good evaluation is one that has explicit theory behind it and the findings of the evaluation in one or more local settings bears not just on program quality, implementation or effects, but also on the theory the underlies the program evaluated. Indeed, the U.S. National Research Council noted difficulties with distinguishing the two and concluded the distinction was not worthwhile (Shavelson & Towne, 2002).

However, such a distinction seemed useful within the context of this project. The highest priorities for evaluating IBSE programs are to improve them, to establish the quality of their implementation, and to compare them against alternatives of interest to policy makers, educators and the public. With such clear foci, the possibility for “piggybacking” research onto evaluation might be missed. Here we briefly point out the kinds of research that might be piggybacked onto IBSE evaluations according to three types questions:²² (1) What’s happening? (2) Is there a systematic effect? (3) What is the causal mechanism and how does it work? (Shavelson & Towne, 2002).

What’s happening?

The question, what’s happening?, calls for description. What does inquiry teaching that impacts students’ learning look like in a fifth-grade classroom? what doesn’t it look like? Crawford (2000), for example, provided a detailed view of the day-to-day instructional processes of a high school biology teacher who successfully developed and sustained an inquiry-based classroom. By observing and analyzing the teacher’s instructional processes, Crawford identified six key aspects of inquiry-based instruction. They were: 1) using authentic problems; 2) grappling with real-world data; 3) teacher-student collaboration; 4) connecting learning to society; 5) teacher-modeling of the behaviors of a scientist; and 6) student ownership. According to Crawford, inquiry-based instruction necessitates teachers to take on more active and complex participation and requires teachers to assume a more expansive role (e.g. as a motivator, mentor, and scientist-modeler), relative to traditional instruction. Also, students need to learn by collaborative interactions and shared experiences, instead of receiving information passively. Research that examines the nature of inquiry teaching and whether it varies as to the nature of student characteristics or country differences would prove immensely valuable as we move to identifying effective IBSE practices

Wallace and Kang (2004) studied six experienced high school science teachers for their beliefs about: 1) what was successful science learning; 2) what were the purposes of laboratory work in science teaching; and 3) how inquiry was implemented in their

²² I am indebted to Jason Tan and Erin Furtak, science education graduate students studying inquiry teaching and learning, for some of the descriptions provided here.

classrooms. The study's results supported previous research findings that teacher beliefs about students and students learning, teacher understanding of the nature of science, and teachers' cultural perceptions about school science education (e.g. covering the curriculum, examination preparation) influenced the way they taught science and could promote or impede instructional methods suggested by professional developers, in this instance, the implementation of inquiry-based teaching. Beliefs that limited inquiry-based teaching were more public and culturally based (e.g. beliefs about: students' ability or maturity, efficiency in covering the mandated curriculum, and preparing students for examinations). By contrast, beliefs that promoted inquiry were more private and based on teachers' personal views of successful science teaching (e.g. beliefs that inquiry implementation fosters: independent thinking, problem-solving, and conceptual understanding). The authors suggested that because these private goals or views were not officially sanctioned, teachers struggled in their teaching to reconcile their goals with those that were culturally imposed on them. Such studies lead to research questions that might accompany evaluations: To what extent is this description of teacher beliefs country specific or cross-national? To what extent do such beliefs correspond with teaching practices and student outcomes? Do they depend on characteristics of the students?

Is there a causal effect?

The question, Is there a systematic effect?, refers to causal relationships. Research that attempts to establish a causal effect (e.g., randomized experiments, quasi-experiments) can quite naturally be piggybacked onto summative evaluations. But-it is also possible to carry out such experiments during formative evaluation as well, when questions arise as to how to best improve inquiry teaching. For example, during the development of an IBSE program, the effectiveness of different approaches to teacher enhancement for an IBSE program might be examined. Of particular import might be to see if effectiveness varies for teachers who differ in their beliefs about science teaching (see below).

Furtak (2006) noted that views of inquiry science teaching ranged from open-ended discovery to highly guided inquiry. Open-ended discovery could be characterized by students identifying questions about the natural world they were interested in and devising methods on their own to explore those questions with teacher support but no intervention. Highly guided inquiry refers to teachers closely directing the activities of students toward clear goals, such as enabling them to design well-controlled experiments. While the research data base seems fairly clear on the issue, what counts as inquiry teaching continues to range the gamut creating considerable confusion.

Mayer (2004) has pointed out that discovery approaches to science inquiry have produced negative effects when compared to guided discovery or direct instruction. He points out the discovery teaching fallacy—teachers' belief that keeping students active will lead to greater learning than inactivity. This turns out to be true if students are mentally active; physically activity is no substitute. In a series of randomized experiments, Klahr (e.g. Klahr & Nigam, 2004) found that guided inquiry produced more effective learning and transfer than did discovery.

The IBSE summative evaluations provide an opportunity to compare inquiry teaching practices. For example, the effect of embedding and using assessments in inquiry curriculum on student learning compared to inquiry teaching alone or traditional

instruction might be studied (see, e.g., Shavelson, 2006). The conceptual framework for such a study of embedded (“formative”) assessment has been provided by Black & Wiliam (1998). IBSE with embedded assessments, if used by teachers to develop students understanding and use of evidence to make knowledge claims, might improve students’ achievement.

Finally, drawing on the Wallace and Kang (2004) research, it appears that for inquiry science teaching to correspond to the description given in this document, teachers’ beliefs need to be aligned with those underlying inquiry. That is, it may not be sufficient for teachers to acquire the skills of inquiry; without the belief, the skills fall short of full implementation (e.g., Furtak, 2006; Yin, 2006). If teacher beliefs impact the effectiveness of professional develop on the quality of their implementation of IBSE, one possible study would be to compare alternative teacher enhancement programs, one of which is explicitly focused on conceptual change and the other the normal program for inquiry teaching. In such a study, teachers’ initial beliefs need to be measured and teachers need to be randomly allocated to enhancement condition so as to insure balanced comparison groups.

What is the causal mechanism?

Establishing a causal effect typically does not isolate the mechanism by which the effect occurs. For example, research in the United States established the effect of class-size reduction in a large-scale randomized trial (e.g., Finn & Achilles, 1990); smaller class-size produced greater student learning. However, the mechanism by which reduced class size caused greater achievement was (and is largely) not known. This opens up research of two kinds. The first takes the results of a summative evaluation that shows IBSE increasing students’ achievement, on average, over traditional teaching and attempting to determine what facet or facets of inquiry teaching were responsible for that increase. The second approach to addressing the mechanism question is to posit a causal mechanism, such as assessment conversations (Duschl & Gitomer, 1997) that have students justify their beliefs about the natural world with empirical evidence, and build the mechanism into inquiry based teaching to see the effects (e.g., Yin, 2005).

References

- Black, P., & Wiliam, D. (1998). Assessment and Classroom Learning. *Assessment in Education*, 5(1), 7-74.
- Crawford, B. A. (2000), Embracing the essence of inquiry: New roles for science teachers, *Journal of research in science teaching*, 37(9), 916-937.
- Duschl, R.A and Gitomer, D. (1997). Strategies and challenges to changing the focus of assessment and instruction in science classrooms. *Educational Assessment*, 4(1): 337-73.
- Finn, J.D., & Achilles, C.M. (1990). Answers and questions about class size: A statewide experiment. *American Educational Research Journal*, 27, 557-577.
- Furtak, E. (2006). The problem with answers: An exploration of guided scientific inquiry teaching. *Issues and Trends*, 453-467.
- Klahr, D., & Nigam, M. (2004). The Equivalence of Learning Paths in Early Science Instruction: Effects of Direct Instruction and Discovery Learning. *Psychological Science*, 15(10), 661-667.
- Mayer, R. E. (2004). Should there be a three-strikes rule against pure discovery

- learning? The case for guided methods of instruction. *American Psychologist*, 59(1), 14-19.
- Shavelson, R.J. (2006). On The Integration of Formative Assessment in Teaching and Learning: Implications for New Pathways in Teacher Education. In F. Oser, F. Achtenhagen, & U. Renold (Eds.), Competence-Oriented Teacher Training: Old Research Demands and New Pathways. Utrecht, The Netherlands: Sense Publishers.
- Shavelson, R.J., & Towne, L. (Eds.) (2002). Scientific research in education. Washington, DC: National Academy Press.
- Wallace, C. S. & Kang, N. H. (2004), An investigation of experienced secondary science teachers' beliefs about inquiry: An examination of competing belief sets, *Journal of research in science teaching*, 41(9), 936-960.
- Yin, Y. (2005). *The influence of formative assessment on student motivation, achievement, and conceptual change*. Unpublished doctoral dissertation, Stanford University.

Classroom Observation

Contributed by Pat Rowell

There are two main aspects of a lesson that may be observed; the framework or structure of the lesson, and the nature of participant engagement (or interactions).

Lesson framework

Topic/Objective/Activities
 Real-life Context for Activities
 Tasks which are investigation/evidence oriented?
 Generation of explanations?
 Communicating ideas?
 Use assessment strategies supportive of inquiry activities?

Classroom interactions

Classroom interactions are integral to the inquiry process.

The teacher is key to shaping these interactions. The teacher is responsible for introducing students to a discourse of inquiry. The teacher scaffolds the inquiry process by orienting student attention to the following aspects of inquiry:

- Students' beginning ideas
- Evidence supporting those ideas
- Asking questions that can be answered by investigation
- Tentative explanations, predictions
- Planning to gather evidence
- Gathering data
- Critiquing evidence
- Constructing explanations based on evidence/arguments supporting explanation
- Revising beginning ideas

An observation protocol could be designed to record observations of the following:

- Teacher talk
 - Does the teacher give instructions, lecture, illustrate explanations with demonstrations?
 - Provide explanations?
 - What view of science is portrayed in the teacher talk?
 - Does the teacher introduce and scaffold a discourse of inquiry?
- Teacher-student talk
 - Does the teacher ask for student ideas? Ask for student clarification? Ask for student explanations?
 - Does the teacher scaffold the planning of investigations by considering student suggestions for the procedures?
 - Does the teacher support student generation of explanations?
 - Does the teacher provide opportunities for students to share ideas?
- Student-student talk
 - Do students' share ideas in groups? Engage in active questioning? Use prior knowledge?
 - Do students show interest? Are students focused? Enthusiastic?

APPENDIX E2

Questionnaires for teachers, administrators and students

Contributed by Wynne Harlen

Views of how practice agrees with intentions

A useful form of questionnaire, which can be adapted for students, teachers or administrators, is to ask for two judgements in relation to a list of practices. The practices are set out in the centre of the page and respondents are asked to indicate:

- on the left hand side whether the statement is ‘never true’, ‘rarely true’, ‘often true’ or ‘mostly true’
- on the right hand side how important they consider the practice to be in helping students’ learning. This indicates how much they value the practice even if they are not able to use it very often.

The form for teachers will look something like this:

How true in your class?				Teaching practices	How important for students’ learning in science?			
Never true	Rarely true	Often true	Mostly true		Not at all	Limited	Important	Essential
				Statement 1				
				Statement 2				
				Statement 3				
				etc				

For a teachers’ questionnaire the statements could take a variety of forms, and include both ‘positive’ and ‘negative’ statements in relation to IBSE practice, for example:

Positive

- Students discuss in groups how to conduct investigations before they start
- Students are encouraged to ask questions that lead to further investigations
- I make sure that the student understanding the goals of the work
- Students have opportunities to assess their own work
- Students review each others’ work using given criteria

Negative

- Students work independently
- I make sure what they write in their notebooks is correct
- Students follow instructions carefully in any hands-on activity
- I ask questions mainly to see that knowledge they have

Comparing the reported practice and the judgement of values provides rich information about what is valued and thus may indicate further training needs, what is happening in practice, and whether there is any conflict between what teachers would like to do and what they are able to practice.

For administrators, the left hand judgements may be ‘How true in the classes you observe?’

For a students' questionnaire, the right hand side judgments would be 'how important is this for helping you to learn science?' Statements can echo those for teachers, suitably rephrased:

- The teacher makes sure that we understand the goals of our work
- We are encouraged to ask questions about our work

Teachers' and administrators' understanding of the meaning of inquiry in science

Teachers and administrators can be asked to express their reaction on a four point scale (disagree strongly, disagree to some extent, agree to some extent, strongly agree) to statements about inquiry, such as:

- Inquiry means learning from practical investigations
- Inquiry means collecting evidence to test ideas
- Inquiry means answering questions
- Inquiry always involves hands-on manipulation of materials
- Inquiry involves using information from various sources including books

Teachers' and administrators' understanding of inquiry teaching

Teachers and administrators respond to a series of statements about the teacher's role in IBSE. In each case they rate the importance on a scale from 1-4 (1.= not at all important; 2.= of a little importance; 3.= quite important; 4.= very important)

- Explaining things to students
- Finding students' ideas
- Asking students questions about what they know
- Encouraging students to ask questions
- Giving students accurate information
- Demonstrating how to do investigations
- Getting students to make predictions about what they might find
- Ensuring that students get scientifically correct answers
- Asking students to compare what they found with what they predicted
- Arranging for students to share what they have done with each other
- Ensuring students base their conclusions on evidence

Teachers' confidence in teaching science

In relation to the following statements, teachers rate the statements below according to these criteria:

- 1 = I need help to develop my knowledge and skills in this area
- 2 = I can manage but depend on advice from others
- 3 = I feel confident with a little guidance from others
- 4 = I feel fully confident in my knowledge and skills

Confidence in ability to help *students* to do the following:

- Suggest relevant questions for investigation
- Provide explanations related to scientific knowledge
- Make and test hypotheses
- Make predictions using knowledge and information gained in several different contexts

- Recognize the significance of variables in hands on investigations
- Provide reasons for planning decisions
- Select appropriate measurement devices
- Make an appropriate series of accurate measurements
- Record findings in tables, databases, bar charts and line graphs
- Use a computer for constructing tables, databases, charts and graphs
- Draw conclusions consistent with the findings
- Suggest ways of improving the reliability of the results
- Link results to the original hypothesis
- Write about an investigation giving details and evidence and using appropriate scientific language

Students' attitudes to IBSE

A questionnaire format that is particularly useful where change is to be measured on two different occasions, is to have pairs of statements on either side of a six point scale (Black et al, 2006)²³.

Possible statements relevant to inquiry based learning are:

I like to work out how to solve problems for myself	- - - - -	I like to be told what to do so that I get the answer right
It's useful to talk to others in the class about the work	- - - - -	I prefer not to talk when I'm working
I like to find things out for myself	- - - - -	I like to learn from what I am told by the teacher
I enjoy science lessons	= = = = =	I don't enjoy science lessons
I like to try different ways of doing things	- - - - -	I prefer to keep to one way of doing things
I learn science best by thinking things out	- - - - -	I learn science best by memorising answers
In science I like to think what might happen and then see if it does	- - - - -	In science I like to know what will happen before trying something
In science don't mind if I don't get the result I expect	- - - - -	In science its important for me to get the answer I expect

Some statements can be reversed in the questionnaire so that students don't automatically answers in one column.

²³ Black, P.Swann, J. and Wiliam, D. (2006) School pupils' beliefs about learning. *Research Papers in Education*. 21,(2)151-170

Interview guide for teachers and pupils

Contributed by Glenn Hultman

These interviews begin with one or several observations (often participant observations) and the evaluator/researcher participates in the classroom events and the IBSE-lesson. He or she takes notes for two purposes: to document what happens and to remember certain events that will be talked about in section three of this interview (the contextual interview). The observer should be a part of the whole situation but have a low profile, placed close to one of the groups of pupils or close in some other way. It is vital to avoid the role of someone who overtly is observing or appears to be controlling the procedures.

The interview starts with rather traditional questions and then moves to something like a conversation between two persons with the same interest, “what happens in the classroom/how do pupils learn science concepts,” with the focus on what took place during that lesson.

The following interview has three parts: background, opinions about IBSE and events in the classroom. Everything should be recorded (tape recorder, disc, MP3) and transcribed exactly (no language corrections made due to jargon or the like). A transcription from a video should document the no-talk parts, documenting what happens when pupils interact without words, when they manipulate materials and engage in a variety of hands-on activities.

Teacher guide:

Section one

1. What is your experience as a teacher (years, classes, colleagues ...)?
2. Why are you teaching science? How did you get interested in science?
3. Your training in science?

Section two

4. Why did you start with IBSE?
5. What's good about IBSE?
6. What's bad about it? Its weaknesses?
7. Has your way of teaching changed since you started with IBSE?
8. Has classroom conversation another dynamic now (do you experience any difference)?
9. Do you spend more time on planning?
10. Do pupils take part in the planning process?
11. Do you make your plans together with colleagues in your school?
12. Is there an increase in science discussions, now, in the school?
13. What is your opinion on the material in the kit (if appropriate)? Do you (or your pupils) have problems, using it?
14. Do you document your work, how?

15. Do you stick to the program guidelines, strictly? Or do you make some modifications? (*this question can result in very elaborate and different responses from teachers – some of them stick to the guide and others have their own way of using it*).
16. How do you “know” that pupils learn and understand science concepts? What kind of “evidence” do you use? What are you looking for, listening for? (*this is a transition or “rites-of-passage” question getting in to a conversation*)

Section three

17. And now the third part, which is a conversation drawn from observations made before this interview. The interviewer uses his/her notes and goes back to what happened in class and starts a discussion in an atmosphere of interest and curiosity – Why do you think it happened? What was your interpretation of event X? Did you see that group of boys in the back ...? Two girls sat on the floor next to me, did you notice that (do you know anything about them)? And so on ...

Pupil guide (adapt to age-level):

Section one

1. Do you like your school? In general.
2. What’s best about school? Do you have a favourite subject?

Section two

3. Do you like IBSE?
4. Do you now what it is about? What does it mean (IBSE/inquiry-based)?
5. What is meant by science? What is science?
6. Do you watch the TV-program “Brain office” (the equivalent local TV-program taking up science matters)
7. Do you watch other science-programs on TV?
8. Do you have favourite programs that you watch all the time?
9. Has your interest in science increased this year?
10. What will you become when you grow up, occupation?
11. What’s best about IBSE?
12. Are there some things that you do not like about IBSE?
13. What areas/concepts/modules have you worked with in the IBSE-program?
14. What did you like best?
15. Did you find anything complicated?
16. Did you learn anything from the IBSE-program/science-lesson? (*it is sometimes hard to answer a question about what you learn, but if the conversation tries different ways to express this, it may be fruitful*)
17. What did you learn today? (*question 16 and 17 is transition or “rites-of-passage” questions getting in to a conversation*)

Section three

18. And now the third part, which is a conversation drawn from observations made before this interview. The interviewer use his/her notes and goes back to what happened in class and starts a discussion in an atmosphere of interest and curiosity – Why did you and your friend use the blue and red piece? What happened when you did that? What did you think when event X happened? You were one of the boys in that group in the back ..., what did you do? You and your friend sat on the floor next to me, did you manage to solve the problem?

Did you understand what happened when you got that reaction? Can you explain to me about the experiment you did on Y? And so on ...

Contextual interviews

The third part of the interview is designed *to create a conversation* between “two observers,” the teacher and the evaluator/researcher. They talk about what happens in the classroom when they work with IBSE--understanding IBSE as perceived by teachers and pupils and the social construction of that reality.

The contextual interview can be done just as an interview but also in *combination* with observations, with the interviews as follow-up (“observings”). The same rules apply to these interviews as conventional ones: avoid leading questions, let interviewees talk most of the time, listen with an open mind and show interest (do not evaluate!) in the everyday-processes, do not limit the conversation to formal goals and policies (blinkers). Not too many questions about who said what; focus on what they did/what happened/the way it happened, and how they understand that event. Can they explain and give their intentions regarding what happened/what they did?

Contextual interviews can be described as conversational interviews, *in-situ*, and can take place in the classroom immediately after the lesson is over, or in a closed room in the school when the teacher and the evaluator discuss different events that happened when the evaluator participated in the teaching process. The evaluator participated (not as a quiet observer in the back of the room) and made small notes in order to remember situations of interest.

When using *video*, observe that the transcription only captures the communications on the tape/disc. Try to illustrate and write the silent parts. Many things happen in the quiet corners of the classroom. When the camera focuses on one corner, other things may be taking place in a different corner.

The following interview guides were used to structure the suggested type of conversation. One needs to formulate follow-up questions and explore alternatives depending on the nature of the interaction.

The quotes and vignettes, below, are examples from interviews and observations. The description of IBSE can be used to understand the “ethno”-version of IBSE. Reading between the lines, one can get an alternative picture of the process and IBSE as practiced in the observed classroom. The following quotes from interviews help in understanding the situation when IBSE was implemented at the classroom level.

The research-based approach is discussed in an interview (Girl, grade 4)

I - What did you learn today?

P - I didn't learn anything. I couldn't do it.

I - You couldn't?

P - No, we had a wire, which was supposed to glow, or I didn't know if it was supposed to smoke or glow or burn and ours just smoked.

I - Why do you think the wire would burn?

P - I don't know. Current came from the battery maybe.

I - That's right. Current came from the battery and you were supposed to see how a light bulb works.

P - But I didn't understand anything because it didn't glow. But in this it talks about the aim [points to a notebook] and we're supposed to write up the aim of the task. Then we know what it's all about and what we're to do.

Interview, the pupils learn about science I (Girl, grade 1)

P - I've learned how light bulbs should be connected. I've learned that if you have a light bulb and two batteries the bulb glows very brightly.

I - If you have a bulb and only one battery?

P - It glows weakly.

I - Do you have to think about anything special when you're connecting bulbs and batteries?

P - Yes, you have to do it right.

I - What do you have to keep in mind?

P - You have to think ...[draws a diagram] that this is a bulb and this is a battery. A round ring with a cross is a bulb and a line is a wire and you put a wire here and a wire here.

I - If you cut this what happens?

P - Oh, well then the bulb can't glow because the wire to the battery is gone and the current has to go round.

Interview, the pupils learn about science II (Boy, year 3)

P - I've learned that a flower has to receive pollen to be able to go on.

I - Do you know anything about the parts of a flower?

P - Yes, petals, stamens and stigma and pistils, stalk and these, for example, are seedpods.

I - Why are you drawing bees today?

P - I don't know ... oh, it was because we were working on pollination.

Conclusions from one implementation of an IBSE program

In a research study of IBSE, the following conclusions emerged about the implemented version of IBSE:

- Too much doing-- need more time for reflection
- Experimental work does not always provide for the development of concepts
- Open-ended questions can be confusing
- Pieces of fragmentary knowledge must be put together
- The task must be embedded in the whole lesson
- Introduction and summary are important
- Balance between learning science and learning processes, such as how to handle the material
- A context that gives pupils motivation and a desire to learn
- The material is not sufficient, the teacher is very important

The teacher's role

- Important to understand the content and purpose of the theme
- Explanation of the task so that the pupils can see the whole and the context
- Important to be familiar with the opportunities and difficulties of the theme
- Introduction of the theme is important
- Discussions and summaries of the pupils' results is a vital part of the whole
- The students need help in drawing conclusions

These conclusions, with a focus on the teacher's role, indicate that, in the reality of the classroom, there were many problems arising in the whole process from introduction to the closing of the theme.

