Report from the FP7 project:

Assess Inquiry in Science, Technology and Mathematics Education

ASSISTME

ASSIST-ME Proposal

Jens Dolin

Indhold

Executive Summary

ASSIST-ME is a high level research project with a societal impact that will investigate formative and summative assessment methods to support and to improve inquiry-based approaches in European science, technology and mathematics (STM) education.

Based on an analysis of what is known about summative and formative assessment of knowledge, skills and attitudes related to key STM competences and an analysis of European educational systems, the project will design a range of combined assessment methods. These methods will be tested in primary and secondary schools in different educational cultures in Europe in order to analyse the conditions that support or undermine the uptake of formative assessment related to inquiry processes.

The resulting synthesis of opportunities and restrictions for implementing an assessment culture using both formative and summative approaches will be evaluated and discussed in relevant forums in order to formulate guidelines and recommendations for policy makers, curriculum developers, teacher trainers and other stakeholders in the different European educational systems.

1. Scientific and technical quality

1.1 Concepts and Objectives

1.1.1 Project Overview

The overall aim of ASSIST-ME is to provide a research base on effective uptake of formative and summative assessment for inquiry-based, competence oriented Science, Technology and Mathematics (STM) education in primary and secondary education in different educational contexts in Europe and to use this research base to give policy makers and other stakeholders guidelines for ensuring that assessment enhances learning in STM education. In order to do this, the project will go through three phases shown in **Fejl! Henvisningskilde ikke fundet.**.

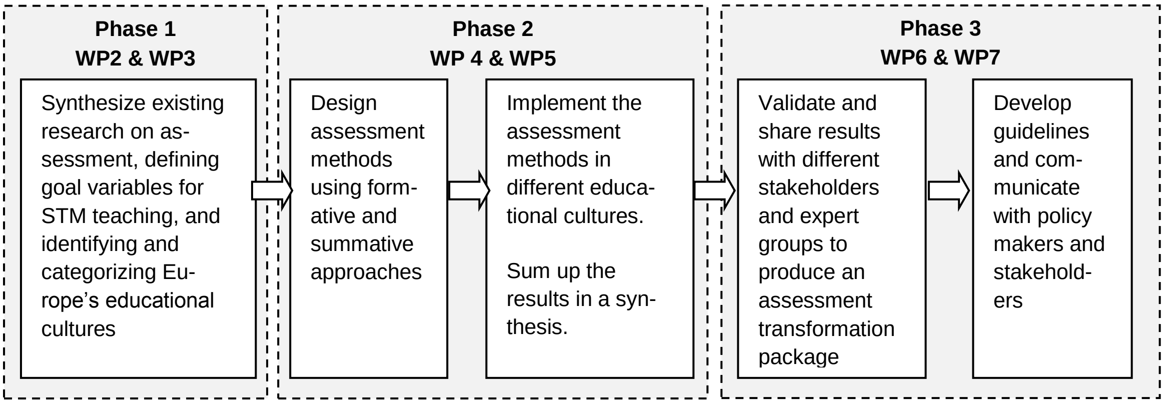

Figure 1. Overall project flow. The project starts in Phase 1, laying out the conceptual basis of the project and ends with Phase 3, validating and communicating the project results.

ASSIST-ME will develop **formative assessment methods** that (1) fit into **everyday classroom practice**, (2) provide qualitatively oriented descriptions and monitoring of **competence-oriented, inquiry-based** learning processes, and (3) can be **combined with existing summative assessment** requirements and methods used in different educational systems. The assessment methods will be developed to capture both general competences and disciplinary process competences such as science investigations and authentic problem solving.

The development and design of these methods will be based on existing research on formative and summative assessment, on current research-based understandings of competences in STM, and on previous and on-going EU projects on inquiry-based education (IBE). The project will characterise the educational systems in Europe based on existing data, such as the Eurydice statistics, and our own research, so that the chosen **set of assessment methods** will represent a relevant variety of methods and be adaptable to all European educational systems. The assessment methods will be tested by **teachers in Local Working Groups** and further developed during the project in cooperation between teachers and researchers in different educational systems. This testing will be at the core of the project as it will provide results from classroom practice in various educational cultures, subjects and educational levels. The results of the testing will be summed up in a **synthesis** of what works and which conditions and factors promote or undermine the uptake of various assessment methods in different educational settings.

From the beginning of the project, this whole process will take place in dialogue with established N**ational Stakeholder Panels** to secure the practical relevance of the results and an effective communication between shareholders and researchers. The assessment synthesis will be used in an extended sharing of results with policy makers, teachers and other stakeholders through forum discussions, workshops and seminars, and the collected feedback will expand the synthesis into an **assessment transformation package**. This package and focused policy oriented conferences will enable the creation of **guidelines and recommendations** for implementing both formative and summative assessment methods in inquiry-based STM all over Europe.

1.1.2 Objectives

The objectives of the project are:

- To describe STM competences, including both key competences in STM (European Commission, 2009) and transversal competences associated with STM (and often with other subjects as well), thereby offering a coherent framework of competences that can also depict those related to inquiry-based education which are currently not being assessed through typical testing.
- To characterise the aspects of educational systems in Europe relevant to producing guidelines for implementing various formative assessment formats.
- To develop a baseline definition of inquiry-based STM education and a set of assessment methods suitable for enhancing inquiry-based learning of STM related competences.
- To identify strategies for formative and summative assessment of competences in STM adaptable to various European educational systems.
- Draw up a set of guidelines and recommendations aimed at policy makers and other stakeholders to improve the uptake of formative assessment in combination with summative assessment to support inquiry-based education in STM.

1.1.3 Concepts of competence

The research into formative assessment of competences relevant for inquiry-based STM education will be based on an understanding of the concepts of competence, inquiry-based education and formative vs. summative assessment. The following first understandings will be refined during the project:

A **competence** is understood as a combination of skills, knowledge, characteristics, and traits that contribute to performances in particular domains. There is not a universal agreement on the terminology of competence. The OECD project ‘Defining and Selecting Key Competencies’ (Rychen & Salganik, 2001) distinguishes between a competence, referring to the concept in general and a level of ability, and a competency, referring to a particular demand that a person may or may not be able to meet. In the plural, they only use the term competencies. In this project we will use the word competence for both, and the plural form competences, to reflect an integration of understanding and attitude into the concept. Hartig, Klieme and Leutner (2008) describe a competence as a complex ability that is closely related to performance in real life situations. With respect to science education, this definition is not far from what PISA describes as scientific literacy referring to an individual's scientific knowledge and use of that knowledge to identify questions, to acquire new knowledge, to explain scientific phenomena, and to draw evidence-based conclusions about science-related issues (OECD, 2006). Following this understanding of competence, Shavelson considers that:

> "a competence measure should tap complex physical and/or intellectual skills, produce observable performance on a common, standardized set of tasks with high fidelity to the performances observed in the "real world" ("criterion") situations to which inferences of competence are to be drawn, with scores reflecting the level of performance (mastery or continuous) on tasks where improvement can be made through deliberative practice" (Shavelson, 2011).

The more complex the learning goals, the more difficult they are to measure. The understanding of competences as the ability to cope with relatively complex challenges in everyday life means that assessment methods necessarily have to be relatively advanced, flexible and process oriented. Thus, artificial tasks such as multiple choice test items that might test simple skills can hardly measure competences.

Iquiry-based education (IBE) can promote achievement of two different types of objectives:

1. Objectives related to domain specific competences, i.e. key competences in mathematics, science and technology.
2. Objectives related to transversal competences, i.e. cross curricular competences (we use the terms transversal competences and cross curricular competences synonymously).

The line of separation between these two sets of competence goals is not sharp and even within each type of objective there might be considerable overlap. The following lists are therefore a preliminary selection and will be adjusted according to country characteristics.

Domain specific competences in *science* include:

- To observe and describe natural and technical phenomena accurately,
- to ask questions and to generate hypotheses,
- to plan, perform, and evaluate experiments,
- to reflect on the plan, performance, and evaluation of the experiments,
- to label, systematise and arrange phenomena, materials, and living beings by characteristics and functions,
- to consider claims and conclusions critically in relation to available evidence, and
- to ask and answer questions related to science and its impact on society.

Domain specific competences in *technology* include:

- To choose and use suitable tools, instruments, and materials well adapted to the research questions and hypotheses,
- to construct and produce technical tools and instruments, i.e. to plan, project, manufacture, evaluate optimize, examine and test technical solutions,
- to reflect on products and solutions and use of materials, and
- to consider questions related to technology and its impact on society.

Domain specific competences in *mathematics* include:

- To test a presumption or speculation in order to discover an adequate procedure that can be generalized,
- to explore a mathematical problem, to formulate hypotheses and to prove or disprove them by systematic experiments,
- to develop and test hypotheses about functional relations that have been observed in the reality (from data) or in mathematical settings (e.g. number patterns),
- to solve algebraic equations of certain types and given in certain forms, possibly with specific technological tools
- to explain and justify steps of reasoning in a formally correct way, for instance in the context of classical (plane) geometry

Transversal competences that should be fostered by IBE are:

- To ask questions and to work on problems independently and with others,
- to plan stages of work and to perform projects,
- to gather, select and judge information in different media (e.g. journals, books, internet),
- to process and interpret data and results,
- to describe and explain results, facts, and circumstances from different perspectives,
- to argue and communicate with peers and experts,
- to reflect on her/his own learning, to control and steer it, and
- to develop a sense for responsibility and to become a responsible-minded citizen.

In some European countries in the last ten years competence models and standards have been developed and become law, (e.g. in Austria, Germany and Switzerland). Their models and standards primarily contain domain specific objectives and only – if at all – secondarily transversal competences. Other European countries (e.g. France and Czech Republic) have developed curricula or frameworks for curriculum development which contain both domain specific and transversal competences.

It will be an important task for ASSIST-ME to analyse the significance of national competence orientations for the implementation of formative assessments within inquiry processes and of summative assessment of competences.

1.1.4 The concept of inquiry-based STM education

Inquiry-based education (IBE) is an umbrella term, encompassing a wide range of teaching approaches that can enhance student motivation and has a potential for also enhancing the learning outcomes.

Inquiry-based STM education includes students' involvement in questioning, reasoning, searching for relevant documents, observing, conjecturing, data gathering and interpreting, investigative practical work and collaborative discussions, and working with problems from and applicable to real-life contexts (Anderson, 2002). Inquiry-based STM-education is not a new teaching method but it is often used as a contrast to more traditional teaching approaches such as those where the teacher presents results and methods which the students are then trained to "apply". Giving students an active part in learning is in accordance with many teachers seeing the pedagogical principles of constructivism as the foundation for understanding and implementing inquiry-based learning (Llewellyn, 2007).

The current emphasis on **inquiry-based science education** originates from the so-called Rocard report (European Commission, 2007) arguing for a renewed pedagogy that encourages inquiry-based methods. The recommendations were followed up by a series of FP7 funded projects such as Mind The Gap, S-TEAM, ESTABLISH, Fibonacci, INQUIRE, PROFILES etc., in which most of the partners in ASSIST-ME have participated. The present consortium behind ASSIST-ME therefore has first-hand knowledge of the results of these projects that we will draw upon. ASSIST-ME is also inspired by recent research that suggests that "…effective informal formative assessment practices may be associated with student learning in scientific inquiry classrooms." (Ruiz-Primo & Furtak, 2007).

Inquiry-based technology education is not sharply separated from inquiry-based science education. Many of the characteristics are the same, particularly in view of increasing emphasis in science on reflective thinking, cooperative learning, and the development of critical and creative thinking, previously ascribed to technology (Adams & Hamm, 1998), probably due to the design and team work focus of technology education. Lewis (2006) identified the conceptual parallels of design and inquiry as providing an ideal "border crossing" for technology education and science education – with mathematics being seen as a bridge between the two.

Inquiry mathematics means mathematics teaching and learning in which students solve nonstandard mathematical problems designed with the potential to bring forth mathematical ideas related to the topic at hand while the teacher supports the students' reasoning and orchestrates the classroom discussion. These kinds of teaching approaches include, for example, inquiry mathematics (Cobb et al., 1992), didactical engineering (Artigue, 1994), open approach lessons (Nohda, 2000) and problem-centered learning (Schoenfeld, 1985). Indeed, most of design oriented research on mathematics education, from the early seventies, can be described roughly as aiming towards developing inquiry situations in which students meet mathematics as a "live subject" to explore and construct, rather than as an inventory of dry standard methods which are simply to be acquired through "training" practices. Also in this field, the difficulty of assessing inquiry practices – often involving some level of creativity on the parts of students – remains a tough challenge, and the outcomes of mathematics teaching is still mostly assessed (in standardized tests) based on mastery of standard techniques.

According to previous research, inquiry mathematics enhances students' understanding and mathematical thinking (Fennema et al., 1996; Wood & Sellers, 1997) as well as creativity and problem solving skills (Kwon et al., 2006). This kind of learning is durable and the knowledge can be applied to new contexts (Francisco & Maher, 2005). Inquiry mathematics also develops positive attitudes and beliefs in students (Wood & Sellers, 1997). According to Sullivan et al. (2006), inquiry mathematics enhances the involvement of all kinds of students. Despite the research based evidence on the benefits of inquiry mathematics it is not applied very often in schools.

The ASSIST-ME approach to IBE

Both inquiry-based and traditional teaching cover a wide range of teaching methods and often proponents of each contrast them with caricatures of alternative modes of instruction of the other. The research synthesis by Minner et al. (2010) noted the poor quality of much of the research in this area and the resulting difficulties in concluding cogently. So, in drawing conclusions from empirical research one must be specific about what the criterion measures are and what the basis is for judging success.

It is therefore important for the project to have a well defined approach to inquiry-based instruction including some parameters that are operational and comparable across different educational settings. The ASSIST-ME definition and these parameters will be established in connection with the characterization of the educational cultures in the participating countries. The definition will to a large degree be based on elements from existing FP7 projects with an understanding of science based on open questioning and model building linked to experimental hypothesis testing and a pedagogical approach emphasising students' investigative and innovative processes, and a specific understanding of learning as involving linguistic processes. Together with frequently quoted definitions of IBE this means that students are

- building their understanding of fundamental scientific ideas (NSF, 1997);
- finding answers to their own and others' questions (AAAS, 2000);
- exploring and using practices employed by scientists (Osborne, 2011; IAP, 2011);
- taking charge of their learning as needed for genuine understanding (Harlen, 2009)

While research says inquiry teaching can produce positive results, it does not, by itself, tell teachers exactly how to do it (Anderson, 2002). Inquiry-based STM approaches can initially be difficult for teachers to adopt and for good reasons they are sometimes met with opposition from teachers. This emphasizes the importance of adapting any proposed change in teaching to existing local teaching culture and for involving teacher expert panels and teaching associates in the research of the project.

The project sees inquiry-based approaches as a very powerful way of improving science, techno-logy and math education. We acknowledge the diversity of inquiry forms needed to enhance the learning in different disciplines, levels and cultures. Various meanings and forms of "inquiry" are possible in different disciplines. In particular, the specificity of inquiry in mathematics (by far the largest school subject involved) needs to be addressed and pursued in the development of assessment methodologies. But it is equally important that these different approaches to inquiry are all aligned with the general definition of inquiry produced by the project.

1.1.5 Concepts of assessment

An assessment method is to be understood here as a package of information, procedures and instructions aimed at collecting and interpreting information about students' competences within STM. Typically, an assessment method will include a hands-on activity, paper & pencil-activity, peer-to-peer feedback or on-screen activity for students. The instructions will describe possible uses of this activity to assess students – for formative purposes or as a contribution to summative assessment, or both, as appropriate – and will give ideas on how to use the results.

The specific assessment methods chosen in ASSIST-ME for classroom testing will depend on the preceding research. But it is likely that they will include a process monitoring tool (e.g. a portfolio format), a method for assessing practical work, a computer based test system with built-in feedback, and a structured dialogue tool (such as the assessment conversation described by Ruiz-Primo 2011).

Science, technology and mathematics (STM) are the central areas to be addressed in this project. At secondary levels science is often taught within specific subjects (biology, physics, chemistry etc.), and depending on the results from the characterising of the educational systems, these subjects will be dealt with separately whenever relevant. Mathematics on the other hand has different practices and experiences with inquiry teaching at all educational levels.

Because several competences in STM can be measured with one assessment method, it is possible to cover a wide range of disciplines and levels with a limited number of assessment methods, where possible variations within a method will enable different levels or achievement to be accommodated. In general for the sciences the levels of interest will be during the final years of primary and lower-secondary schools and the middle of upper-secondary. Methods will also vary according to whether the sciences are integrated or separate. Mathematics assessment trials at the same levels will be conducted separately with a specific view on compatibility with high stakes exams in the subject which strongly influence the shape of this discipline in many countries, particularly at secondary levels.

Assessment is one of the most important drivers in education and a defining aspect of any educational system. Formative and summative assessments are similar in that they involve the collection, interpretation and use of data for some purpose. They are mainly identified and distinguished from each other by the purpose of the assessment but often also in the way data is collected. Thus:

- Formative assessment has the *purpose* of assisting learning and for that reason is also called 'assessment *for* learning'. It involves *processes* of "seeking and interpreting evidence for use by learners and their teachers to decide where the learners are in their learning and where they need to go and how best to get there" (ARG 2002).
- Summative assessment has the *purpose* of summarising and reporting learning at a particular time and for that reason is also called 'assessment *of* learning'. It involves *processes* of summing up by reviewing learning over a period of time or checking-up by testing learning at a particular time

For the purpose of improving student learning, formative formats must be used and they must be intimately connected with the processes of instruction and learning as well as with the goals for this instruction (Black, 1998). An effective formative assessment depends on a notion of progression in the competences that are the learning goals and an understanding of the inquiry processes that lead to their acquisition – thus linking the major elements of this project.

The importance of what is included in summative assessment, and how the assessment is conducted cannot be expressed too strongly. Summative assessment has an undeniably strong impact on what is taught and how it is taught (Harlen, 2007). This influence is positive if the learning aims are fully reflected in the assessment, but all too often this is not the case. Particularly when the assessment results have important consequences (high stakes) for students and/or teachers, the assessment will tend to determine what is taught and indeed what students pay attention to. Thus in order to promote IBE outcomes, it is essential to ensure that these outcomes are included in what is assessed. This applies both to formative assessment by teachers and to the summative assessment determined by teachers or by institutions external to the school, which are responsible for producing assessment instruments and procedures. This also explains the importance of the policy level of the project, as the change of high stakes (summative) assessment procedures must usually be decided at this level and cannot be changed by individual teachers.

However, assessing the goals of IBE – understanding of powerful scientific and mathematical ideas, building inquiry competences, developing understanding of scientific and mathematical activity and fostering corresponding attitudes – is not easy. It is far easier to assess straight-forward recall of facts and principles, or standard techniques. Where this continues, the spread of better practices in STM education – as widely advocated by the OECD and the EU – will be held back. What is assessed and how it is assessed has to be brought into line with the competences identified in this proposal as important outcomes of science, technology and mathematics education. The ASSIST-ME product, the assessment transformation package, will therefore seek to enable the development of effective summative assessment procedures and policies which are consistent with the aims of IBE in STM.

The **assessment transformation package** produced by ASSIST-ME will offer users examples of assessment guidelines and methods focusing on inquiry learning environments and accommodate differences in educational systems across Europe. This can be used by three main target groups:

- By policy makers to inform decision making on curriculum design, teacher training and assessment strategies at institutional, regional and national levels taking into account relevant system characteristics and variables.
- By teachers and teacher educators to develop effective combinations of formative and summative assessment in daily practice in primary and secondary schools.
- By researchers to study formative and summative assessment methodologies and practices in different educational systems.

The assessment transformation packages will be accessible through SCIENTIX and other the ASSIST-ME project website, free to use and maintained for at least five years after the project's completion. In the last part of the project there will be dissemination activities involving active engagement with key personal throughout Europe (see also Work Package 7 description in Table 1.3a and Section 3.2 on p. 40).

1.1.6 Variables used to characterise educational systems in Europe

It is a main goal of this project to find effective combinations of formative and summative assessment methodologies. We will use the existing assessment procedures and methodologies used in a range of educational systems in Europe as a starting point. The differences between European countries are described in Eurydice (2011) and from this it is possible to group educational systems according to their assessment in STM. The project will use this data and combine it with data on variables that are found, as part of the project, to have an effect and impact on the promotion of inquiry-based STM education, to characterize educational systems in Europe.

The variables in Table 1.1.6 will serve as a starting point. Our consortium represents a broad spectrum of educational systems. For almost each of the ten variables at least one of our countries is a paradigmatic example for one end or for the other of the dimension. For example: 1st variable: France represents a typical centralized system, Switzerland a typical de-centralized system. 2nd variable: Germany has very strong streaming from the beginning of grade 5 (in 13 out of 16 federal states); Finland has no streaming at all until after the end of grade 9.

Table 1.1.6a Variables related to inquiry-based STM and useful to characterize educational systems in Europe

Paradigmatic countries represent the particular antipodes of a variable.

Variable	Antipode I	***Paradigmatic country for I***	***Paradigmatic country for II***	Antipode II
1. Centralization	Centralized System	*France, Finland*	*Switzerland*	De-centralized system
2. Streaming in lower secondary level	Strong streaming (more than one track to follow)	*Germany*	*Finland, Denmark*	No streaming at all (only one track, comprehensive schools)

Variable	Antipode I	*Paradigmatic country for I*	*Paradigmatic country for II*	Antipode II
3. Standardised tests vs. teacher autonomy	Standardized Tests required for assessment	*England, France*	*Finland, Switzerland, Czech Republic*	Teachers have full autonomy in assessment
4. Existence of Competence model	Explicit (model written in the curriculum)	*France, Germany, Switzerland, Czech Republic*	*England, Denmark*	Implicit (no model explicitly stated)
5. Integration of science subjects in primary and lower secondary	Separate subjects: Biology, Chemistry, Physics	*France, Germany, Finland, Denmark (lower secondary)*	*Switzerland, Denmark (primary)*	Integrated science
6. Autonomy of schools	Low (schools follow rules/are govern in detail)	*France*	*Finland, Czech Republic*	High (schools are responsible for school profile, parts of curriculum, teacher development, etc.)
7. Professional teacher development	Low (PTD is not common for every STM teacher)	*Germany, Finland, Czech Republic*	*Switzerland*	High (teachers have the right and duty to participate in PTD)
8. Textbook	Prescribed	*Germany (Math)*	*Denmark, Switzerland, England, Czech Republic*	Not prescribed
9. Practical work	Not common (less than 25% of teaching time)	*Czech Republic*	*Germany, England*	Common (more than 25% of teaching time)
10. Status of Inquiry-based STM education	Low (IBE is not mentioned in STM curriculum or not used often in STM education	*Finland (Math), Czech Republic*	*England*	High (IBE is mentioned in STM curriculum and used often in STM education)

1.2 Progress beyond the state-of-the-art

Whilst a good deal is known about many of the general obstacles to implementing inquiry processes in science, technology and mathematics (STM) in classrooms (e.g. Harlen, 2009a), there is less clarity about how these obstacles play together with the assessment traditions and competence orientation in different educational settings. Moreover the research into implementation has not yet been disseminated effectively at the action and policy levels due to a missing links between teachers, researchers and policy makers. Consequently, ASSIST-ME will investigate the implementation processes of various aspects of formative assessment related to inquiry teaching in STM education and their relationship to variables of educational systems in Europe. We will look at relations between formative and summative assessment procedures and the problems of assessing competences not normally captured by traditional assessment methods. We will research the classroom implementation of these formative assessment methods and we will specifically look into the role of ICT as a lever for enhancing formative assessment of complex competences. To secure a political impact we will set up models for involving policy makers and other stakeholders in formulating recommendations and guidelines for change and its realization.

These goals lead to the following research questions which ASSIST-ME will address:

1. What are the main challenges related to the uptake of formative assessment in the daily practices in science, technology and mathematics in primary and secondary schools in different European educational systems?
   1.1. In their efforts to enact innovative inquiry-based teaching-learning sequences, how do teachers approach the need to monitor student learning as it develops? To what extent does they use structured formative assessment and in what formats?
   1.2. What systemic support measures and what tools do teachers need in order to integrate formative assessment of student learning in their classroom practice?
2. What changes are needed in summative assessment practices?
   2.1. To bring them into consistency with the learning aims of IBE in STM?
   2.2. To ensure that they support and do not inhibit the practice of formative assessment?
3. How can formative and summative assessment methods including use of ICT be used together to promote learning in inquiry-based STM?
4. How can research-based strategies for the use of formative/summative assessment be adapted to various European educational traditions to ensure their use and avoid hindrances?
   4.1. How can the diverse roles of summative and formative assessment be clearly delineated for teachers and what strategies can help them make appropriate use of both, each to fit its own purposes?
   4.2. How can relevant stakeholders be invited to take co-ownership to the research results and how can a partnership between researchers, policy makers, and teachers be established in order to secure relevant actions following implementation guidelines?

1.2.1 The influence and shortcomings of (high stakes) summative tests

The widely available products of recent FP7 projects aimed at promoting inquiry teaching methods have not been accompanied by equally innovative assessment methods appropriate for these inquiry approaches. Instead many teachers have been left with traditional testing instruments which do not capture the full range of processes and outcomes related to inquiry learning. The lack of such supportive tools can diminish the implementation and the effects of the newly developed FP7 products. Like all learning processes, inquiry learning is strongly enhanced by formative assessment (Black and Wiliam, 1998). Formative assessment, by providing frequent feedback to students, can also significantly enhance student self-efficacy in STM and therefore increase the potential for greater learning outcomes.

There is strong opposition to summative assessments when used for high-stake decisions affecting individual students, teacher and schools (Linn, 2000). International comparative tests such as TIMSS and PISA provide data at the system level and potentially have greater validity than tests used to report on individual students since they can include a greater range of test items. PISA differentiates itself from TIMSS with an assessment framework based on universal definitions of disciplinary literacies in modern economies, not the shared science curricula of TIMSS, and these definitions and the PISA results which address them have sometimes become determining factors in establishing national policies (Dolin and Krogh, 2010). The PISA Framework has an intention of going beyond traditional standardized tests, not checking students' knowledge but rather testing whether students can apply their knowledge to new situations, which is consistent with a competence model of assessment. The problem is that PISA results cannot be used by teachers as tools for monitoring and improving their own practice. Even if PISA has achieved its objectives of providing policy makers with relevant information, it still leaves teachers with a lack of tools for pedagogical use that relate their practice to the PISA results. Moreover, the competences assessed by PISA, designed to describe "common ground", do not capture the full breadth of goals and needs to be met in a given educational context. Consequently, there can be a tension between the teachers' use of this information for pedagogical purposes in the classroom and the legitimate need for information about the performance of an educational system to inform policy makers.

In particular, because of their use of paper and pencil testing formats, large-scale international as well as national tests are often unable to fully assess many STM competences such as the ability to design experiments, collect data, hypothesize, predict, argue, conclude, and validate; not to mention the cross curricular competences such as working together, search for relevant information, directing one's own learning etc. These 'blind-spots' in most large scale testing results hits the assessment of FP7 inquiry teaching particularly hard since much inquiry-based teaching emphasizes the acquisition of just these competences which are so essential for education in the 21st century (OECD, 2000). Consequently, without test-data to support these methods, they may not have the institutional and stakeholder support to be maintained and supported.

1.2.2 The role of formative and summative assessment

The role of formative assessment in inquiry-based STM education

Ensuring that students have the kinds of opportunities needed for real progress requires a broad interpretation of inquiry-based STM education: more than using skills for exploring and finding out, and more than providing first-hand experiences of materials and phenomena – even though these are important. Development of understanding involves taking into account students' existing ideas and skills and promoting progression by adjusting challenge to match these starting ideas (Bransford et al., 1999). The practice of FA, through collecting data about learning as it takes place and feeding this back to teachers and students, helps to regulate the teaching and learning process and promote progression. It also supports ownership of their learning through promoting student self-assessment and participation in decisions about next steps, helping students to take some responsibility for their learning.

The role of summative assessment in inquiry-based STM education

One must also be concerned with SA for two main reasons, both firmly supported by research evidence. Firstly, the strong influence that SA has on the curriculum and pedagogy means that SA must reflect the full range of attitudes and competences if students are to have the opportunity to achieve the goals of inquiry-based STM education. Secondly, when summative assessment data are used for high stakes accountability this has the effect of focusing attention on reliability rather than on the validity of the measures used. Tests are preferred to other methods of assessment since they give the impression of precision and the provision of 'hard evidence'. The demands of high reliability in test construction further reduce the range of learning outcomes that are assessed and encourage 'teaching to the test'. It also leads to a distortion of teachers' own assessment, towards constantly checking on summative achievement in ways that copy formal tests rather than using assessment to help learning (James, 2000). Frequent summative testing sets up "a performance orientation that ultimately may decrease motivation" (Brookhart and DeVoge, 1999: 423). The review of research on motivation and testing by Harlen and Deakin Crick (2003) gave clear evidence of the negative impact of tests on motivation for learning.

The influence that SA can have on FA means that giving attention to FA only would be likely to have little effect. Indeed the experience of the obstacles to introducing genuine FA in countries where there exists a strong dependence on external high stakes tests, bears evidence to this. Thus if we are to improve inquiry-based learning in science through the use of FA it is necessary also to ensure that the SA is congruent with the learning aims of inquiry-based STM. As suggested in the next section, there are opportunities to do this through a system that links the two, ensuring consistency between the information used to help learning and that used to report on it.

1.2.3 Effective combination of formative and summative assessment

The research into and development of an effective combination of – and alignment between – formative assessment (FA) and summative assessment (SA) aims at fostering major progress beyond the state-of-the-art of assessment in STM. As noted earlier, the main distinction between FA and SA is in terms of uses and purposes. For FA there is one main use – to support learning (both directly by the student him/herself and by the teacher's improvement of teaching). If the information about student learning is not used to facilitate this learning, then the process cannot be described as formative assessment. By contrast, the data from SA can be used in several ways, some relating to individual students and some to the aggregated results of groups or populations.

Linking formative and summative assessment

Although a broad distinction can be made between FA and SA in terms of use of information, there is no sharp discontinuity between them, for it is possible to use some data both formatively and summatively. This may be either a) by making formative use of data gathered for a summative purpose or b) making summative use of data gathered for a formative purpose.

a) *Formative use of summative data.* Black et al. (2003) include formative use of summative data from tests as one of four practices that teachers found were effective ways of implementing FA (the others being questioning, feedback through marking student work and student self-assessment). In order to use tests formatively the teachers involved the students in creating and marking the tests and used the results diagnostically.

b) *Summative use of formative data.* Summative assessment is an occasional, if regular, event. In between classroom tests or grading of course work, there are innumerable opportunities for teachers to gather data about students' on-going learning. The evidence these data provide may be used immediately to guide students or considered later and used in planning subsequent teaching. The evidence will relate to the particular goals of the lesson or sequence of lessons, but this rich and varied evidence which is "tied to students' instructional experiences over the course of a year, [but] can also be evaluated in a standardized ways for the purposes of summative assessment" (Chudowsky and Pellegrino, 2003). Harlen and James (1997) suggest that this means taking care to distinguish between evidence and the interpretation of the evidence. For formative purposes the interpretation is against the goals of particular work and how to help further progress. For summative purposes the evidence is interpreted in terms of more general criteria or standards. Since FA has to be conducted by the teacher, the resulting SA will be derived from classroom-based assessment and steps have to be taken to ensure confidence in the outcomes. Both Chudowsky and Pellegrino (2003) and Harlen and James (1997) provide examples of dual use of data from regular classroom work.

c) *Combining summative and formative assessment.* In both a) and b) there are limitations in the dual use of the evidence, but of rather different kinds. The limitation of using evidence which has initially been gathered for a summative purpose to help learning bears on the validity of the evidence; it is just not sufficiently rich and readily available to be fully adequate for formative use. The limitation of using evidence which has initially been gathered by teachers to help learning in order to report on achievement bears on the reliability of the evidence and its interpretation. The richness of formative data means that it has great potential for encompassing the full range of goals of inquiry-based STM, providing strong arguments for developing procedures that enhance the reliability of classroom-based approaches. In order to overcome these limitations we need "to integrate summative and formative assessment more closely" (Looney, 2011). This point of view leads to building an integrated assessment system which combines formative and summative assessments (Birenbaum et al, n.d.).

As noted earlier, it is generally recognized among educators and policy makers that assessment plays a crucial role in enhancing the outcome of education. Many politicians still have a strong belief in tests as a lever for raising standards and feel a need for comparable data based on national and international tests while they at the same time advocate for pursuing more advanced competences in the school system, competences that cannot be captured by traditional tests. This contradiction between the existing assessment systems and politicians' need for hard data on the one side and the necessity for assessment methods that can enhance a competence oriented learning process on the other side, can only be overcome through alignment of formative and summative assessment. An assessment *of*, *for* and *as* learning rather than seeing testing as an isolated element in education, gives opportunities for enhancing the learning of competences for the 21st century.

A combination of formative and summative assessment can be defined through six dimensions (Birenbaum et al., n.d.), which will be the founding elements of the ASSIST-ME assessment methods:

- **alignment to instruction**: the topic of assessment is what the learners know and are able to do (it is not focusing on gaps in learner knowledge and/or performance);
- **authenticity**: it reflects real life situations and includes the development of problem solving skills which represent a basic requirement of modern societies;
- **multidimensional approach**: it presents non-conventional tasks and stores students' performances on these and other classroom tasks;
- **transparency**: assessment criteria are made explicit to individual learners by teachers;
- **assisting learning**: it allows both students and teachers to gain information about learning progression in order to plan future learning appropriately;
- **multilevel reporting**: it generates reports at various levels of aggregation, addressing both students' progress and students' level of achievement.

Engagement in developing and implementing this combination of FA and SA transforms teachers' practices so that they will use the results in feedback to the students. Furthermore, this combination impacts also students and reinforce their motivation since they are encouraged to participate in the process and will get information about their learning progression.

1.2.4 The role of ICT

ICT will play an important role in the developed methods. The use of computer based learning and assessment packages opens up the potential to explore science and mathematics constructs more fully. It offers improved representation and responds to the ways that science and mathematics are presented to students in classrooms and in the modern world. It is also an advantage that through computer based delivery, fast, formative and diagnostic feedback for teachers and students can be designed and built into learning and assessment systems.

Simulations can actively engage students in a range of science or mathematics content or inquiry skill areas and allow for many different types of inquiry in science, from the active manipulation of variables within a system to the use of articles and data to demonstrate how science is evaluated and communicated. Development of science conceptual understanding through modelling is a significant feature of simulated environments, allowing students to create multiple variable models, test, run and evaluate them. These approaches support a constructivist view to the learning and assessment of science. Similarly in mathematics, simulations allow modelling of more abstract areas, making hypotheses and (inductively) proving their ideas in ways not possible on paper.

Pearson has carried out extensive research and development programmes in the use of new technologies to support learning and assessment in science and mathematics (Clesham, 2011, 2010, 2009; Dimitrova, 2006; Dimitrova et al., 2004) and can offer a range of formative and summative assessment approaches. Pearson can thus provide the project with knowledge, expertise and experience in the opportunities that computer-based learning and assessment can afford in terms of systematic efficiency and reliability, increased levels of engagement, motivation and accessibility for students and the potential to enhance the areas of science and mathematics constructs that can be learnt and assessed through this medium.

1.2.5 The need for an understanding of progression in competences

In section 1.1.3 we identified competence as including skills and knowledge, as well as other traits that contribute to performance. In order to find out where students are in the development of ideas, skills and traits and to use this information to identify next steps, through the use of formative assessment, it is necessary to know what the course of development is. This assumes that there is a typical and identifiable progression in development. However, this is not universally agreed upon (Millar and Driver,1987) the main point at issue being whether this progression is determined by teaching or is a 'natural' one, reflecting underlying cognitive development (Adey, 1997). The issue is illustrated in relation to skills.

Since skills are always applied in relation to some content, the nature of this content influences their deployment. Because of this content dependency, it can be argued that skills do not develop, but that students become progressively more able to deploy them in relation to more complex content and a wider range of situations (Perkins and Saloman, 1989). However, evidence supporting the notion of development of inquiry skills has come from research, from classroom experience where students are using skills with familiar content and from national (Russell, 1988; NAEP, 2008) and international surveys such as the PISA in which results are reported in terms of levels of achievement.

There is undoubtedly a need for more systematic study to determine more reliably the course of progression in skills. Meanwhile, various attempts to describe progression in skills and conceptual understanding have been made drawing on existing knowledge. Masters and Forster (1996) have produced a series of 'progress maps', describing dimensions of achievement in aspects of mathematics and Harlen (2006) has produced a series of indicators of progression in science inquiry skills. Other descriptions of progression are expressed in various national curricula, as in the levels of attainment of the national curriculum for England (DfEE, 1999), in the Swiss competence model (Konsortium HarmoS Naturwissenschaften, 2008), in the US National Science Education Standards (NRC, 1996), and in the Benchmarks of the AAAS (AAAS, 1992, 2000) and associated Atlas for Science Literacy (AAAS, 2001). However these are coarse-grained statements suitable for summarising attainment. For helping learning, finer-grained statements are needed, describing progression in more detail without necessarily relating to levels or grades.

ASSIST-ME will review what is known about progression in competences relevant to IBE in STM and use this in Phase 2 to define and describe progression in terms suitable for guiding formative assessment and for summative reporting of achievement.

1.2.6 Linking research and policy through National Stakeholder Panels

One of the main goals for ASSISST-ME is to provide policy makers with data and guidelines for an informed decision making about encouraging and improving formative assessment in STM education. But research tells us that it is not easy for research to affect policy (Fensham, 2009) due to different logics and different discourses. In his article “Speaking Truth to Power with Powerful Results: Impacting Public Awareness and Public Policy” Mack Shelley II (2009) underlines the need for eclecticism in the research and its interface with expertise and policy. The point is that to transmit understanding to decision makers you need to break down the barriers between the research world and the policy world through better communication and an understandable and usable message. This can only be achieved if the two parts meet to exchange ideas and understandings and accept each other’ respective capacities and influence.

The same argumentation is valid for the relation between researchers and teachers and the relation between teachers and policy makers. So, even if the main axis in this project is along researchers and policy makers, teachers will be included in the meetings described below to secure the legitimacy for practice.

In order to maximise the project’s engagement with user communities including policy makers, invited experts will be brought in for meetings with the project management board and national stakeholder panels will be established to advise and provide professional development on how best to impact on policy and practice.

A ‘Developing Links’ strand of work will be undertaken with the specific aim of bringing policy and practice users into dialogue with the project and its outcomes throughout the project’s lifetime. This programme of work will involve a proactive identification of key partners with whom the project would benefit from establishing links and dialogue.

On a national level **National Stakeholder Panels** will be established in all partner countries at the project start. University of Copenhagen will facilitate the identification of relevant stakeholders via a social network analysis method developed at the university based on Knoke (2011). These panels will work throughout the project linking research, policy and practice and securing a meaningful communication.

King’s College London will organise a series of **meetings and roundtable discussions on EU-level** with key organisations such as the education ministries, professional societies, science and mathematics education organisations and representatives from the teaching profession and related education professionals (e.g. professional developers) on key issues.

Publicity and marketing of the project, its findings and associated resources will be targeted broadly, e.g. through the national centres for science and mathematics education, via articles in education newspapers (such as the *Times Education Supplement* in the UK) through professional teacher publications and via a network of links, such as the websites and publications associated with relevant EU projects and advisory committee members.

A carefully considered and planned **engagement with the media and press** will also be essential to the delivery of the work package. A detailed plan will be drawn up in conjunction with the scientific advisory board, which includes policy makers, but is likely to involve a 'drip effect' approach, with for instance, press releases every six months to heighten awareness of the project and to publicise key findings and invite dialogue with wider publics.

1.3 Scientific methodology and associated work plan

1.3.1 Overall strategy

The project proceeds through different phases as outlined in Figure 2 on page 24, involving various research approaches and interacting processes depending on the tasks formulated. Figure 3 also on page 24 shows the timing of work packages.

Phase 1 – Building the foundation (Work Package 2 and 3): The first phase focuses on producing the knowledge base for making a research-based design and trial of assessment methods (Phase 2) and for creating strong relations to stakeholders in every participating country (Phase 3). This knowledge base secures strong relations between ongoing and ended projects on STM-education and a common reference for the project participant. The work consists of four parts:

- Analyse existing research on how summative and formative assessment of knowledge, competences and attitudes in STM can be coupled to inquiry-based teaching.
- Establish a competence model for the STM specific and the transversal competences.
- Describe and characterise educational systems in Europe with respect to variables and factors relevant to both formative and summative assessment in STM.
- Map and analyse national networks of relevant stakeholders using social network analysis and establish National Stakeholder Panels.

Phase 2 – Design and test assessment methods (Work Package 4 and 5): In order to support the real large scale uptake of formative assessment it is not enough to give theoretical evidence. Methods must be tested in real classrooms by real teachers and be combined with existing summative assessment. This will produce evidence on how to implement new forms assessment relevant for other countries with similar characteristics. The work is divided into three parts:

- Design a number of assessment methods for formative and summative assessment of competences in different educational settings in Europe – with special attention on inquiry-based teaching methods.
- Research the classroom implementation of the assessment methods in different STM subjects, levels and educational settings selected to represent the variation across Europe.
- Analyse the test results and develop an assessment synthesis that accounts for all relevant factors supporting or undermining the effective uptake of formative assessment and the combination of formative and summative assessment.

Phase 3 – Transform and communicate research results (Work Package 6 and 7): Research results from the classroom are not policy. The results need to be interpreted and transformed into a discourse of policy in order to be used for decision making. Phase three does this transformation and disseminate the outcome in two steps:

- Use the National Stakeholder Panels to transform Phase 2 results into a transformation package of data, guidelines and recommendations adaptable for large scale implementation of the assessment methods in different educational contexts.
- Communicate the transformation package to policy makers, decision makers and other stakeholders in partner countries and representatives from EU member states through dissemination activities to facilitate a large scale adoption of the competence based formative assessment.

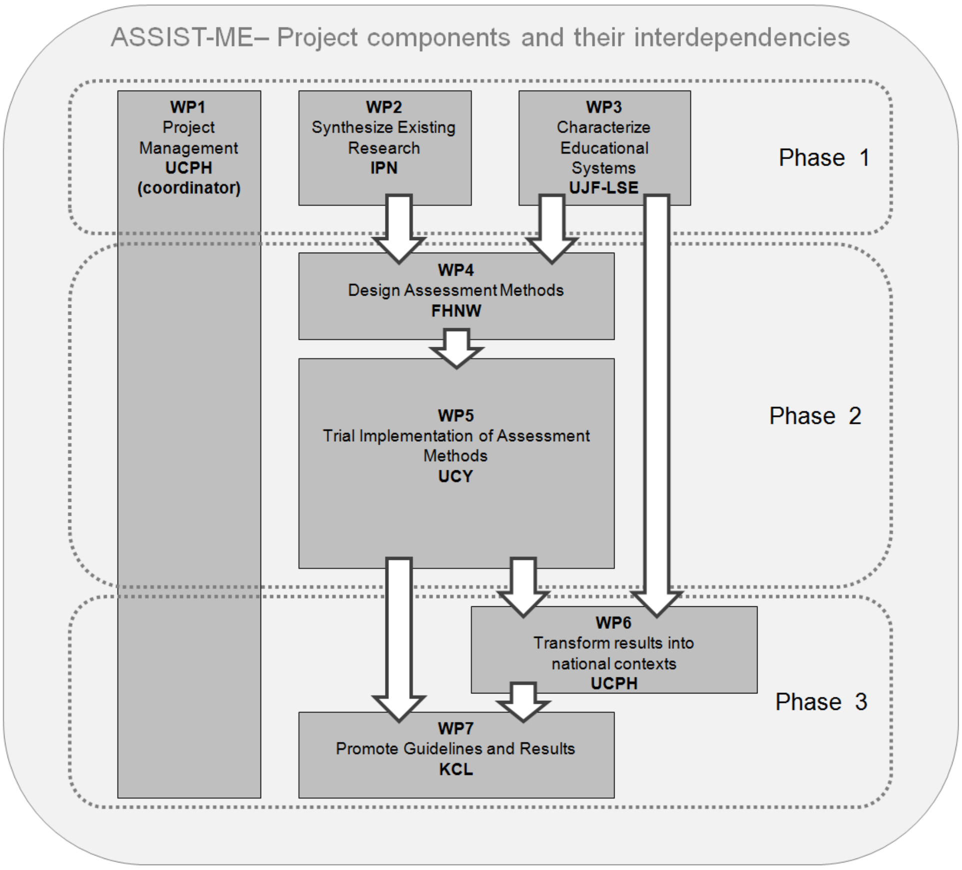

Figure 2. ASSIST-ME overview and work packages.

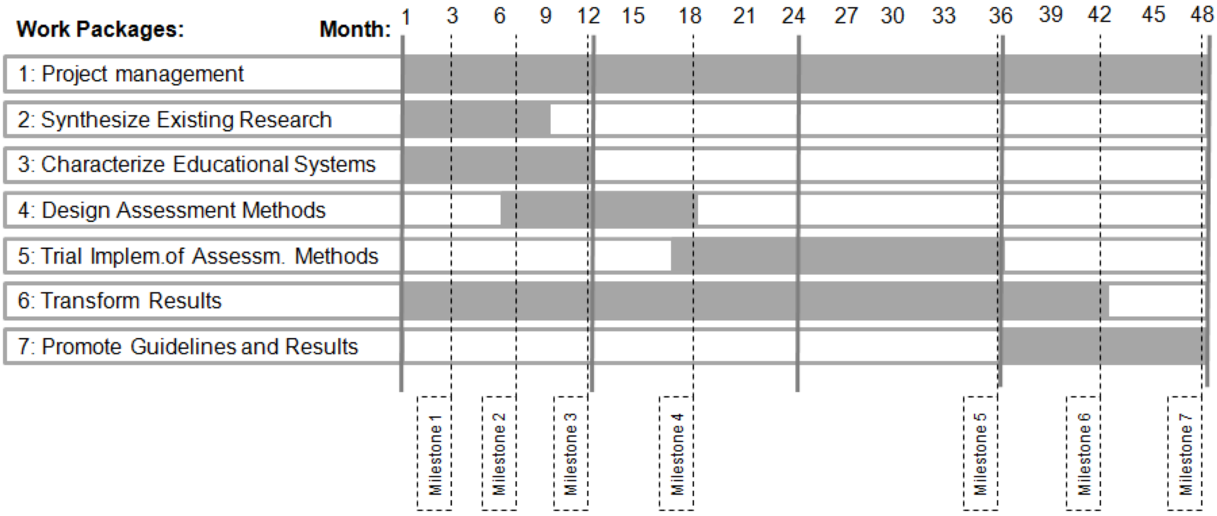

Figure 3. The timing of the different work packages and milestones.

1.3.2 Research methods

The research in **Phase 1** will be *desktop research* combined with *field data collection and validation*. The constituent models will be formulated by the researchers based on the research reviews and feedback from relevant stakeholders. The stakeholder networks will be established using *network theory*. The research in **Phase 2** will mainly be classroom based *action research* with the trial implementation of the assessment models and methods. The focus will not be on the student learning outcomes nor the chosen assessment methods' ability to measure such outcomes. The focus will be on the teachers' trial implementation of the various assessment methods (and thus indirectly taking the efficacy of the assessment methods into account). Experienced teachers will form **Local Working Groups** together with researchers. For comparability reasons all teachers will participate in teaching profile questionnaires to clarify their approaches to teaching, to their discipline, their current use of formative assessment methods and their familiarity with inquiry-based education.

The trial implementation of each assessment method will take place in four month periods and before each implementation period the teachers will be familiarized with an assessment method. The implementation will be monitored at

- Individual teacher levels,
- Local Working Group levels, and
- National levels (to generalize beyond the two LWGs).

For each assessment method a teacher *self assessment form* will be developed focusing on the quality of the method, the implementation process, the applicability of the method (e.g. its ability to assess complexity, progress, various competences etc.), method refinement ideas and other relevant information.

The individual teacher evaluations will be discussed with the researchers in the Local Working Groups, *focus group interviews* will be performed and the results will be aggregated at national levels.

Parallel with the teacher self-reporting, researchers will *observe* the teachers' use of the assessment methods and *video-record* the implementation process for extraction of the key elements of the process. The purpose of the videos will be to illustrate strategies for use of the formative assessment methods in order to facilitate teacher conversation about the processes.

A Steering Group will *collect and organize the data and analyse* it for results at national levels. The results will be discussed with and qualified by the National Stakeholder Panels.

The research in **Phase 3** will be a *transformation* of the accumulated feedback from the National Stakeholder Panels and *production of recommendations and guidelines* based on theories of communication and theories of impact.

Table 1.3f: Significant risks and contingency plans

Risk	Contingency plan
Changes in national educational systems and consequently any national regulation of assessment during project	Since ASSIST-ME has chosen partners to represent a diverse sample of European educational systems, any changes in the national system of any of our partners will not hinder the project since we are developing assessment methods for a large variety of systems. Such changes may actually be useful if they include new approaches to teaching and learning which we don't already have represented in the project.
We may find that the educational systems represented by our partners do not include a key system variable determined by our WP2 and WP3 research which needs to be addressed.	We assume that even if an educational system at the official national level does not include a variant of interest, with our close connection to schools we can find these variations in local and individual schools where we can test the assessment methods.
Some teacher associates (LWGs) may not fulfil their obligations to fully test an assessment in their classroom according to the project protocol.	Since we have budgeting significant funds to pay teachers for their active participation, we will make those payments contingent upon complete fulfilment of trial protocols and reporting. With three rounds of assessment trials, we can replace teachers after the first round, and with the motivation of remuneration, readily recruit new teachers.
Finding teachers who use IBE in their teaching may be difficult.	Since we are intentionally not attempting to sample national teaching populations and because our partners have close connections to schooling in their own countries, we are confident, based on conversations we have already had, that we can find enough LWG teachers who use inquiry for the trial implementations.
Permission to use video-recording in some classrooms may not be obtainable.	Depending on the situation: we can agree not to use a recording for any purpose other than for collecting data; avoid recording students who withhold permission; record only teachers and not students; use only audio-recording and occasionally, simply make personal observations.
Even with our best efforts at engaging stakeholders from early in the project, we may find it difficult to motivate them to take action based on our research results.	By including two media representative in the Stakeholder Panels from the beginning, the panels will have public pressure generated by the media to take action based on the project's findings. We will encourage the media to promote our outcomes to the general public and thereby make it difficult for our stakeholders to neglect project results.
The media reporting from those representatives on our Stakeholder Panels may be critical of aspects of our project, processes and results.	We will welcome such criticism since it would serve as an external check of our efforts. Since we will have reporters involved from early-on in the project, their feedback (via reports and articles) will formatively allow us to be responsive to the stakeholder community through-out the project.

Risk	Contingency plan
Local resources for some assessment methods which require materials (e.g. for hands-on testing) or equipment (e.g. computers for screen-based assessment) may be scarce.	Both of these limitations are part of the diversity of educational system variance which the project addresses. So dealing with them in the trial implementations will allow us to create solutions for them which will be part of our product methods. Making such adjustments will provide precisely the kinds of adaptations to different educational systems which will make our final assessment transformative package widely useable.

2. Implementation

2.1 Management structure and procedures

The management structure of ASSIST-ME has three layers. The General Assembly is responsible for the overall administrative and financial aspects of the project. The Management board is responsible for the scientific aspects, and lastly the coordinator is responsible for the day-to-day management of the project. Due to the scope and expected impact of the project, an additional Scientific and Policy advisory board has been created. Each of these, supporting arrangements will be described in detail in the following sections. An overview of the project management is presented in Figure 4.

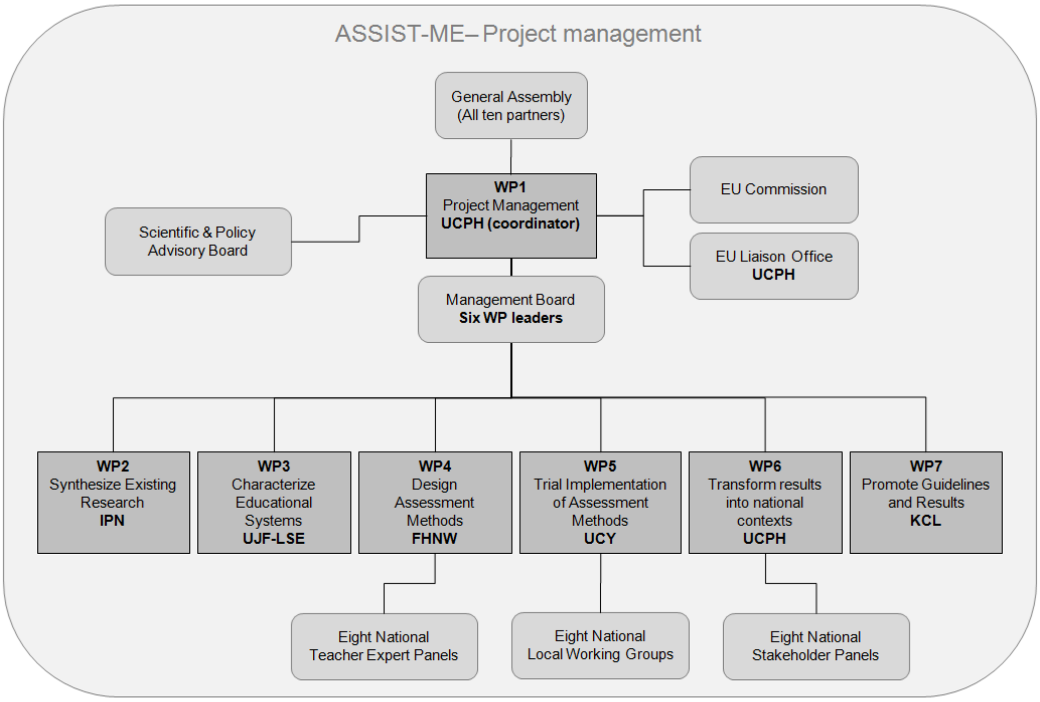

Figure 4. Overview of ASSIST-ME project management and relations between work packages, Advisory Board, Panels and the Local Working Groups of teachers.

2.1.1 Project Coordinator

The University of Copenhagen (UCPH) will act as coordinator for the ASSIST-ME project. UCPH has been selected for their strong scientific position in the area, the strong network and competent leadership of Jens Dolin and management team (described in detail below).

The project will be anchored with **Head of Department Dr. Jens Dolin**, Department of Science Education, University of Copenhagen. He will ensure the overall coordination and leading of the project. He has worked for six years as the Head of Department, has extensive experience in international cooperation including recently serving as a member of the Executive Board of ESERA (European Science Education Research Association) and has participated in a large number of international research projects and EU framework projects on Science and Technology education (S-TEAM, Mind the Gap and more).

As project coordinator, University of Copenhagen, represented by Mr Dolin, will be responsible for contact with the Commission. UCPH will ensure the drafting and negotiation of the Consortium Agreement, represent the project externally, and monitor the progress of the project. As the coordinator UCPH will also be in close contact with the Scientific and Policy advisory Board. As coordinator UCPH has accepted to take on the contractual responsibilities such as; the financial and administrative management of the project; monitoring the delivery of all deliverables and dissemination activities; organizing meetings, general assemblies and project conferences. Besides these, as coordinator UCPH will also be responsible for setting up a project website, and will for this task, team up with both a professional web-designer and the UCPH (in-house) graphics designers. The Website will contain both public spaces for dissemination of results and closed work spaces for the project partners to work and share documents. The graphics designers will also be utilized for the production of written material on the project (booklet, flyers, posters, brief, etc.).

Professor Dr Carl Winsløw has been appointed project manager for the ASSIST-ME project. He will be a driving force in both the scientific progress of the project and he will undertake of the day-to-day communication between partners on development and research issues. Carl Winsløw has a core expertise in mathematics education as well as project management from a number of international projects.

The experiences of Jens Dolin and Carl Winsløw cover the wide spectrum of STM education and inquiry-based methods at different educational level, and also the high level administration and day-to-day management skills needed to ensure the success of the project. The project will also receive the full support of the University of Copenhagen, including the resources at the **EU Liaison Office**. The EU Office at UCPH have vast experiences in running EU projects, with participation in more than 265 ongoing FP7 projects, of which around 20 are as coordinator. The EU office will be available to the Coordinator and the project partners with assistance on any issues related to the European Commission. The extensive participation in the project has ensured smooth procedures regarding payments, both financial and scientific reporting and more.

2.1.2 General Assembly

The General Assembly (GA) is the **overall governing body of the project** and consists of one representative of each partner. The GA will be a main forum for information transfer and debate on project progress and impact of the planned dissemination activities. The GA will decide on all major consortium and financial issues of the project. Each partner will have one vote and a simple majority will make all decisions. In case of a draw, the coordinator will hold the deciding vote. The assembly will hold four face-to-face meetings of only GA members. The first will be at the kick-off meeting, then again at months 18, 36 and 48, at different locations chosen in order to minimize travel expenses. Virtual meetings and decisions will be held as needed. As often as possible, the General Assembly will also meet in connection with other project meetings and/or international conferences, both to enhance the synergies to other partners and stakeholders, but also for the reduced expenditure. GA meetings will be chaired by the Coordinator, Jens Dolin. The minutes from the General Assembly meetings will be distributed to all partners and to the Commission.

2.1.3 Management Board

The Management Board will consist of the Work Packages leaders, i.e. seven people, and will be chaired by the project manager, Carl Winsløw. They will meet twice a year, and half of these meetings are expected to be virtual meetings using Skype or Adobe Connect, in order to reduce travel costs.

The daily on-site management of the project will be handled by the work package leaders, who will be **responsible for attaining milestones and submitting deliverables** to the consortium on time. They will also ensure liaison between the project management and the work packages as well as facilitating discussions and collaboration within and between the work packages. The work package leaders and their expertise are listed in Table 2.3a on page **Fejl! Bogmærke er ikke defineret.**.

2.1.4 Scientific & Policy Advisory Board

The consortium has created a Scientific & Policy Advisory Board. Each of the experts has a **strong international expertise relevant to the Call** we will utilize in the project. We expect this board to consist of five members covering the field, and so far three named experts have been invited and accepted to participate. We will invite two high-level policy delegates on a European level.

Name	Expertise	Focus Area in ASSIST-ME
Wynne Harlen (Professor of Science Education)	Formerly Director of the Scottish Council for Research in Education and Chair of the science expert group for the OECD's PISA project. Chaired the International Oversight Committee of the IAO (InterAcademies Panel) Science Education Programme on the development of inquiry-based science education in pre-secondary schools. President of the British Educational Research Association in 1993/4. Editor of Primary Science Review, for the ASE from 1999 to 2004. President of the ASE in 2008. Member of Scientific Committee of the Fibonacci project.	Science and technology educations and assessment
Michèle Artigue (Emeritus Professor at the University Paris Diderot, France)	President of ICMI from 01/012007 to 31/12/2009, ex-officio member of the ICMI Executive Committee as past-President since 01/01/2010, and currently in charge of relationships with UNESCO both for ICMI and IMU (International Mathematical Union) Involved in the STREP ReMath (Project Number: IST4-26751) in charge of WP1, and scientific advisor for the European projects Fibonacci and PRIMAS focusing on Inquiry-based Learning in Mathematics and Science of the European Commission PRD7 since 2010.	Mathematics education and assessment
Doris Jorde (Professor in Science Education)	Member of the European Commission's High Level Group on Science Education (The Rocard Report), (European Commission, 2007). Director of the Norwegian Centre for Science Education, former President of ESERA. Leader of the FP7 projects Mind the Gap and S-TEAM, contributor to Science Education in Europe. Critical Reflections (Osborne and Dillon,	Science in Society, linking STM educational research with policy making

	2008).	

The role of the Scientific and Policy Advisory Board is to ensure the optimal direction of the project both with regards to the scientific scope of the project and the impact the project will have on both national and European policy development of the educational system and assessment of competences across Europe. The board will participate in face-to-face and virtual meetings with the management according to needs and at least to the four General Assembly meetings. They will receive the recent status, progress reports and any relevant deliverables prior to the project meetings. They will then have time to read them and will be expected to provide advice and counsel on project decisions regarding the direction the project is taking, but to also provide practical advice from their experiences in the field. UCPH will take care of the contact to the Scientific and Policy Advisory Board.

2.1.5 Management and communication procedures

Communication regarding development, dissemination and management issues will be carried out using various forms of communication throughout the project.

Day-to-day communication will be addressed through normal means of communication, and UCPH as the coordinator will seek to implement an "open door" policy where partners are welcome and encouraged to seek answers and guidance immediately, instead of waiting until the next planned meeting.

The coordinator will be in contact with all partners throughout the project, both to ensure that the project is progressing as planned but also to get feedback on progress. **Quarterly Virtual Conferences** will be held using Skype, phone, or Web conference tools. The meetings will be chaired by the coordinator, who will send out an agenda for each meeting no later than two weeks before the meeting. Partners will also be welcome to add points under the agenda point "Other business". Minutes and slides from the meetings will be made available to all partners and the commission via the public side of the ASSIST-ME Webpage. In quarters where an annual meeting is to be held, these will replace the quarterly conferences.

Semi-annual Work Package Reports from all current work packages will be the basis of the planning and progress monitoring. Work package leaders will generate reports every six months that will commence at the start of a work package describing the plan the next six months, including staff assigned and tasks to be undertaken. These reports will be used in Management Board meetings and to inform the rest of the consortium of progress, obstacles and potential issues to be addressed. This will allow a large degree of transparency throughout the project. The reports will be submitted to the coordinator and the private project Wiki 14 days prior to the deliverable deadline, and will subsequently be distributed to all partners. Due to the desired reporting period length of 18 months, we will send semi-annual reports to the commission to allow greater transparency and progress monitoring. These reports will also be used as the foundation for the periodic reports.

Table 2.1a: Overview of reporting procedures

Activity	Responsibility	Timing
Annual Report to Commission (Activity report, Management report, Total cost statement (Form C)	Coordinator	Months 12, 24, 36 and 48
Reporting to Coordinator (WP report)	Work package leaders	Semi-annual
Progress report to WP leaders (Progress reports)	Partners	Quarterly

To ensure constant and clear reporting, the reports will be made using a template containing all relevant information, such as pending deliverables and deadlines; these will allow challenges to be identified in order to be resolved in a timely manner. The work package reports will be amended according to feedback from the conference and submitted to the Commission.

In case of delays or inadequate quality in reporting and deliverables, the WP-leader will be responsible for contacting the responsible partner within one week of the agreed submission date. A new realistic date for submission is then agreed upon and any necessary corrections of the reports are made. The WP-leader will inform the Coordinator of the delay and of the new submission date. If the Coordinator cannot approve a report or deliverable, these are returned to the responsible WP-leader and/or partner within one week of the decision with precise directions for the needed corrections. A new realistic deadline for submission is agreed upon.

If there are further delays or if the submission cannot be completed in a satisfactory manner, the WP-leader and the Coordinator will organize an (Internet) meeting with the partner(s) responsible for the submission. If it is agreed that the submission will only be possible with considerable delays or not possible at all, a contingency plan will be drawn up for how to resolve the issue – either by setting a new realistic timetable, designing new activities that can solve the problem (to the extent the budget allows) or deciding to cancel the activity. In any case, the decision will be submitted to the EU-Commission for approval as this may affect the overall plan of the project. Moreover, the Coordinator will together with WP-leaders ensure a high level of data sharing in the project. In addition to all of the communication and contact with stakeholder expert panels in eight countries, project results will be made available externally mainly through the SCIENTIX platform, but will also be accessible on the project Website.

2.1.6 Timing of meetings

The partners will primarily share information by the use of e-mail, phone meetings and the Internet or during regional meetings in the networks or dissemination regions, see Table 2.1b. The venues for the General Assembly will be held in locations where the project meetings can be supported by other international conferences or other FP7 meetings where more of the project partners are present. This will help reduce travel costs and to ensure and enhance the synergies between the various projects.

3. Impact

It is an overall goal of the Science in Society Work Programme to make research and innovation more attractive for developing careers, and enable citizens and to be better informed, to better understand and to participate more comprehensively and efficiently in the research and innovation processes. In order to reach this goal, citizens need to build – starting from early age – competences of scientific reasoning, as well as transversal competences such as problem solving, creativity, critical thinking, teamwork and communication skills.

A teaching methodology such as the **inquiry-based education can strongly contribute to the development of these competences**. But not without the large scale uptake of effective assessment methods that can be used for both formative and summative purposes.

ASSIST-ME will **provide policy makers with evidence based guidelines for changing educational policies** in directions that promotes excellence in education and skills development within the field of Science, Technology and Mathematics. This way ASSIST-ME will **support the "Innovation Union" Flagship Initiative** under Europe 2020 and be a part of the effort on making STM careers more attractive for young people.

Education researchers primarily report their results to the education community, which is most often isolated from the political processes, not because researchers have little knowledge about their national policy matters but to secure their academic independence. This dissociation between research and policy goes back to the building of the independent universities and their insistence on free research.

One of the consequences of this has been the development of different logics and discourses within these separated areas of activity – with the result that educational research sometimes has little influence on policy. To change this, it is **necessary to understand how policy is made and how it is implemented**. It is also important to know the discourse of policy, to speak the language and communicate results in a way politicians can understand and use. As formulated by Jonathan Osborne (2011b) "… for science educators to have influence on science education policy, they must step out of their scholarly worlds and engage directly with the practical world of policy makers."

Although this project is not aimed at changing educational practice directly, the same kind of communication challenges can be found between researchers and teachers. Other communication challenges and even disagreements exist between practitioners and policy makers. It is fair to say that **research form a triangle with policy and practice** and that research has the potential to influence both, giving it can communicate its results in a discourse they can understand (see Figure 5.

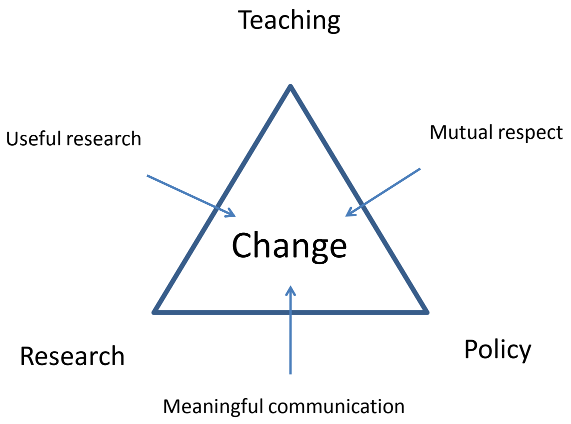

Figure 5. Research form a triangle with policy and practice in order to make change happen.

Fensham (2009) puts forward this role of research for affecting policy to change practice:

1. Establish research-based cases of the inadequacies of current policy.
2. Argue for change.
3. Propose a changed policy and offer research-based cases for its positive prospects.
4. Formulate the conditions needed for implementation.

ASSIST-ME will follow this course in providing policy makers with research evidence and strategies to encourage successful inquiry-based instruction through supporting assessment methods.

3.1 Expected impacts listed in the work programme

3.1.1 An understanding of how assessment strategies influences STM teaching

The research literature science, mathematics and technology education is replete with the use of formative assessment instruments that have been found to be variously useful in providing teaching and learning feedback. As well, recent FP7 inquiry development projects such as Mind The Gap, Comenius, S-TEAM, INSPIRE and Fibonacci, have also provided useful formative assessment strategies as part of their inquiry materials. In addition, various reviews of research have attempted summaries of viable assessment methods in support of inquiry learning. However, no meta-analysis of the research literature in mathematics, science and technology education to collect patterns of formative feedback which are efficacious has been done.

ASSIST-ME will prepare such a **synthesis of the formative assessment methods which support aspects of inquiry teaching and learning**, based on research literatures in science, mathematics and technology education (WP2). In addition to international journal reports and FP7 project results, ASSIST-ME will review the relevant literature in at least the seven European languages represented by the project partner countries (Czech, Danish, English, Finnish, French, German, and Greek) in order to include these findings in the overall review and add educational system variance to the resulting perspectives. In addition, the project will provide a report on the current state of the art in formative and summative assessment in IBE in STM, including the use of ICT based tools, for use in the other work packages.

Summative assessment is a necessary part of STM assessment and has as important role in education. In theory it can provide explicit operational examples of the meaning of learning goals and provide some motivation for the effort required to learn (Kellaghan et al., 1996). However it is the mismatch between what is assessed by tests and examinations and the important goals of education – such as currently found in the assessment of inquiry-based STM – that distorts the curriculum when test items are taken as indicators of what to learn. Crooks (1988) concluded from a review of research into the impact of assessment practices on students that it was vital that assessment must emphasise the skills, knowledge and attitudes perceived to be most important however difficult the technical problems with designing such an assessment method. It follows that it is essential to create and evaluate methods of summative assessment that are capable of providing information about the competences developed through inquiry-based teaching and learning.

ASSIST-ME will rectify mismatches between educational goals and assessment practices by discovering (WP2), designing (WP4) and testing (WP5) methods that address the special process needs of inquiry learning. The project will examine and **promote ways in which existing summative methods can be used to enhance overall assessment**. Where current summative methods interfere with inquiry learning by failing to reinforce and reward skills and content special to inquiry lessons, strategies for modifying those methods will be devised. These new insights will be spread through the many ASSIST-ME promotion and dissemination activities and inspire educational development projects implementing new assessment strategies.

3.1.2 Support the uptake of IBE in different European contexts

One of the outcomes of the FP7 projects such as Mind The Gap, Fibonacci, PRIMUS, S-TEAM has been a renewed appreciation for how the varied European educational systems require educational change to be uniquely adapted for each context. Maps of scientific literacy statements for seven European countries from Mind The Gap and S-TEAM showed large differences in both content and process objectives from country to country (Bruun, J., Dolin, J. and Evans, R. H., 2009).

ASSIST-ME will **synthesize educational system variables throughout Europe with relevance to formative or summative assessment and IBE in STM** and map out the participating countries with respect to these variables so it is clear from early on which system characteristics are represented in our project countries, i.e. Cyprus, Czech Republic, Denmark, Finland, France, Germany, Poland, Switzerland, and the United Kingdom (WP3). Since we have partially chosen partners based on the diversity of their educational systems using analyses such as Eurydice (2011) and OECD databases, our own systems represent a large measure of the European educational system variance. We will also use existing surveys on uptake of IBE in STM from EU projects such as Mind The Gap, Fibonacci, PRIMUS, S-TEAM to validate relevant educational system variables (WP3). Our product of this research will be **maps of European countries with respect to factors facilitating or obstructing the uptake of IBE and formative assessment strategies** and to what extent these factors are represented in our project partner sample. This will provide a guiding matrix useful in adjusting assessment methods to different systems as well as a generalizability perspective for assessment method deliverables. Such a matrix will be a strong tool for educational developers that will implement new assessment methods in their own context.

3.1.3 Practice-near research resulting in high usability

European projects such as Mind The Gap, Fibonacci, PRIMUS, S-TEAM have developed a plethora of pathways to increased use of inquiry-based education in science, mathematics and technology classrooms but have not been as successful at associating these strategies with assessment methods which are able to capture the effects of inquiry-based teaching. Consequently, traditional summative assessment methods which have difficulty identifying the unique process skills from inquiry have reduced teacher and student motivation to use inquiry. Being able to adequately assess the products of inquiry teaching methods will allow educational system personnel to verify that inquiry methods have a measurable impact on learning.

The main emphasis of the tests and trials of assessment methods (WP5) will be on **engaging teachers in a process of testing and proofing formative assessment tasks in their daily practice** in authentic classroom environments, and with an additional focus on use of ICT. It will also be on how to abstract a set of guidelines for teachers, teacher educators and school administrators on how to adopt coherent combinations of formative and summative assessment strategies.

In order to obtain a strong alignment with current teacher practices Work Package 4 will organize **Teacher Expert Panels** for each of the eight partner countries. These panels will on the basis of the research into educational system differences by Work Package 3, be integrated in the work on assessment methods to validate the methods produced and prepare for the testing of the methods during WP5, in which they will be participants. The teacher expert panels will in WP5 evolve into the **Local Working Groups** that will test the assessment methods, both ICT based and non-ICT based. The already participating teachers will be supplemented with a sufficient amount of teachers in order to cover primary and secondary levels and different subjects. This grounding of the research into real teaching practices to secure the relevance and adaptability of the research findings of WP5 – meeting both teachers' and politicians' interests and thus facilitates high impact.

3.1.4 Supporting a learning oriented STM teaching

At the school and classroom levels, assessment efforts often tend to focus on grading students, informing them about errors committed and determining their future pathways. This is more acute in secondary education where the emphasis on summative assessment tends to be more pronounced. Assessment information is often coded based on memorization of facts, laws and principles or on solving quantitative exercises. It is widely interpreted as an indicator of student intellectual capability and future potential, a predictor of performance in the higher stakes national assessments. In a situation where educational change is desirable this practice becomes problematic. It represents a continuous missed opportunity because summative assessment does not commonly yield valuable information about how to refine the change or learning processes and make them more effective. The feedback needed for both students and teachers in order to act as self-regulating intellectual members of a teaching-learning community does not often come from summative assessment tasks.

ASSIST-ME will seek to investigate this problem and develop tools and mechanisms for addressing it (WP2, 4 and 5). Where traditional summative methods are continued to be used for IBE, students realize that many of the things they are learning, particularly the process goals of IBE, are never assessed: this mismatch reduces the importance of inquiry. Teachers are also discouraged since their adoption of contemporary inquiry methods are not only non-rewarding by student performance on traditional end-of-course tests, but student performance may actually be seen as declining when IBE is used.

Specifically, we intend to map out teacher and school priorities in existing assessment practices in Europe taking care to identify those schools and teachers that have engaged with efforts to implement inquiry-oriented teaching and learning in STM. We expect that such teachers will have already adapted formative assessment methods that guide students during unfamiliar processes and provide them with the feedback necessary to maintain their self-efficacies and success. We will build on those methods using our project research and development and test them for dissemination when adapted to various educational systems.

We will also **develop alignment templates for engaging schools and teachers** in a process of integrating formative assessment into existing classroom practices in ways that encourage both evidence-based feedback to the teaching and learning process and closer alignment between classroom practice and the main features of inquiry-oriented education, including authenticity, epistemic anchoring and emergent autonomy (WP5).

3.1.5 Validated methods designed to overcome negative effects of "teaching to the test"

There is extensive evidence that what is taught is influenced, both in content and pedagogy, by what is assessed and how it is assessed. Some of the negative effects of testing are evident in the review of research by Harlen and Deakin Crick (2003). They cite, for instance, the tendency for teachers to teach to the test, particularly when test results have high stakes for the students or the teacher. This means that the scope and depth of learning are seriously undermined. Even when not directly teaching to the test, it was found that teachers change their approach to one that is perceived as necessary to meet the requirements of the test. Johnston and McClune (2000) reported teachers spending most of the time in science in direct instruction and less in providing opportunities for students to learn through inquiry, even though they preferred the latter.

ASSIST-ME will intentionally discover (WP2), reformulate (WP4), test (WP5) and promote (WP6 and 7) formative assessment methods which because of their position in a course (during the learning rather than at its conclusion) and role in student feedback will naturally force them to be in closer alignment with educational goals. Simultaneously, ASSIST-ME will work to **align project tested formative assessments with existing summative forms** so that they work together for teaching and learning and make full use of ICT potentials. We will provide a "strategies for alignment" template which teachers can use to match their educational goals and teaching content with formative and summative assessment practices (WP5). Such concordances will actually allow teachers to "teach to the test" since the good alignment will enhance both teacher and student success.

3.1.6 High impact through clear communication

Educational policy varies from country to country depending on whether they have highly centralized educational systems, are more decentralized or according to other characteristics as will be discovered by Work Package 3 of this project. But because policy development involves the same issues such as knowledge, values, and stakeholder interests, there are also similarities in how educational policy is enacted and implemented across nations. Especially in Europe the shared participation in the Enlightenment period has given a common commitment to empirical evidence and rational argumentation. These beliefs in the value of evidence give research a relatively high status in Europe and provide educational researchers with the opportunity to influence policy and practice throughout all of Europe, despite the national differences (Osborne, 2011b):

> "For the researcher, empirical evidence remains a tool that can be used to challenge deeply held beliefs. Although the practice of science might be secured by its commitment to evidence, the practice of science education often falls well short of such a standard of rationality. Where evidence is thin or questionable, value will predominate. And even where evidence does exist, both policy makers and practitioners will be selective in choosing the evidence they will attend to. Communicating and influencing practitioners and policy makers therefore requires an understanding of their values and how deeply they are held" (Osborne, 2011b).

ASSIST-ME will **secure such a communication by cooperate directly during the project with policy makers** and other stakeholders. Reports for policy makers will address them in a language that is clear and convincing. Among our models for good communication will be two research based reports which have had a strong influence on educational policy: *Beyond 2000: Science Education for the Future* (Millar & Osborne, 1998) and *Science Education in Europe: Critical Reflections* (Osborne & Dillon, 2008). They are both characterized by reporting the findings in expressive and powerful quotations and by focusing on the implications of the study through a set of recommendations. Four ASSIST-ME researchers have contributed to the last report, and the ASSIST-ME reports from WP6 and WP7 will be worked out and communicated according to the principles behind these two reports.

3.1.7 Influential networks will secure change

With research evidence for a variety of educational systems we will work together with stakeholders in the **National Stakeholder Panels** (WP6) on how to formulate recommendations for change in educational policy. These panels will be made up of professional education representatives, policy experts and media experts. Policy experts included on each of our eight panels will act as representatives of the main target group of this project. The media experts will help transform research findings into more readable and publically understandable summaries.

Finding the relevant stakeholders will be a crucial part of the project. We will apply a method developed at UCPH, based on network theory and social network analysis. Applying this method to the participating countries will help finding the relevant stakeholders in each country based on the same principles.

The involvement of these panels is starting in the beginning of the project because all of these stakeholder groups have relevant perspectives to contribute to the on-going research. The panels will meet three times during the project to learn about, validate and begin to disseminate the accumulating findings of the on-going project. This inclusiveness will **assure national teacher and stakeholder involvement throughout the project** and serve to align the research processes with the understanding and perspectives of the panels for enhancing the feasibility of the results. The meetings of each group will secure meaningful communication between researchers, practitioners, and policy makers and serve to refine the project's assessment methods and procedures based on feedback.

The strong connections between researchers and policy makers will give a shared ownership to the research findings and secure a higher impact of the research on practice than normal.

3.2 Dissemination and exploitation of project results, and management of intellectual property

ASSIST-ME recognises that the rollout and dissemination of a project across Europe is a challenging task, especially in light of the social and cultural diversity of the schools and social settings that exists. ASSIST-ME welcomes this challenge and will work collaboratively as a group to encourage and support partners in developing and adapting materials to suit national requirements to ensure largest possible impact across the cultural boundaries of Europe.

All dissemination activities of the ASSIST-ME project will be led and managed by KCL in Work Package 7. These activities will compliment and support each other and will allow project results to be readily accessible by the stakeholders of the project.

The main deliverables that will be prepared as part of the dissemination will be:

- To develop practical guidelines for implementing and using the assessment methods in different educational contexts, using expert panels and forums from different educational systems.
- To develop recommendations for policy makers and other stakeholders to improve decisions on curricula design, teacher training and alignment with traditional summative assessment indicators. Communicate these with stakeholders in partner countries and representatives from every EU member state through dissemination activities to facilitate the adoption of the competence based assessment and formative assessment by diverse European educational systems.

3.2.1 Key Stakeholders

The project has identified the following as key stakeholders of the project that are crucial to the further exploitation of the project results:

- **Government and municipalities (central as well as local levels):** It is important that governments drive policy development in the area of innovative science engagement strategies. Dissemination will be directed towards government from within the project, as well as through hosting targeted seminars with both government departments and policy makers present. Depending on the educational system and culture municipalities and local authorities can be crucial in facilitation cultural change in STM education. Therefore ASSIST-ME will involve and communicate with policy makers on local levels.
- **Media:** Journalist in leading printed media and broadcasting corporations has substantial influence on policy and decision makers, especially with respect to agenda setting. ASSIST-ME will communicate during the project with media representatives in all participating countries in order to facilitate the public discussions of how assessment strategies influence outcomes of teaching. This will support policy makers in working with cultural change in STM education.
- **Business and industry:** With the key role STM education plays in the economic development of all EU countries, organizations representing business interests will have a strong incentive to be involved in the project.
- **Teachers:** It is vitally important that teachers are both personally and professionally involved in this project as they are our agents for implementing real changes in the classroom. Dissemination towards teachers will emphasise the value of assessment of key skills and competencies associated with an inquiry-based approach across the STM subjects and its contribution to teaching and learning through inquiry, as well as the promotion of how to integrate the outputs from this project's work into existing national curricula, so that the day-to-day teaching of the prescribed curriculum is incorporated.
- **Teacher Trainers and Professional Development Providers:** For the future development of project ideas and findings, it is essential that both teacher train-

ers and professional development providers are conversant and fully aware of the project objectives and findings so that they can adopt and adapt those aspects that will encourage more widespread uptake of inquiry-based methods across the STM subjects in schools. This will be ensured through the established links that partners already have with these stakeholders, plus a strong communication link through the ASSIST-ME newsletter.

- **Research Communities of STM Education:** The new pedagogical methodologies and the results achieved with those methodologies within the context of the ASSIST-ME project needs to influence future areas of study in this area. The project has strategic goals to contribute to the open-access policy of publishing the reports of the project as well as participating in peer-reviewed conferences and publications. The Scientific and Policy Advisory Board of ASSIST-ME will include respected and renowned pedagogical experts, which will further strengthen our ability to disseminate the pedagogical results of the project to this community. The results of the project will be disseminated by publishing in conference and journal papers, with particular attention being paid in getting published in the conference proceedings and journals outlined in Table 3.2a on page 43. Through dissemination in specialised journals and congresses, ASSIST-ME expects to highlight the impact of its contribution among this scientific community.

The concrete composition of stakeholders will depend on the political, historical and cultural conditions in each country. We will apply the network analysis method to the participating countries to find the relevant stakeholders in each country based on the same principles.

3.2.2 Dissemination Activities

All ASSIST-ME partners are committed to the development, dissemination and use of project results and will be actively involved in maximising the impact of ASSIST-ME, both nationally and internationally. An adaptive framework for dissemination plan will be prepared, as part of WP7. This will design a general roadmap for dissemination and awareness activities to take place as part of the project activities. The partners will then adapt this general roadmap for individual national markets.

Specific dissemination activities have been presented in the descriptor for WP7, but in general the dissemination activities of ASSIST-ME will include the following:

- Organising an international conference where key invited stakeholders will have opportunity to explore, critique and evaluate project findings.
- Organising national conferences and workshops to disseminate project findings.
- Project specific workshops and special sessions in major European conferences.
- Publications at commercial and scientific conferences and exhibitions.
- Publication of reports to generate awareness within national strategic networks. In addition to this, the progress of the ASSIST-ME project will be disseminated within the Science in Society projects of the European Commission.

- Communication through the regular ASSIST-ME newsletter to facilitate the creation of content appropriate for dissemination, while it would also motivate partners to communicate their research results and work. ASSIST-ME will produce a quarterly newsletter, in a scheme that would involve all of the partners during the entire project lifetime.
- Communication through project liaisons by undertaking collaboration activities with other EU projects and initiatives outside ASSIST-ME, in order to get inputs and provide outputs that could enrich all parties. This will be done partly with the use of SCIENTIX – the community for science education in Europe – that facilitates regular dissemination and sharing of know-how and best practices in science education across the European Union

The project has identified a list (non-exhaustive) of relevant international journals, conferences and symposia, to which contributions will be made during the duration of the project, see Table 3.2a.

Table 3.2a International Journals and Conferences to be subject for dissemination

International Journals: • Science Education International (ICASE) • International Journal of Science Education • Research in Science Education • International Journal of Science and Mathematics Education • Assessment in Education: Principles, Policy and Practice • Educational Research and Evaluation • School Science Review
International Conferences: • IOSTE (International Organization for Science and Technology Education) conference, 2014, 2016. Up to 200 science and technology teachers, curriculum developers. • ICASE (International Council of Associations for Science Education) conference, 2014, 2016. • ICCE (International Conference on Computers in Education) 2013, 2014, 2015, 2016. • ESERA (European Science Education Research Association) conferences 2013, 2015, 2017. Up to 1000 science teachers, researchers, doctoral students. • GIREP (Groupe International de Recherche sur l'Enseignement de la Physique) seminars and conferences, 2012, 2013, 2014, 2015, 2016. • NARST (National Association for Research in Science Teaching) conference 2013, 2014, 2015, 2016. • ECER (European Conference on Educational Research) conference 2013, 2014, 2015, 2016.

3.2.3 Handling of intellectual property

All written deliverables – including assessment methods, guidelines and recommendations – will be public in their final state and can be exploited by everybody.

We have to protect the existing intellectual property rights of our commercial partner, Pearson. This implies that the specific programming done by Pearson in already existing software used for ICT-based tools for assessment belongs to Pearson. The consortium will secure that sufficient descriptions of the ICT-based tools will be published in public reports so that others programmers of ICT-based assessment tools can produce similar ICT-tools with the use of other software.

This way the results of ASSIST-ME are not subject to restrictions regarding the management of knowledge.

4. Ethics Issues

ASSIST-ME research will involve school children and their teachers and include both virtual and real-time observations.

Children and their teachers are involved in our research since we will conduct three rounds of trials of the assessment methods in classrooms in all partner countries. Administrators, teachers and pupils all need to be aware of any risks associated with participation. We will gather written informed consent from all the teachers and school administration for each participating school with consent statements that describe exactly the kinds of activities in which the teachers will participate and assure administrators and teachers that all data will be confidentially held, will remain anonymous and only reported in aggregate with all identifiable factors removed. In addition, specific results of classroom trials will not be reported to nor available to school administrators from the project. Only aggregate project results with school and personal identifiers removed will be publicly shared. Since the assessment methods used in the trials will be within the range of normal teacher classroom methods, we will not gather informed consent from the pupils for participating in these classroom activities.

Since in some classes we will make video-recordings of the trials, we will for those classes get written informed consent from both the teachers and the parents of the pupils. Teachers will be assured of protection from any administrative reactions to the success or failure of trials in their classrooms since we will not share identifiable data with anyone. In cases where we want to collect video data solely for research purposes (and not for dissemination) we will request parental permission to do that. Where we may want to share video records more broadly, we will assure parents of anonymity for their children and only use video images of their children with full parental consent and understanding of how the material will be shared. IThe informed consent forms will follow EU and national regulations.

The management of the consortium will secure that all contact with schools, teachers and children is in accordance with the regulations of the involved countries, as well as EU regulations.

5. Consideration of gender aspects

ASSIST-ME will work explicitly to make assessment methods in STM support the goal of achieving gender equity in STM education. The innovation to the classroom practices through inquiry-based education techniques and combinations of formative and summative assessment methods will be an important steps toward enhanced learning for all pupils. By improving formative and summative assessment methods and consequently increasing the quality of student and teacher feedback, the mutual success with inquiry-based science will be of equal benefit to all.

There is a broad spectrum of clear research results indicating inadequacies in STM education in regard to girls and young women. This concerns specifically physics, technology and chemistry education, less so for mathematics and least for biology education, where the issues may be one of attracting more boys. Girls, particularly from the start of puberty and onwards, show less interest in physics, chemistry and technology, less self-efficacy and in many cases also less knowledge than boys (Schreiner and Sjøberg, 2004).

The project ASSIST-ME can contribute to addressing the gender issues in physics, technology, chemistry and mathematics education. When choosing the assessment methods and when developing paradigmatic examples, gender ‘friendliness’ will be an important criterion. This and the development of formative assessment itself will have a positive impact on the interest of girls and young women in STM, their self-efficacy and their knowledge.

The consortium behind ASSIST-ME consists of an almost equal number of men and women. Of the ten persons leading each partner group six are male and four are female. We will also seek to ensure gender balance in the various local working groups that assist in the development of the project’s outcomes.

6. References

AAAS (American Association for the Advancement of Science) (1992): Project 2061 *Benchmarks for Science Literacy.* New York: Oxford University Press.

AAAS (American Association for the Advancement of Science) (2000): Project 2061 *Designs for Science Literacy*. New York: Oxford University Press.

AAAS (American Association for the Advancement of Science) (2001): *Atlas of Science Literacy.* Project 2016. Washington DC: AAAS and NSTA.

Adams, D. & Hamm, M. (1998): *Collaborative Inquiry in Science, Math, and Technology*. Portsmouth: Heinemann.

Adey, P. (1997): It all depends on context, doesn't it? Searching for general, educable dragons. *Studies in Science Education,* 29, 45-91.

Anderson, R. (2002): Reforming Science Teaching: What Research says about Inquiry. *Journal of Science Teacher Education*, 13(1): 1-12.

ARG (Assessment Reform Group) (2002): *Assessment for Learning: 10 Principles.* Available on http://www.assessment-reform-group.org.

Artigue, M. & Winsløw, C. (2010): International comparative studies on mathematics education: a viewpoint from the anthropological theory of didactics. *Recherches en Didactique des Mathématiques,* 30(1), 47-82.

Artigue, M. (1994): Didactical engineering as a framework for the conception of teaching products. In: R. Biehler et al. (Ed.), *Didactics of Mathematics as a scientific discipline* (27-39). Dordrecht: Kluwer Academic Publishers.

Benke, G., Hošpesová, A. & Tichá, M. (2008): The Use of Action Research in Teacher Education. In: K. Krainer, T. Wood (Eds.) *The International Handbook of Mathematics Teacher Education (volume 3), Participants in Mathematics Teacher Education: Individuals, Teams, Communities and Networks* (283-307). Rotterdam: Sense Publishers.

Bianco, M., Bressoux, P., Doyen, A.L., Lambert, E., Lima, L., Pellenq, C., & Zorman, M. (2010): Early training in oral comprehension and phonological skills: Results of a three-year longitudinal study. *Scientific Studies of Reading,* 14, 3, 211-246.

Birenbaum, M., Breuer, K., Cascallar, E., Dochy, F., Dori, Y., Ridgway, J. & Wiesemes, R. (no date): A *learning integrated assessment system.* EARLI position paper retrieved from http://www.earli.org/resources/1st%20EARLI%20Position%20Paper%20A%20Learning%20Integrated%20Assessment%20System.pdf

Black, P. & Wiliam, D. (1998): Assessment and classroom learning. *Assessment in Education: Principles Policy and Practice,* 5(1), 7-73.

Black, P. & Wiliam, D. (1998a): *Inside the Black Box*. London: NFER.

Black, P. & Wiliam, D. (1998b): Developing a theory of formative assessment. In: Gardener, J. (Ed.): *Assessment and Learning* (81-100). London: Sage.

Black, P. (1998): Assessment by Teachers and the Improvement of Students' Learning. In: Fraser & Tobin (Eds.), *International Handbook of Science Education*. Kluwer Academic Press.

Black, P., Harrison, C, Hodgen, J., Marshall, B., & Serret, N. (2011): Can teachers' summative assessments produce dependable results and also enhance classroom learning? *Assessment in Education: Principles, Policy & Practice*, 18(4), 451-469.

Black, P., Harrison, C., Hodgen, J., Marshall, B., & Serret, N. (2010): Validity in teachers' summative assessments. *Assessment in Education: Principles, Policy & Practice*, 17(2), 217–34.

Black, P., Harrison, C., Lee, C., Marshall, B. & Dylan, W. (2002): *Working inside the black box: assessment for learning in the classroom.* London, UK: King's College London Department of Education and Professional Studies.

Black, P., Harrison, C., Lee, C., Marshall, B. & William, D. (2003): *Assessment for Learning: Putting it into Practice*. Maidenhead: Open University Press.

Borasi, R., Fonzi, J., Smith, C. F. & Rose, B. J. (1999): Beginning the process of rethinking mathematics instruction: A professional development program. *Journal of Mathematics Teacher Education,* 2, 49-78.

Bransford, J. D., Brown, A. & Cocking, R. R. (Eds.) (1999): *How People Learn. Brain, Mind, Experience and School.* Washington, D.C.: National Academy Press.

Bressoux, P., Kramarz, F., & Prost, C. (2009): Teachers' training, class size and students' outcomes : learning from administrative forecasting mistakes. *Economic Journal*, 119, 540-561

Brookhart, S. & DeVoge, J. (1999): Testing a theory about the role of classroom assessments in pupil motivation and achievement, *Applied Measurement in Education,* 12, 409-425.

Bruun, J., Dolin, J. & Evans, R. H. (2009): *Minding the Gap Between Policy and Practice of Inquiry Based Science Teaching in Seven European Countries: Norway, Germany, Denmark, Spain, Framce, Hungary, and the United Kingdom*. University of Copenhagen.

Chudowsky, N. & Pellegrino, J.W. (2003): Large-scale assessments that support learning: What will it take? *Theory into Practice*, 42(1), 75-83.

Clesham, R. (2009): *Making it real: Interactive assessment of national curriculum science process skills.* Proceeding of BERA, Edinburgh.

Clesham, R. (2010): *Onscreen Assessment: the effects of mode on student performance and perceptions.* Proceeding of the CIEA, October 2011, London

Clesham, R. (2011): *Comparability of computer and paper-based assessments: how do they measure up?* Proceeding of European Science Education Research Association (ESERA) Lyon, France, September 2011.

Cobb, P., Wood, T., Yackel, E., & McNeal, B. (1992): Characteristics of classroom mathematics traditions: an interactional analysis. *American Educational Research Journal, 29*(3), 573-604.

Coppé, S. (2011): Travail collaboratif d'enseignants de mathématiques pour la production et la diffusion de ressources pour les professeurs et les formateurs. Actes du colloque de l'AREF, Genève, septembre 2010. Actes électroniques https://plone2.unige.ch/aref2010.

Crooks, T. J.(1988): The impact of classroom evaluation practices on students, *Review of Educational Research*, 58, 438-81.

DfEE (Department for Education and Employment (1999): *Science. The National Curriculum for England.* London: HMSO

Dimitrova, M. (2006): Evaluating the Flexibility of Learning Processes in E-Learning Environments. Book chapter in *Flexible Learning in an Information Society, Khan B. H., 2006*

Dimitrova, M., Mimirinis, M., Murphy, A. (2004): *Evaluating the Flexibility of a Pedagogical Framework for e-Learning*. In Proceedings of ICALT 2004.

Dolin, J. (2007): PISA - An Example of the Use and Misuse of Large-Scale Comparative Tests. In S. T. Hopman, G. Brinek & M. Retzl (Eds.), *PISA zufolge PISA - PISA according to PISA* (93-125). Wien, Berlin: LIT Verlag.

Dolin, J. (2007a): Science education standards and science assessment in Denmark. In D. Waddington, P. Nentwig & S. Schanze (Eds.), *Standards in Science Education* (71-82). Waxmann Verlag.

Dolin, J., Evans, R. & Bruun, J. (2010): Report on understanding of IBST and Scientific Literacy, In University of Oslo, *Mind The Gap*. Seventh Framework Programme, Commission of the European Communities.

Dolin, J., & Krogh, L. B. (2010): The Relevance and Consequences of Pisa Science in a Danish Context. *International Journal of Science and Mathematics Education,* 8, 565-592.

Durand-Guerrier, V. & Winsløw, C. (2007): Education of lower secondary mathematics teachers in Denmark and France: A comparative study of characteristics of the systems and their products. *Nordic Studies in Mathematics education*, 12(2), 5-32.

European Commission (2007): Science Education Now: A Renewed Pedagogy for the Future of Europe. Brussels: European Commission.

European Commission (2009): *Key competences for a changing world. Draft 2010 joint progress report of the Council and the Commission on the implementation of the "Education & Training 2010 work programme"*. Retrieved from http://ec.europa.eu/education/lifelong-learning-policy/doc1532_en.htm.

Eurydice (2011): *Science Education in Europe: National Policies, Practices and Research.* Education, Audiovisual and Culture Executive Agency, European Commission. (http://eacea.ec.europa.eu/education/eurydice).

Fennema, E., Carpenter, T., Franke, M., Levi, L., Jacobs, V. & Empson, S. (1996): A longitudinal study of learning to use children's thinking in mathematics instruction. *Journal for Research in Mathematics Education,* 27(4), 403–434.

Fensham, P. (2009): The Link Between Policy and Practice in Science Education: The Role of Research. Wiley InterScience.

Francisco, J. & Hähkiöniemi, M. (2011): Students' ways of reasoning about nonlinear functions in Guess-My-Rule games. *International Journal of Science and Mathematics Education*, Online First.

Francisco, J. & Maher, C. (2005): Conditions for promoting reasoning in problem solving: Insights from a longitudinal study. *Journal of Mathematical Behavior,* 24(3–4), 361–372.

Grangeat, M. & Gray, P. (2007): Factors influencing teachers' professional competence development. *Journal of Vocational Education and Training,* 59(4), 485-501.

Grangeat, M. & Gray, P. (2008): Teaching as a collective activity: analysis, current research and implications for teacher education. *Journal of Education for Teaching*, 34(3), 177-189.

Grangeat, M. (2008): *Teachers' knowledge: a synthesis between personal goals, collective culture and conceptual knowledge.* Paper presented at the European Conference on Educational Research (ECER), Gothenburg, Sweden. London*. Education-line*.

Grangeat, M. (Ed.). (2011): *Les démarches d'investigation dans l'enseignement scientifique Pratiques de classe, travail collectif enseignant, acquisitions des élèves*. Lyon : Ecole Normale Supérieure.

Harlen, W. & Deakin Crick, R. E. (2003): Testing and motivation for learning. *Assessment in Education,* 10(2), 169-207. http://sohs.pbs.uam.es/webjesus/motiv_ev_autorr/lects%20extranjeras/efecto%20ev.pdf

Harlen, W. & James, M. (1997): Assessment and learning: differences and relationships between formative and summative assessment. *Assessment in Education,* 4(3), 365-379.

Harlen, W. (2006): *Teaching, Learning and Assessing Science 5-12.* 4th edition. London: Sage.

Harlen, W. (2007): The Quality of Learning: Assessment alternatives for primary education (Primary Review Research Survey 3/4). Cambridge: University of Cambridge Faculty of Education.

Harlen, W. (2009): Teaching and learning science for a better future. *School Science Review*, 90(933), 33-41.

Harlen, W. (2009a): *IAP – International Conference: Taking Inquiry - Based Science Education (IBSE) Into Secondary Education.* York: National Science Learning Centre.

Harrison, C. & Howard, S. (2009): *Inside the Primary Black Box: Assessment for Learning in primary and early years classrooms*. London: GLAssessent.

Harrison, C. (2005): Teachers developing assessment for learning: mapping teacher change. *International Journal of Teacher Development*, 9(2), 255-263.

Harrison, C. (2011): Classroom assessment. In: Dillon, J. and Maguire, M.: *Becoming a Teacher: Issues in Secondary Teaching* (fourth edition). London: OUP.

Harrison, C., Hofstein, A., Eylon, B.S. & Simon, S. (2008): Evidence-Based Professional Development of Science Teachers in Two Countries. *International Journal of Science Education,* 30(5), 577-591.

Hartig, J., Klieme, E., & Leutner, D. (2008): *Assessment of competences in educational contexts*. Göttingen: Hogrefe.

Heinze, A., Ufer, S., Rach, S. & Reiss, K. (2011): The Student Perspective on Dealing with Errors in Mathematics Class. In: Wuttke, E. & Seifried, J. (Eds.): *Learning from errors at School and Work* (Research in Vocational Education), (65-79). Opladen: Barbara Budrich.

Hošpesová, A. & Tichá, M. (2010): Reflexion der Aufgabenbildung als Weg zu Erhöhung der Lehrerprofesionalität. In Böttinger, C., Bräuning, K., Nührenbörger, M., Schwarzkopf, R., Söbbeke, E. (Eds.): *Mathematik im Denken der Kinder; Anregung zur mathematik-didaktischen Reflexion,* (122-126). Seelze: Klett/Kalmeyer.

IAP (Interacademy panel of international issues) (2011): *Taking IBSE into Secondary Education. Report of a conference held in York, UK, October 2010.* http://www.fasas.org.au/downloads/YorkIBSEConference2010Report.pdf

James, M. (2000): Measured lives: the rise of assessment as the engine of change in English schools, *The Curriculum Journal*, Vol 11 No 3 343-364.

Joët, G., Usher, E.L., & Bressoux, P. (2011): Sources of Self-Efficacy: An Investigation of Elementary School Students in France. *Journal of Educational Psychology*. *Advance online publicatio*n June 20.

Johnston, J. & McClune, W. (2000): Pupil motivation and attitudes – self-esteem, locus of control, learning dispositions and the impact of selection on teaching and learning, in *The Effects of the Selective System of Secondary Education in Northern Ireland:* Research Papers Vol II. Bangor, County Down: Department of Education (1-37).

Kellaghan, T., Madaus, G. F. & Raczek, A. (1996): *The Use of External Examinations to Improve Student Motivation*. Washington DC: American Educational Research Association.

Knoke, D. (2011): Policy Networks. In: Scott, J. and Carrington, P. J. (Eds.): *The SAGE Handbook of Social Network Analysis*. London: Sage Publications Ltd.

Köller, O. & Parchmann, I. (2012): Competencies: The German Notion of Learning Outcomes. In: Bernholt, S., Neumann, K., & Nentwig, P. (Eds.): *Making It Tangible - Learning Outcomes in Science Education,* (165-185). Münster: Waxmann.

Köller, O. (2011): Improving pupils' mathematics and science skills in the 21st century. Public Service Review. *European Science and Technology*, 12, 104-105.

Konsortium HarmoS Naturwissenschaften (2008): *HarmoS Naturwissenschaften - Wissenschaftlicher Schlussbericht.* Bern: EDK.

Kwon, O. N., Park, J. S. & Park, J. H. (2006).: Cultivating divergent thinking in mathematics through an open-ended approach. *Asia Pacific Educational Review, 7*(1), 51-61.

Labudde, P. (2000): Constructivism in physics instruction at the upper secondary level (*Konstruktivismus im Physikunterricht der Sekundarstufe II*). Bern: Haupt.

Labudde, P. (2007): How to Develop, Implement and Assess Standards in Science Education? 12 Challenges from a Swiss Perspective. In: D. Waddington; Nentwig, P. & Schanze, S. (Eds.): *Making it Possible: Standards in Science Education* (277-301). Münster: Waxmann.

Labudde, P. (Ed.) (2010): *Fachdidaktik Naturwissenschaften 1.-9. Schuljahr.* Bern: Haupt.

Labudde, P. et al. (2012): The development, validation, and implementation of standards in science education: chances and difficulties in the Swiss project HarmoS. In: Schanze, S. et al. (eds.), 197-222. Münster: Waxmann.

Le Hebel, F., Montpied P. & A. Tiberghien (2011): Which Effective competencies do students use in PISA assessment of scientific literacy. ESERA Lyon, France. (*to appear in e proceedings*).

Lehesvuori, S., Ratinen, I., Kulhomäki, O., Lappi, J., & Viiri, J. (2011): Enriching primary student teachers' conceptions about science teaching: Towards dialogic inquiry-based teaching. *NorDiNa - Nordic Studies in Science Education*, 7(2),

140-159.
Lewis, T. (2006): Design and inquiry: Bases for an accommodation between science and technology education in the curriculum? *Journal of Research in Science Teaching,* 43(3), 255-281.
Linn, R.L. (2000): Assessment and Accountability, *Educational Researcher,* 29 (2) 4 - 16.
Llewellyn, D., (2007): *Inquiry within: Implementing inquiry-based science standards in Grades 3–8*. Thousand Oaks, CA: Corwin Press.
Looney, J. W. (2011): Integrating Formative and Summative Assessment: Progress Toward a Seamless System? *OECD Education Working Papers*, No. 58, OECD Publishing. http://dx.doi.org/10.1787/5kghx3kbl734-en.
Masters, G. & Forster, M. (1996): *Progress Maps.* Camberwell, Victoria: Australian Council for Research in Education.
Masters, G. & Forster, M. (1996a): *Development Assessment.* Camberwell, Victoria: Australian Council for Research in Education.
Millar, R. & Driver, R. (1987): Beyond processes, *Studies in Science Education,* 14, 33-62.
Millar, R. & Osborne, J. F. (Eds.) (1998): *Beyond 2000: Science Education for the Future*. London: King's College London.
Minner, D.D., Levy, A.J. and Century, J. (2010): Inquiry-Based Science Instruction – What Is It and Does It Matter? Results from a Research Synthesis Years 1984 to 2002. *Journal of Research in Science Teaching*, 47, 474–496.
Miyakawa, T. & Winsløw, C. (2009): Didactical designs for students' proportional reasoning: An "open approach" lesson and a "fundamental situation". *Educational Studies in Mathematics,* 72(2), 199-218.
NAEP (2008): *Science Framework for the 2009 National Assessment of Educational Progress.* Washington DC: US Govt Printing Office.
Nieminen, P., Savinainen, A., & Viiri, J. (2010): Force Concept Inventory-based multiple-choice test for investigating students' representational consistency. *Physical Review Special Topics - Physics Education Research*, 6 (020109).
Nohda, N. (2000): Teaching by open-approach method in Japanese mathematics classroom. In T. Nakahara & M. Koyama (Eds.), *Proc. of the 24th Conf. of the Int. Group for the Psychology of Mathematics Education*, 1, 39-53. Hiroshima, Japan: PME.
Novotná, J. & Hošpesová, A. (2009): Linking in teaching linear equations - forms and purposes. The case of the Czech republic. In: Shimizu, Y., Kaur, B., Huang, R., Clarke, D. (Eds.): *Mathematical Tasks in Classrooms around the World*, 103-118. Rotterdam: Sense Publishers.
NRC (National Research Council) (1996): *National Science Education Standards*. Washington DC National Academy Press
NSF (National Science Foundation) (1997): *The challenge and Promise of k-8 Science Education Reform*. Foundations 1. Arlington VA: NSF
OECD, 2000: *21st Century Learning: Research, Innovation and Policy. Directions from recent OECD analyses*. Centre for Educational Research and Innovation. Retrieved from: http://www.oecd.org/dataoecd/39/8/40554299.pdf.

OECD, 2006: *Assessing Scientific, Reading and Mathematical Literacy: A Framework for PISA 2006.* Paris: OECD Publishing.

Osborne, J. (2011): Science teaching methods: a rationale for practices. *School Science Review*, 93(343), 93-103.

Osborne, J. (2011b): Science Education Policy and Its Relationship with Research and Practice. In: DeBoer, G. (Ed.): *The Role of public Policy in K-12 Science Education*, 13-46. Charlotte: Information Age Publishing.

Osborne, J. F. & Dillon, J. (2008): *Science Education in Europe: Critical Reflections*. London: Nuffield Foundation.

Papacek, M. (2010): Badatelsky orientované přírodovědné vyučování – cesta pro biologické vzdělávání generací Y, Z a alfa? [Inquiry-based science education: A way for the biology education of generations Y, Z, and alpha?]. *Scientia in educatione*, 1(1), 33-49.

Perkins, D.N. & Saloman, G. (1989): Are cognitive skills context bound? *Educational Researcher*, 18(1) 16-25.

Petr, J. (2010): Biologická olympiáda - inspirace pro badatelsky orientované vyučování přírodopisu a jeho didaktiku. [The Biological Olympics – the Inspiration for Inquiry Based Science Teaching and Education and it's Didactics]. In: Papáček M. (Ed.): *Didaktika biologie v České republice 2010 a badatelsky orientované vyučování. DiBi 2010,* 136-144. České Budějovice: JU.

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

Ruiz-Primo, M. A. & Furtak, E. M. (2007): Exploring teacher's informal formative assessment practices and student's understanding of the context of scientific inquiry. *Journal of Research in Science Teaching,* 44, 57-84.

Russell, T. J. (Ed.) (1988): *Science at Age 11. A Review of APU Survey Findings 1980-84*. London: HMSO.

Rychen, D. S. & Salganik, L. H. (Eds.) (2001): *Defining and Selecting Key Competencies.* Hogrefe & Huber Publishers.

Savinainen, A. & Viiri, J. (2008): The force concept inventory as a measure of students' conceptual coherence. *International Journal of Science and Mathematics Education,* 6, 719-740.

Schoenfeld, A.H. (1985): *Mathematical problem solving.* San Diego, CA: Academic Press.

Schreiner, C. & Sjøberg, S. (2004). *Sowing the Seeds of Rose.* University of Oslo: Acta Didactica.

Shavelson, R. (2011): An approach to testing and modeling competence. Paper presented at the Bad Honnef Conference on Teachers' Professional Knowledge, May 6-7.

Shelley II, M. C. (2009): Speaking Truth to Power with Powerful Results: Impacting Public Awareness and Public Policy. In: Shelley II, M.C. et al. (eds.), *Quality Research in Literacy and Science Education.* Springer Science + Business Media.

Stuchlíková, I., (2010): Concept of Inquiry- Based Learning in teachers' professional development. In: Janik, T., Knech, P. (Eds.): *New pathways in Professional De-*

velopment of Teachers, 195-201*.* Vienna: LIT Verlag.

Sullivan, P., Mousley, J. & Zevenberger, R. (2006): Teacher actions to maximize mathematics learning opportunities in heterogeneous classrooms. *International Journal of Science and Mathematics Education,* 4, 117-143.

Tiberghien, A. (2009): The Importance of Aligning Teaching and Assessment. In: Bybee, R. W. & McCrae, B. (Eds.): *In PISA science 2006,* 133-138. Arlington: NSTA Press.

Tiberghien, A., Vince, J., & Gaidioz, P. (2009): Design-based Research: Case of a teaching sequence on mechanics. *International Journal of Science Education,* 31(17), 2275-2314.

Winsløw, C. & Emori, H. (2006): Comparative research on secondary mathematics education: a semiotic approach. In: Graf, K.-D. et al. (Eds.): *Mathematics education in different cultural traditions: A comparative study of East Asia and the West,* 553-566. Berlin: Springer, 2006.

Winsløw, C. (2009): The first years of teaching. In: Evens, R. & Ball, D. (Eds.): *The Professional Education and Development of Teachers of Mathematics*, 93-101. New York: Springer US.

Wood, T. & Sellers, P. (1997): Deepening the analysis: Longitudinal assessment of a problem-centered mathematics program. *Journal for Research in Mathematics Education,* 28(2), 163-186.