Report from the FP7 project:

Assess Inquiry in Science, Technology and Mathematics Education

Synopsis of the literature review

Silke Rönnebeck & Olaf Köller

Delivery date	28.02.2013
Deliverable number	D2.2
Lead participant	Leibniz-Institute for Science and Mathematics Education (IPN), Kiel, Germany
Contact person	Silke Rönnebeck, roennebeck@ipn.uni-kiel.de
Dissemination level	PP

Table of Contents

Summary

ASSIST-ME aims at developing assessment methods suitable for enhancing inquiry-based learning in science, technology and mathematics (STM) education. The development process will be based on an understanding of the concepts of competences (both domain specific and transversal), inquiry-based education and formative versus summative assessment. To help providing this understanding, WP2 in ASSIST-ME will carry out a review of the existing research literature on inquiry-based education in STM with a particular focus on the role that formative and summative assessment play for successful learning in STM classes.

This report first summarizes the basic definitions of IBE, competences and formative and summative assessment, respectively, that are building the foundation for the literature review. It then gives an overview of the intended strategy for the literature review and summarizes preliminary results from the first two months of the project. The appendix contains a table with the present content of the literature database. An extended version of this table (including abstracts) will be uploaded to the internal project website and updated in regular intervals.

Introduction

One mayor objective of ASSIST-ME is to develop a set of assessment methods suitable for enhancing inquiry-based learning of STM related competences. Based on these methods, strategies for formative and summative assessment of competences in STM that are adaptable to various European educational systems will then be identified (Dolin, 2012). The research into formative and summative assessment of competences relevant to inquiry-based STM education will be based on an understanding of the concepts of competences (both domain specific and transversal), inquiry-based education and formative versus summative assessment.

To help providing this understanding, WP2 in ASSIST-ME will carry out a review of the existing research literature on inquiry-based education in STM with a particular focus on the role that formative and summative assessment play for successful learning in STM classes.

The objectives of WP2 in ASSIST-ME are to

- review what we know about formative and summative assessment of competences in STM and what methods can work to improve student outcomes, including use of ICT tools,
- collect recent work that has been conducted in the partner countries on assessment in STM,
- synthesize the outcomes of other EU-projects on IBE, e.g. Mind the Gap, S-TEAM, ESTABLISH, Fibonacci, INQUIRE, PRIMAS, and SAILS,
- formulate an operational definition of IBE related to STM and give guidelines for measuring the degree of IBE,
- provide recommendations for WP4 and WP5.

The literature review is supposed to answer the following research questions:

- What are specific and transversal competences and attitudes in IBE in STM?
- What formative and summative assessment methods are used in STM with respect to IBE, how are they used and what is their role in STM teaching and learning?
- What are key aspects of efforts to promote formative assessment?
- How do assessment methods influence the uptake of IBE in STM?
- Which protocols and instruments (including ICT tools) for summative and formative assessment of IBE in STM exist and how are they integrated into the classroom?

One mayor challenge for this literature review is that the field of interest is not clearly defined. With respect to science education, there is still disagreement among researchers about what features define the instructional approach of inquiry-based teaching (Furtak, Shavelson, & Shemwell, 2012; Hmelo-Silver, Duncan, & Chinn, 2007). In the most recent meta-analysis of inquiry-based science teaching (IBST), Furtak and colleagues (Furtak, Seidel, Iverson, & Briggs, 2012) analyzed earlier meta-analyses with regard to their definitions of IBST. They found that the majority of meta-analyses relied upon expansive definitions of inquiry-based teaching. In some cases explicit definitions of some aspects of IBST were given (e.g. innovative, activity-based, process oriented or discovery oriented), in others the meaning of the term remained less clear. Moreover, the authors found that researchers and educators use a rich vocabulary to describe inquiry-based approaches to teaching and learning like e.g. inquiry-based teaching and learning, authentic inquiry, model-based inquiry, modeling and argumentation, project-based science, hands-on science, and constructivist science. These approaches might include characteristics of IBE to a varying degree but they are not necessarily synonyms of IBE. One example of a common misunderstanding with respect to IBE especially in science education is that inquiry is equated with "hands-on" whereas real inquiry not only requires "hands-on" but also "minds-on" as it is stated in the EU-report Europe needs more scientists (European Commission, 2004).

A similar situation is described by Black & Williams in their meta-analysis of formative assessment in the classroom (Black & Wiliam, 1998). They found a literature search by keywords in the ERIC database inefficient for their purposes because of "a lack of terms used in a uniform way". As in the case of IBE, formative assessment may be described under a variety of names like e.g. classroom evaluation, curriculum-based assessment, feedback or formative evaluation (Black & Wiliam, 1998, p. 53)

With respect to the research questions, this has some consequences that have to be kept in mind while designing the search strategies. Most of the research questions are related to the term IBE. As can be seen above, one objective of WP2 is to formulate an operational definition of IBE related to STM. This will be an outcome of the literature review but in some respect it is also a prerequisite for doing it in the first place. It is thus necessary to establish a "working definition" of IBE – and formative/summative assessment and competences, respectively – to use as a starting point for the literature research that will thus be refined during the process.

Working definitions of IBE in STM and formative/summative assessment

The working definitions of competences, IBE, and formative and summative assessment, respectively are based on the definitions given in the ASSIST-ME application (Dolin, 2012) and definitions used by other EU projects like S-TEAM, Fibonacci, or PRIMAS.

Competences

Within ASSIST-ME, a competence is understood as a complex ability that is based on a combination of skills, knowledge, characteristics, and traits. It contributes to performances in particular domains but also in real-life situations and is thus closely related to the concept of scientific literacy (OECD, 2006). ASSIST-ME explicitly dispenses with the distinction between competence and competency as found e.g. in (OECD, 2002) to reflect an integration of demand-orientation and a conceptualization of internal mental structures (abilities, dispositions, and attitudes) into the concept (Dolin, 2012). Although the lines of separation are not sharp and may vary from country to country, one generally distinguishes domain specific competences, i.e. key competences that are specific for a certain subject and transversal or cross-curricular competences that are needed in different subjects and are thus important across domain boundaries. Examples of domain-specific competences are e.g. to observe and describe natural phenomena accurately and to plan, perform, evaluate and reflect on experiments (science); to construct and produce technical tools and instruments e.g. to plan, manufacture, evaluate, and optimize technical solutions (technology); to develop and test hypotheses about functional relations that have been observed in the reality (from data) or in mathematical settings (mathematics). Examples for transversal competences are e.g. to process and interpret data and results, to argue and communicate with peers and experts, and to develop a sense of responsibility and become a responsible-minded citizen (Dolin, 2012).

Inquiry-based education in science, technology and mathematics

According to Anderson (Anderson, 2002) – whose definition forms the basis of the ASSIST-ME application – 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. Whereas these characteristics generally apply to all three subject areas – science, technology and mathematics – the ASSIST-ME application explicitly acknowledges that various meanings and forms of inquiry are possible in different disciplines and need to be addressed in the project. These different approaches to inquiry, however, need to be aligned with a general definition of the construct that will be produced by the project.

Looking into the literature, IBE has mainly been investigated in the field of science education. Performing a basic search in Web of Science for the period 1996 to 2012 using the keywords "science/scientific" crossed with "teaching", "learning", "education" and "instruction" and crossed with "inquiry" gave 2034 entries. Replacing "science/scientific" by ″mathematics" reduced the number of results to 218, by "technology" to 567 with most of the entries in technology dealing with the use of technology in inquiry-based (science) education and not with inquiry in technology education (search performed in November 2012). We will thus start with characteristics and definitions of inquiry-based science education in various FP7 projects and then explicitly discuss commonalities and differences in technology and especially in mathematics education.

Inquiry-based science education

A lot of former and on-going EU projects in the field of IBE (e.g. Mind the Gap, S-TEAM, and Fibonacci) have based their understanding of inquiry-based science education on a definition from Linn, Davis and Bell (Linn et al., 2004): "[inquiry is] the intentional process of diagnosing problems, critiquing experiments, and distinguishing alternatives, planning investigations, researching conjectures, searching for information, constructing models, debating with peers and forming coherent arguments". In IBSE, students should be able to identify relevant evidence and use critical thinking and logical reasoning to reflect on its interpretation. They should develop the skills necessary for inquiry and the understanding of science concepts through their own activity and reasoning. This involves exploration and hands-on experiments (Fibonacci project).

IBSE should foster critical and creative minds, it should encourage students to engage, explore, explain, extend, and evaluate in real-life situations and in collaboration and cooperation with their peers (PRIMAS project). It is thus based on a specific understanding of learning as deliberately involving linguistic processes like e.g. argumentation (Dolin, 2012) and requires students' to take charge of their own learning for genuine understanding (Harlen, 2009). With respect to teacher support and classroom climate, IBSE should lead to a shared sense of purpose and ownership in the classroom where contributions and mistakes are valued as learning opportunities. Teachers should be open-minded and dialogic; they should foster and value students' reasoning and scaffold learning by drawing upon students' experience (PRIMAS project).

Inquiry-based technology education

A lot of the characteristics of IBSE can be transferred to technology education. Inquiry in science and technology education is not sharply separated from each other particularly in view of the increasing emphasis in science on reflective thinking, cooperative learning, and the development of critical and creative thinking formerly ascribed to technology (Adams & Hamm, 1998). In the NRC Framework for K-12 Science Education, "asking questions" and "constructing explanations" in science is replaced by "defining problems" and "designing solutions" in technology – the other characteristics of IBSE apply in the same way for technology (National Research Council, 2012).

Inquiry-based mathematics education

Differences and similarities between inquiry-based science and mathematics education have been investigated and discussed within the Fibonacci-project. In the Fibonacci Background Resource Booklets *Learning through Inquiry* (Artigue, Dillon, Harlen, & Léna, 2012) and *Inquiry in Mathematics Education* (Artigue & Baptist, 2012), the authors state communalities and specificities of mathematical inquiry compared to scientific inquiry:

"Like scientific inquiry, mathematical inquiry starts from a question or a problem, and answers are sought through observation and exploration; mental, material or virtual experiments are conducted; connections are made to questions offering interesting similarities with the one in hand and already answered; known mathematical techniques are brought into play and adapted when necessary. This inquiry process is led by, or leads to, hypothetical answers – often called conjectures – that are subject to validation." (Artigue & Baptist, 2012, p. 4)

The main differences between mathematical and scientific inquiry are based on the type of questions (or problems) they address and the processes they rely on for answering or solving them. Whereas scientific inquiry is mostly motivated by questions arising from the natural or human-built world, the questions in mathematical inquiry are more diverse - including technical problems, human artifacts or art - and influence the inquiry process. Some specificities of mathematical inquiry are the pragmatism and non-linearity of the process, the dialectic interplay between proof and refutation, the definitive nature of the results obtained and the principal aim for generalization (Artigue & Baptist, 2012). As in scientific inquiry, students are seen as active participants in their own learning process but also in a community of learners. They should work collaboratively and engage in mathematical discussions by active listening, questioning, constructing and using arguments, communicating, and analyzing and evaluating the thinking of their peers (Hunter & Anthony, 2011). Specific to mathematics is that students should be able to articulate or elaborate questions in order to make them accessible to mathematical work (Artigue et al., 2012) and design procedures to find answers (American Association for the Advancement of Science, 1993). They should experiment with mathematical objects (e.g. numbers, patterns, forms, etc.) and use modeling and mathematizing (Artigue & Baptist, 2012). Since mathematical ideas are not directly accessible to our physical senses, students need to construct representations. If they find a solution to a specific problem, they are searching for generalizations (Hunter & Anthony, 2011). They construct meaning through the use of logic, evidence, and reflection (American Association for the Advancement of Science, 1993). In addition to the aspects mentioned for science and technology, inquiry mathematics puts special emphasis on enhancing students' problem-solving skills (Kwon, Park, & Park, 2006) and mathematical thinking (American Association for the Advancement of Science, 2009).

Formative and Summative Assessment

Assessment is one of the most important drivers in education and a defining aspect of any educational system. It can be regarded, however, from different perspectives as it is stated in the European report *Europe needs more scientists* (European Commission, 2004, p. 137): "Assessment has to be seen from at least three perspectives: (a) the traditional function is the evaluation of students' achievement to put them on a certain 'career track' by giving marks and reports; (b) assessment should also be used as an instrument for diagnosis to give students and teachers permanent feedback about learning outcomes and difficulties, and therefore the need for support; and (c) in recent years, international comparative and large-scale assessment studies have become more popular as they should enable broader knowledge about the conditions and influences on students' understanding and competence." Assessment always involves the collection, interpretation and use of data for some purpose. The purpose and often also the way of data collection, however, may differ. In ASSIST-ME, we summarize these different purposes under the expressions of summative and formative assessment.

Summative assessment has the purpose of summarizing 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. Summative assessment has an undeniably strong impact on teaching methods and content (Harlen, 2007). It is perceived to be highly visible, especially if high stakes are connected to it that can easily undermine innovative approaches to teaching (Looney, 2011). This is also emphasized in the above mentioned European report: "Although the results [of large international assessments like PISA and TIMSS] may be used to identify strengths and weaknesses in each country, there is a danger that these studies may trivialise the purpose of schooling by its implicit definition of how educational 'quality' might be understood, defined and measured. It is likely that national school authorities put undue emphasis on these comparative studies, and that curricula, teaching and assessment will be 'PISA-driven' in the years to come (European Commission, 2004, p. ix). The dominance of external summative assessment leads to situations where testing remains distinct from learning in the minds of most students and teachers (American Association for the Advancement of Science, 1998) thus when teachers are required to undertake their own assessments they tend to imitate external assessments and think only in terms of frequent summative assessment (American Association for the Advancement of Science, 1998; Black & Wiliam, 1998).

Formative assessment, in contrast, 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, where they need to go and how best to get there" (Assessment Reform Group, 2002). Or, as Bell & Cowie put it: "Formative assessment is the process used by teachers and students to recognize and respond to student learning in order to enhance that learning, during the learning" (Bell & Cowie, 2001). The term formative with respect to evaluation and assessment was first used by Scriven (Scriven, 1967) and Bloom (Bloom, 1969) in the late 1960s. It is explicit in these early uses that the term formative cannot be a property of an assessment (Wiliam, 2006). In principle, the same test could be used for formative or summative purposes (Bloom, 1969). Assessments are formative, however, if, and only if, something is contingent on their outcome and the information is actually used to alter what would have happened in the absence of that information (Wiliam, 2006; Black & Wiliam, 1998). There are different instruments and techniques that are used for formative assessment like e.g. portfolios, discourse, questioning, student observations, quizzes, interviews, etc.. In their 1998 review on formative assessment, Black and Wiliam (Black & Wiliam, 1998) could show that formative assessment methods and techniques produce significant learning gains that are among the largest ever identified for educational interventions (Looney, 2011). Despite the variety of methods, techniques, and instruments, however, formative assessment possesses some common characteristics. Among these are that it has to be an integral part of teaching and learning (Bell & Cowie, 2001). It has to be frequent, actively engages students by peer- and self-assessment and provides feedback and guidance to learners on how to improve by focusing on the learning process (Looney, 2011; Wilson & Sloane, 2000). Recent OECD publications stress the importance of formative assessment and the integration of formative and summative assessment, respectively (Looney, 2011; OECD, 2005). They also realize, however, that many logistical barriers to making formative assessment a regular part of the teaching practice exist, rooted in organizational issues such as large classes, extensive curriculum requirements, or the difficulty of meeting diverse and challenging student needs, but also in teachers' often deeply held pedagogical beliefs of assessment as a tool for teacher use and accountability rather than as a method to involve students in a constructivist assessment environment (Wilson & Sloane, 2000).

The ASSIST-ME project aims at developing 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 (Dolin, 2012). Within ASSIST-ME, an assessment method is understood as a package of information, procedures and instructions aimed at collecting and interpreting information about students' competences within STM. Typically, an assessment package will include a hands-on activity, paper & pencil-activity, peer-to-peer feedback or on-screen activity for students; assessment tools might include a process monitoring tool (e.g. portfolios), a performance assessment tool, a computer-based system with built-in feedback, and a structured dialogue tool (e.g. assessment conversations (Ruiz-Primo, 2011)).

Methods

To answer the research questions mentioned in the introduction, three different search strategies will be used:

1. For the journals that are considered to be the most important in STM education and assessment, the contents of all issues published during the last 10 years will be scanned
2. Expert knowledge of relevant publications in the field will be used and the reference lists of these publications will be scanned and followed up
3. Searches by key-words in relevant databases (PsychInfo, ERIC, Web of Science or Google Scholar) will be performed (including citation searches e.g. on the review by Black and Wiliams (1998) on formative assessment)

Following discussions with experts in the specific domains and taking into account impact factors, it is planned to include the following journals in the first strategy:

- Science Education: Journal of Research in Science Teaching, Science Education, International Journal of Science Education
- Technology Education: Journal of Technology Education, Research in Science & Technological Education
- Mathematics Education: Journal of Mathematics Education, Educational Studies in Mathematics, International Journal of Science and Mathematics Education
- ICT-assisted Learning: International Journal of Computer-Supported Collaborative Learning, Computer & Education, Interactive Learning Environments
- Assessment: Assessment in Education, Educational Assessment, Applied Measurement in Education

If during the work progress it seems sensible or necessary, further journals in the specific science subject domains (i.e. in biology, chemistry, and physics education) or journals in the field of general education like e.g. Review of Educational Research or Learning & Instruction might be included in the search.

The systematic search by key-words (strategy 3) faces some challenges that have already been mentioned in the introduction. These challenges are mainly based on the diversity of terms and instructional or teaching approaches that include characteristics of inquiry-based education. A literature search just using “inquiry” as keyword is thus very likely to miss a lot of relevant publications. Looking into earlier reviews in the field of inquiry-based science education, one finds that quite general key-words have been used to ensure that as many as possible of the relevant publications will be found. In their meta-analysis of inquiry based science teaching, Furtak et al. (Furtak et al., 2012) e.g. used the keywords: effective instruction, instructional effectiveness, direct instruction, teacher effectiveness, mastery learning, constructivist teaching, science instruction, classrooms, science teaching, and inquiry. Each of these keywords was then crossed with the output keywords achievement, competencies, interest, motivation, engagement, and attainment leading to an initial sample of 1625 entries for the years 1996-2006. Another review published in the context of the S-TEAM project (Heinz, 2012) used the keywords: inquiry based science teaching, science teaching and learning, science literacy and scientific literacy, collaborative science learning, argumentation in science education, heuristic in science education, and inquiry-based instruction. These keywords were then crossed with the target group keywords policy/stakeholders, teacher educators/teacher education, teachers, and students/pupils. They limited their search to the years 2005-2009 and found around 600 publications. Using such open keyword approaches – although we clearly admit the benefits they have - seems not feasible for our work in ASSIST-ME for several reasons. First, the research questions we are supposed to answer span a broad range from definitions of IBE to competences in IBE in STM to the state-of-the-art in formative and summative assessment of IBE in STM (including efforts to foster formative assessment, the interdependence between IBE and formative assessment, and instruments and protocols for summative and formative assessment of IBE). Second, not only IBE comes under a variety of terms and approaches, but also some of our outcome variables like formative assessment or competences. And third, the alternative key-words for IBE are to a significant extend domain-specific which means different combinations of key-words have to be used for at least science and mathematics. Using very open key-words would thus lead to an initial sample of publications that is impossible to accurately analyze within the limited timeframe of WP2. The definition of sensible key-words is thus of crucial importance.

As a starting point we compared the sets of key-words used in recent reviews of inquiry-based science education with alternative expressions of IBE found in the thesaurus of the ERIC and PsychInfo databases. With these keywords we then conducted exploratory searches in PsychInfo to get a feeling how open the key-words could be to still yield feasible sample sizes and which combinations of key-words lead to publications that are useful with respect to our research questions. From these searches, a first set of key-words emerged. With these key-words we performed a search in the PsychInfo database looking for publication in books or peer-reviewed journals that were written in English and published between 1996 and 2012 in the field of science education. 1996 was chosen as the starting date since in that year the National Science Education Standards (National Research Council, 1996) were published and introduced a time period during which the reform spotlight was intensely focused upon inquiry-based education, especially in science. As key-words for the search we used *inquiry, collaborative learning, mastery learning, discovery learning, cooperative learn-*

ing, discovery teaching method, constructivist teaching, and *problem based learning*. These keywords were crossed with *science education, science instruction, science teaching and learning*, and *classroom/teachers*. The resulting set of publications is considered to represent the concept of "inquiry in science education". It was then crossed with the output variables *competences, competencies, attitudes, beliefs, assessment, achievement, interest, motivation*, and *engagement*. This search was supplemented by two further searches using very specific search terms: *inquiry-based science education, inquiry-based science instruction, inquiry-based science teaching, inquiry-based science learning*, and *scientific inquiry* crossed with *classroom*. These searches yielded approximately 250 results that were exported into a CITAVI database.

Since WP2 is one of the WPs in "Phase 1 – Building the Foundation" in ASSIST-ME, it is important to ensure that the information we produce really complies with the needs of WP4 and WP5. The preliminary set of key-words was thus discussed with the other project partners at the WP2-workshop during the ASSIST-ME kick-off conference in Copenhagen on January 26. Following the discussions and further literature work, the key-words will be extended by *skills*, *feedback*, and *classroom evaluation*, and *assessment* replaced by *formative* and *summative assessment*, respectively.

Results

Progress of work

The results presented here represent the stage of work after one and a half months of ASSIST-ME and can thus only be preliminary. Based on the PsychInfo search described in the previous section, publications that were known to us or that we found by following up reference lists or through our exploratory searches or through journal-based searches in *the International Journal of Science and Mathematics Education* and *Assessment in Education*, a CITAVI database has been built up that initially contained about 500 publications. At the moment, we are in the process of screening these publications by reading abstracts or, if necessary, full texts to decide whether they contribute to our field of interest or not. By doing so, up to now approx.. 220 of the initial publications has been excluded. At the same time, we are trying to categorize the results with respect to our research questions (i.e. are they dealing with definitions of IBE, competences, formative/summative assessment, or instruments), the grade level they address (primary, lower secondary, and upper secondary following, in general, the German grading system of primary meaning grades 1-4 (or 6), lower secondary meaning grades 5 (or 7) to 9/10 and upper secondary meaning grades beyond 9/10) and the domain (science, mathematics or technology). The underlying idea is, in addition to the reports WP2 will provide, to use the results of our literature review as the starting point of a research base within ASSIST-ME that can be used by the other partners to support their work. This could e.g. be achieved by uploading the database to the internal project website in a format that allows the partners to search it according to the above mentioned criteria (for the start, this will be an Excel-file). The file will then be updated in regular intervals as we proceed in our work. In the appendix of this report, you will find a table containing the current content of the database.

In order to extend the key-word search to mathematics and technology, appropriate key-words have to be defined. In the case of mathematics, teaching approaches that include characteristics of mathematical inquiry are – as named in the ASSIST-ME application – inquiry mathematics (Cobb, Wood, Yackel, & McNeal, 1992) didactical engineering (Artigue, 1994), open approach lessons (Nohda, 2000), and problem-centered learning (Schoenfeld, 1985). The Fibonacci-project (Artigue & Baptist, 2012) extends this list towards didactical learning (Gallin, 2012), the Dutch approach of realistic mathematics education (Freudenthal, 1973) and the French theory of didactical situations (Brousseau, Brousseau, & Balacheff, 1997). Another approach of inquiry in mathematics education is the concept of "problem-based learning" that is also mentioned in the well-known Rocard report (European Commission, 2007, p. 9): "In mathematics teaching, the education community often refers to "Problem-Based Learning (PBL)" rather than to IBSE. In fact, mathematics education may easily use a problem-based approach while, in many cases, the use of experiments is more difficult. Problem-based learning describes a learning environment where problems drive the learning."

As shown in the definition section, inquiry in technology education is quite closely related to the concept in science education. However, to ensure that no relevant publications are missed, it is planned to look more detailed for teaching approaches in technology that include characteristics of IBE.

Definition of IBE in STM

With respect to the development of a feasible definition of IBE in STM within the ASSIST-ME-project, existing definitions in other FP7-projects have been analyzed. As already mentioned in the introduction, a general agreement in the definitions of inquiry-based science education that is based upon Linn, Davis, and Bell (Linn et al., 2004) could be observed. This definition is in alignment with the definition of inquiry in the National Science Education Standards (National Research Council, 1996): "Inquiry is a multifaceted activity that involves making observations; posing questions; examining books and other sources of information to see what is already known; planning investigations; reviewing what is already known in light of experimental evidence; using tools to gather, analyze, and interpret data; proposing answers, explanations, and predictions; and communicating the results. Inquiry requires identification of assumptions, use of critical and logical thinking, and consideration of alternative explanations." This notion is also shared by Hmelo-Silver and colleagues who understand inquiry learning as the process of learning content as well as discipline-specific reasoning skills and practices by collaboratively engaging in investigations. It is organized around relevant, authentic problems or questions and engages students cognitively in sense-making, developing evidence-based explanations, and communicating their ideas (Hmelo-Silver et al., 2007). Inquiry requires students to combine 'hands-on' with 'minds-on' (European Commission, 2004). The situation of IBE, however, seems to become more complicated recently since e.g. in the US the field of science education moves away from using the term inquiry and call it now "scientific and engineering practices" (National Research Council, 2012) that are defined as: Asking questions (for science) and defining problems (for engineering), developing and using models, planning and carrying out investigations, analyzing and interpreting data, using mathematics and computational thinking, constructing explanations (for science) and designing solutions (for engineering), engaging in argument from evidence, and obtaining, evaluating, and communicating information.

Competences
An overview of domain-specific and transversal competences that are ascribed to inquiry can be found e.g. in the AAAS Benchmarks for Science Literacy (American Association for the Advancement of Science, 2009). They describe inquiry as part of the nature of science and mathematics and divide it into the aspects of problem-solving, understanding and conducting investigations, and reasoning. Domain-specific competences in mathematics include:

- to identify the underlying problem and formulate questions which are accessible to mathematical investigations
- to use and manipulate mathematical objects and representations
- to use modeling and mathematization
- to choose appropriate rules or methods
- to make simplifying assumptions or approximations
- to evaluate the results with respect to their sense and usefulness
- to realize limitations of models and develop criteria for “good solutions”
- to use cycles of trial, evaluation, and revision that should lead to valid mathematical ideas
- to understand the nature of logic
- to use logic in making and judging arguments
- to distinguish between example and proof, sufficient and necessary conditions, good and bad logic, analogies/similarities and the “real thing”
- to organize evidence
- to formulate arguments, debate with peers, respond to criticism, offer criticism
- to generalize (accuracy, limits)
- to deal with uncertainty
- to distinguish between mistakes and reasonable choices that turned out to be unsuccessful
- to realize that more than one mathematical description might be equally appropriate and that there is not always a clear-cut answer

Domain-specific competences in science include:

- to identify scientific questions
- to formulate research questions and hypotheses
- to design experiments and conduct trials
- to collect data by observations
- to develop, use, and interpret different representations of the data
- to evaluate data
- to formulate conclusions based on evidence
- to communicate results, debate with peers, respond to criticism
- to realize the importance of and dangers to objectivity
- to use logic and make logical interferences
- to distinguish theory from evidence
- to understand the necessity of controlled variables

Domain-specific competences in technology include (Eckersall, 1987):

- to clarify the problem and propose solutions
- to assess the appropriateness of the proposed solutions
- to implement and evaluate the selected solution
- to apply technical thinking and acting (including using drawings, measuring devices, tools, equipment, and machines)
- to reflect on the social consequences of the technology with which the students are involved

With respect to transversal competences, the Benchmarks pay special attention to the so-called “habit of mind” which describes problem-solving skills that are relevant in all subjects. These skills are computation and estimation, manipulation and observation, communication and quantitative thinking, and critical response skills (evaluating evidence and claims). Moreover, however, a habit of mind includes also values and attitudes like honesty, curiosity, open-mindedness and skepticism. Other transversal competences mentioned in the Benchmarks include making accurate measurements, using mathematics and data analysis to solve problems, being creative in designing experiments and solving mathematical or scientific problems; the competence of the students is reflected in the quality of questions they pursue and the rigor of their methodology (American Association for the Advancement of Science, 1998). The key competences for lifelong learning described in the **Recommendation of the European Parliament (European Parliament, 2006) can also be understood as transversal competences:**

- **communication (**express and interpret concepts, thoughts, feelings, facts, and opinions**)**
- **learning to learn (**the ability to pursue and organize one's own learning, either individually or in groups)
- **sense of initiative and entrepreneurship (**creativity, innovation and risk-taking, as well as the ability to plan and manage projects in order to achieve objectives)

Moreover, **mathematical competence and basic competences in science and technology are seen as transversal competences.**

State of the art of formative and summative assessment

As stated in the AAAS Blueprints for Reform, any effort at nationwide science education reform must include reform of student assessment as a major goal (American Association for the Advancement of Science, 1998). This is also realized in the European report *Europe needs more Scientists* (European Commission, 2004, p. 137): “The enlargement of curricula towards achieving a greater relevance for everyone also requires changes in assessment strategies and instruments. If science no longer consists of theoretical knowledge about concepts and processes only, other competencies have to be assessed and their importance pointed out to students and parents.” At least since the review by Black & Wiliam (Black & Wiliam, 1998), formative assessment is seen as one means to achieve a better alignment between learning goals and assessment. This is also supported by more recent European documents on formative assessment or a possible integration of formative and summative assessment, respectively (OECD, 2005; Centre for Educational Research and Innovation (CERI), 2008; Looney, 2011).

Although a sound research base about the benefits of formative assessment exists, however, assessment in many countries still seems to be dominated by summative assessment. Looney attributes this, among others, to a perceived tension between formative and highly-visible summative assessments. Especially high stakes connected to summative assessment often undermine innovative approaches to teaching (Looney, 2011; American Association for the Advancement of Science, 1998). The target-driven approach of external testing leads to problems, including 'teaching to the test', the detriment of the wider curriculum, and motivational problems (Gardner, 2010). In addition, high stakes has little influence on teaching strategies, addressing tensions and fostering constructive cultures of evaluation (OECD, 2005).

A prerequisite to foster the use of formative assessment is to enable teachers to change their deeply held pedagogical beliefs of assessment as a tool for teacher use and accountability rather than as a method to involve students in a constructivist assessment environment. The understanding and acceptance of innovations by the teachers is crucial to the ultimate success of change (Wilson & Sloane, 2000). This can be supported by:

- **Integrating assessment and instruction** (Shepard, 2000; Pellegrino, Baxter, & Glaser, 1999). Assessment still often remains distinct from learning in the minds of most students and teachers (American Association for the Advancement of Science, 1998).
- **Embedding formative assessment in the curriculum.** For assessment to become fully and meaningfully integrated into the teaching and learning process, it must be curriculum dependent i.e. linked to a specific curriculum (Wilson & Sloane, 2000). The effectiveness of an assessment depends, in large part, on how well it aligns with the curriculum and instruction to reinforce common learning goals (Pellegrino, Chudowsky, & Glaser, 2001) and to clarify inconsistencies within and between lessons (Shavelson et al., 2008) .
- **Fostering the collaboration between curriculum and assessment experts as well as teachers.** Building stronger bridges between research, policy and practice is essential for the success but also challenging (Shavelson et al., 2008).
- **Enhancing accountability**. Teachers must feel confident that new assessment methods will be accepted for accountability purposes by school administrators and the public at large (American Association for the Advancement of Science, 1998). Fears regarding accountability might be mitigated if teachers can follow state curriculum frameworks.
- **Supporting teachers by teacher professional development (TPD)** (Pedder, 2006). Wiliam considers “the task of improving formative assessment [to be] substantially, if not mainly, about TPD” (Wiliam, 2006). The provision of tools for formative assessment – although a necessary condition – will only improve formative assessment practices if teachers can integrate them into their regular classroom activities. To reach this goal teachers need time because it is difficult for them to change practices which are closely embedded within their pattern of pedagogy, thus the pace of change is slow (Black & Wiliam, 1998). During this process, they need help to change the perception of their own role **(**American Association for the Advancement of Science, 1998). If teachers are comfortable with new testing procedures, student opinions are likely to follow. Moreover,

TPD could foster the integration of assessment and instruction by combining work on assessment with work on instruction and materials.

In her report about the integration of formative and summative assessment, Looney (Looney, 2011) names barriers to an implementation of formative assessment as well as policies that might support it. As barriers are regarded:

- large classes,
- extensive curriculum requirements,
- the difficulty of meeting diverse and challenging student needs,
- fears that formative assessment is too resource-intensive and time consuming to be practical,
- a perceived tension between formative assessment and highly visible summative assessment (see above). Within the "Learning How to Learn" project , Pedder (Pedder, 2006) found that classroom assessment practices are underpinned by conflicting and quite separate principles, namely assessment for learning principles (making learning explicit and promoting learning autonomy) and assessment of learning principles (performance orientation). Teachers' assessment practices were often out of step with their teaching values.
- lack of coherence between assessments and evaluations at the policy, school and classroom level,
- perception of formative assessment methods as "soft", non-quantifiable assessments by policy makers/administrators

Although ASSIST-ME is primarily interested in approaches or policies to foster the implementation of formative assessment, the perceived barriers can give valuable information that has to be kept in mind when developing assessment methods.

Policies supporting formative assessment are seen in (Looney, 2011; OECD, 2005):

- building teachers' and school leaders' *assessment literacy* (i.e. an awareness of the different factors that may influence the validity and reliability of results, the capacity to make sense of data, to identify appropriate actions and to track process (American Federation of Teachers, National Council on Measurement in Education, & National Education Association, 1990; Brookhart, 2011; Alkharusi, 2011 and references therein)
- investing in training and support for formative assessment
- encouraging innovation and creating opportunities for teachers to innovate
- providing guidelines and tools to facilitate formative assessment practice
- providing curriculum guidelines to assist teachers in a more systematic integration of formative assessment
- keeping the focus on teaching and learning
- developing clear definitions of learning goals and a theoretical framework of how that learning is expected to unfold as the student progresses through the instructional material (ensures construct and instructional validity (Wilson & Sloane, 2000)). This can be supported by developing "progress variables" that mediate between the level of detail present in the content of a specific curriculum and more vague contents of e.g. state standards. The assessments are then aligned with these variables that define goals of student growth.

- actively involving students and parents in the formative process (American Association for the Advancement of Science, 1998)
- building stronger bridges between research, policy and practice
- convincing policy makers/administrators that formative assessment methods are not “soft” but measuring the development of higher order thinking skills (American Association for the Advancement of Science, 1998)
- ensuring that classroom, school and system level evaluations are linked and are used formatively to shape improvements at every level of the system

The links between formative and summative assessment could be strengthened by (Looney, 2011):

- strengthening teachers’ assessment roles (see assessment literacy above). Heritage et al. (Heritage, Kim, Vendlinski, & Herman, 2009) found that teachers are quite literate at identifying the key mathematical principles being assessed and characterizing the student's level of understanding but had problems determining appropriate next instructional steps.
- strengthening teacher appraisal (there are, however, a number of challenges to developing coherent and valid measures of formative assessment practice as it involves several steps including the assessment process, interpretation of evidence of students learning, and the development of next steps for instruction (Herman, Osmundson, & Silver, 2010).
- drawing on advances in the cognitive sciences to strengthen the quality of formative and summative assessment (Shepard, 2000 and references therein)
- developing curriculum-embedded or “on-demand” assessments (Chudowsky & Pellegrino, 2003)
- developing complementary diagnostic assessments for students at lower proficiency levels to identify specific learning difficulties
- taking advantage of technology (Chudowsky & Pellegrino, 2003)
- using population instead of census sampling (Chudowsky & Pellegrino, 2003)
- meet standards of validity, reliability, feasibility, and equity (American Association for the Advancement of Science, 1998)

Looney (Looney, 2011) argues that large-scale tests often do not reflect the promoted development of higher-order skills such as problem-solving, reasoning, and collaboration – which are key competences in IBE. This leads to technical barriers to a more closely integration of formative and summative assessment because large-scale summative assessment data are often not detailed enough to diagnose individual student needs or they are not delivered in timely manner to have impact on the students assessed. Moreover, creating reliable measures of higher-order skills is still a challenge. Relating to this, she sees three major challenges: 1. Developing assessments that measure not only “what” but also “how to”, 2. reporting results “criterion-referenced” instead of “norm-referenced” including the development of focused reporting scales in criterion-referenced systems to provide diagnostic information (especially for weak students), and 3. finding a balance between generalizability, reliability, and validity (issue of ensuring fairness of classroom-based approaches also mentioned in (Wilson & Sloane, 2000)).

Formative assessment methods (including instruments and protocols)
As said before, within ASSIST-ME, an assessment method is understood as a package of information, procedures and instructions aimed at collecting and interpreting information about students' competences within STM. Typically, an assessment package will include a hands-on activity, paper & pencil-activity, peer-to-peer feedback or on-screen activity for students; assessment tools might include a process monitoring tool (e.g. portfolios), a performance assessment tool, a computer-based system with built-in feedback, and a structured dialogue tool (e.g. assessment conversations (Ruiz-Primo, 2011)).

In the following, some examples of methods and instruments we have found so far in the literature are given. This strand of research will be expanded in the following months.

- Feedback (e.g. Black & Wiliam, 1998; Wilson & Sloane, 2000)
- Discourse (e.g. Anderson, Zuiker, Taasoobshirazi, & Hickey, 2007)
- Effective questioning (e.g. Williams & Ryan, 2000; OECD, 2005)
- Self- and peer assessment (e.g. Sadler, 1989)
- Portfolios or learn logs (e.g. Johnson, Mims-Cox, & Doyle-Nichols, op. 2006; Gläser-Zikuda, 2007)
- Performance assessments (e.g. Gallavan, 2009, Looney, 2011)
- Mind-mapping (e.g. Goodnough & Long, 2006)
- Concept maps (e.g. Stoddart, Abrams, Gasper, & Canaday, 2000)
- Rubrics (e.g. assessment rubrics used in *Inquiry in Action* (Kessler & Galvan, 2007); (Siegel, Hynds, Siciliano, & Nagle, 2006))
- Embedded assessment (e.g. BEAR - Berkeley Evaluation and Assessment research - assessment system (Wilson & Sloane, 2000))

References

Alkharusi, H. (2011). Psychometric properties of the teacher assessment literacy questionnaire for preservice teachers in Oman. *Procedia - Social and Behavioral Sciences*, *29*, 1614–1624.

Adams, D. & Hamm, M. (2008). Helping students who struggle with math and science: A collaborative approach for elementary and middle schools. Blue Ridge Summit: Rowman &Littlefield Education.

American Association for the Advancement of Science. (1993). *Benchmarks for science literacy: Project 2061*. New York ; Oxford: Oxford University Press.

American Association for the Advancement of Science. (1998). *Blueprints for Reform - Project 2061: Chapter 8: Assessment*. Retrieved from http://www.project2061.org/publications/bfr/online/blpintro.htm

American Association for the Advancement of Science. (2009). *Benchmarks for Science Literacy*. Retrieved from http://www.project2061.org/publications/bsl/online/index.php

American Federation of Teachers, National Council on Measurement in Education, & National Education Association. (1990). Standards for teacher competences in educational assessment of students. *Educational Measurement: Issues and Practice*, *9*(4), 30.

Anderson, K. T., Zuiker, S. J., Taasoobshirazi, G., & Hickey, D. T. (2007). Classroom Discourse as a Tool to Enhance Formative Assessment and Practise in Science. *International Journal of Science Education*, *29*(14), 1721–1744.

Anderson, R. D. (2002). Reforming Science Teaching: What Research Says About Inquiry. *Journal of Science Teacher Education*, *13*(1), 1–12.

Artigue, M. (1994). Didactical Engineering as a framework for the conception of teaching products. In R. Biehler (Ed.), *Didactics of mathematics as a scientific discipline* (pp. 27–39). Dordrecht ; Boston: Kluwer Academic Publishers.

Artigue, M. & Baptist, P. (2012). *Inquiry in Mathematics Education* (Resources for Implementing Inquiry in Science and in Mathematics at). Retrieved from http://www.fibonacci-project.eu/

Artigue, M., Dillon, J., Harlen, W., & Léna, P. (2012). *Learning through inquiry* (Resources for Implementing Inquiry in Science and in Mathematics at). Retrieved from http://www.fibonacci-project.eu/resources

Assessment Reform Group. (2002). *Assessment for Learning: 10 principles*. Retrieved from http://assessmentreformgroup.files.wordpress.com/2012/01/10principles_english.pdf

Bell, B., & Cowie, B. (2001). The characteristics of formative assessment in science education. *Science Education*, *85*(5), 536–553.

Black, P., & Wiliam, D. (1998). Assessment and Classroom Learning. *Assessment in Education: Principles, Policy & Practice*, *5*(1), 7–74.

Bloom, B. S. (1969). Some theoretical issues relating to educational evaluation. In R. W. Tyler (Ed.), *National Society for the Study of Education Yearbook: 68 (2). Educational evaluation: New roles, new means* (pp. 26–50). Chicago: University of Chicago Press.

Brookhart, S. M. (2011). Educational Assessment Knowledge and Skills for Teachers. *Educational Measurement: Issues and Practice*, *30*(1), 3–12.

Brousseau, G., Brousseau, G., & Balacheff, N. (1997). *Theory of didactical situations in mathematics: Didactique des mathématiques, 1970-1990*. Dordrecht: Kluwer Academic Publishers.

Centre for Educational Research and Innovation (CERI). (2008). *Assessment for Learning: Formative Assessment*. Retrieved from http://www.oecd.org/site/educeri21st/40600533.pdf

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

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.

Dolin, J. (2012). *Assess Inquiry in Science, Technology and Mathematics Education: ASSIST-ME proposal.*

Eckersall, K. (1987). Some developments in technology education in Australia. In K. Riquards (Ed.), *Science and Technology Education and the Quality of Life - Vol. 2. Papers submitted to the 4th International Symposium on World Trends in Science and Technology Education, Kiel, 1987* (pp. 371–376). Kiel: IPN.

European Commission. (2004). *Increasing human resources for science and technology in Europe: Report of the High Level Group on Human Resources for Science and Technology in Europe, chaired by Prof. José Mariano Gago*. Luxembourg: Office for Official Publications of the European Communities.

European Commission. (2007). *Science education now: A renewed pedagogy for the future of Europe*. Luxembourg: Office for Official Publications of the European Communities.

European Parliament, C. (2006). *Key competences for lifelong learning: Summary of the recommendation 2006/962/EC of the European Parliament and of the Council of 18 December 2006 on key competences for lifelong learning*. Retrieved from http://europa.eu/legislation_summaries/education_training_youth/lifelong_learning/c11090_en.htm

Fibonacci project. *Disseminating inquiry-based science and mathematics education in Europe: Principles*. Retrieved from http://www.fibonacci-project.eu/project/principles

Freudenthal, H. (1973). *Mathematics as an educational task*. Dordrecht: Kluwer Academic Publishers.

Furtak, E. M., Seidel, T., Iverson, H., & Briggs, D. C. (2012). Experimental and Quasi-Experimental Studies of Inquiry-Based Science Teaching: A Meta-Analysis. *Review of Educational Research*, *82*(3), 300–329.

Furtak, E. M., Shavelson, R. J., & Shemwell, J. &. F. M. (2012). To teach or not to teach through inquiry: Is that the question? In S. M. Carver & J. Shrager (Eds.), *The journey from child to scientist. Integrating cognitive development and the education sciences* (1st ed., pp. 227–244). Washington, DC: American Psychological Association.

Gallavan, N. P. (2009). *Developing performance-based assessments: Grades K-5*. Thousand Oaks, Calif: Corwin Press.

Gallin, P. (2012). Dialogic learning - from an educational concept to daily classroom teaching. In Fibonacci project (Ed.), *Implementing Inquiry in Mathematics Education* (Resources for Implementing Inquiry in Science and in Mathematics at, pp. 23–33). Retrieved from http://www.fibonacci-project.eu/resources

Gardner, J. (2010). *Developing teacher assessment*. Maidenhead: Open University Press.

Gläser-Zikuda, M. (Ed.). (2007). *Lernprozesse dokumentieren, reflektieren und beurteilen: Lerntagebuch und Portfolio in Bildungsforschung und Bildungspraxis*. Bad Heilbrunn: Klinkhardt.

Goodnough, K., & Long, R. (2006). Mind mapping as a flexible assessment tool. In M. McMahon, P. Simmons, R. Sommers, D. DeBeats, & F. Crawley (Eds.), *Assessment in science. Practical experiences and education research* (pp. 219–228). Arlington, Va: NSTA Press.

Harlen, W. (2007). *The Quality of Learning: Assessment Alternatives for Peimary Education* (Primary Review Research Survey No. 3/4). Retrieved from http://image.guardian.co.uk/sys-files/Education/documents/2007/11/01/assessment.pdf

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

Heinz, J. (2012). *Indicators and instruments in the context of inquiry-based science education*. Münster: Waxmann.

Heritage, M., Kim, J., Vendlinski, T., & Herman, J. (2009). From Evidence to Action: A Seamless Process in Formative Assessment? *Educational Measurement: Issues and Practice*, *28*(3), 24–31.

Herman, J. L., Osmundson, E., & Silver, D. (2010). *Capturing quality in formative assessment practice: Measurement challenges: CRESST Report 770*. Los Angeles.

HMELO-SILVER, C. E., DUNCAN, R. G., & CHINN, C. A. (2007). Scaffolding and Achievement in Problem-Based and Inquiry Learning: A Response to Kirschner, Sweller, and Clark (2006). *Educational Psychologist*, *42*(2), 99–107.

Hunter, R., & Anthony, G. (2011). Forging Mathematical Relationships in Inquiry-Based Classrooms With Pasifika Students. *Journal of Urban Mathematics Education*, *4*(1), 98–119.

Johnson, R. S., Mims-Cox, J. S., & Doyle-Nichols, A. (op. 2006). *Developing portfolios in education: A guide to reflection, inquiry, and assessment*. Thousand Oaks, Calif. [etc.]: Sage Publications.

Kessler, J. H., & Galvan, P. M. (2007). *Inquiry in Action: Investigating Matter through Inquiry*. A project of the American Chemical Society Education Division, Office of K–8 Science.: American Chemical Society. Retrieved from http://www.inquiryinaction.org/download/

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

Linn, M. C. [., Davis, E. A. [., Bell, P. [., Linn, M. C., Davis, E. A., & Bell, P. (Eds.). (2004). *Internet environments for science education*. Mahwah: Lawrence Erlbaum Associates Publishers; US.

Looney, J. W. (2011). *Integrating Formative and Summative Assessment: Progress Toward a Seamless System?* (OECD Education Working Papers No. 58). Retrieved from http://dx.doi.org/10.1787/5kghx3kbl734-en

National Research Council. (1996). *National Science Education Standards*. Wasching-ton, D.C.: The National Academies Press.

National Research Council. (2012). *A framework for K-12 science education: Practices, crosscutting concepts, and core ideas*. Washington, D.C: The National Academies Press.

Nohda, N. (2000). Teaching by Open-Approach Method in Japanese Mathematics Classroom. *Proceedings of the Conference of the International Group for the Psychology of Mathematics Education (PME)*, (1), 39–53.

OECD. (2002). *Definition and selection of competences (DeSeCo): Theoretical and conceptual foundations*. Strategy paper. Retrieved from http://www.deseco.admin.ch/bfs/deseco/en/index/02.parsys.34116.downloadList.87902.DownloadFile.tmp/oecddesecostrategypaperdeelsaedcericd20029.pdf

OECD. (2005). *Formative Assessment: Improving Learning in Secondary Classrooms*. Paris: OECD Publishing and Centre for Educational Research and Innovation.

OECD. (2006). *Assessing scientific, reading and mathematical literacy: A framework for PISA 2006*. Paris: Organisation for Economic Co-operation and Development.

Pedder, D. (2006). Organizational conditions that foster successful classroom promotion of Learning How to Learn. *Research Papers in Education*, *21*(2), 171–200.

Pellegrino, J. W., BAXTER, G. P., & Glaser, R. (1999). Chapter 9: Addressing the "Two Disciplines" Problem: Linking Theories of Cognition and Learning With Assessment and Instructional Practice. *Review of Research in Education*, *24*(1), 307–353.

Pellegrino, J. W., Chudowsky, N., & Glaser, R. (Eds.). (2001). *Knowing what students know: The science and design of educational assessment*. Washington, DC: National Academies Press.

PRIMAS project. *Promoting inquiry in science and mathematics education across Europe: What does inquiry-based learning mean?* Retrieved from http://www.primas-project.eu/artikel/en/1302/What+exactly+does+inquiry-based+learning+mean/view.do?lang=en

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.

Sadler, R. (1989). Formative assessment and the design of instructional systems. *Instructional Science*, *18*(2), 119–144.

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

Scriven, M. (1967). The methodology of evaluation. In R. W. Tyler, R. M. Gagné, & M. Scriven (Eds.), *Monograph Series on Educational Evaluation: Vol. 1. Perspectives of curriculum evaluation* (pp. 39–83). Chicago: Rand McNally.

Shavelson, R. J., Young, D. B., Ayala, C. C., Brandon, P. R., Furtak, E. M., Ruiz-Primo, M. A., … (2008). On the Impact of Curriculum-Embedded Formative Assessment on Learning: A Collaboration between Curriculum and Assessment Developers. *Applied Measurement in Education*, *21*(4), 295–314.

Shepard, L. A. (2000). The Role of Assessment in a Learning Culture. *Educational Researcher*, *29*(7), 4–14.

Siegel, M. A., Hynds, P., Siciliano, M., & Nagle, B. (2006). Using rubrics to foster meaningful learning. In M. McMahon, P. Simmons, R. Sommers, D. DeBeats, & F. Crawley (Eds.), *Assessment in science. Practical experiences and education research* (pp. 89–106). Arlington, Va: NSTA Press.

Stoddart, T., Abrams, R., Gasper, E., & Canaday, D. (2000). Concept maps as assessment in science inquiry learning - a report of methodology. *International Journal of Science Education*, *22*(12), 1221–1246.

Wiliam, D. (2006). Formative assessment: Getting the focus right. *Educational Assessment*, *11*(3-4), 283–289.

Williams, J., & Ryan, J. (2000). National Testing and the Improvement of Classroom Teaching: Can they coexist? *British Educational Research Journal*, *26*(1), 49–73.

Wilson, M., & Sloane, K. (2000). From Principles to Practice: An Embedded Assessment System. *Applied Measurement in Education*, *13*(2), 181–208.

Appendix 1: List of publications found so far (26.02.2013)

Author/Editor	Year	Title	Place	Publisher	Journal	Vol.	No.	pages
		Web of Inquiry (WOI)						
		WISE web-based inquiry science environment		WISE 4.0 open source technology				
	2008	FAST Legacy index						
	2011	Mathematics education in Europe	Brussels	Education, Audiovisual and Culture Executive Agency				
	2011	Science education in Europe	Brussels	Education, Audiovisual and Culture Executive Agency				
Abd El Khalick, Fouad et al.	2004	Inquiry in Science Education: International Perspectives.			Science Education	88	3	397–419
Adams, Dennis; Hamm, Mary	2008	Helping students who struggle with math and science: A collaborative approach for elementary and middle schools.	Blue Ridge Summit	Rowman & Littlefield Education				

Akerson, Valarie L.; Hanuscin, Deborah L.	2007	Teaching nature of science through inquiry: Results of a 3-year professional development program.			Journal of Research in Science Teaching	44	5	653–680
Alkharusi, Hussain	2011	Psychometric properties of the teacher assessment literacy questionnaire for preservice teachers in Oman			Procedia - Social and Behavioral Sciences	29		1614–1624
American Association for the Advancement of Science	1993	Benchmarks for science literacy: Project 2061	New York	Oxford University Press				
American Association for the Advancement of Science	1998	Blueprints for Reform - Project 2061	Washington, DC					
American Association for the Advancement of Science	2009	Benchmarks for Science Literacy						
American Federation of Teachers	1990	Standards for teacher competences in educational assessment of students			Educational Measurement: Issues and Practice	9	4	30
Anderson, Kate T. et al.	2007	Classroom Discourse as a Tool to Enhance Formative Assessment and Practise in Science			International Journal of Science Education	29	14	1721–1744
Anderson, Ronald D.	2002	Reforming Science Teaching: What Research Says About Inquiry			Journal of Science Teacher Education	13	1	1–12
Artigue, Michèle	1994	Didactical Engineering as a framework for the conception of teaching products						27–39
Artigue, Michèle; Baptist, Peter	2012	Inquiry in Mathematics Education						

Artigue, Michèle et al.	2012	Learning through inquiry						
Assessment Reform Group	2002	Assessment for Learning						
Atkin, J. Myron; Coffey, Janet E.	2003	Everyday Assessment in the Science Classroom	Arlington	NSTA Press				
Avraamidou, Lucy; Zembal Saul, Carla	2010	In search of well-started beginning science teachers: Insights from two first-year elementary teachers.			Journal of Research in Science Teaching	47	6	661–686
Ayala, Carlos C. et al.	2008	From Formal Embedded Assessments to Reflective Lessons: The Development of Formative Assessment Studies			Applied Measurement in Education	21	4	315–334
Aydeniz, Mehmet; Southerland, Sherry A.	2012	A National Survey of Middle and High School Science Teachers' Responses to Standardized Testing: Is Science Being Devalued in Schools?			Journal of Science Teacher Education	23	3	233–257
Bangert-Drowns, R. L. et al.	1991	The Instructional Effect of Feedback in Test-Like Events			Review of Educational Research	61	2	213–238
Barron, Brigid J.S. et al.	1998	Doing With Understanding: Lessons From Research on Problem- and Project-Based Learning			Journal of the Learning Sciences	7	3-4	271–311
Bedford, Simon; Legg, Serena	2007	Formative peer and self feedback as a catalyst for change within schience teaching			Chemistry Education Research and Practice	8	1	80–92
Bell, Beverley; Cowie, Bronwen	2001	The characteristics of formative assessment in science education			Science Education	85	5	536–553

Bell, Philip; Linn, Marcia C.	2002	Beliefs about science: How does science instruction contribute?						321–346
Bell, Thorsten et al.	2010	Collaborative Inquiry Learning: Models, tools, and challenges			International Journal of Science Education	32	3	349–377
Bennett, Randy Elliot	2011	Formative assessment: a critical review			Assessment in Education: Principles, Policy & Practice	18	1	5–25
Berland, Leema Kuhn; Reiser, Brian J.	2009	Making sense of argumentation and explanation. [References]			Science Education	93	1	
Berry, Rita	2011	Assessment trends in Hong Kong: seeking to establish formative assessment in an examination culture			Assessment in Education: Principles, Policy & Practice	18	2	199–211
Biehler, Rolf	1994	Didactics of mathematics as a scientific discipline	Dordrecht	Kluwer Academic Publishers				
Binkley, Marilyn et al.	2012	Defining twenty-first century skills						17–66
Birenbaum, Menucha et al.	2006	A learning integrated assessment system			Educational Research Review	1		61–67
Birenbaum, Menucha	2002	Assessing Self-directed Active Learning in Primary Schools			Assessment in Education: Principles, Policy & Practice	9	1	119–138
Black, Paul	2003	Assessment for learning	Maidenhead	Open University Press				
Black, Paul et al.	2004	Working inside the Black Box: Assessment for Learning in the Classroom			Phi Delta Kappan	86	1	8–21

Black, Paul	1995	Curriculum and assessment in science education: the policy interface			International Journal of Science Education	17	4	453–469
Black, Paul	2003	The Importance of Everyday Assessment						1–12
Black, Paul et al.	2010	Validity in teachers' summative assessments	Abingdon	Routledge, Taylor & Francis Group	Assessment in education : principles, policy & practice	17	2	215–232
Black, Paul et al.	2011	Can teachers' summative assessments produce dependable results and also enhance classroom learning?			Assessment in Education: Principles, Policy & Practice	18	4	451–469
Black, Paul et al.	2006	Learning How to Learn and Assessment for Learning: a theoretical inquiry			Research Papers in Education	21	2	119–132
Black, Paul; Wiliam, Dylan		Inside the Black Box	London					
Black, Paul; Wiliam, Dylan	1998	Assessment and Classroom Learning			Assessment in Education: Principles, Policy & Practice	5	1	7–74
Black, Paul; Wiliam, Dylan	2005	Lessons from around the world: how policies, politics and cultures constrain and afford assessment practices			Curriculum Journal	16	2	249–261
Black, Paul; Wiliam, Dylan	2006	Developing a theory of formative assessment						81–100
Blanchard, Margaret R. et al.	2009	No silver bullet for inquiry: Making sense of teacher change following an inquiry-based research experience for teachers.			Science Education	93	2	322–360

Blanchard, Margaret R. et al.	2010	Is inquiry possible in light of accountability?: A quantitative comparison of the relative effectiveness of guided inquiry and verification laboratory instruction	Sci. Ed. (Science Education)	94	4	577–616
Bloom, B. S.	1969	Some theoretical issues relating to educational evaluation		68 (2)		26–50
Brandon, Paul R. et al.	2009	THE INQUIRY SCIENCE IMPLEMENTATION SCALE: DEVELOPMENT AND APPLICATIONS	Int J of Sci and Math Educ (International Journal of Science and Mathematics Education)	7	6	1135–1147
Brandon, Paul R. et al.	2008	Lessons Learned from the Process of Curriculum Developers' and Assessment Developers' Collaboration on the Development of Embedded Formative Assessments	Applied Measurement in Education	21	4	390–402
Breslyn, Wayne; McGinnis, J. Randy	2012	A comparison of exemplary biology, chemistry, earth science, and physics teachers' conceptions and enactment of inquiry.	Science Education	96	1	48–77
Briggs, Michael et al.	2011	Qualitative Assessment of Inquiry-Based Teaching Methods	J. Chem. Educ. (Journal of Chemical Education)	88	8	1034–1040
Broadfoot, Patricia; Black, Paul	2004	Redefining assessment? The first ten years of assessment in education	Assessment in Education: Principles, Policy & Practice	11	1	7–26
Brookhart, Susan M.	2011	Educational Assessment Knowledge and Skills for Teachers	Educational Measurement: Issues and Practice	30	1	3–12

Brousseau, G. et al.	1997	Theory of didactical situations in mathematics	Dordrecht	Kluwer Academic Publishers				
Brown, Evelyn	2008	Removing the grade from a formative assessment						
Buck, Gayle A. et al.	2010	Making formative assessment discernable to pre-service teachers of science.			Journal of Research in Science Teaching	47	4	402–421
Buckley, B. C. et al.	2004	Model-based teaching and learning with BioLogica: What do they learn? How do they learn? How do we know?			Journal of Science Education and Technology	13	1	23–41
Buczynski, Sandy; Hansen, C. Bobbi	2010	Impact of professional development on teacher practice: Uncovering connections.			Teaching and Teacher Education	26	3	599–607
Burns, Joseph C. et al.	1985	Development of an integrated process skill test: TIPS II			Journal of Research in Science Teaching	22	2	169–177
Cáceres, M.J. et al.	2010	Analysis of the revisions that pre-service teachers of mathematics make of their own project included in their learning portfolio			Teaching and Teacher Education	26	5	1186–1195
Cambridge Assessment Network	2006	EXCELLENCE IN ASSESSMENT: Assessment for Learning						
Carless, David	2005	Prospects for the implementation of assessment for learning			Assessment in Education: Principles, Policy & Practice	12	1	39–54
Carlson, Maura O'Brien et al.	2003	Weaving science inquiry and continuous assessment: Using formative assessment to improve learning	Thousand Oaks	Corwin Press				

Carver, Sharon M.; Shrager, Jeff	2012	The journey from child to scientist	Washington, DC	American Psychological Association				
Centre for Educational Research and Innovation (CERI)	2008	Assessment for Learning						
Chang, Huey-Por et al.	2011	The development of a competence scale for learning science: Inquiry and communication			International Journal of Science and Mathematics Education	9	5	1213–1233
Chen, Chang-Hua et al.	2012	Eight grade mathematics teachers' formative assessment practices in SES-different classrooms: A Taiwan study			International Journal of Science and Mathematics Education	10	3	553–579
Chin, Christine; Teou, Lay-Yen	2010	Formative assessment: Using concept cartoon, pupils' drawings, and group discussions to tackle children's ideas about biological inheritance			Journal of Biological Education	44	3	108–115
Chinn, Clark A.; Malhotra, Betina A.	2002	Epistemologically authentic inquiry in schools: A theoretical framework for evaluating inquiry tasks			Science Education	86	2	175–218
Chudowsky, N.; Pellegrino, J.W	2003	Large-scale assessments that support learning: what will it take?			Theory into Practice	42	1	75–83
Clark, Ian	2010	Formative Assessment: 'There is nothing so practical as a good theory'			Australian Journal of Education	54	3	
Cobb, Paul et al.	1992	Characteristics of Classroom Mathematics Traditions: An Interactional Analysis			American Educational Research Journal	29	3	573–604
Cobern, William W. et al.	2010	Experimental comparison of inquiry and direct instruction in science.			Research in Science & Technological Education	28	1	

Coffey, Janet et al.	2008	Assessing science learning	Arlington	National Science Teachers Association (NSTA)				
Coffey, Janet E. et al.	2011	The missing disciplinary substance of formative assessment			Journal of Research in Science Teaching	48	10	1109–1136
Co-Lab EU project		Co-Lab		Co-Lab EU project				
Cowie, Bronwen; Bell, Beverley	1999	A Model of Formative Assessment in Science Education			Assessment in Education: Principles, Policy & Practice	6	1	101–116
Crawford, Barbara A.	2007	Learning to teach science as inquiry in the rough and tumble of practice.			Journal of Research in Science Teaching	44	4	613–642
Crooks, Terence J.	1988	The Impact of Classroom Evaluation Practices on Students			Review of Educational Research	58	4	438–481
Delandshere, Ginette	2002	Assessment as Inquiry			Teachers College Record	104	7	1461–1484
Denzin, Norman K.; Lincoln, Yvonna S.	2005	The Sage handbook of qualitative research	Thousand Oaks	Sage Publications Ltd				
Dillon, Justin; Maguire, Meg	2011	Becoming a teacher	Maidenhead	Open University; McGraw-Hill Education				
Dolin, Jens	2012	Assess Inquiry in Science, Technology and Mathematics Education						
Duschl, Richard A.	2003	Assessment of Inquiry						41–60

Ebenezer, Jazlin et al.	2011	Engaging students in environmental research projects: Perceptions of fluency with innovative technologies and levels of scientific inquiry abilities.			Journal of Research in Science Teaching	48	1	
Eckersall, Ken	1987	Some developments in technology education in Australia						371–376
EDUCERI21ST	2008	Assessment for Learning: Formative Assessment						
Elwood, Jannette	2006	Formative assessment: possibilities, boundaries and limitations			Assessment in Education: Principles, Policy & Practice	13	2	215–232
Erdogan, Ibrahim et al.	2011	The Student Actions Coding Sheet (SACS): An instrument for illuminating the shifts toward student-centered science classrooms			International Journal of Science Education	33	10	1313–1336
European Commission	2004	Increasing human resources for science and technology in Europe	Luxembourg	Office for Official Publications of the European Communities				
European Commission	2007	Key competences for lifelong learning. European reference framework	Luxembourg					
European Commission	2007	Science education now	Luxembourg	Office for Official Publications of the European Communities				

European Parliament, Council	2006	Key competences for lifelong learning				
Falk, Andrew	2012	Teachers learning from professional development in elementary science: Reciprocal relations between formative assessment and pedagogical content knowledge	Science Education	96	2	265–290
Fan, Ya-Ching et al.	2011	A Web-based model for developing assessment literacy of secondary in-service teachers	Computers & Education	57	2	1727–1740
Ferreira, Maria M.; Trudel, Anthony R.	2012	The impact of problem-based learning (PBL) on students attitudes towards science, problem-solving skills, and sense of community in the classroom.	Journal of Classroom Interaction	47	1	23–30
Fibonacci project		Disseminating inquiry-based science and mathematics education in Europe				
Firestone, William A. et al.	2000	Different Assessments, Common Practice? Mathematics testing and teaching in the USA and England and Wales	Assessment in Education: Principles, Policy & Practice	7	1	13–37
Fogleman, Jay et al.	2011	Examining the effect of teachers' adaptations of a middle school science inquiry-oriented curriculum unit on student learning. [References]	Journal of Research in Science Teaching	48	2	149–169
Forbes, Ellen White	2007	Improving the knowledge and use of formative assessment: A case study of formative assessment in a K-3 science curriculum				

Freudenthal, H.	1973	Mathematics as an educational task	Dordrecht	Kluwer Academic Publishers				
Freudenthal Institute	1998	Realistic Mathematics Education						
Furtak, Erin Marie	2006	The problem with answers: An exploration of guided scientific inquiry teaching			Science Education	90	3	453–467
Furtak, Erin Marie et al.	2008	On the Fidelity of Implementing Embedded Formative Assessments and Its Relation to Student Learning			Applied Measurement in Education	21	4	360–389
Furtak, Erin Marie et al.	2012	Experimental and Quasi-Experimental Studies of Inquiry-Based Science Teaching: A Meta-Analysis			Review of Educational Research	82	3	300–329
Furtak, Erin Marie et al.	2012	To teach or not to teach through inquiry: Is that the question?						227–244
Gallavan, Nancy P.	2009	Developing performance-based assessments	Thousand Oaks	Corwin Press				
Gardner, John	2006	Assessment and Learning	London	Sage Publications				
Gardner, John et al.	2010	Developing teacher assessment	Maidenhead	Open University Press				
Gee, Kevin A.; Wong, Kenneth K.	2012	A cross national examination of inquiry and its relationship to student performance in science: Evidence from the Program for International Student Assessment (PISA) 2006			International Journal of Science Education	53		303–318

Gerard, Libby F. et al.	2010	Teacher use of evidence to customize inquiry science instruction			Journal of Research in Science Teaching	47	9	1037–1063
Gersten, Russell et al.	2009	Mathematics Instruction for Students With Learning Disabilities: A Meta-Analysis of Instructional Components			Review of Educational Research	79	3	1202–1242
Gioka, Olga	2009	Teacher or Examiner? The Tensions between Formative and Summative Assessment in the Case of Science Coursework			Research in Science Education	39	4	411–428
Gitomer, D. H.; Duschl, R. A.	1995	Moving toward a portfolio culture in science education						299–326
Gläser-Zikuda, Michaela	2007	Lernprozesse dokumentieren, reflektieren und beurteilen: Lerntagebuch und Portfolio in Bildungsforschung und Bildungspraxis	Bad Heilbrunn	Klinkhardt				
Glasgow, Neal A. et al.	2010	What successful science teachers do: 75 research-based strategies	Thousand Oaks	Corwin Press				
Glynn, Shawn M.; Duit, Reinders	1995	Learning science in the schools: Research reforming practice	Mahwah	Erlbaum				
Gobert, Janice D. et al.	2010	Unpacking inquiry skills from content knowledge in geoscience: a research and development study with implications for assessment design			International Journal of Learning Technology	5	3	310
Goldston, M. Jenice et al.	2010	Psychometric analysis of a 5E learning cycle lesson plan assessment instrument			International Journal of Science and Mathematics Education	8	4	633–648

Goodnough, K.; Long, R.	2006	Mind mapping as a flexible assessment tool						219–228
Griffin, Patrick E. et al.	2012	Assessment and teaching of 21st century skills	Dordrecht	Springer				
Gröschner, Alexander et al.	2010	Baseline report and indicators review on science teaching methods and attitudes	München					
Hammerman, Elizabeth et al.	2008	Integrating science with mathematics & literacy: New visions for learning and assessment	Thousand Oaks	Corwin Press				
Hanauer, David I. et al.	2009	Active assessment: Assessing scientific inquiry	New York	Springer				
Harlen, Wynne	2007	The Quality of Learning: Assessment Alternatives for Peimary Education					3/4	
Harlen, Wynne	2009	Teaching and learning science for a better future			School Science Review	90	333	33–41
Harlen, Wynne; Deakin Crick, Ruth	2003	Testing and Motivation for Learning			Assessment in Education: Principles, Policy & Practice	10	2	169–207
Harlen, Wynne; James, Mary	1997	Assessment and Learning: differences and relationships between formative and summative assessment			Assessment in Education: Principles, Policy & Practice	4	3	365–379
Harris, C.J et al.	2006	Usable assessments for teaching science content and inquiry standards						67–87
Harrison, Christine	2011	Making assessments work in the classroom						222–235

Hattie, John; Timperley, Helen	2007	The Power of Feedback			Review of Educational Research	77	1	81–112
Heinz, Jana	2012	Indicators and instruments in the context of inquiry-based science education	Münster	Waxmann				
Heinz, Jana et al.	2010	Review of indicators and instruments in the context of inquiry-based science teaching	Trondheim					
Heritage, Margaret et al.	2009	From Evidence to Action: A Seamless Process in Formative Assessment?			Educational Measurement: Issues and Practice	28	3	24–31
Heritage, Margaret; Niemi, David	2006	Toward a Framework for Using Student Mathematical Representations as Formative Assessments			Educational Assessment	11	3-4	265–282
Herman, Joan L. et al.	2010	Capturing quality in formative assessment practice: Measurement challenges	Los Angeles					
Herrenkohl, Leslie Rupert et al.	2011	Pedagogical practices to support classroom cultures of scientific inquiry			Cognition and Instruction	29	1	
Hiebert, James et al.	1996	Problem Solving as a Basis for Reform in Curriculum and Instruction: The Case of Mathematics			Educational Researcher	25	4	12–21
Hill, Mary F.	2011	'Getting traction': enablers and barriers to implementing Assessment for Learning in secondary schools			Assessment in Education: Principles, Policy & Practice	18	4	347–364
Hmelo-Silver, Cindy E. et al.	2007	Scaffolding and Achievement in Problem-Based and Inquiry Learning: A Response to Kirschner, Sweller, and Clark (2006)			Educational Psychologist	42	2	99–107

Hofer, Barbara K. [Ed] et al.	2002	Personal epistemology: The psychology of beliefs about knowledge and knowing	Mahwah	Lawrence Erlbaum Associates Publishers				
Huang, Chenn Jung et al.	2011	Performance evaluation of an online argumentation learning assistance agent.			Computers & Education	57	1	1270–1280
Hume, Anne; Coll, Richard	2010	Authentic student inquiry: The mismatch between the intended curriculum and the student-experienced curriculum.			Research in Science & Technological Education	28	1	43–62
Hunter, Roberta; Anthony, Glenda	2011	Forging Mathematical Relationships in Inquiry-Based Classrooms With Pasifika Students			Journal of Urban Mathematics Education	4	1	98–119
INQUIRE project	2010	Taking IBSE into secondary education	York					
Inquiry in Motion	2009	EQUIP (Electronioc Quality of Inquiry Protocol)						
International Technology Education Association	1996	Technology for all Americans:						
International Technology Education Association	2007	Standards for Technological Literacy		International Technology Education Association				
J. Nelissen; W. Tomic	1993	Learning and thought processes in realistic mathematics instruction			Curriculum and Teaching	8	1	
Johnson, Ruth S. et al.	2006	Developing portfolios in education	Thousand Oaks	Sage Publications				

Jonsson, Anders; Svingby, Gunilla	2007	The use of scoring rubrics: Reliability, validity and educational consequences			Educational Research Review	2	2	130–144
Kaberman, Zvia; Dori, Yehudit Judy	2009	QUESTION POSING, INQUIRY, AND MODELING SKILLS OF CHEMISTRY STUDENTS IN THE CASE-BASED COMPUTERIZED LABORATORY EN-VIRONMENT			International Journal of Science and Mathematics Education	7	3	597–625
Kalantzis, Mary et al.	2003	Assessing Multiliteracies and the New Basics			Assessment in Education: Principles, Policy & Prac-tice	10	1	15–26
Kessler, James H.; Galvan, Patricia M.	2007	Inquiry in Action		American Chemical Soci-ety				
Ketelhut, Diane Jass; Nelson, Brian C.	2010	Designing for real-world scientific inquiry in virtual environments. [References]			Educational Research	52	2	151–167
Kirschner, Paul A. et al.	2006	Why Minimal Guidance During Instruc-tion Does Not Work: An Analysis of the Failure of Constructivist, Discovery, Problem-Based, Experiential, and In-quiry-Based Teaching			Educational Psychologist	41	2	75–86
Klahr, David et al.	2007	Hands on what? The relative effective-ness of physical versus virtual materials in an engineering design project by middle school children			Journal of Research in Science Teaching	44	1	183–203
Klieme, Eckhart et al.	2008	Assessment of competencies in educa-tional settings	Cambridge	Hogrefe & Huber Publish-ers				

Kluger, Avraham N.; DeNisi, Angelo	1996	The effects of feedback interventions on performance: a historical review, a meta-analysis, and a preliminary feedback intervention theory		Psychological Bulletin	119	2	254–284
Koretz, Daniel	1998	Large-scale Portfolio Assessments in the US: evidence pertaining to the quality of measurement		Assessment in Education: Principles, Policy & Practice	5	3	309–334
Krajcik, Joseph et al.	2008	Learning-goals-driven design model: Developing curriculum materials that align with national standards and incorporate project-based pedagogy		Science Education	92	1	1–32
Kwon, Oh Nam et al.	2006	Cultivating divergent thinking in mathematics through an open-ended approach		Asia Pacific Educational Review	7	1	51–61
Learning how to Learn Project	2002	Learning how to learn Homepage	Cambridge				
Lederman, N. et al.	1998	Assessing understanding of the nature of science: A historical perspective					331–350
Lin, Fou-Lai et al.	2004	THE FEATURES AND RELATIONSHIPS OF REASONING, PROVING AND UNDERSTANDING PROOF IN NUMBER PATTERNS			2	2	227–256
Linn, M. C.	2006	INQUIRY LEARNING: Teaching and Assessing Knowledge Integration in Science		Science	313	5790	1049–1050
Linn, Marcia C. et al.	2003	WISE design for knowledge integration		Science Education	87	4	517–538
Linn, R. L.	2000	Assessments and Accountability		Educational Researcher	29	2	4–16

Liu, Ou Lydia et al.	2010	An investigation of teacher impact on student inquiry science performance using a hierarchical linear model. [References]			Journal of Research in Science Teaching	47	7	807–819
Liu, Ou Lydia et al.	2010	Multifaceted assessment of inquiry-based science learning. [References]			Educational Assessment	15	2	
Liu, Ou Lydia et al.	2008	Assessing Knowledge Integration in Science: Construct, Measures, and Evidence			Educational Assessment	13	1	33–55
Liu, Tzu-Chien et al.	2009	IDENTIFYING SENIOR HIGH SCHOOL STUDENTS' MISCONCEPTIONS ABOUT STATISTICAL CORRELATION, AND THEIR POSSIBLE CAUSES: AN EXPLORATORY STUDY USING CONCEPT MAPPING WITH INTERVIEWS			International Journal of Science and Mathematics Education	7	4	791–820
Llewellyn, Douglas	2002	Inquire within: Implementing inquiry-based science standards. [References]	Thousand Oaks	Corwin Press				
Looney, Janet W.	2011	Integrating Formative and Summative Assessment: Progress Toward a Seamless System?					58	
Lorenzo, Mercedes	2005	THE DEVELOPMENT, IMPLEMENTATION, AND EVALUATION OF A PROBLEM SOLVING HEURISTIC			International Journal of Science and Mathematics Education	3	1	33–58
Maier, Uwe	2010	Formative Assessment – Ein erfolgversprechendes Konzept zur Reform von Unterricht und Leistungsmessung?			Zeitschrift für Erziehungswissenschaft	13	2	293–308

Marshall, Jeff C. et al.	2010	THE DESIGN AND VALIDATION OF EQUIP: AN INSTRUMENT TO ASSESS INQUIRY-BASED INSTRUCTION			International Journal of Science and Mathematics Education	8	2	299–321
McComas, W. F.	1998	The nature of science in science education	Dordrecht	Kluwer Academic Publishers				
McMahon, M. et al.	2006	Assessment in science: Practical experiences and education research	Arlington	NSTA Press				
McNeill, Katherine L.	2009	Teachers' use of curriculum to support students in writing scientific arguments to explain phenomena. [References]			Science Education	93	2	233–268
Meier, Sherry L. et al.	2006	Teachers' use of rubrics to score non-traditional tasks: factors related to discrepancies in scoring			Assessment in Education: Principles, Policy & Practice	13	1	69–95
Mertler, C. A.	2004	Secondary teachers’ assessment literacy: Does classroom experience make a difference?			American Secondary Education	33	1	49–64
Mertler, C. A.; Campbell, C.	2005	Measuring Teachers' Knowledge & Application of Classroom Assessment Concepts: Development of the "Assessment Literacy Inventory"	Montreal					
Mertler, C.A		Classroom Assessment Literacy Inventory						
Minner, Daphne D. et al.	2010	Inquiry-based science instruction-What is it and does it matter? Results from a research synthesis years 1984 to 2002. [References]			Journal of Research in Science Teaching	47	4	474–496

Mislevy, R. J. et al.	2001	Levering point for improving educational assessment	Los Angeles				534	
Mislevy, Robert J. et al.	2003	Design Patterns for Assessing Science Inquiry	Menlo Park	SRI International, Center for Technology in Learning				
Mislevy, Robert J.; Haertel, Geneva D.	2006	Implications of Evidence-Centered Design for Educational Testing			Educational Measurement: Issues and Practice	25	4	6–20
Mistler Jackson, Megan; Songer, Nancy Butler	2000	Student motivation and internet technology: Are students empowered to learn science?			Journal of Research in Science Teaching	37	5	459–479
Morrison, Judith A. et al.	2005	PRESERVICE TEACHERS' DEVELOPMENT AND IMPLEMENTATION OF SCIENCE PERFORMANCE ASSESSMENT TASKS			International Journal of Science and Mathematics Education	3	3	379–406
Mui So, Winnie Wing; Hoi Lee, Theodore Tai	2011	Influence of teachers' perceptions of teaching and learning on the implementation of Assessment for Learning in inquiry study			Assessment in Education: Principles, Policy & Practice	18	4	417–432
Mullis, I. V. A. et al.	2009	TIMSS 2011 Assessment Frameworks	Boston	TIMSS & PIRLS International Study Center				
Munns, Geoff; Woodward, Helen	2006	Student engagement and student self-assessment: the REAL framework			Assessment in Education: Principles, Policy & Practice	13	2	193–213

NAEP 2014	2012	2014 NAEP Technology and Engineering Literacy Framework						
National Research Council	1996	National Science Education Standards	Washington, DC	The National Academies Press				
National Research Council	2000	Inquiry and the National Science Education Standards	Washington, DC	National Academy Press				
National Research Council	2012	A framework for K-12 science education	Washington, DC	The National Academies Press				
Neber, Heinz; Schommer-Aikins, Marlene	2002	Self-regulated Science Learning with Highly Gifted Students: The role of cognitive, motivational, epistemological, and environmental variables			High Ability Studies	13	1	59–74
Nichols, Paul D. et al.	2009	A Framework for Evaluating and Planning Assessments Intended to Improve Student Achievement			Educational Measurement: Issues and Practice	28	3	14–23
Nohda, Nobuhiko	2000	Teaching by Open-Approach Method in Japanese Mathematics Classroom			Proceedings of the Conference of the International Group for the Psychology of Mathematics Education (PME)		1	39–53
OECD	2002	Definition and selection of competences (DeSeCo)						
OECD	2005	Formative Assessment	Paris	OECD				
OECD	2006	Assessing scientific, reading and mathematical literacy	Paris	OECD				

Offerdahl, Erika G.; Tomanek, Debra	2011	Changes in instructors' assessment thinking related to experimentation with new strategies			Assessment & Evaluation in Higher Education	36	7	781–795
O'Neil, Harold F. et al.	2003	Issues in the Computer-based Assessment of Collaborative Problem Solving			Assessment in Education: Principles, Policy & Practice	10	3	361–373
Pedder, David	2006	Organizational conditions that foster successful classroom promotion of Learning How to Learn			Research Papers in Education	21	2	171–200
Pellegrino, J. W. et al.	1999	Chapter 9: Addressing the "Two Disciplines" Problem: Linking Theories of Cognition and Learning With Assessment and Instructional Practice			Review of Research in Education	24	1	307–353
Pellegrino, J. W. et al.	2001	Knowing what students know: The science and design of educational assessment	Washington, DC	National Academies Press				
Phelan, Julia C. et al.	2012	The effects of POWERSOURCE © assessments on middle-school students’ math performance			Assessment in Education: Principles, Policy & Practice	19	2	211–230
Pine, Jerome et al.	2006	Fifth graders' science inquiry abilities: A comparative study of students in hands-on and textbook curricula			Journal of Research in Science Teaching	43	5	467–484
PRIMAS project	2013	Promoting inquiry in science and mathematics education across Europe						
Program in Education, Afterschool and Resiliency (PEAR)	2013	Discovery Inquiry Test in Science (DIT)						

Program in Education, Afterschool and Resiliency (PEAR)	2013	Test of Science Related Attitudes (TOSRA)						
Program in Education, Afterschool and Resiliency (PEAR)	2013	Views of Scientific Inquiry, Primary School Version (VOSI-P)						
Quellmalz, Edys S.; Pellegrino, James W.	2009	Technology and Testing			Science	323		75–79
Quellmalz, Edys S. et al.	2012	Science assessments for all: Integrating science simulations into balanced state science assessment systems			Journal of Research in Science Teaching	49	3	363–393
R. Hunter; G. Anthony	2011	Forging Mathematical Relationships in Inquiry-Based Classrooms With Pasifika Students			Journal of Urban Mathematics Education	4	1	98–119
Ridgway, Jim; McCusker, Sean	2003	Using Computers to Assess New Educational Goals			Assessment in Education: Principles, Policy & Practice	10	3	309–328
Riquards, Kurt	1987	Science and Technology Education and the Quality of Life - Vol. 2	Kiel	IPN				
Rivet, Ann E.; Krajcik, Joseph S.	2004	Achieving Standards in Urban Systemic Reform: An Example of a Sixth Grade Project-Based Science Curriculum. [References]			Journal of Research in Science Teaching	41	7	669–692
Ruiz Primo, Maria Araceli; Furtak, Erin Marie	2007	Exploring Teachers' Informal Formative Assessment Practices and Students' Understanding in the Context of Scientific Inquiry. [References]			Journal of Research in Science Teaching	44	1	

Ruiz Primo, Maria Araceli et al.	2010	Testing one premise of scientific inquiry in science classrooms: Examining students' scientific explanations and student learning. [References]	Journal of Research in Science Teaching	47	5	583–608
Ruiz-Primo, Maria Araceli	2011	Informal formative assessment: The role of instructional dialogues in assessing students' learning	Studies in Educational Evaluation	37	1	15–24
Ruiz-Primo, Maria Araceli; Furtak, Erin Marie	2006	Informal Formative Assessment and Scientific Inquiry: Exploring Teachers' Practices and Student Learning	Educational Assessment	11	3-4	205–235
Ruiz-Primo, Maria Araceli; Furtak, Erin Marie	2007	Exploring teachers' informal formative assessment practices and students' understanding in the context of scientific inquiry	Journal of Research in Science Teaching	44	1	57–84
Ruiz-Primo, Maria Araceli et al.	2010	Testing one premise of scientific inquiry in science classrooms: Examining students' scientific explanations and student learning	Journal of Research in Science Teaching	47	5	583–608
Ruiz-Primo, Maria Araceli et al.	2012	Developing and evaluating instructionally sensitive assessments in science	Journal of Research in Science Teaching	49	6	691–712
Ruiz-Primo, Maria A. et al.	2004	Evaluating students' science notebooks as an assessment tool	International Journal of Science Education	26	12	1477–1506
Sadler, D. Royce	1989	Formative assessment and the design of instructional systems	Instructional Science	18	2	119–144
Sadler, D. Royce	1998	Formative Assessment: revisiting the territory	Assessment in Education: Principles, Policy & Practice	5	1	77–84

Samarapungavan, Ala et al.	2011	What kindergarten students learn in inquiry-based science classrooms. [References]			Cognition and Instruction	29	4	416–470
Santau, Alexandra O. et al.	2011	Science achievement of English language learners in urban elementary schools: Fourth-grade student achievement results from a professional development intervention. [References]			Science Education	95	5	771–793
Saunders-Stewart, K. S. et al.	2012	Student Outcomes in Inquiry Instruction: A Literature-Derived Inventory			Journal of Advanced Academics	23	1	5–31
Scardamalia, Marlene et al.	2012	New Assessments and Environments for Knowledge Building						231–300
Schoenfeld, A. H.	1985	Mathematical problem solving	San Diego	Academic Press				
Schwartz, R. S. et al.	2008	An Instrument To Assess Views Of Scientific Inquiry: The VOSI Questionnaire	Baltimore					
Schwarz, Christina V.; White, Barbara Y.	2005	Metamodeling Knowledge: Developing Students' Understanding of Scientific Modeling			Cognition and Instruction	23	2	165–205
Scriven, M.	1967	The methodology of evaluation				1		39–83
Shavelson, Richard J. et al.	2008	On the Impact of Curriculum-Embedded Formative Assessment on Learning: A Collaboration between Curriculum and Assessment Developers			Applied Measurement in Education	21	4	295–314
Shepard, L. A.	2000	The Role of Assessment in a Learning Culture			Educational Researcher	29	7	4–14

Shepard, Lorrie A.	2003	Reconsidering Large-Scale Assessment to Heighten Its Relevance to Learning				121–146
Shepard, Lorrie A.	2009	Commentary: Evaluating the Validity of Formative and Interim Assessment	Educational Measurement: Issues and Practice	28	3	32–37
Siegel, M. A. et al.	2006	Using rubrics to foster meaningful learn-ing				89–106
Smith, M. L.	1991	Put to the Test: The Effects of External Testing on Teachers	Educational Researcher	20	5	8–11
Stecher, Brain	1998	The Local Benefits and Burdens of Large-scale Portfolio Assessment	Assessment in Education: Principles, Policy & Prac-tice	5	3	335–351
Stecher, Brian M. et al.	2000	The Effects of Content, Format, and Inquiry Level on Science Performance Assessment Scores	Applied Measurement in Education	13	2	139–160
Stieff, Mike	2011	Improving representational competence using molecular simulations embedded in inquiry activities. [References]	Journal of Research in Science Teaching	48	10	1137–1158
Stoddart, Trish et al.	2000	Concept maps as assessment in sci-ence inquiry learning - a report of meth-odology	International Journal of Science Education	22	12	1221–1246
Stokking, Karel et al.	2004	Teachers' assessment of students' research skills	British Educational Re-search Journal	30	1	93–116
Suurtamm, Christine et al.	2010	Teachers' assessment practices in mathematics: classrooms in the context of reform	Assessment in Education: Principles, Policy & Prac-tice	17	4	399–417

Swaffield, Sue	2011	Getting to the heart of authentic Assessment for Learning			Assessment in Education: Principles, Policy & Practice	18	4	433–449
Taylor, Amy R. et al.	2008	Creativity, inquiry, or accountability? Scientists' and teachers' perceptions of science education. [References]			Science Education	92	6	1058–1075
Thomas, Lynn et al.	2011	Elementary teachers' formative evaluation practices in an era of curricular reform in Quebec, Canada			Assessment in Education: Principles, Policy & Practice	18	4	381–398
Tierney, Robin D.	2006	Changing practices: influences on classroom assessment			Assessment in Education: Principles, Policy & Practice	13	3	239–264
Torrance, Harry	2007	Assessment as learning? How the use of explicit learning objectives, assessment criteria and feedback in post-secondary education and training can come to dominate learning.	Abingdon	Routledge, Taylor & Francis Group	Assessment in education : principles, policy & practice	14	3	281–294
Toth, Eva Erdosne et al.	2002	"Mapping to know": The effects of representational guidance and reflective assessment on scientific inquiry. [References]			Science Education	86	2	264–286
Trumbull, Deborah J. et al.	2005	Developing materials to promote inquiry: Lessons learned. [References]			Science Education	89	6	879–900
Tyler, R. W.	1969	Educational evaluation: New roles, new means	Chicago	University of Chicago Press		68 (2)		
Tyler, R. W. et al.	1967	Perspectives of curriculum evaluation	Chicago	Rand McNally		1		

Urhahne, Detlef et al.	2010	Role of the Teacher in Computer-supported Collaborative Inquiry Learning	International Journal of Science Education	32	2	221–243
van Joolingen, W.R. et al.	2005	Co-Lab: Research and development of an online learning environment for collaborative scientific discovery learning	Computers in Human Behavior	21	4	671–688
W. Harlen	2012	Inquiry in Science Education				
Wallace, Carolyn S.; Priestley, Mark	2011	Teacher beliefs and the mediation of curriculum innovation in Scotland: A socio-cultural perspective on professional development and change	Journal of Curriculum Studies	43	3	357–381
Webb, Mary; Jones, Jane	2009	Exploring tensions in developing assessment for learning	Assessment in Education: Principles, Policy & Practice	16	2	165–184
White, Barbara Y.; Frederiksen, John R.	1998	Inquiry, Modeling, and Metacognition: Making Science Accessible to All Students	Cognition and Instruction	16	1	3–118
Wiliam, D.	2006	Formative assessment: Getting the focus right	Educational Assessment	11	3-4	283–289
Wiliam, Dylan	2008	International comparisons and sensitivity to instruction	Assessment in Education: Principles, Policy & Practice	15	3	253–257
Wiliam, Dylan et al.	2004	Teachers developing assessment for learning: impact on student achievement	Assessment in Education: Principles, Policy & Practice	11	1	49–65
Williams, Anthony; Williams, P. John	1997	Problem-based learning: An appropriate methodology for technology education	Research in Science & Technological Education	15	1	91–103

Williams, Julian; Ryan, Julie	2000	National Testing and the Improvement of Classroom Teaching: Can they coexist?	British Educational Research Journal	26	1	49–73
Wilson, Christopher D. et al.	2010	The relative effects and equity of inquiry-based and commonplace science teaching on students' knowledge, reasoning, and argumentationargumentation. [References]	Journal of Research in Science Teaching	47	3	276–301
Wilson, Mark; Scalise, Kathleen	2003	Reporting Progress to Parents and Others: Beyond Grades				89–108
Wilson, Mark; Sloane, Kathryn	2000	From Principles to Practice: An Embedded Assessment System	Applied Measurement in Education	13	2	181–208
Winnie Wing-Mui So	2003	Learning Science through ivestigations: An experience with Hong Kong primary school children	International Journal of Science and Mathematics Education	1	2	175–200
Wirth, Joachim; Klieme, Eckhard	2003	Computer-based Assessment of Problem Solving Competence	Assessment in Education: Principles, Policy & Practice	10	3	329–345
Yin, Yue et al.	2008	On the impact of formative assessment on student motivation, achievement, and conceptual change. [References]	Applied Measurement in Education	21	4	335–359
Yoon, Cho Hee	2009	Self-regulated learning and instructional factors in the scientific inquiry of scientifically gifted Korean middle school students. [References]	Gifted Child Quarterly	53	3	203–216

Young, Betty J.; Lee, Sharon K.	2005	The effects of a kit-based science curriculum and intensive science professional development on elementary student science achievement	Journal of Science Education and Technology	14	05. Jun	471–481
Zhang, Baohui et al.	2005	Opportunities and challenges of China's inquiry-based education reform in middle and high schools: perspectives of science teachers and teacher educators	International Journal of Science and Mathematics Education	1	4	477–503
Zion, Michal et al.	2005	The effects of metacognitive instruction embedded within an asynchronous learning network on scientific inquiry skills	International Journal of Science Education	27	8	957–983