



Thinking Deeper Research Paper No.1 - Part 3

Making it Happen: Formative Assessment and Educational Technologies

Author: Janet Looney

Sponsored by: Promethean Education Strategy Group

Promethean Thinking Deeper Research Papers provide sponsorship for global experts to conduct and report on cutting-edge thinking and state-of-the-art global academic/policy knowledge with respect to the role of new technologies (innovative tools and organization) in meeting the challenges of learning in the 21st Century.



Introduction

No one component in education is more important for bringing together all of the elements needed for a leap in educational productivity than assessment. Assessment is perhaps the most cross cutting enabler of transformations in learning processes and outcomes. It touches all aspects of the learning process. Rich, multi-dimensional and timely feedback as one learns provides a way to refine and direct both the tools and objectives of learning. This Promethean Thinking Deeper paper provides a general overview of the nature and uses of assessment to support learning in education today. The research reviewed below shows not only why and how assessment plays a pivotal role in improving the overall productivity of learning processes but also the growing role of new technologies and the new organisational approaches these tools enable.

The Wired Classroom: Supporting Formative Assessment Through Technology

In many ways, this is an exciting time in the field of education. We are learning more about learning. We know more, for instance, about how students progress from novice to more expert levels of performance in different subject domains, and about how they develop sophisticated skills for problem solving and collaboration. We also know more about the importance of assessment as an integral part of the learning process (Bransford et al., 1999; Pellegrino et al., 1999).

This report focuses on the key role of formative assessment in supporting learning. Formative assessment refers to the frequent assessment of learner progress to identify learning needs and shape teaching. Innovations featuring formative assessment may lead to substantial learning gains – according to empirical evidence reported by Black and Wiliam (1998) “among the largest ever reported for educational interventions” (p. 140). Moreover, formative assessment methods are, in some cases, particularly effective for lower-achieving students, so it is possible to both reduce inequity of student outcomes and raise overall levels of achievement.

Yet a number of studies have found that effective classroom assessment is rare. Teachers are more likely



to emphasise rote learning and develop superficial questions (Black, 1993; Black and Wiliam, 1998; Stiggins et al., 1989). Teachers may have difficulty developing strategies to elicit information on student understanding of conceptual content, or to respond to identified needs. An OECD (2005) study on implementation of formative assessment in international classrooms found that while the concept of formative assessment may resonate with teachers, many protest that it is too difficult to put into regular practice. Teachers note the difficulty of tailoring learning for individual students in large classes, of working with students they consider as more challenging, and of meeting extensive curriculum requirements within limited time periods.

External assessments developed for purposes of school monitoring and accountability may also undermine innovative teaching and assessment. While high stakes associated with assessment results – such as the threat of school closure or financial sanctions for poorly performing schools – are intended to provide incentives for teacher to focus on meeting high standards, they may also discourage innovation and risk taking. Highly competitive university entrance examinations or certification programmes may also have a powerful impact on classroom-based assessment (See definition of terms in Annex 1).

Several new educational technologies designed to support formative assessment may help to address these barriers. These new technologies enhance learning and assessment, for example, by enabling more frequent feedback, creating immersive learning environments that highlight problem-solving processes and make



student thinking visible, and by providing opportunities for independent and collaborative learning. Teachers, students and parents are able to track learning over time, to identify patterns in learning, and highlight progress. New ICT-based examinations, while still in the early stages of development, have the potential to improve the integration of summative and formative assessments.

As of yet, many teachers do not take advantage of the potential of new technologies to enhance classroom assessment. They may not be aware of how to integrate technologies into classroom assessment, or may not know how to respond to student needs identified in the assessment process (Becta, 2010). Or they may be using new technologies, but to reinforce more traditional approaches to assessment – losing out on the potential to deepen classroom interactions and strengthen inquiry-based learning.

In the following pages, this report:

- Sets out a model describing the elements of effective formative assessment.
- Provides an overview of classroom technologies that may facilitate classroom assessment, provide rich records of student learning and progress, and promote higher order learning.
- Explores the knowledge base on the implementation and impact of new technologies, and suggests directions for future research and innovation.
- Briefly addresses the role of policy in promoting broader changes.

The Elements of Formative Assessment

The model of formative assessment presented in this report blends elements from the English- and French-language research.

In their review of the English-language literature, Black and Wiliam (1998) defined formative assessment as involving:

- Establishment of a standard or expected level for student performance.
- Gathering of information on the students' current level of performance.
- Development of a mechanism to compare current and expected levels of performance.
- Development of a mechanism to address the gap.

Assessment is “formative”, they note, when the information gathered is actually used to alter the performance gap.

Allal and Mottier-Lopez (2005) refer to the Francophone literature on formative assessment to extend these definitions, placing a strong emphasis on how teachers organise and orchestrate learning. This includes:

- The actions that teachers and students actually carry out to alter a learning gap or to arrive at a shared vision of learning objectives.
- The degree of student involvement in the assessment process.
- The meaning attributed by teachers and students to assessment practices and to their effects.

Thus, the effectiveness of formative assessment depends upon teacher's knowledge and skills – including deep knowledge of the subject being taught and of the core principles students must learn. Theories of “how students learn” within different domains, based in part on evidence of common obstacles to learning, also help to shape more effective assessment and strategies to meet student needs.

Drawing on these different approaches the blended model defines the elements of formative assessment as:

- Integral to classroom culture



- Oriented toward clear learning goals
- Incremental and interactive
- Providing feedback that is timely and specific
- Focused on the process of learning as much as the outcome

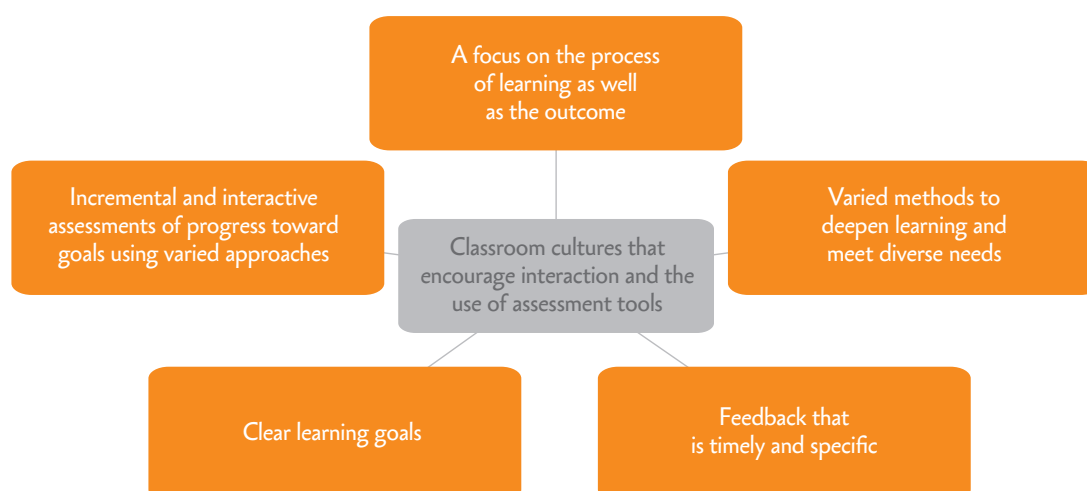
- Using varied methods to deepen learning and meet diverse student needs

Each of these elements is described in the following pages.

Assessment as an integral part of classroom culture

Formative assessment is most effective when practiced systematically – that is, when it is integral to the

The Elements of Formative Assessment



Source: Adapted from OECD (Looney, Ed.), 2005

teaching and learning process. Teachers may develop classroom cultures that encourage interaction and regularly incorporate the use of assessment tools such as rubrics outlining criteria for quality work, or interactive ICT-based tools.

The most effective assessment is grounded in theories of cognition – that is, how students learn – in different subject domains. These theories shape the choice of learning goals, the learning artifacts teachers choose to convey and represent knowledge, the forms of questioning or tasks teachers use to explore whether

and how well students have grasped new concepts, an understanding of typical learning progressions from novice to expert levels and landmarks in learning progress, and knowledge of the typical obstacles many students encounter in different domains as well as an understanding of the variety of ways in which students learn new subject matter. As Black (2002) notes, formative assessment is “...only a shell into which effective learning strategies need to be fitted” (p. 410). Thus, formative assessment is an integral part of the teaching and learning process.



Assessment tools, which might also be considered as learning artifacts, are an important feature in the formative assessment process. These tools enable and mediate assessment processes, shape the way in which teachers and students judge the quality of performances and view progress, and affect the quality of classroom interactions. They also serve as templates for self- and peer-assessment – essential to the skills for “learning to learn”.

Clear learning goals

Effective formative assessment focuses attention on students learning goals, and what students need to do to reach goals. Teaching and learning oriented toward specific, clear learning goals and which supports a “mastery” approach to learning as opposed to a “performance” approach is more effective. Students who take a goal-oriented approach are more likely to believe that learning is based on effort (mastery) rather than innate ability (performance), and are also more likely to persist in their learning (Dweck, 2002). With clear learning goals, students do not need to guess at what counts as a competent performance, and are better able to assess their own and their peers’ progress toward goals.

Classroom cultures integrating formative assessment help students to understand that they can learn from their mistakes, and that the most effective learning is based on willingness to address learning gaps and to persist toward goals. Students are more likely to take the risks necessary for further learning in such environments (OECD, 2005).

An incremental and interactive approach

Once the learning goal has been identified, teachers and student assess current understanding, and then track progress toward the goal. In classrooms featuring formative assessment, teachers interact regularly with students to draw out evidence as to whether and how well they understand new concepts. They may engage students with sophisticated questioning techniques and extended dialogues to explore thinking, identify possible misconceptions, and deepen understanding. Teachers gather new evidence of student progress and understanding at successive stages.

Mischo and Rheinberg (1995) and Köller (2001) found positive effects in several experimental and field studies where teachers referred to student progress over time. These included students’ intrinsic motivation, self esteem, academic self concept, attribution of progress to effort, and attainment (see also Krampen, 1987).

Effective questions help to reveal students’ level of understanding and identify possible misconceptions (in contrast to questions that are designed to elicit a “yes” or “no” response or that stress recall rather than reasoning processes provide little information on the student’s level of understanding and may hide errors in thinking). Through extended dialogues and series of questions aimed at building understanding, teachers may guide students toward deeper levels of understanding (OECD, 2005). Teachers may also gain insight into student thinking through observation, review of written work products and portfolios, student presentations, activities and projects, interviews, tests and quizzes (Shepard, 2006). These varied views on student work over time and in different contexts allow teachers to identify patterns in thinking and problem solving (Harlen and James, 1997; Williams and Ryan, 2000).

Feedback that is timely and specific

Studies suggest that feedback is most effective when provided in a timely manner. Wiliam and colleagues (2004) found that in classrooms where teachers provided formative feedback within or between teaching units – for instance, during an in-class interaction or over the course of a month-long teaching unit – the rate of student progress over the year was approximately double that found in the control classrooms. At the same time, feedback should not be provided too rapidly. Students may need some time to work out problems before referring to teachers or peers.

Several studies show that learning gains are stronger when feedback is focused on the task at hand, rather than the student’s ego – even in the form of praise (Boulet et al., 1990). Task-based feedback includes specific suggestions for improvement. Feedback such as “need to work harder”, for example, does not provide guidelines as to how the student may go about improving his or her work. Good feedback is



based upon explicit criteria regarding expectations for performance. It also includes as much or as little information as the student may need to make progress (that is, teachers “scaffold” feedback¹).

A focus on the process of learning as well as the outcome

A number of researchers have emphasized the importance of engaging students in the process of learning so that they are actively building their understanding rather than simply absorbing information (see for example Bransford et al., 1999). They also learn to judge the quality of their own and their peers’ work against well-defined criteria and clear learning goals. Students thus build their skills for “learning to learn”. This focus on the process of learning is well aligned with national education standards across OECD countries emphasising development of students’ higher order skills.

Studies have shown that students focused on the learning process, rather than just “getting through a test”, perform better and retain information longer [see, for example Nuthall and Alton-Lee (1995) and Wiliam et al. (2004) for research on the impact of teaching, learning and assessment approaches on long-term performance]. Students also obtain better results when they are working toward process goals rather than product goals, and when tracking progress toward overall goals of learning (Schunk, 1996). This focus on the process of learning is well aligned with national education standards across OECD countries emphasising development of students’ higher order skills.

Parents also play an important part in the learning and assessment process. In several countries (e.g. Denmark, Finland, Scotland, Sweden) students work with their teachers to develop individual development plans (IDPs – or, alternatively, personal learning plans or PLPs). Students set their own goals and track progress. Teachers and parents provide support. In addition, there are examples of schools that have chosen to develop “comments-only” or rubric-based marking portfolios. The student may also keep a record of the assessments they have made of their own work. Many parents have found these different approaches to assessment provide them with a better idea as to how they might be able to help their child with schoolwork (Becta, 2010; OECD, 2005).

Varied methods to deepen learning and meet diverse student needs

Teachers may also elicit more information on student learning through varied tasks, applying new skills in different contexts, and using a variety of media. In this way, students also have the opportunity to explore new subjects from different angles and to reinforce and deepen their understanding, and strengthen their ability to transfer knowledge to new situations.

The use of varied methods is also important for meeting needs of diverse students within a classroom. Each student brings different experiences, perceptions and ways of communicating and processing ideas to the classroom. Through varied methods, teachers may enliven classroom learning and reach more students. Teachers thus draw upon a broad repertoire to help different students reach learning goals.





Enhancing Formative Assessment through technology:

The state of the art

Several new technologies incorporate different ways to assess student performance. These technologies are designed to support:

- Rapid assessment of student understanding
- Timely and targeted feedback, scaffolding of learning
- Interactive learning and assessment of higher-order skills.
- Tracking of student learning in different contexts and over time.

Some of these technologies – such as simulations with real-time feedback – blend instruction and assessment in new ways. Others are intended to facilitate student collaboration and to encourage peer- and self-assessment.

While a range of surveys and evaluations have found that a majority of teachers tend to use new technologies to reinforce traditional approaches to assessment (e.g. through superficial questioning) (Langworthy et al, 2010; Selwyn, 2010; Voogt, 2009), there is also some evidence that, as teachers embed new technologies in their pedagogy and gain confidence, their goals for teaching and assessment begin to shift. Teachers who are already engaged in innovative approaches to classroom assessment may deepen practice with new technologies (Somekh et al., 2007; Yarnall et al., 2006).

The following discussion provides a brief overview of some of the technologies and software programmes that have been developed to facilitate formative assessment. It is not intended to be an exhaustive survey of technologies in this area. As will be discussed later in the report, research and evaluation on education technologies is still fairly limited. Thus, the focus here is on highlighting selected studies that provide more detailed views of how teachers are using different ICT-based tools to facilitate classroom assessment, their impact on the assessment process and on student outcomes. Some of the technologies described

(e.g. simulations and video games to assess students' problem-solving processes) are in the very early stages of development, and in some cases may not have yet moved beyond the pilot stage. For these new technologies, we can only discuss their potential impacts on assessment practices.

Rapid assessment of student understanding

Polling tools, sometimes referred to as learner response systems (LRSs), allow teachers to conduct on-the-spot surveys. The LRS consists of handheld "clickers", or input devices, which communicate with software on the teacher's computer. Students use the LRS devices to respond to questions posed by the teacher, and responses are aggregated and displayed on the teachers' computer in the form of bar charts or graphs. The devices allow students to respond to yes/no or multiple-choice questions. Some devices also accept free text or numeric answers. Using these polling devices, teachers are able engage all students, including those who are less likely to speak up during class, in active classroom discussions.

The evidence regarding the impact of polling technologies on formative assessment practice is still limited. Two recent U.S.-based studies explore how teachers are using the polling devices to facilitate formative assessment practices². Both studies point to the importance of teachers' skills for setting up assessment situations and for responding to the results as the prime determinant of impact. At the same time, the tools appear to have helped teachers to deepen practices over time.

In the first of these studies, Yarnall and colleagues (2006) worked in partnership with teachers from a school district in California between 2001 and 2004 to develop a set of software applications for handheld computers to support formative assessment in primary and lower secondary school science classrooms³. At the start of the project (known as Project WHIRL), the researchers set out to learn more about the range of instructional contexts in which LRSs might support more frequent assessment, and whether they supported



more active roles for students in assessment. They also wanted to know if the use of LRSs changed teachers' goals for assessment and what factors accounted for any differences in how teachers used software to support different assessment practices.

Over the course of the project, the researchers found that the polling technologies tended to reinforce teachers' existing approaches to assessment. For example, they found that teachers who had been more focused on "accountability" goals for assessment (e.g. keeping students on task), tended to use the new technologies to reinforce these narrower approaches. Teachers who had "thinking-focused" goals for assessment (e.g. assessing students' approaches to problems solving, understanding of new concepts, and so on) were more likely to use the new technologies to deepen their assessment practices, to focus student reflection on critical elements of learning science and to foster student skills for self-assessment.

Some of the "thinking-focused" teachers used the polling tools to engage students in improving their own questioning skills. In one example, the teacher involved students in judging the quality of questions – i.e. questions that required more thought vs. those that required only a simple answer. In another example, teachers used software to gather data on students' thinking during hands-on activities, such as laboratory exercises, and then analysed the results with students.

According to Yarnall and colleagues, teachers who began the study with stronger "thinking-focused" approaches also tended already to have stronger skills for classroom management and more experience in structuring and scaffolding science inquiry. At the same time, the researchers found that over the course of the project, the participating teachers had increased their interest in "assessment for learning", and had started to blend instructional and assessment uses of tools (Yarnall et al., 2006).

A second study on the use of polling devices, the Technology-Enhanced Formative Assessment (TEFA) project, was initiated in 2005 and will end in 2010.

Project leaders provided intensive professional development for more than 40 secondary school science and mathematics teachers⁴, focused on building their skills for:

- developing effective questions,
- orchestrating effective classroom dialogues to draw out thoughtful contributions and multiple points of view,
- understanding and responding to student needs, and
- managing the technical requirements of LRS devices and linked software.

Thus, the study focuses primarily in building teachers' formative assessment skills, with use of technology as a secondary aspect. The "question cycle" was used as prototype for using LRSs to enhance formative assessment. During the question cycle, the teacher presents a question or problem to the class. Students are given time to reflect on the question, either individually or in small groups. The students enter their responses through the LRSs, and the teacher shares the results with the class and then guides a whole-class discussion, exploring the different responses. The teacher concludes the discussion with a micro-lecture or by highlighting major points (Beatty et al., 2008).

The project researchers and participating teachers also developed a range of ways to use LRSs to assess student understanding and guide subsequent instruction. For example, teachers might poll students regarding how well they believe they understand a topic, assess prior knowledge, ask about areas where students would like to spend more time in future sessions, open new topics with thought-provoking questions, elicit misconceptions, challenge students' interpretations and ideas, and so on.

In describing the project, Beatty and Gerace (2009) note that the basic goals of the project could have been realised without LRS technology. However, the LRS did enhance teachers' formative assessment practice by: providing both student anonymity and accountability, allowing teachers to analyse the distribution of



responses and to organise discussions to explore the range of responses, to analyse individual and whole class student needs, and to evaluate the impact of instruction.

In their preliminary observations, the researchers note that teachers tend to concentrate on the use of technology first, then move to question design, followed by efforts to improve management of whole class discussion, and later on interpretation of student responses. Those teachers who were the most self-reflective were the most likely to improve. Beatty and Gerace note that their preliminary findings indicate that “TEFA can be highly effective, even transformative, for secondary school science and mathematics instruction” p. 149.

The two studies provide contrasting models for integrating technologies to support formative assessment in regular classroom practices. The study led by Yarnall, Project WHIRL, invested less time in developing teachers’ skills for formative assessment up front, preferring to investigate whether and how the use of LRS tools affected their practices. The study led by Beatty, TEFA, involved teachers in building sophisticated skills for formative assessment at the outset. The Beatty study is just now coming to a conclusion, so results are not yet available. However, based on the tentative observations, the researchers found that teachers needed time to embed the new technologies. In both studies, teachers’ pedagogical skills, including skills for classroom management, were the most important determinant of success. In other words, the technologies did not automatically relieve teachers of logistical burdens. Once teachers had developed greater facility with the tools, however, they were able to engage more students in classroom discussions, and to tailor instruction to better meet needs.

Timely and targeted feedback, scaffolding of learning

Increasingly, ICT-based learning programmes are able to provide rapid and targeted feedback for students working independently. Some online learning programmes use tutors to provide real-time support for learners. Other programmes provide automated feedback. Although in some cases this feedback may

be fairly generic, some programmes search for patterns in student work to better target feedback and to then adjust the level of difficulty in subsequent exercises according to needs. In other words, the programmes scaffold learning. Tuomi (2006) has coined the term “pedagogical veils” to describe this process.

As an example, the “e-gramm”, an ICT-based programme developed by Lawley and Beltrán at the Universidad Nacional de Educación a Distancia in Spain for learners of English as a Foreign Language provides detailed feedback on written compositions. The developers analysed common mistakes across hundreds of student compositions, encoded them, and developed feedback to address these kinds of mistakes and to support learners in modifying their own writing. By 2008, the programme was able to detect 60 to 70% of common mistakes made by Spanish mother tongue students. The developers believe that students get much more feedback than they would from teachers, who have little time to provide detailed comments or suggestions (Sanz, 2008).

Programmes such as SuccessMaker’s Reader’s Workshop (NCS Learn), and the Accelerated Reader, which are popular in the U.S., provide ICT-based instruction with animations and game-like scenarios. Assessments are embedded within the programmes, and feedback is tailored according to student performance on prior exercises, so learning is scaffolded. These programmes have been evaluated as having positive impacts on learning⁵ (Murphy et al., 2002; see the full study for a review of evidence on the impact of a range of educational software in use in U.S. schools.)

Experts in assessment and psychometrics are also exploring ways to improve the timeliness and specificity of feedback from large-scale, external assessments. These national or regional-level assessments, which are increasingly common across OECD countries, are used primarily as a means for monitoring school performance. Schools may be held accountable for meeting national or regional standards as measured by



these assessments. The data are also used for policy decisions regarding allocation of resources, and to guide investment strategies to address areas where students are not performing as well. However, with current assessment technologies, the results generally are not delivered in a timely enough manner and do not provide the detailed information necessary to guide instruction or address the individual needs of the test takers (although the results may inform more general strategies for future students).

Interactive learning and assessment of higher-order skills

Several new technologies allow teachers to blend instruction and assessment. The technologies briefly reviewed here include **interactive white boards (IWBs)**, **simulations**, **video games**, and **social networking tools**. Different technologies enable teachers to follow students' reasoning and problem solving approaches, thus providing a window on student thinking.

IWBs consist of large interactive displays connected to a computer or projector. The computer may be linked to the Internet, providing easy access to a range of tools for teaching. The IWBs provide a space for "co-learning" as students and teachers use the board as a shared space. Students may demonstrate their skills or their understanding of processes, so teachers may identify and address misconceptions. With access to a range of materials, teachers are better able to take advantage of "moments of contingency", that is spontaneous teaching moments that may arise in the course of a lesson.

Somekh and colleagues (2007) conducted an evaluation of the implementation and impact of IWBs in U.K. primary schools. The study, which took place between 2003 and 2007, did not have a specific focus on ways in which teachers use the IWBs to support formative assessment, so it is difficult to discern impact in this area. However, it is worth noting that students' attainment levels increased in classrooms with IWBs. The length of time students had been taught with an IWB was the major factor associated with gains. The researchers found that teachers needed at least two years to embed whiteboard use in their pedagogy⁶.

A more recent evaluation by Haystead and Marzano (2010) also found a significant correlation between teachers' self-reported confidence in the use of IWBs and LRSs and the length of time they had been using technologies (reported in months).

ICT-based **simulations** also encourage interactivity and help to make students' thinking processes visible. One of the key advantages of simulations is that they create opportunities for students to develop and apply skills and knowledge in more realistic contexts and provide feedback in real time. Some simulations create immersive environments, with detailed descriptions of the context and the problem that needs to be resolved. For example, simulations may involve mini-laboratory investigations, or "predict-observe-explain" demonstrations (students predict what will happen in a given set of circumstances, observe a demonstration, and then explain what happened and why) (Shavelson et al., 2003). The programmes usually provide opportunities for students to reflect on their own actions and patterns they may detect in the responses provided by the simulation programmes.

Alternatively, students may develop concept maps using online tools to show their understanding of processes. Computer-based maps are scored against expert maps. Evaluators have found that students respond positively to Internet tasks (searching, and judging the quality of on-line information) and are engaged, even in difficult, open-ended tasks (Osmundson et al., 1999). In another example, a programme known as Sketchy enables students to create animations of complex processes in the sciences (e.g. electrical circuits, lunar cycles, etc.). Teachers are able to identify student misconceptions more rapidly with these visuals. They may also use student drawings to generate discussions with individual students or with the class (Yarnall et al., 2006).

ICT-based assessments that take advantage of simulations are still in the early stages of development. Part of the challenge involves the need to develop appropriate models for automated assessment of problem-solving skills and for learning progression in different domains. For example, Mislevy and colleagues (2001) describe the development of a computer-based



simulation and assessment of problem solving in dental hygiene (assessing the status of a new patient). The simulation and assessment are modeled on a theory of learning in this field, including appropriate tasks (the behaviours or performances expected to reveal reasoning and levels of expertise in a domain), and situations to elicit those behaviours and demonstrate knowledge and skills. An automated assessment may apply probability distributions on levels of knowledge and competence for different levels of expertise. The model integrates new information as the student proceeds through tasks of the simulation. West and colleagues (2010) describe their efforts to develop a model for learning progressions and task performance in the field of computer networking. They are exploring how to develop more sophisticated approaches to modeling learning progressions in order to improve task design, simulation design, modeling of complex performance and analysis of assessment data in different domains.

Gaming (serious games) that draw on more advanced cognitive capabilities may also serve valuable pedagogical purposes. Learning through play is highly effective for improving motivation, and engagement, and many teachers see adventure games and simulations as useful ways to develop students' strategic skills (Gros, 2003).

However, video games have not typically been used as tools for assessment. Game developers from the University of Southern California and researchers at the National Center for the Research on Evaluation, Standards and Student Testing (CRESST) have piloted a rational numbers game, Puppetman, to learn whether it would be possible to develop a valid assessment context. The game increases in challenge and complexity as it progresses, with later levels requiring the most mathematical knowledge for success. Potentially, teachers will also be able to examine process data from game play for formative assessment (e.g. time spent on each level, strategies used). Importantly, the game and assessment designs were built around a theory of learning (i.e. how students learn rational numbers concepts). Findings from the pilot suggest that games

can provide a useful assessment environment, while many students do not perceive the game as a test (Delacruz et al., 2010; Vendlinski et al., 2010).

Social networking tools (Web 2.0) are already sufficiently developed for widespread use in schools. These tools enable students to work collaboratively on projects, conducting joint research, sharing and structuring information, developing group reports using weblogs, wikis, discussion forums, online chat groups as well as software programmes that allow task sharing. Web 2.0 tools may also support collaborative "quiz" communities. Students using Web 2.0 have opportunities to assess their own and their peers' work and to adapt and improve products over time. In this way, they may take ownership of the assessment process.

Selwyn (2010) notes that various evaluations of the use of Web 2.0 technologies in schools have shown they are not being used to promote the types of interactions described above (see, for example, evaluations in the U.K., by Grant, 2009; the U.S., by Knobel and Lankshear, 2006; and in Norway, by Lund and Smørdal, 2005). But these critiques do not point to a failure of Web 2.0 or other ICT-based programmes as tools for learning so much as a need to help students to develop their skills for critical inquiry, collaboration and for self- and peer-assessment. Teachers will need a clear understanding of why they are using Web 2.0 and other ICT-based tools, and how to best ensure that they are used to enhance learning and peer- and self-assessment.

Tracking of student learning over time

No single assessment can provide enough information to understand how well instructional strategies are working and whether students are learning. Moreover, any single assessment, whether formative or summative, should not be considered as more than a snapshot of student performance on a given day. Assessment systems will provide more valid results if they measure student learning over time, and track learning in different contexts, and using different assessment tools. Parents might also refer to these data in order to follow their child's progress and lend support for schoolwork.



ICT-based assessments have the potential to improve the integration of formative and summative approaches. For example, ICT-based assessments might be administered at several points throughout the school year. Test results would be delivered to schools immediately following administration (and in some cases students would react to feedback during the test) and could be used to adapt instruction in a timely manner. Because schools teachers would track student learning over time and would cover more aspects of the curriculum, they would also have a better view as to the effectiveness of instructional strategies. Scoring may be purely automated or may blend scores assigned by human raters and automated scores⁷.

The primary barrier to developing these assessments is in ensuring that measures of complex performances are both valid (measure what they are intended to measure) and reliable (results are consistent). A secondary barrier is that the cost of developing these sophisticated ICT-based assessments is still relatively high (Means, 2010)⁸. There are also a number of concerns regarding privacy, equity and practicality that need to be addressed before such a system can be implemented. But there are also many potential benefits to integrating formative assessment in this manner.

ICT-based assessments would also provide means to aggregate data at different levels – for individual students, specific classes, grade levels, schools, and so on. Policy officials and school leaders would be able to analyse data showing patterns of performance across groups of students (e.g. by gender or socio-economic status), and in specific subjects. Teachers would be able to create complex streams of data about student learning, enhancing their ability to analyse patterns in their approaches to learning, possible misconceptions, and how they have progressed over time (Chudowsky and Pellegrino, 2003).

Policy makers have also discussed the possibility of developing digital records of individuals' learning histories. Miller and colleagues (2008) refer to "personal digital spaces", which would provide a lifelong record of individual achievements and aspirations for

further learning. The personal digital spaces, which would include information from a wide range of assessments, would move well beyond the familiar formal transcripts, final grades and certificates earned.

Digital records are also promising for adult learners, who may have irregular patterns of participation in formal learning programmes, given the demands of their daily lives. The JISC, which works with colleges and universities in the UK to develop and implement new digital technologies, is exploring "personal development planning" to facilitate transfer of learner records across institutions. These records might also track learning in work-based or community-based settings. Learners would not be required to repeat earlier course work, and would also have greater control over their own learning pathways. Thus, systems would more effectively facilitate lifelong learning.





Learning more about how and why teachers use technology to enhance formative assessment, and how technologies may affect practice

The discussion above points to the complexities of the relationship between teaching and technology. Our understanding of this relationship – how and why teachers use technology to enhance assessment, and in turn, how these technologies may influence their practice over time – is still limited. There is a need to strengthen and deepen research in this area. There is also a need to consider where technological innovations may support deeper changes.

Strengthening methodological rigour, asking the right questions

Programme monitoring and evaluation are essential to innovation. Yet formal evaluations of ICT-based innovations in schools are rare. Different commentators have also pointed to the need to strengthen the methodological rigour in this area (Cox and Marshall, 2007; Murphy et al., 2002; OECD, 2008). Many studies, these commentators observe, do not include adequate sample sizes, comparison groups, or data that would allow calculation of effect sizes. Many also exclude information on study design and hypotheses to be tested.

For those studies that do meet standards for methodological rigour, researchers have generally found positive associations between the use of educational software and student achievement in subjects such as mathematics and reading⁹. Yet these studies often do not provide sufficient explanations as to why programmes are effective, so it is not always possible to rule out alternative explanations for impact. Studies rarely include analysis of implementation – that is, how teachers use different tools for different assessments, how their instructional philosophies shape use, the impact of technologies on classroom management (and vice versa), or the kinds of technical and professional development support teachers have received. Very few studies provide longitudinal data on how teachers take on and integrate technologies over time, or whether the frequency of technology use has any impact on effectiveness. Yet all these factors are vital to programme impact.

In general, there is a need for more quasi-experimental evaluation to deepen understanding of programme implementation and effectiveness – i.e. what works and why (or why not), for whom and under what circumstances. The studies cited above begin to address some of these important questions. But they are relatively small in scope and scale. Beatty and Gerace (2009), for instance, call for more systematic research to “define, ground, justify and thoroughly explicate coherent pedagogies” for teaching and assessment with technologies (p. 148).

A separate but related set of research questions has to do with whether and how increasingly sophisticated technologies may better serve pedagogical needs. Assessment technologies now under development build on models of student learning and progression in different domains, enabling more accurate and timely measurement of students’ higher order skills. Perhaps the most promising aspect of this research is related to the possibility of bringing formative and summative assessment into closer alignment. Whether and how these new technologies will influence teachers’ approaches to formative assessment remains to be seen.

Developing teacher expertise

While the incoming generation of teachers may be assumed to be “digital natives”, and therefore much more comfortable with the use of technology for teaching than previous generations, according to an OECD (2008) report, this is not necessarily the case. New teachers appear to be affected by the existing culture of the teaching profession, in which research-based approaches to teaching, learning and assessment are not widely disseminated. In addition, while new teachers may be conversant with technology in their daily lives, they are not exposed to ideas about how to integrate technology in classroom settings – let alone how to use technology to support formative assessment practices.



There are at least four areas where education systems might sharpen the focus on developing or improving incentives for teachers to develop their expertise in formative assessment and technology-enhanced formative assessment. They include: **teacher training, teacher appraisal, professional development, and Web 2.0 to support teacher communities of expertise.**

Teacher training institutions may be missing important opportunities to help new teachers to transfer “native” expertise with technology to support the learning of next generation of digital natives. More research will need to be conducted in this area to understand how teacher-training institutions may better prepare their students.

Teacher appraisal systems are another area where education systems may strengthen the focus on formative assessment and technology. According to the OECD’s (2009) Teaching and Learning International Survey (TALIS)¹⁰, teachers and school leaders responding to the survey reported that the greater the emphasis on a specific aspect of practice in the appraisal system, the more likely teachers were to make efforts to improve their practice. At the same time, TALIS also found that about three-quarters of teachers participating in the study reported that they would not be recognised for being more innovative with their teaching.

The OECD TALIS also explored teacher participation in **professional development**. Three-quarters of teachers reported that they would like more opportunities to participate in professional development related to technology for instruction. Thus, teachers report that they are interested in learning more about how they can integrate new technologies into teaching – but do not always have the training to help them do so.

Web 2.0 to support teacher communities of expertise may support wider, systematic dissemination of innovative practices (Hargreaves, 2003). There are a number of sites where teachers may access materials and interact with colleagues. Tuomi (2001,2002) notes that individuals participating in Internet developer communities have developed high-level competences

through participation in online communities of practice surprisingly quickly. We do not know, however, the impact of equivalent teacher online communities of practice on teacher development.

Concluding Observations

As discussed at the outset of this paper, assessment touches all aspects of the learning process and the role of technology is central. There can be little doubt that a leap in educational productivity will not happen without significant strides in the development of assessment methods and tools. Equally there is a growing awareness on the part of policy makers and researchers that we now have the capacity to dramatically improve assessment systems. To do so will require a wide range of initiatives, aimed at:

- **Developing coherent strategies for strengthening ICT in education.** While many countries have made significant investments in ICT and simultaneously promote student-centred learning as a priority, these priorities may not be linked. By explicitly bringing the two domains together, policy makers may strengthen efforts to support innovation and improve impact.
- **Investing in research and evaluation.** The above discussion highlights the need for more rigorous research to address a range of questions regarding the impact of technology on teaching, learning and assessment. Evaluation systems will be most effective if they include feedback loops to shape improvements. Greater insight on the process of change and innovation in complex education systems, which as Hylén (2010) points out are not fully understood, will also be important.
- **Disseminating new knowledge.** Systems may facilitate adoption of effective practices through wider dissemination. Adopters may, in turn, support further innovation as they share their experiences with implementation, and ways in which they have adapted new approaches.



- **Developing guidelines and exemplars.** Education systems may support teachers in their efforts to strengthen formative assessment and use of technology-enhanced formative assessment with guidelines and exemplars – in effect, providing concrete examples of effective practice. Guidelines should be based on empirical evidence of what works and how.
- **Considering formative assessment and technologies to support assessment within broader frameworks for assessment and evaluation.** Assessment and evaluation frameworks encompass standards for performance, classroom-based formative and summative assessment; external summative assessments for monitoring and accountability, or for student certification; inspection systems, and teacher appraisal. The coherence and alignment of these different parts of the system are vital to effectiveness. As discussed above, new technologies have the potential to promote closer integration of formative and summative assessments and to ensure that assessments measure students' higher order skills. Indeed, investments in new ICT-based summative assessment of higher order skills and to provide better and more timely information must be a priority area for research and development. At the same time, education inspectorates need to rethink criteria for evaluating the use of technologies in classrooms and schools. While teacher appraisal systems need to be strengthened in ways that highlight effective classroom-assessment and use of ICT.
- **Encouraging partnerships between and among ICT product developers, educators and policy makers.** In many cases, these different actors are moving in separate spheres, and missing opportunities to create more powerful products and strategies for education and ICT. Through productive partnerships, these different players may define effective approaches to integrating teaching, assessment and effective ICT-based programmes (Pedró, 2010).
- **Involving parents.** As technologies provide opportunities for education to spread beyond

classroom walls, they also create opportunities for parents to become more involved in their child's education. In this way, parents are empowered to act as partners in schooling and learning.

Taken together, these measures could support and enable a leap in educational productivity. As subsequent papers in the Promethean Deeper Thinking Papers series will explore, there is a strong inter-dependency of changes in one part of the learning process on changes in the other parts, from the way data are used to make decisions and how learning processes are personalized to programmes that enhance teacher effectiveness and introduce the most appropriate tools for different kinds of learning. Learning is an inter-connected, infinitely diverse voyage of permanent discovery. The way we assess what we learn in everyday life, how that assessment becomes part of the creation of value, of identity and of shared knowledge, is changing both what and how we learn.



Endnotes - Paper 1, Part 3

¹ The idea of “scaffolding” learning is based on Vygotsky’s (1978) “zone of proximal development, sometimes also referred to as the “ZPD”. Vygotsky defined the ZPD as “the distance between the actual developmental level as determined by independent problem solving and the level of potential development as determined through problem solving under adult guidance, or in collaboration with more capable peers” (p. 86). Educators should provide students with experiences that are within their ZPD, guiding them toward increasing levels of independence.

² Beatty and Gerace (2009) found that only one prior study, A2L (Dufresne et al., 2000), had explicitly linked LRS technology to a pedagogical perspective. Others have attempted to identify significant features of instruction with LRSs (Roschelle et al., 2004), and then connected them to established pedagogical constructs.

³ The project recruited 18 teachers and reached 802 socially and economically diverse students in grades 4 to 9. Teachers participated in training workshops and were also supported by a local technology coordinator. Researchers followed teacher use of the LRSs during the 2003 – 04 school year.

⁴ The study involved lower and upper secondary school teachers from 3 geographically diverse school districts in the state of Massachusetts. The project involved intensive, sustained, on-site professional development programmes.

⁵ Murphy et al. report an effect size of 0.19 for the Accelerated Reader, and +0.66 for SuccessMaker. While the meta-analysis was conducted in 2002, these programmes are still in wide use.

⁶ A more recent evaluation of the Promethean ActivClassroom (incorporating IWBs and LRSs) by Marzano (2010) sought to determine impact on students’ mastery of academic content. The evaluation did not focus on specific pedagogical uses of technologies, such as formative assessment. In line with earlier research, teachers reported increased confidence in their use of technologies over time (a period of months).

⁷ Increasingly sophisticated ICT programmes with the capacity to score “open-ended performances” are now under development (see Chung and Baker, 2003; Chung et al., 2001; Herl et al., 1999). These programmes use natural-language processing, artificial intelligence and/or information retrieval technologies to detect of textual features of essays (for example, variety in use of syntax, quality of content and organisation of ideas). These models are still in the relatively early stages of development and cannot replace human raters entirely. Further studies are also needed to determine the validity and generalisability of different automated essay scoring tools (Wang and Brown, 2007).

⁸ The U.S. Department of Education has invested \$350 million in research to develop ICT-based assessments that will measure higher-order skills and that may be administered several times throughout the year, so results may be used formatively. The aim is to implement new assessments by the 2014 – 15 school year (Dillon, S, 2010).

⁹ Murphy and colleagues (2002), identified 31 studies that met criteria for inclusion in their synthesis. They found a positive association between the use of discrete educational software products and student achievement in reading and mathematics. The overall weighted effect size for all studies was $d = +0.38$ ($d = +0.35$ for reading and $+0.45$ for mathematics applications). They note that this is in line with, and slightly larger than effect sizes calculated in earlier research syntheses. The study does not include a specific focus on formative assessment uses of different products.

¹⁰ TALIS surveyed school leaders and teachers in lower secondary education in both the public and private sectors. The survey examined aspects of teachers’ professional development; beliefs, attitudes and practices; appraisal and feedback; and school leadership in the 23 participating countries. The survey included 200 schools per country, and 20 teachers in each school. Survey designers aimed for a 75% response rate in all participating countries.

Terms - Annex 1

Accountability—The public reporting of student, program, or institutional data for the purpose of program improvement.

Achievement—The attainment of new knowledge or understanding as the result of an educational activity.

Assessment—The systematic process of determining educational objectives, gathering, analyzing and using information about student learning outcomes to make decisions about programs, individual student progress, and/or accountability.

Assessment (Formative)—Purposeful, ongoing collection of

information regarding how students are learning while there is still the opportunity to make improvements. Both teachers and students use the information to guide continuous improvement toward the intended learning.

Assessment (Summative)—An assessment measure of achievement at the end of an instructional unit, course of study, or program.

Benchmark (Standard, Cut-score)—A criterion-referenced objective. Performance data that are used to set a level for comparison, either between different programs or over time for the same program. If data from another exemplary program are chosen as benchmarks, it becomes a target to strive for, rather than a baseline to improve upon.



Bloom's Taxonomy—Learning modes that can be applied to content.

Curriculum—A collection of statements of the objectives, or learning goals, to be taught in each subject at an individual grade level. Curriculum guides what is taught and how the subject is taught in the classroom.

Evaluation—The process of collecting information from multiple sources to make judgments (assign a grade, for example) about how well students have learned and about program effectiveness.

Extended Response—An assessment item a student answers with writing or drawing. Often evaluated with a rubric.

Formative (Evaluation)—An assessment that is used for making improvements (individual or program level) rather than for making final decisions or for accountability purposes.

Gridded Response—A paper entry mechanism for numbers, by bubbling each number sequentially in a grid.

Internal Assessment—Proctored and scored by the school or learning agency.

Item (Also question)—A term used to refer to a question on a written assessment or a job and task in a performance assessment.

Item Bank—A collection of items; generally topical or surrounding a specific passage.

Learning Map—See Topic Map.

Learning Outcomes—The specific knowledge or skills students actually acquire/develop through their educational experience.

Learning Progression—Longitudinal motion across a topic map. Often used in place of Topic Map or Learning Map.

Learning Response Device—A device, generally wireless, which allows input of assessment answers. Some devices also allow the display of questions or more advanced interactions.

Likert Scale—A specialized form of multiple choice used as a rating system, for instance Dislike, No Opinion, Like.

Modality—Different group structures of learning such as classroom, large group, small group and individual.

Multiple Choice—An item which has multiple answers, and the student selects one.

Objectives—Statements of the specific knowledge, skills, or attitudes that students are expected to achieve through their educational experience.

Observational Assessment—An assessment that is done while observing learning, either by the teacher or other education professional.

Passage—A piece of content about which questions are asked (such as a short poem).

Performance Band—A mechanism for grouping scores into sets. For instance: 90-100 for Far Above, 80-90 for Above, 70-80 for Met, 60-70 for Below and 0-60 for Far Below.

Power Standard—A pedagogically important Topic or Standard for teaching, assessing or both.

Response—A student answer to an assessment item, such as a selected answer to a multiple choice item.

Results for Findings—The outcomes of an assessment study or the experiment.

Rubric—A set of scoring guidelines that can be used by raters to evaluate student's work.

Short Answer—An item type which expects the student to fill in a blank or write a single-word response.

Skills—Observable behaviors that document levels of competence (i.e., knowledge, comprehension, application, analysis, synthesis, and evaluation).

Standards—The broadest of a family of terms referring to statements of expectations for student learning, including content standards, performance standards, and benchmarks.

Stem—The part of an objective test question or item that poses the question to be answered or the problem to be solved.

Subject—A coarse grouping for topics, often things like "Mathematics," "History" and "Literacy."

Test—Measurement instrument, procedure or device that requires scored responses from examinees; may include both selective and constructed response formats.

Topic—An atomic knowledge particle, such as "Measurement and Geometry 1.0: Students understand that measurement is accomplished by identifying a unit of measure, iterating (repeating) that unit, and comparing it to the item to be measured."

Topic Map—A collection of topics with precursor skills necessary for full understanding program. If data from another exemplary program are chosen as benchmarks, it becomes a target to strive for, rather than a baseline to improve upon.

References

- Beatty, I.D., et al. (2008)**, "Teacher Learning of Technology-Enhanced Formative Assessment", National Association for Research in Science Teaching, Baltimore, Maryland, April 1.
- Beatty, I.D. and W.J. Gerace (2009)**, "Technology-Enhanced Formative Assessment: A Research-Based Pedagogy for Teaching Science with Classroom Response Technology", *Journal of Science Education and Technology*, Vol. 18, pp. 146 – 162.
- Becta (2010)**, "Messages from the Evidence: Assessment Using Technology", Spring, Becta, United Kingdom. 03/09-10-WR497/TT23257/PDF.
- Black, P. (1993)**, "Formative and Summative Assessments by Teachers", *Studies in Science Education*, Vol. 21, pp. 49 – 97
- Black, P. (2002)**, "Research and Development of Educational Assessment", *Oxford Review of Education*, Vol. 26, pp. 407- 419.
- Black P. and D. Wiliam (1998)**, "Assessment and Classroom learning", *Assessment in Education: Principles, Policy and Practice*, CARFAX, Oxfordshire, Vol. 5, No.1, pp. 7-74.
- Boulet, M.M. et al. (1999)**, "Formative Evaluation Effects on Learning Music", *Journal of Educational Research*, Vol. 84, pp. 119-125.
- Bransford, J.D., A.L. Brown and R.R. Cocking (1999)**, *How People Learn: Brain, Mind, Experience, and School*, National Research Council, Washington, D.C.
- Brown, J.S., A. Collins and P. Duguid (1989)**, "Situated Cognition and the Culture of Learning", *Educational Researcher*, vol. 18, no. 1, pp. 32 – 42.
- Chudowsky, N. and J.W. Pellegrino (2003)**, "Large-Scale Assessments That Support Learning: What Will It Take?", *Theory into Practice*, Vol. 42, pp. 75-83.
- Chung, G.K.W.K. and E.L. Baker (2003)**, "Issues in the Reliability and Validity of Automated Scoring of Constructed Responses" in M.D. Shermis and J. Burstein (eds.), *Automated Essay Scoring: A Cross-Disciplinary Perspective*, pp. 23-40, Lawrence Erlbaum, Mahwah, NJ.
- Chung, G.K.W.K., E.L. Baker and A.M. Cheak (2001)**, "Knowledge Mapper Authoring System Prototype", University of California, National Center for Research on Evaluation, Standards, and Student Testing (CRESST), Los Angeles.
- Cox, M. J., and G. Marshall (2007)**, "Effects of ICT: Do We Know What We Should Know?", *Education and Information Technology*, Vol. 12, pp. 59-70.
- Delacruz, G.C., G.K.W.K. Chung and E.L. Baker (2010)**, "Validity Evidence for Games as Assessment Environments", University of California, National Center for Research on Evaluation, Standards, and Student Testing (CRESST), Los Angeles.
- Dillon, S. (2010)**, "U.S. Asks Educators to Reinvent Student Tests, and How They Are Given", *The New York Times*, September 2.
- Dweck, C. S. (2002)**, "Beliefs that Make Smart People Dumb", In R. J. Sternberg (Ed.), *Why Smart People Do Stupid Things*, New Haven, CT: Yale University Press.
- Dufresne, R.J., W.J. Gerace, J.P. Mestre, W.J. Leonard (2001)**, "ASK-IT/A2L: Assessing Student Knowledge with Instructional Technology", (Tech. Rep. dufresne-2000ask), University of Massachusetts Amherst Scientific Reasoning Research Institute.
- Grant, L. (2009)**, "HEATHER I DON'T CARE DO UR OWN PAGE! A Case Study of Using Wikis for Collaborative Inquiry in School", *Learning, Media and Technology*, Vol. 34, No. 2, pp. 105-117.
- Green, C. and D. Bavelier (2003)**, "Action Video Game Modifies Visual Selective Attention", *Nature*, Vol. 423, pp. 534-537.
- Gros, B. (2003)**, "The Impact of Digital Games in Education", *First Monday – Peer-reviewed Journal*.
- Hargreaves, D. (2003)**, "Education Epidemic: Transforming Secondary Schools Through Innovation Networks", Demos, London.
- Harlen, W. and M. James (1997)**, "Assessment and Learning: Differences and Relationships between Formative and Summative Assessment", *Assessment in Education*, Vol. 4, pp. 365- 379.
- Haystead, M.W. and R.J. Marzaon (2010)**, "Preliminary Report: A Second Year Evaluation Study of Promethean ActivClassroom", Marzano Research Laboratory, Englewood, CO.
- Herl, H.D., H.F. O'Neil, Jr., G.K.W.K. Chung and J. Schacter (1999)**, "Reliability and Validity of a Computer-Based Knowledge Mapping System to Measure Content Understanding", *Computers in Human Behavior*, Vol. 15, pp. 315-334.
- Honig, B. (2001)**, Comment on Chester E. Finn, Jr., "State Academic Standards" *Brookings Papers on Education Policy*, No. 4, pp. 174 - 179
- Hylén, J., (2010)**, "Can Digital Learning Resources Spur Innovation?" in OECD (ed.), *Inspired by Pedagogy, Driven by Technology*, OECD, Paris, pp. 45 – 64.
- Knobel, M. and C. Lankshear (2006)**, "Weblog worlds and constructions of effective and powerful writing: cross with care and only where signs permit", in Pahl, K. and J. Roswell (eds.), *Travel Notes From the New*



Literacy Studies: Instances of Practice, Multilingual Matters, Clevedon.

Köller, O. (2001), "Mathematical World Views and Achievement in Advanced Mathematics: Findings from the TIMSS Population III", *Studies in Educational Evaluation*, Vol. 27, pp. 65-78.

Krampen, G. (1987), "Differential Effects of Teacher Comments", *Journal of Educational Psychology*, Vol. 79, No. 2, pp. 137-146.

Langworthy, M., L. Shear and B. Means (2010), "The Third Lever: Innovative Teaching and Learning Research" in OECD (ed.), "Inspired by Technology, Driven by Pedagogy", OECD, Paris, pp. 105 – 124.

Lave, J. and E. Wenger (1991), "Situated Learning: Legitimate Peripheral Participation", Cambridge University Press, Cambridge.

Lawley, J. (2003), E-gramm, <http://www.innova.uned.es>

Lewis, A. (1998), "Assessing Student Achievement: Search for Validity and Balance", University of California, National Center for Research on Evaluation, Standards, and Student Testing (CRESST), Los Angeles.

Looney, J. (2009), "Assessment and Innovation in Education", OECD Education Working Papers No. 24, OECD Publishing, Paris.

Lund, A. and O. Smørdal (2006), "Is There a Space for the Teacher in a Wiki?", in *Proceedings of the 2006 International Symposium on Wikis*, pp. 37-46, Odense, Denmark.

Means, B. (2010), "What Should We Assess with Technology?", Center for Technology in Learning, SRI International, presented at A Workshop for the University of Hong Kong (<http://www.slideshare.net/citechku/day1-workshop-means>, accessed December 2010).

Miller, R., H. Shapiro and K.E. Hilding-Hamann (2008), "School's Over: the Rise of Learning Spaces: Report on Learning Spaces in Europe in 2020 – A Rigorous Imagining Exercise", commissioned by the Institute for Prospective Technological Studies, in cooperation with the Danish Technological Institute, October, EUR 23532 EN – 2008.

Mischo, C. and F. Rheinberg (1995), "Erziehungsziele von Lehrern und individuelle Bezugsnormen der Leistungsbewertung", *Zeitschrift für Pädagogische Psychologie*, Vol. 9, pp. 139-151.

Mislevy, R.J. et al. (2001), "Making Sense of Data From Complex Assessments", University of California, National Center for Research on Evaluation, Standards, and Student Testing (CRESST), Los Angeles.

Motluk, A. (2005), "Do Games Prime Brain (Sic) for Violence?", *New Scientist*, Vol. 186, No. 2505, 25 June, p. 10.

Murphy, R.F., W.R. Penuel, B. Means, C. Korbak, A. Whaley and J.E. Allen (2002), "E-DESK: A Review of Recent Evidence On the Effectiveness of Discrete Educational Software", SRI International, Menlo Park, CA.

Nuthall, G. and A. Alton Lee (1995), "Assessing Classroom Learning: How Students Use Their Knowledge and Experience to Answer Classroom Achievement Test Questions in Science and Social Studies", *American Educational Research Journal*, Vol. 32, No.1, pp. 185-223.

OECD (2005), *Formative Assessment: Improving Learning in Secondary Classrooms*, J. Looney (Ed.) OECD, Paris.

OECD (2007), *Understanding the Brain: Birth of a Learning Science*, OECD, Paris.

OECD (2008), "New Millennium Learners: Initial Findings on the Effects of Digital Technologies on School-age Learners", OECD, Paris.

OECD (2009), "Creating Effective Teaching and Learning Environments: First Results from TALIS", OECD, Paris.

OECD (2010), *PISA 2009 Results: What Students Know and Can Do*, OECD Paris

Osmundson, E., G.K.W.K. Chung, H.E. Herl and D.C.D. Klein (1999), "Knowledge Mapping in the Classroom: A Tool for Examining the Development of Students' Conceptual Understandings", University of California, National Center for Research on Evaluation, Standards, and Student Testing (CRESST), Los Angeles.

Pedré, (2010), "Inspired by Technology, Driven by Pedagogy: A Systemic Approach to Technology-based School Innovations", OECD, Paris

Pellegrino, J.W., G.P. Baxter and R. Glaser (1999), "Addressing the 'Two Disciplines' Problem: Linking Theories of Cognition and Learning with Assessment and Instructional Practice", *Review of Research in Education*, Vol. 24, pp. 307 – 353.

Roschelle, J., W.R. Penuel and L.A. Abrahamson (2004), "Classroom Response and Communication Systems: Research Review and theory", Paper presented at the annual meeting of the American Educational Research Association, San Diego, CA.

Sanz, F. (2008), "Spanish-language Literature Review", In Looney, Teaching, Learning and Assessment for Adults: Improving Foundation Skills, OECD.

Schunk, D.H. (1996), "Goal and Self-evaluative Influences during Children's Cognitive Skill Learning", *American Educational Research Journal*, 33, pp. 359-382.



Selwyn, N. (2010), "Web 2.0 and the School of the Future, Today", in OECD (ed.), OECD, Paris.

Shavelson, R. M.A. Ruiz-Primo, M. Li and C. Cuahtemoc Ayala (2003), "Evaluating New Approaches to Assessing Learning", University of California, National Center for Research on Evaluation, Standards, and Student Testing (CRESST), Los Angeles.

Shepard, L.A. (2006), "Classroom Assessment", in R.L. Brennan (ed.), Educational Measurement (4th ed.), American Council on Education/Praeger Publishers, Westport, CT, pp.623-646.

Stiggins, R. J., M.M. Griswold, and K.R. Wikelund (1989), "Measuring Thinking Skills through Classroom Assessment", Journal of Educational Measurement, Vol. 26, pp. 233-246.

Stiggins, R. (1999), "Student Confidence and School Success", The Phi Delta Kappan, Vol. 81, pp. 191 –198.

Tuomi, I. (2001), "Internet, Innovation, and Open Source: Actors in the Network", First Monday, Vol. 6, No. 1, http://firstmonday.org/issues/issue6_1/tuomi/index.html

Tuomi, I. (2002), Networks of Innovation: Change and Meaning in the Age of the Internet, Oxford University Press, Oxford.

Tuomi, I. (2006), "The Future of Learning in the Knowledge Society: Disruptive Changes for Europe by 2020", in: Punie, Y. & M. Cabrera (eds.), The Future of ICT and Learning in the Knowledge Society, European Commission, Luxembourg, pp. 47-85.

Vendlinski, T.P., G. C. Delacruz, R.E. Buschang, G.K.W.K. Chung and E.L. Baker (2010), "Developing High-quality Assessments that Align with Instructional Video Games", University of California, National Center for Research on Evaluation, Standards, and Student Testing (CRESST), Los Angeles.

Voogt, J. (2009), "How Different Are ICT-supported Pedagogical Practices from Extensive and Non-extensive ICT-using Science Teachers?", Education and Information Technologies, Vol. 14, pp 325 – 343.

Vygotsky, L (1978), Mind in Society: Development of Higher Psychological Processes, 14th edition, Harvard University Press, Cambridge, MA.

Wang, J. and M.S. Brown (2007), "Automated Essay Scoring Versus Human Scoring: A Comparative Study", The Journal of Technology, Learning, and Assessment", Vol. 6, No. 2, retrieved November 2010 from <http://escholarship.bc.edu/jtla/>

West, P. et al (2010), "A Bayesian Network Approach to Modeling Learning Progressions and Task Performance", University of California, National Center for Research on Evaluation, Standards, and Student Testing (CRESST), Los Angeles.

William, D. (2006), "Formative Assessment: Getting the Focus Right", Educational Assessment, Vol. 11, pp. 283 – 289.

William, D. et al. (2004), "Teachers Developing Assessment for Learning: Impact on Student Achievement", Assessment in Education: Principles, Policy and Practice, Vol. 11, No. 1, pp. 49-65.

Williams, J. and J. Ryan (2000), "National Testing and the Improvement of Classroom Teaching: Can the Coexist?", British Educational Research Journal, Vol. 26, pp. 49 -73.

World Innovation Summit for Education – WISE, (2010), Session 1.6 Assessment and Education Improvement, Doha, December 6-8, 2010 <http://www.wise-qatar.org>

Yarnall, L, N. Shechtman and W.R. Penuel (2006), "Using Hand-held Computers to Support Improved Classroom Assessment in Science: Results from a Field Trial", Journal of Science Education and Technology, Vol. 15, No. 2, pp. 142 – 158, DOI: 10.1007/s10956-006-9008-4.