Chapter 2
Defining Twenty-First Century Skills

Marilyn Binkley, Ola Erstad, Joan Herman, Senta Raizen, Martin Ripley, May Miller-Ricci, and Mike Rumble

Abstract As the previous chapter indicates, there has been a significant shift in advanced economies from manufacturing to information and knowledge services. Knowledge itself is growing ever more specialized and expanding exponentially. Information and communication technology is transforming the nature of how work is conducted and the meaning of social relationships. Decentralized decision making, information sharing, teamwork, and innovation are key in today's enterprises. No longer can students look forward to middle class success in the conduct of manual labor or use of routine skills – work that can be accomplished by machines. Rather, whether a technician or a professional person, success lies in being able to communicate, share, and use information to solve complex problems, in being able to adapt and innovate in response to new demands and changing circumstances, in being able to marshal and expand the power of technology to create new knowledge, and in expanding human capacity and productivity.

Research during the last decade has shown how new social practices evolve due to increased use of new digital technologies, especially among young people (Buckingham and Willett 2006). Such practices create reconceptions of

M. Binkley (✉)
University of Luxembourg
e-mail: marilyn.binkley@uni.lu
O. Erstad
University of Oslo
J. Herman
University of California
S. Raizen (retired)
M. Ripley
World Class Arena Limited
M. Miller-Ricci
WestEd, San Francisco, California
M. Rumble
World Class Arena Limited

P. Griffin et al. (eds.), *Assessment and Teaching of 21st Century Skills*,
DOI 10.1007/978-94-007-2324-5_2,

key competencies and skills, not defined from a systems level but from the everyday lives of people in our societies. One example is research done on computer games and online communities (Gee 2007), where problem solving is defined as a key component of such practices. Such experiences of problem solving among young people need to inform us in the way we design assessment tasks and define key competencies. Hence, new standards for what students should be able to do must replace the basic skills and knowledge expectations of the past. To meet this challenge, schools must be transformed in ways that will enable students to acquire the sophisticated thinking, flexible problem solving, and collaboration and communication skills they will need to be successful in work and life. New conceptions of educational standards and assessment, the subject of this chapter, are a key strategy for accomplishing the necessary transformation. Such standards and assessment can both focus attention on necessary capacities and provide data to leverage and evaluate system change. Technology too serves as both a driver and lever for the transformation.

In the sections that follow, we

- synthesize research on the role of standards and assessment in promoting learning,
- describe the nature of assessment systems that can support changes in practice and use these to develop guiding principles for the design of next generation assessments,
- illustrate the use of technology to transform assessment systems and learning, and
- propose a MODEL for assessing twenty-first century skills.

Our intent is to learn from the past as we prepare for new futures in educational standards and assessment. While we provide a list of twenty-first century skills based on our analysis of twelve relevant frameworks drawn from a number of countries, these serve as an example of how to think about assessing twenty-first century skills. We expect that educators, as they consider our model, may need to make adaptations that fit their own contexts as they design assessments appropriate for their schools and students.

We have organized the ten skills we have identified into four groupings:

Ways of Thinking

1. Creativity and innovation
2. Critical thinking, problem solving, decision making
3. Learning to learn, Metacognition

Ways of Working

4. Communication
5. Collaboration (teamwork)

Tools for Working

6. Information literacy
7. ICT literacy

Living in the World

8. Citizenship – local and global
9. Life and career
10. Personal and social responsibility – including cultural awareness and competence

The Role of Standards and Assessment in Promoting Learning

The Importance of Standards That Promote Learning

Worldwide research has established the significant role that curriculum standards and assessment can play in molding new expectations for learning. Although the terminology of standards-led reform may have been initially associated with accountability and improvement initiatives in the USA (e.g., National Center on Education and the Economy 1998; No Child Left Behind Act 2001), the approach has widespread currency in educational systems as divergent as England, Germany, Norway, Singapore, and Australia, to name just a few. The basic ideas followed by these accountability and school improvement systems have rested on three principles:

- Be clear about expectations by establishing standards
- Develop high visibility (sometimes referred to as high stakes) assessments based on the standards
- Use the assessments to communicate what is expected to hold relevant stakeholders accountable and to publish data to inform decisions.

Such standards-based assessments provide empirical evidence for judging performance and can serve a variety of decision-making purposes (accountability, selection, placement, evaluation, diagnosis, or improvement), but the very existence of the assessments and the attention they engender carry important social, motivational, and political consequences.

Researchers around the globe studying such assessments have found fairly uniform effects. This is documented by a number of examples: studies of state accountability assessments in more than a dozen states in the USA, of A- or GCSE or Key Stage Exams in England, and of language and higher education admissions testing programs in countries such as Australia, China, Israel, Japan, New Zealand, and Sri Lanka, and areas such as Central and Eastern Europe (see, for example, Cheng et al. 2004; Herman 2008; Wall 2005). In summary:

- *Assessments signal priorities for curriculum and instruction; high visibility tests serve to focus the content of instruction*. School administrators and teachers pay attention to what is tested, analyze test results, and adapt curriculum and teaching accordingly.
- *Teachers tend to model the pedagogical approach reflected on high visibility tests.* When high visibility assessments are composed of multiple-choice items, teachers tend to rely heavily on multiple-choice worksheets in their classroom

instruction and emphasize lower level cognitive skills. However, when the assessments use extended writing and/or performance assessments, teachers incorporate similar activities in their classroom practice.

- *Curriculum developers, particularly commercial interests, respond* to important tests by modifying existing textbooks and other instructional materials and/or developing and marketing new ones to address test expectations. These products in turn may become primary resources that influence practice and also influence teachers' understandings of test expectations. At the same time research documents effects that can propel productive changes in practice. Thus, it too shows the potential for substantial negative consequences.
- *Schools and teachers tend to focus on what is tested rather than on what the underlying standards or learning goals* are and to ignore what is not tested. Both the broader domain of the tested disciplines and important subjects that are not tested may get short shrift. In the USA, England, and other countries, tests tend to give relatively little attention to complex thinking and problem solving and focus on lower levels of learning, which can lead to similar emphases in classroom practice.
- *Focusing on the test, rather than underlying learning, may encourage a one-time performance orientation and transmission-type teaching.* When doing well on the test, rather than learning, becomes the goal, schools may unwittingly promote a performance orientation in students, which in turn can work against students' engagement and persistence in learning, metacognition, and self-regulation. Especially for high visibility multiple-choice tests, teachers may concentrate on helping students acquire specific content rather than helping students build conceptual understandings and problem-solving capabilities.
- *Instructional/teaching time is diverted to specific test preparation activities*. Schools provide students with practice on the specific types of tasks and formats that are expected on the test through commercial test preparation packages, special classes, and homework. Such activities aim specifically to help students do well on the test, rather than promoting students' learning, and depending on the school and the pressure to improve test scores, can divert weeks or more of instructional time.

These consequences and caveats underscore an important challenge in using assessments to promote twenty-first century skills. The research clearly shows that whatever is measured matters and that educators tend to model and mimic the content and format of high visibility assessments in their curriculum and instruction and use a significant amount of classroom time for special test preparation activities. In some countries, however, testing has become dominated by routine and highly predictable items, which are also often short and highly scaffolded, thus reducing the expectation that students should apply knowledge, skills, and broader capabilities demanded by today's world. For example, analyses of annual state, standards-based tests in the USA show a preponderance of items addressing lower level cognitive demand to the detriment of complex thinking and problem-solving applications (see Webb 1999). Other countries provide more promising examples.

For instance, end of secondary school/university access examinations such as the Baccalaureate, the Matura, Abitur, etc. probe in depth the content and skills that students are expected to acquire and call on students to demonstrate their knowledge and skills in a wide variety of oral and written formats and project-based work. In the Nordic countries, there is a tradition of integrating project work into the curriculum promoting more locally adapted and general standards for assessment. Such examples involve students in important, authentic performances. Even so, the assessment standards for these exams have not yet been fully updated to reflect the demands of an information and innovation age, nor do they take advantage of twenty-first century technology. Just as students need to be literate in new media and be able to harness their power, so too technology can open up new, cost-effective possibilities for the design and use of a new generation of assessments.

Assessment Systems That Promote Learning

The contrast between US-type accountability exams and promising, secondary and university access examinations is also noteworthy in that the latter are embedded in coursework rather than external to it, where they can become an integral part of the teaching and learning process. The exams establish meaningful goals on which course assignments and assessments can be built and are used regularly to assess and respond to student progress. Research shows the powerful effect that ongoing assessment, so-called formative assessment, has on student learning, particularly for low-ability students (Black and Wiliam 1998); OECD 2005).

The use of assessment information is key to the idea: To be considered formative, assessment evidence must be *acted upon* to inform subsequent instruction. Rather than focusing backward on what has been learned, formative assessment helps to chart the learning road forward, by identifying and providing information to fill any gaps between the learners' current status and goals for learning. Moreover, more than solely a source of evidence that informs subsequent teaching and learning, carefully crafted formative assessments can directly support the learning process by incorporating principles of learning and cognition (Herman and Baker 2009; Bennett and Gitomer 2009). For example, by asking students to make public their thinking, formative probes can provide scaffolding that helps students confront their misconceptions, refine and deepen their understandings, and move to more sophisticated levels of expertise (Shepard et al. 2005; Herman and Baker 2005). By asking students for explanations and providing practice over multiple and authentic contexts, assessment tasks can help students to connect new knowledge to their existing structures and build transfer capability (see, for example, Sweller 2003; Holyoak 2005; Ericsson 2002; Gick and Holyoak 1983). By making learning goals explicit and involving students in self-assessment, formative assessment also can promote students as agents in their own learning, increasing student motivation, autonomy, and metacognition, as well as learning (Black et al. 2006; Shepard 2007; Harlen 2006; Gardner 2006). Such characteristics can be similarly incorporated into accountability assessments to increase their learning value.

The Nature of Quality Assessment Systems

Learning-Based Assessment Systems

Assessment design and development must bring together the rich, existing research base on student learning and how it develops with state-of-the-art psychometric theory to produce a new generation of assessments. As a prominent panel in the USA stated:

> Every assessment […] rests on three pillars: a model of how students represent knowledge and develop competence in a subject matter domain; tasks or situations that allow one to observe students' performance; and an interpretation method for drawing inferences from the performance evidence thus obtained (Pellegrino et al. 2001, p. 2).

Adopting this general model, Fig. 2.1 is intended to communicate that quality assessment starts, and ends with clearly specified and meaningful goals for student learning (see also Baker 2007; Forster and Masters 2004; Wilson and Sloane 2000). The assessment task vertex signals that any learning-based assessment must elicit responses that can reveal the quality of student understandings and/or where students are relative to the knowledge and skills that comprise intended learning goals. The interpretation link reinforces the idea that responses from assessment tasks must be specially analyzed and synthesized in ways that reveal and support valid inferences

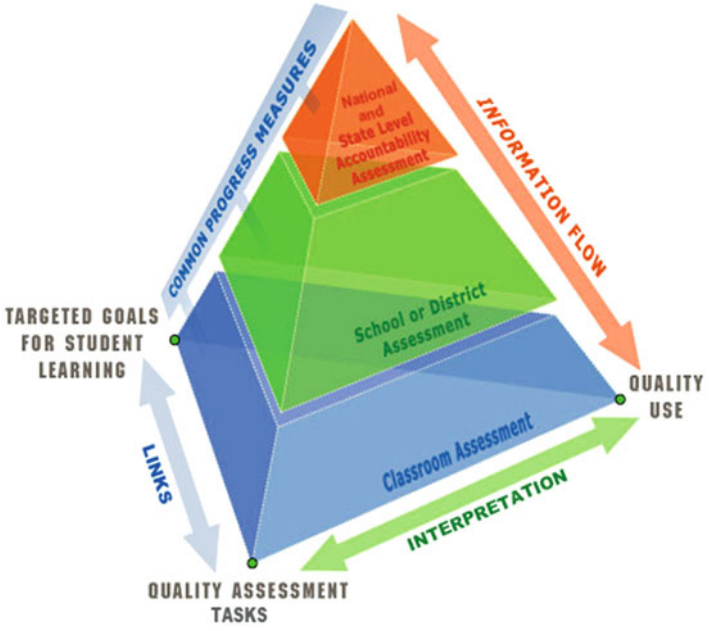

Fig. 2.1 Integrated assessment system

that connect to intended uses of the assessment. The use vertex highlights that results must be used for student learning relative to initial goals. Assessment quality then resides in the nature of the relationships between and among all three vertices and their connections — in the relationship between learning goals and tasks used to assess their development, in how well the analysis and scoring schemes capture important dimensions of intended understandings and skills, and in how well they support use and are used to improve learning. Inherent here too are the more traditional dimensions of validity, accuracy, and fairness of interpretations of student learning and — particularly for external and higher stakes tests — evidence that interpretations and inferences are justified (see Chap. 3).

As Fig. 2.1 shows, there are multiple levels for which data may be gathered and used for various decision-making purposes, from ongoing data to inform and enrich classroom teaching and learning (see Chap. 5), to periodic data to support policy and practical decision-making at higher levels of the educational system — e.g., school, district, province, state, and national. Importantly, large-scale international, national, and/or state or provincial assessments, for example, may provide policymakers a general barometer for judging and responding to schools' progress in promoting student learning, for allocating resources, and identifying locales that need help, etc. Schools and teachers may use the same data to evaluate their programs, refine their curricula, frame improvement plans, and/or identify individual students who need special attention. But to fuel ongoing decisions to optimize teaching and learning, teachers need a more continuous flow of data. Figure 2.1 implies a system of assessments, grounded in a common, well-specified set of learning goals that is purposively designed to satisfy the decision-making needs of all actors within and across the educational enterprise. Such a system needs to be aligned with the twenty-first century skills that will enable students' future success. Large-scale assessments can serve an important function in communicating and signaling what these skills are, as well as provide important models of how they can be assessed.

Improving the Quality of Assessment Systems

This system perspective also requires a different vantage point for considering assessment quality. Rather than focusing only on a single test, we need to consider the quality of the system for providing valid evidence to support the varied decision-making needs at multiple levels of the educational system. Balanced assessment seems an overriding criterion (Bell et al. 1992). Pellegrino et al. (2001), for example, argued for the development of balanced assessment systems to serve both accountability and policy purposes, as well as those of improving classroom teaching and learning. A balanced system, in their view, incorporates three critical principles: *coherence, comprehensiveness, and continuity*.

- A *coherent* assessment system is built on a well-structured conceptual base — an expected learning progression, which serves as the foundation both for large-scale and classroom assessments. That foundation should be consistent and

complementary both across administrative or bureaucratic levels of the education system and across grades.

- A *comprehensive* assessment system uses a range of assessment methods to ensure adequate measurement of intended constructs and measures of different grain size to serve decision-making needs at different levels of the education system. Inherently, a comprehensive assessment system is also useful in providing productive feedback, at appropriate levels of detail, to fuel accountability and improvement decisions at multiple levels.
- *Continuity* captures the principle that assessment at all levels is conceived as part of a continuous stream of evidence that tracks the progress of both individual students and educational programs over time. This can only be possible when there is consistency in the definition of the constructs across time, e.g., from the beginning to the end of the year and across grades.

While inherent in the above formulation, *fairness* is also a fundamental principle for assessment systems. All assessments should be designed to enable the broadest possible population of students to show what they know, without being unfairly hampered by individual characteristics that are irrelevant to what is being assessed. For example, students who are not proficient in the language of the test and test items may well find it difficult to show their mathematics capability; and students from one culture may lack the background knowledge to deal with a reading passage about a context with which they are unfamiliar. Disabled or very-low-ability students may be below the learning threshold on which a test is based. A fair system of assessment offers accommodations for students who may need them and is sensitive to the range of student abilities and developmental levels likely in the assessed population.

Principles for Twenty-First Century Standards and Assessments

While it should be clear that large-scale state, national, regional, or international assessments should be conceived as only part of any system to support student learning, assessments at each level represent a significant opportunity to signal the important learning goals that should be the target of the broader system as well as to provide valuable, actionable data for policy and practice. Moreover, carefully crafted, they can model next generation assessments that, through design and use, can support learning. To do so, our review to this point suggests that twenty-first century standards and assessments should:

- *Be aligned with the development of significant, twenty-first century goals*. Assessments that support learning must explicitly communicate the nature of expected learning. Standards and assessments must fully specify the rich range of twenty-first knowledge and skills students are expected to understand and apply. In addition, the standards and assessments should ideally represent how that knowledge and set of skills is expected to develop from novice to expert performance.

- *Incorporate adaptability and unpredictability*. One hallmark of twenty-first century demands is the need to adapt to evolving circumstances and to make decisions and take action in situations where prior actions may stimulate unpredictable reactions that in turn influence subsequent strategies and options. Dealing with such uncertainty is essential, but represents a new challenge for curriculum and assessment
- *Be largely performance-based.* The crux of twenty-first century skills is the need to integrate, synthesize, and creatively apply content knowledge in novel situations. Consequently, twenty-first century assessments must systematically ask students to apply content knowledge to critical thinking, problem solving, and analytical tasks throughout their education, so that we can help them hone this ability and come to understand that successful learning is as much about the process as it is about facts and figures.
- *Add value for teaching and learning*. The process of responding to assessments can enhance student learning if assessment tasks are crafted to incorporate principles of learning and cognition. For example, assessment tasks can incorporate transfer and authentic applications and can provide opportunities for students to organize and deepen their understanding through explanation and use of multiple representations.
- *Make students' thinking visible*. The assessments should provide a window into students' understandings and the conceptual strategies a student uses to solve a problem. Further, by making students' thinking visible, assessments thus provide a model for quality practice.
- *Be fair.* Fair assessments enable all students to show what they know and provide accommodations for students who would otherwise have difficulty accessing and responding to test items for reasons other than the target of the assessment.
- *Be technically sound.* Assessment data must provide accurate and reliable information for the decision-making purposes for which they are intended to be used. In the absence of reasonable measurement precision, inferences from results, and decisions based on them may well be faulty. The requirement for precision relative to intended purposes means both that intended uses and users must be clearly specified and evidence of technical quality must be established for each intended purpose. Establishing evidence of quality for innovative approaches to assessing twenty-first century skills may well require new psychometric approaches.
- *Valid for purpose.* To the extent an assessment is intended to serve as an indicator of schools' success in helping students acquire twenty-first century skills, skills and test results must be both instructionally sensitive and generalizable. That is, instructionally sensitive tests are influenced by the quality of instruction. Students who receive high-quality instruction should out-perform those who do not. The alternative is that students' basic ability or general intelligence, which are not under a school's control, are the reason for performance. A generalizable result transfers to other real-life applications.
- *Generate information that can be acted upon and provides productive and usable feedback for all intended users.* Teachers need to be able to understand what the assessment reveals about students' thinking. School administrators, policymakers,

and teachers need to be able to use this assessment information to determine how to create better opportunities for student learning.
- *Provide productive and usable feedback for all intended users.* It seems axiomatic that if stakeholders such as teachers, administrators, students, parents, and the public are expected to use the results of an assessment, they must have access to reports that are accurate, understandable, and usable.
- *Build capacity for educators and students.* Feedback from assessments can help students, teachers, administrators, and other providers to understand the nature of student performance and the learning issues that may be impeding progress. Teachers and students should be able to learn from the process.
- *Be part of a comprehensive and well-aligned system of assessments designed to support the improvement of learning at all levels of the educational hierarchy.*

Using Technology to Transform Assessment and Learning

The following sections of this paper address large-scale assessments. Chapter 5 deals more explicitly with classroom assessments.

Assessment Priorities Enabled by Information and Communication Technology

In this section, we draw attention to three areas where ICT has greatly increased the potential for assessing twenty-first century skills. ICT can be thought of not only as a tool for traditional assessments but also as presenting new possibilities for assessing skills formerly difficult to measure. ICT also develops new skills of importance for the twenty-first century. As much as we need to specify the skills needed, we also need to specify approaches that might measure the extent to which students have acquired them. During the last decade, several initiatives have explored how ICT might be used for assessment purposes in different ways in different subject domains. The discussion below is based on a review of relevant research in this area.

Although assessment in education is a substantial research field, it has only been during the last decade that ICT-based assessment has been growing as a research field (McFarlane 2003). This is partly due to an increase in developments of the ICT infrastructure in schools with expanded access to hardware, software, and broadband internet connections for students and teachers. Existing research has examined both the impact of ICT on traditional assessment methods and how ICT raises new issues of assessment and skills. For example, as part of the Second International Technology in Education Study (Kozma 2003), innovative ICT-supported pedagogical practices were analyzed. In several countries, some of these practices demonstrated a shift toward more use of formative assessment methods when ICT was

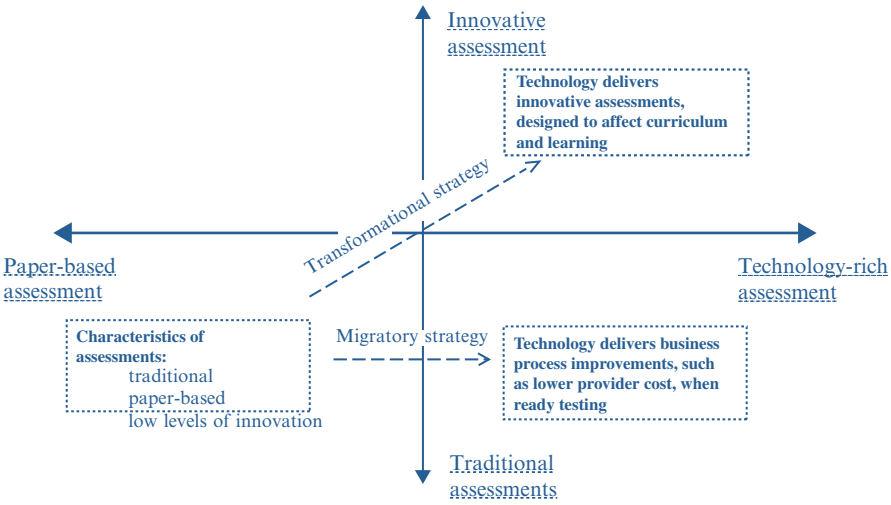

Fig. 2.2 The dimensions of e-assessment innovations

introduced (Voogt and Pelgrum 2003). However, in most practices, often new and old assessment methods coexisted because schools had to relate to national standards and systems over which they had no control, while they were simultaneously developing alternative assessment methods for their own purposes.

The use of the term e-assessment has gained acceptance in recent years. Advocates of e-assessment frequently point to the efficiency benefits and gains that can be realized. These benefits might have to do with the costs of test production, the ability to reuse items extensively or to create power and adaptive tests, or to build system improvements such as test administration systems, which are able to provide tests whenever students want to take them. However, in the report *Effective practice with e-assessment* (Whitelock et al. 2007), the writers conclude that e-assessment is "much more than just an alternative way of doing what we already do." Through evidence and case studies, the report provides examples of e-assessment widening the range of skills and knowledge being assessed, providing unprecedented diagnostic information, and supporting personalization (Ripley 2007). Thus, we argue that e-assessment has the potential of using technology to support educational innovation and the development of twenty-first century skills, such as complex problem solving, communication, team work, creativity and innovation.

Figure 2.2 provides a representation of the contrast between the two drivers: the business efficiency gains versus the educational transformation gains. The lower-left quadrant represents traditional assessments, typically paper-based and similar year-on-year. Most school- and college-based assessments are of this type. Moving from the lower-left to the lower-right quadrant represents a migratory strategy in which paper-based assessments are migrated to a screen-based environment. Delivery is more efficient, but assessments are qualitatively unchanged. In contrast,

moving to the upper-right quadrant represents a transformational strategy in which technology is used to support innovative assessment designed to influence (or minimally to reflect) innovation in curriculum design and learning.

The Migratory Strategy with ICT

Conceptions of twenty-first century skills include some familiar skills that have been central in school learning for many years, such as information processing, reasoning, enquiry, critical thinking, and problem solving. The question is: To what extent does ICT enhance or change these skills and their measurement? Indeed, during the last decade most of the research on the use of ICT for assessment has dealt with the improvement of assessment of traditional skills — improvement in the sense that ICT has potential for large-scale delivery of tests and scoring procedures, easily giving the learner accessible feedback on performances. For example, many multiple-choice tests within different subject domains are now online. The focus is then on traditional testing of reasoning skills and information processing among students, on memorization, and on reproduction of facts and information. Using online tests will make this more cost-effective and less time-consuming. However, there are several concerns raised about assessment of traditional skills in an online setting, especially regarding security, cheating, validity, and reliability.

Many countries and states have adopted a "dual" program of both computer-based and paper-and-pencil tests. Raikes and Harding (2003) mention examples of such dual programs in some states in the U.S. where students switch between answering computer-based and paper-and-pencil tests. The authors argue that assessments need to be fair to students regardless of their schools' technological capabilities and the need to avoid sudden discontinuities so that performance can be compared over time. This may require a transitional period during which computer and paper versions of conventional external examinations run in parallel. They sketch some of the issues (costs, equivalence of test forms, security, diversity of school cultures and environments, technical reliability) that must be solved before conventional examinations can be computerized. In a meta-evaluation of initiatives in different states in the US, Bennett (2002) shows that the majority of these states have begun the transition from paper-and-pencil tests to computer-based testing with simple assessment tasks. However, he concludes, "If all we do is put multiple-choice tests on computer, we will not have done enough to align assessment with how technology is coming to be used for classroom instruction" (pp. 14–15).

Recent developments in assessment practices can be seen as a more direct response to the potential of ICT for assessment. An example of such developments is the effort to use computers in standardized national exams in the Netherlands, going beyond simple multiple-choice tests. The domain for the assessment is science, where exams contain 40% physics assignments which have to be solved with computer tools such as modeling, data video, data processing, and automated control technique (Boeijen and Uijlings 2004).

Several studies comparing specific paper-and-pencil testing with computer-based testing have described the latter as highly problematic, especially concerning issues of test validity (Russell et al. 2003). Findings from these studies, however, show little difference in student performance (Poggio et al. 2005), even though there are indications of enough differences in performance at the individual question level to warrant further investigation (Johnson and Green 2004). There are differences in prior computer experience among students, and items from different content areas can be presented and performed on the computer in many different ways, which have different impacts on the validity of test scores (Russell et al. 2003). While some studies provide evidence of score equivalence across the two modes, computerized assessments tend to be more difficult than paper-and-pencil versions of the same test. Pommerich (2004) concludes that the more difficult it is to present a paper-and-pencil test on a computer, the greater the likelihood of mode effects to occur. Previous literature (Russell 1999; Pommerich 2004) seems to indicate that mode differences typically result from the extent to which the presentation of the test and the process of taking the test differ across modes rather than from differences in content. This may imply a need to try to minimize differences between modes. A major concern is whether computer-based testing meets the needs of all students equally and whether some are advantaged while others are disadvantaged by the methodology.

In a recent special issue of the *British Journal of Education Technology* focusing on e-assessment, several studies are presented where students' traditional skills are assessed in different ways (Williams and Wong 2009; Draper 2009; Shephard 2009).

The introduction of ICT has further developed an interest in formative ways of monitoring and assessing student progress. The handling of files and the possibility of using different modes of expression support an increased interest in methods such as project work (Kozma 2003), which can be used for formative assessment. The increased use of digital portfolios in many countries (McFarlane 2003) is an example of how formative assessment is gaining importance. Although the use of portfolio assessments is not new and has been used for some time without ICT (see e.g., special issue in *Assessment in Education*, 1998, on portfolios and records of achievement; Koretz et al. 1998), the use of digital tools seems to have developed this type of assessment further by bringing in some new qualitative dimensions such as possibilities for sending files electronically, hypertexts with links to other documents, and multimodality with written text, animations, simulations, moving images, and so forth. As a tool for formative assessment, and compared to paper-based portfolios, digital portfolios make it easier for teachers to keep track of documents, follow student progress, and comment on student assignments. In addition, digital portfolios are used for summative assessment as documentation of the product students have developed and their progress. This offers greater choice and variety in the reporting and presenting of student learning (Woodward and Nanlohy 2004). This research indicates a strengthening of collaboration (teamwork) and self-regulated learning skills. Related research deals with critical thinking skills, an area of student competency highlighted in curricula in many countries. What is needed in the application of ICT to assessment is to look for new ways of making student

attainment visible in a valid and reliable way (Gipps and Stobart 2003; see also Thai school project, critical thinking skills, Rumpagaporn and Darmawan 2007).

In short, in the matter of measuring more traditional skills, development has been directed toward the delivery of large-scale tests on information handling and mapping levels of knowledge at different stages of schooling. Information literacy in this sense has become an important area of competence in itself, and even more so in relation to information sources on the internet. ICT is seen as an important tool in making assessment more efficient as well as more effective in measuring desired skills in traditional ways.

The Transformational Strategy with ICT

Although there are few instances of transformative e-assessment, the projects that do exist provide us with a compelling case for researching and investing in assessments of this type. There are exciting and effective examples of the use of ICT to transform assessment, and, therefore, learning. What is changing in the e-assessment field is usability. Where previously much of the preparatory work had to be done by third party or other technically expert staff, programs are increasingly providing end users with the tools to implement their own e-assessment. New technologies have created an interest in what some describe as "assessing the inaccessible" (Nunes et al. 2003) such as metacognition, creativity, communication, learning to learn, and lifelong learning skills (Anderson 2009; Deakin Crick et al. 2004). Below, we review the research on assessing complex skills that have been difficult to assess or not assessed at all with traditional tests.

The *review of advanced e-assessment techniques* project — commissioned by the Joint Information Systems Committee (JISC) in the UK — began by considering what constituted an advanced technique. "Advanced" refers to techniques that are used in isolated or restricted domains that have successfully applied technology to create an assessment tool. "Advanced" does not necessarily imply "newness." The project collated a long list of over 100 "advanced" e-assessment projects. It was a surprise how few previously unknown advanced e-assessment projects came to light through the trawl for information. The community of experts using e-assessment is small. This continues to have implications for scaling up e-assessment and for stimulating the growth of additional innovative approaches. A brief description of an advanced e-assessment developed in the UK is provided in Fig. 2.3.

One important aspect about the advances in e-assessment is that ICT brings new dimensions to what is being measured. Consider, for example, multimodality, or what Gunter Kress (2003) describes as multimodal literacy. How might different skills like creativity, problem solving, and critical thinking be expressed in different ways using different modes and modalities that ICT provides? The increased uses of visualization and simulation are examples of how ICT has made an impact on measurement of different skills, though so far the research has been inconclusive (Wegerif and Dawes 2004).

Four ICT skills were assessed:

1. Finding things out – obtaining information well matched to purpose by selecting appropriate sources; or, questioning the plausibility and value of information found.
2. Developing ideas and making things happen – using ICT to measure, record, respond to and control events.
3. Exchanging and sharing information – using ICT to share and exchange information, such as web publishing and video conferencing.
4. Reviewing, modifying and evaluating work as it progresses – reflecting critically on own and others' use of ICT.

The design included a simulated environment in which students complete tests; a desktop environment with software and tools for students; new ways of scoring student performances based on the ICT processes students used to solve problems rather than the products, and new ways of enabling access to tests for all students. In one case, an email ostensibly from the editor of a local news website would request students to research local job vacancies and prepare a vacancies page for the website. To complete this task, students would need to run web searches and email virtual companies to request more information about vacancies. The extent and quality of information available would vary, reflecting real-world web information. While completing the task, a student would receive further requests from the editor, perhaps changing deadlines or adding requirements. A student's work would be graded automatically.

The project provided proof-of-concept and identified the following major obstacles and challenges in developing a simulation-based assessment of 21st century skills

- Developing a psychometric approach to measuring and scaling student responses. Since the assessment is designed to collect information about processes used by students, a method is needed to collect data and create summary descriptions/analyses of those processes.
- Aligning schools' technology infrastructure to support wide-scale, high-stakes, computer-based testing.
- Communicating effectively to introduce new approaches to testing to a world of experts, teachers, students, parents and politicians, all of whom have their own mental models and classical approaches for evaluating tests.

Fig. 2.3 Innovative UK assessment of ICT skills of 14-year-olds

Creativity in particular is an area that has been growing in importance as a key twenty-first century thinking skill (Wegerif and Dawes 2004, p. 57). For example, Web 2.0 technology enables users to produce and share content in new ways: User-generated content creation and "remixing" (Lessig 2008) become creative practices that challenge the traditional relationships between teachers and students in providing information and content for learning and the role of the "school book" (Erstad 2008). The use of new digital media in education has been linked to assessment of creative thinking as different from analytic thinking (Ridgway et al. 2004). Digital camera and different software tools make it easier for students to show their work and reflect on it. However, one of the problems with the discussions around creativity has been the often simplified and naïve notions and romantic conceptions of the creative individual (Banaji and Burn 2007), without clear specifications of what this skill area might entail. Thus, it has proved to be difficult to assess students' creativity. In a systematic review of the impact of the use of ICT on students and teachers for the assessment of creative and critical thinking skills, Harlen and Deakin Crick (2003) argue that the neglect of creative and critical thinking in assessment methods is a cause for concern, given the importance of these skills for lifelong learning and in the preparation for life in a rapidly changing society. Their review documents a lack of substantial research on these issues and argues for more strategic research.

A second area of great interest concerns the way digital tools can support collaboration in problem solving, creative practices, and communication. There are many examples of how computer-based learning environments for collaboration can work to stimulate student learning and the process of inquiry (Wasson et al. 2003; Laurillard 2009). Collaborative problem-solving skills are considered necessary for success in today's world of work and school. Online collaborative problem-solving tasks offer new measurement opportunities when information on what individuals and teams are doing is synthesized along the cognitive dimension. Students can send documents and files to each other and, in this way, work on tasks together. This raises issues both for interface design features that can support online measurement and how to evaluate collaborative problem-solving processes in an online context (O'Neil et al. 2003). There are also examples of web-based peer assessment strategies (Lee et al. 2006). Peer assessment has been defined by some as an innovative assessment method, since students themselves are put in the position of evaluators as well as learners (Lin et al. 2001). It has been used with success in different fields such as writing, business, science, engineering, and medicine.

A third area of research with important implications for how ICT challenges assessment concerns higher-order thinking skills. Ridgway and McCusker (2003) show how computers can make a unique contribution to assessment in the sense that they can present new sorts of tasks, whereby dynamic displays show changes in several variables over time. The authors cite examples from the World Class Arena (www.worldclassarena.org) to demonstrate how these tasks and tools support complex problem solving for different age groups. They show how computers can facilitate the creation of micro-worlds for students to explore in order to discover hidden rules or relationships, such as virtual laboratories for doing experiments or games to explore problem-solving strategies. Computers allow students to work with complex data sets of a sort that would be very difficult to work with on paper. Tools like computer-based simulations can, in this way, give a more nuanced understanding of what students know and can do than traditional testing methods (Bennett et al. 2003). Findings such as those reported by Ridgway and McCusker (2003) are positive in the way students relate to computer-based tasks and the increased performances they exhibit. However, the authors also find that students have problems in adjusting their strategies and skills since the assessment results show that they are still tuned into the old test situation with correct answers rather than explanations and reasoning skills.

An interesting new area associated with what has been presented above is the knowledge-building perspective developed by Scardamalia and Bereiter (2006; see also Chap. 5). In developing the technological platform *Knowledge Forum*, Scardamalia and Bereiter have been able to measure students learning processes that have traditionally been difficult to assess. This platform gives the students the possibility of collective reasoning and problem solving building on each other's notes, often as collaboration between schools in different sites and countries. Some key themes in the research on these skills and their online measurement tools are:

- Knowledge advancement as a community rather than individual achievement
- Knowledge advancement as idea improvement rather than as progress toward true and warranted belief

- Knowledge of, in contrast to knowledge about
- Discourse as collaborative problem solving rather than as argumentation
- Constructive use of authoritative information
- Understanding as emergent

Similar points have been made by Mercer and Wegerif and colleagues in the UK (e.g., Mercer and Littleton 2007) in their research on "thinking together" and how we might build language for thinking, what they term as "exploratory talk." Computers and software have been developed for this purpose together with other resources. Wegerif and Dawes (2004, p. 59) have summarized the "thinking together" approach in four points, each of which assumes the crucial importance of teachers:

- The class undertakes explicit teaching and learning of talk skills that promote thinking
- Computers are used both to scaffold children's use of these skills and to bridge them in curriculum areas
- Introductions and closing plenaries are used to stress aims for talk and for thinking as well as to review progress
- Teacher intervention in group work is used to model exploratory talk

The above examples have shown how ICT represents the transformative strategy in developing assessments, especially formative assessment, and how the complexity of these tools can be used to assess skills that are difficult to assess by paper and pencil. As McFarlane (2001) notes, "It seems that use of ICT can impact favorably on a range of attributes considered desirable in an effective learner: problem-solving capability; critical thinking skill; information-handling ability." (p. 230) Such skills can be said to be more relevant to the needs of an information society and the emphasis on lifelong learning than those which traditional tests and paper-based assessments tend to measure.

Arriving at a Model Twenty-First Century Skills Framework and Assessment

In this section, we provide a framework that could be used as a model for developing large-scale assessments of twenty-first century skills. To arrive at this model framework we compared a number of available curriculum and assessment frameworks for twenty-first century skills and skills that have been developed around the world. We analyzed these frameworks to determine not only the extent to which they differ but also the extent to which these frameworks provide descriptions of twenty-first century learning outcomes in measureable form. Based on our analysis, we identified ten important skills that in our opinion typify those necessary for the twenty-first century. For each of the ten skills we have analyzed the extent to which the identified frameworks provide measurable descriptions of the skill, considering the *K*nowledge, *S*kills, and *A*ttitudes, *V*alues and *E*thics aspects of each skill. This framework is referred to as the *KSAVE* framework and is described in more detail below.

Existing Twenty-First Century Skills Frameworks

A number of organizations around the world have independently developed frameworks for twenty-first century skills. For the purposes of our analysis, we considered the frameworks listed in the chart appearing on the next page. To explore the number and range of modern twenty-first century curricula that are currently in place, wider searches were carried out for national education systems that build aspects of the ten KSAVE skills into their national curricula. Searches were made for "national" curricula, references to "twenty-first century learning," and references to "skills" and "competency-based" standards. A relatively small number of nations define a national curriculum in detail, while a larger number have national aims or goals for their education system. A growing number of countries are undertaking significant reviews of their national curricula. A small number are undertaking the task of developing their first national curriculum. "Twenty-first century learning needs" are frequently included within these new and revised curriculum documents. The sources are listed in Table 2.1.

With very few exceptions, references to twenty-first century knowledge, skills, or the individual attitudes and attributes of learners are contained within overarching statements of goals or educational aims. These are generally brief statements but are supported by justifications for change. For example, there are references to: the need to educate for new industry, commerce, technology, and economic structures; the need for new social interaction and communication skills; the need for imagination, creativity, and initiative; the need to learn and continue to learn throughout employment; the need to maintain national and cultural values; and the need to operate in an increasingly international and global environment. Few of the frameworks and curricula of national systems we have examined provide detailed descriptions or clearly elaborated curriculum standards. Similarly, few include descriptions of what the curriculum experienced by learners will actually look like if the broader aims of its framework are to be realized.

All the curricula reviewed maintain a subject structure. It is this structure that forms the basis for curriculum design. The naming and grouping of learning under subject titles may differ slightly between countries, but the general principles of learning a core curriculum (home language, mathematics, and science) are common. In many national curricula, the skills associated with ICT have been raised in status to this core while history, particularly national history, and indigenous culture, often including religion, form a secondary layer. Other subjects may be described individually or combined, for example as the "Arts" or "Humanities." Thus to date the teaching of twenty-first century skills has been embedded in the subjects that make up the school curriculum. It is not clear whether such skills as critical thinking or creativity have features in common in related subjects such as mathematics and science, let alone across the STEM fields and the arts and humanities. For other skills, however, such as information and ICT literacy, the argument has been made more frequently that these are transferrable. These questions of skill generalizability and transferability remain deep research challenges.

Table 2.1 Sources of documents on twenty-first century skills

Country/region	Document(s)
European Union	*Key Competencies for Lifelong Learning – A European Reference Framework, November 2004* Recommendation of the European Parliament and of the Council of 18 December 2006 on key competences for lifelong learning Implementation of "Education and Training 2010" work programme http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2006:394:0010:0018:en:PDF
OECD	New Millennium Learners Project: Challenging our Views on ICT and Learning www.oecd.org/document/10/0,3343,en_2649_35845581_38358154_1_1_1_1,00.html
USA (partnership for twenty-first century skills)	P21 Framework definitions P21 Framework flyer http://www.p21.org/documents/P21_Framework_Definitions.pdf
Japan	Center for Research on Educational Testing (CRET) www.cret.or.jp/e
Australia	*Melbourne declaration on educational goals for young Australians* www.mceecdya.edu.au/verve/_resources/National_Declaration_on_the_Educational_Goals_for_Young_Australians.pdf
Scotland	A curriculum for excellence – the four capabilities www.ltscotland.org.uk/curriculumforexcellence/index.asp
England	*The learning journey*
England	*Personal learning & thinking skills – the national curriculum for England* http://curriculum.qcda.gov.uk/uploads/PLTS_framework_tcm8-1811.pdf
Northern Ireland	*Assessing the cross curricular skills* http://www.nicurriculum.org.uk/key_stages_1_and_2/assessment/assessing_crosscurricular_skills/index.asp
ISTE	*National educational technology standards for students, second edition, global learning in the digital age* http://www.iste.org/standards.aspx
USA. National Academies, science for the twenty-first century	*Exploring the intersection of science education and the development of* twenty-first *century skills* http://www7.nationalacademies.org/bota/Assessment_of_21st_Century_Skills_Homepage.html
USA, Department of Labor	Competency models: *A review of the literature* *The role of the Employment and Training Administration* (ETA), Michelle R. Ennis

Where the aims and goals of twenty-first century learning are described in the frameworks we examined, they are generally specified as being taught through, within and across the subjects without the detail of how this is to be achieved or

what the responsibilities of each subject might be in achieving them. Without this depth of detail, these national statements of twenty-first century aims and goals are unlikely to be reflected in the actual learning experience of students or in the assessments that are administered. Without highly valued assessments of these twenty-first century aims or goals requiring their teaching, it is difficult to see when or how education systems will change significantly for the majority of learners.

The KSAVE Model

To structure the analysis of twenty-first century skills frameworks, an overall conceptual diagram was created. This diagram defines ten skills grouped into four categories:

Ways of Thinking

1. Creativity and innovation
2. Critical thinking, problem solving, decision making
3. Learning to learn, metacognition

Ways of Working

4. Communication
5. Collaboration (teamwork)

Tools for Working

6. Information literacy (includes research on sources, evidence, biases, etc.)
7. ICT literacy

Living in the World

8. Citizenship – local and global
9. Life and career
10. Personal and social responsibility – including cultural awareness and competence

Although there are significant differences in the ways in which these skills are described and clustered from one framework to another, we consider that the above list of ten is sufficiently broad and comprehensive to accommodate all approaches. At an early stage we found that frameworks for twenty-first century skills differ considerably in terms of the nature of their content. Some seek to define student behaviors; for example, an aspect of creativity might include "openness and responsiveness to new ideas." Other frameworks refer extensively to skills: for example, an aspect of creativity might refer to the ability to "develop innovative and creative ideas." A third category used in some frameworks refers to specific knowledge: for example, an aspect of creativity might be "knowledge of a wide range of idea creation techniques." Some frameworks cover two or more of these categories; few comprehensively cover all three. To accommodate and reflect these differences in

approach, we have designed three categories within the KSAVE model. Keep in mind that the model does not resolve the issue of subject-embedded knowledge, skills, and attitudes versus their generalizability across domains.

Knowledge

This category includes all references to specific knowledge or understanding requirements for each of the ten skills.

Skills

This category includes the abilities, skills, and processes that curriculum frameworks are designed to develop in students and which are a focus for learning.

Attitudes, Values, and Ethics

This category refers to the behaviors and aptitudes that students exhibit in relation to each of the ten skills.

The method used to complete the analysis of twenty-first century skills frameworks was to populate the KSAVE grid with indexes taken from each framework, retaining original wording as far as was sensible. Decisions were made to refine or amalgamate wording taken from frameworks where the intention appeared similar. Decisions were also made on whether to allocate indexes to knowledge, skills, or attitudes/values/ethics. For some of the indexes, the decision whether to allocate them to the skills category or to the attitudes/values/ethics category appeared to be marginal.

In the following pages, we present each group of skills and discuss some of the thinking behind the grouping. In addition, we provide examples of how the skills might be measured in an effort to open our eyes to what is possible. These example assessments really only scratch the surface of what is needed to measure twenty-first century skills.

Ways of Thinking

Together the three categories of skills under "Ways of thinking" represent a push forward in the conceptualization of thinking. These skills emphasize higher order thinking skills, and subsume more straightforward skills such as recall, and drawing inferences. A major characteristic of these skills is that they require greater focus and reflection.

Creativity and Innovation

Operational definitions of creativity and innovation are provided in Table 2.2. While creativity and innovation can logically be grouped together, they originate in two different traditional schools of thought. Creativity is most often the concern of

Table 2.2 Ways of thinking – creativity and innovation

Knowledge	Skills	Attitudes/values/ethics
Think and work creatively and with others • Know a wide range of idea creation techniques (such as brainstorming) • Be aware of invention, creativity, and innovation from the past within and across national boundaries and cultures • Know the real-world limits to adopting new ideas and how to present them in more acceptable forms • Know how to recognize failures and differentiate between terminal failure and difficulties to overcome *Implement innovations* • Be aware of and understand where and how innovation will impact and the field in which the innovation will occur • Be aware of the historical and cultural barriers to innovation and creativity	*Think creatively* • Create new and worthwhile ideas (both incremental and radical concepts) • Be able to elaborate, refine, analyze, and evaluate one's own ideas in order to improve and maximize creative efforts *Work creatively with others* • Develop, implement, and communicate new ideas to others effectively • Be sensitive to the historical and cultural barriers to innovation and creativity *Implement innovations* • Develop innovative and creative ideas into forms that have impact and can be adopted	*Think creatively* • Be open to new and worthwhile ideas (both incremental and radical) *Work creatively with others* • Be open and responsive to new and diverse perspectives; incorporate group input and feedback into the work • View failure as an opportunity to learn; understand that creativity and innovation is a long-term, cyclical process of small successes and frequent mistakes *Implement innovations* • Show persistence in presenting and promoting new ideas

cognitive psychologists. Innovation, on the other hand, is more closely related to economics where the goal is to improve, advance, and implement new products and ideas. Measuring both can be quite challenging. The tasks require an interactive environment, but they frequently cannot be done in the short period of time allocated to a large-scale assessment, nor are there good benchmarks against which respondent output can be evaluated.

Creativity is often described as a thinking skill or at least as an important aspect of thinking that can and should be fostered (Wegerif and Dawes 2004, p. 57). In a review of the connection between technology, learning, and creativity, Loveless (2007) shows how technology allows children to produce high quality finished products quickly and easily in a range of media that provide opportunities for creativity. Loveless argues that to foster creativity in the classroom, teachers need to create a social atmosphere in which children feel secure enough to play with ideas and to take risks.

Although, as noted above, it has proven to be difficult to assess creativity, the use of new digital media has been linked to assessment of creative thinking as different from analytic thinking (Ridgway et al. 2004). Digital cameras and different software tools make it easier for students to show their work and reflect on it. A number

of subjects in the school curriculum ask students to make various kinds of products. (Sefton-Green and Sinker 2000). These might include paintings in art class, creative writing in english, performance in drama, recording in music, videos in media studies, and multimedia "digital creations" in different subjects. There are so far not many examples of how ICT might influence assessment of such student products (Sefton-Green and Sinker 2000).

eSCAPE

The eSCAPE project does not test creativity and innovation, but it does test some aspects of this domain. Specifically it offers a glimpse of how we might test the ability to develop innovative and creative ideas into forms that have impact as well as showing persistence in presenting and promoting new ideas.

For many years, England's school examinations for 16-year-old students have included an optional assessment in Design and Technology. Traditionally, this examination includes a requirement for students to complete a design project of over 100 h duration and to submit a written report on the project. The report is graded.

In 2001, the Qualifications and Curriculum Authority commissioned the Technology Education Research Unit (TERU) at Goldsmiths College in London to undertake to develop a technology-led replacement to this traditional paper-based assessment. The result is an assessment completed in six hours, in a design workshop, with students working in groups of three or four. During the course of the six h, students are given a number of staged assessment instructions and information via a personal, portable device. The handheld device also acts as the tool to capture assessment evidence – via video, camera, voice, sketchpad, and keyboard. During the six hours, each student's design prototype develops, with the handheld device providing a record of progress, interactions, and self-reflections.

At the end of the assessment, the assessment evidence is collated into a short multimedia portfolio. Human raters, who score each student's responses, view this. eSCAPE directors turned to the work of Thurstone (1927) to develop a graded-pairs scoring engine to provide a holistic judgment on the students' work. This engine supports human raters in making a number of paired judgments about students' work. The result is an assessment that exhibits rates of reliability equal to, or slightly in excess of, the levels of reliability achieved on multiple-choice tests.

Critical Thinking, Problem Solving and Decision Making

Operational definitions of critical thinking and problem solving are provided in Table 2.3. Critical thinking and problem solving have become an increasingly important feature of the curriculum in many parts of the world. In the UK there are popular high school qualifications in critical thinking. In the USA, the American Philosophical Association has published the Delphi report on critical thinking (Facione 1990). This report identified six cognitive thinking skills: interpretation, analysis, evaluation, inference, explanation, and self-regulation. This framework was further elaborated

Table 2.3 Ways of thinking – critical thinking, problem solving, and decision making

Knowledge	Skills	Attitudes/values/ethics
Reason effectively, use systematic thinking and evaluate evidence • Understand systems and strategies for tackling unfamiliar problems • Understand the importance of evidence in belief formation. Reevaluate beliefs when presented with conflicting evidence *Solve problems* • Identify gaps in knowledge • Ask significant questions that clarify various points of view and lead to better solutions *Articulation* • Clearly articulate the results of one's inquiry	*Reason effectively* • Use various types of reasoning (inductive, deductive, etc.) as appropriate to the situation *Use systems thinking* • Analyze how parts of a whole interact with each other to produce overall outcomes in complex systems. Examine ideas, identify, and analyze arguments • Synthesize and make connections between information and arguments • Interpret information and draw conclusions based on the best analysis. Categorize, decode, and clarify information • Effectively analyze and evaluate evidence, arguments, claims, and beliefs • Analyze and evaluate major alternative points of view. • Evaluate. Assess claims and arguments • Infer. Query evidence, conjecture alternatives, and draw conclusions • Explain. State results, justify procedures, and present arguments. • Self-regulate, self-examine, and self-correct	*Make reasoned judgments and decisions* • Consider and evaluate major alternative points of view • Reflect critically on learning experiences and processes • Incorporate these reflections into the decision-making process *Solve problems* • Be open to non-familiar, unconventional, and innovative solutions to problems and to ways to solve problems • Ask meaningful questions that clarify various points of view and lead to better solutions *Attitudinal disposition* • Trustful of reason • Inquisitive and concerned to be well informed • Open and fair minded • Flexible and honest • Inquisitiveness and concern to be well informed • Alert to opportunities to use ICT • Trustful of and confident in reason • Open and fair minded, flexible in considering alternative opinions • Honest assessment of one's own biases • Willingness to reconsider or revise one's views where warranted

to include attitudinal- and values- based criteria: Students should be inquisitive, well informed, open-minded, fair, flexible, and honest. Research subsequent to the Delphi Report has shown that being "trustful of reason" (one of the Delphi Report's key findings) plays a vital role in what it means to think critically.

In contrast to creativity and innovation, critical thinking, problem solving, and decision making have been part of large-scale assessments for some time. Critical thinking frequently appears as part of reading, mathematics, and science assessments, with such assessments as the US National Assessment of Educational Progress and the OECD Program for International Student Achievement (PISA).

Problem solving has been a focused area of research for decades, yielding a number of definitions and frameworks. In addition, problem solving has appeared in various forms in a number of large-scale international assessments such as PISA and the Adult Literacy and Lifelong Learning Skills (ALL). These assessments specifically include items that are designed to measure how well students can evaluate evidence, arguments, claims, and warrants; synthesize and make connections between information and arguments; and analyze and evaluate alternative points of view. ALL 2003 focused on problem-solving tasks that were project oriented and most closely resembled analytic reasoning. Problem solving in mathematics and science has been part of the PISA assessment since its inception in 2000. In PISA 2003 a problem-solving scale that included three kinds of problems – decision-making, system analysis and design (and troubleshooting) was developed. For 2012, PISA will move beyond the 2003 scale by including dynamic items that may be linked to the OECD's Program for the International Assessment of Adult Competencies (PIAAC) 2011, where problem solving is in a technology rich environment is measured.

The following examples illustrate the direction of assessments for the twenty-first century. The first, Primum, from the USA, illustrates authentic open-ended tasks that can be machine scored. The second example, World Class Tests, illustrates highly innovative problem solving in mathematics, science, and design and technology that are by design not familiar to the student (much of our current testing is routine and predictable), interesting, motivating, psychologically challenging, and focused on a specific dimension of problem solving, such as optimization or visualization, in a mathematics/science/design context. These tasks offer the hope that it is possible to design lively, 5–10 min long, interactive, and complex problems for students to solve in the context of an on-screen test. The third example, the Virtual Performance Assessment (VPA) project, also from the USA, addresses the feasibility of using immersive technologies to deliver virtual performance assessments that measure science inquiry knowledge and skills, as defined in the U.S. National Science Education Standards (NRC 1996).

Primum

Some advocates of e-assessment point to the potential of computers to support simulation and scenario-based assessment. There are few examples of this category of e-assessment being developed successfully, especially not in high-stakes testing contexts. Primum, which assesses decision making in a very specific context, is an

exception. It provides an assessment of trainee medical practitioners' ability to make medical diagnoses when presented with a fictitious patient exhibiting a number of symptoms. This automated assessment has been designed to provide an authentic and reliable assessment at a price that compares favorably with the alternative – human-scored evaluation at patients' bedsides.

World Class Tests

In 2000, England's Department for Education commissioned the development of new computer-based tests of problem solving in the domains of mathematics, science, and design and technology. These tests are intended for worldwide application and were designed to make creative use of computer technology. Also, they are intended to set new benchmarks in the design of assessments of students' thinking and ability to apply a range of techniques to solve novel and unexpected problems. These tests have become known as World Class Tests and have been adapted for children aged 8–14. These tests are now sold commercially under license in East Asia.

The VPA Project

The Virtual Performance Assessment project utilizes innovations in technology and assessment to address the problem of measuring a student's ability to perform scientific inquiry to solve a problem. The project is developing assessments for use in school settings as a standardized component of an accountability program. The goal is to develop three assessments in the context of life science that appear different on the surface, but all measure the same inquiry process skills. Each assessment will take place in a different type of ecosystem, and students will investigate authentic ecological problems as they engage in the inquiry process.

Learning to Learn and Metacognition

Operational definitions of learning to learn and metacognition are provided in Table 2.4. Learning to learn and metacognition have most frequently been measured by think-aloud protocols that have been administered in one-on-one situations. Clearly this methodology is not amenable to large-scale assessments. However, technology might be used to support and assess learning to learn, which includes self-assessment and self-regulated learning. One interesting example of this is the eVIVA project developed at Ultralab in the UK.

eVIVA

The intention of eVIVA was to create a more flexible method of assessment, taking advantage of the possibilities new technologies such as a mobile phone and web-based formative assessment tools offer. By using such tools, project authors Ultralab promoted self- and peer-assessment as well as dialogue between teachers and students.

Table 2.4 Ways of thinking – learning to learn, metacognition

Knowledge	Skills	Attitudes/values/ethics
• Knowledge and understanding of one's preferred learning methods, the strengths and weaknesses of one's skills and qualifications • Knowledge of available education and training opportunities and how different decisions during the course of education and training lead to different careers	• Effective self-management of learning and careers in general. Ability to dedicate time to learning, autonomy, discipline, perseverance, and information management in the learning process • Ability to concentrate for extended as well as short periods of time • Ability to reflect critically on the object and purpose of learning • Ability to communicate as part of the learning process by using appropriate means (intonation, gesture, mimicry, etc.) to support oral communication as well as by understanding and producing various multimedia messages (written or spoken language, sound, music etc.)	• A self-concept that supports a willingness to change and further develop competencies as well as motivation and confidence in one's capability to succeed • Positive appreciation of learning as a life-enriching activity and a sense of initiative to learn • Adaptability and flexibility • Identification of personal biases

In this project, the students had access to the eVIVA website where they could set up an individual profile of system preferences and record an introductory sound file on their mobile phone or landline. After this, students could carry out a simple self-assessment activity by selecting a series of simple "I Can" statements designed to start them thinking about what they are able to do in ICT. The website consisted of a question bank from which the pupils were asked to select four or five questions for their telephone viva or assessment carried out toward the end of their course, but at a time of their choice. Students were guided in their choice by the system and their teacher. They had their own e-portfolio web space in which they were asked to record significant *milestone* moments of learning and to upload supporting files as evidence. Each milestone was then annotated or described by the pupil to explain what they had learned or why they were proud of a particular piece of work. Once milestones had been published, teachers and pupils could use the annotation and the messaging features to engage in dialogue with each other about the learning. Students were encouraged to add comments to their own and each other's work. The annotations could be sent via phone using SMS or voice messages. When ready, students would dial into eVIVA and record their answers to their selected questions. This gave students the opportunity to explain what they had done and reflect further on their work. Their answers were recorded and sent to the website as separate sound files. The teacher made a holistic assessment of the pupil's ICT capabilities based on the milestones, work submitted in the e-portfolio, student reflections or annotations, the recorded eVIVA answers, any written answers attached to the questions, and classroom observations (see Walton 2005).

Cascade

Cascade, which is under development at the University of Luxembourg and the Center for Public Research Henri Tudor, is an innovative item type that is more amenable to large-scale assessments with limited testing time.

The Cascade test items are designed so that respondents answer a set of questions and are then asked to rate how certain they are about the correctness of their response on each item. Then the respondent is given an opportunity to access multimedia information to verify the correctness of the response. At that point, the respondent once again answers the same set of questions and again rates his/her certainty. Scoring is based on the comparison of the first and second set of responses and tracing the information information paths he/she took in acquiring additional information.

Ways of Working

In business, we are witnessing a rapid shift in the way people work. Outsourcing services across national and continental borders are just one example. Another is having team members telecommute while working on the same project. For instance, a small software consulting team has members located in three continents. They work on developing prototypes using teleconferences and email, with the occasional "sprint" sessions where they gather in a single location and work 24 h a day to develop the product. Similarly, in the large-scale international assessments such as PISA, TIMSS (Trends in Mathematics and Science Study), and PIAAC, teams of researchers and developers across continents and at multiple locations work together to develop the assessments. To support these examples of moves toward globalization, communication and collaboration skills must be more finely honed. Communication must be rapid, concise, and cognizant of cultural differences.

Communication

Operational definitions of communication are provided in Table 2.5. Communication has been a mainstay of assessments in the form of reading, writing, graphing, listening and speaking. However, the assessments have not taken into account the full range of possibilities. At the most minimal, PowerPoint presentations are now ubiquitous. These frequently include graphic displays that, in conjunction with language, can more succinctly deliver a message. Video presentations also require the combination of communication forms in ways that have never before been within the realm of most people's capability. To date, newer modes of communication have rarely been represented in large-scale assessments. However, in light of the developments described below, it is essential that we take these changes into account.

Table 2.5 Ways of working – communication

Knowledge	Skills	Attitudes/values/ethics
Competency in language in mother tongue. • Sound knowledge of basic vocabulary, functional grammar and style, functions of language • Awareness of various types of verbal interaction (conversations, interviews, debates, etc.) and the main features of different styles and registers in spoken language • Understanding the main features of written language (formal, informal, scientific, journalistic, colloquial, etc.) *Competency in additional language/s.* • Sound knowledge of basic vocabulary, functional grammar and style, functions of language • Understanding the paralinguistic features of communication (voice-quality features, facial expressions, postural and gesture systems) • Awareness of societal conventions and cultural aspects and the variability of language in different geographical, social, and communication environments	*Competency in language in mother tongue and additional language/s.* • Ability to communicate, in written or oral form, and understand, or make others understand, various messages in a variety of situations and for different purposes • Communication includes the ability to listen to and understand various spoken messages in a variety of communicative situations and to speak concisely and clearly • Ability to read and understand different texts, adopting strategies appropriate to various reading purposes (reading for information, for study, or for pleasure) and to various text types • Ability to write different types of texts for various purposes and monitor the writing process (from drafting to proofreading) • Ability to formulate one's arguments, in speaking or writing, in a convincing manner and take full account of other viewpoints, whether expressed in written or oral form • Skills needed to use aids (such as notes, schemes, maps) to produce, present, or understand complex texts in written or oral form (speeches, conversations, instructions, interviews, debates)	*Competency in language in mother tongue.* • Development of a positive attitude to the mother tongue, recognizing it as a potential source of personal and cultural enrichment • Disposition to approach the opinions and arguments of others with an open mind and engage in constructive and critical dialogue • Confidence when speaking in public • Willingness to strive for aesthetic quality in expression beyond the technical correctness of a word/phrase • Development of a love of literature • Development of a positive attitude to intercultural communication *Competency in additional language/s.* • Sensitivity to cultural differences and resistance to stereotyping

Consider the use of text messaging. The first commercial text message was sent in December of 1992. Today the number of text messages sent and received everyday exceeds the total population of the planet. Facebook, which started as a communication vehicle for college students, reached a market audience of 50 million people within just two years. In 2010 Facebook had more than 750 million active users, and more than 375 million users were logging on at least once each day. It has

now moved into business applications, with business and interest groups having Facebook pages. It is also increasingly more common to use Facebook as the venue for organizing and conducting conferences.

Why are these communication innovations important? Beginning with text messaging, we need to consider the shift in grammar, syntax, and spelling that pervades these communications. If we consider the proliferation of videos on YouTube, it is important to see how effective different presentation forms of the same information can be. Similarly, Facebook presents even more challenges as it merges registers — here professional and personal communications can exist side-by-side.

One prominent example of incorporating new technologies into measures of communication was developed for PISA 2009. PISA's Electronic Reading Assessment simulated reading in a web environment. In many ways, this step forward represents not only migration to newer innovative assessment items but also a first step in transforming assessments to more authentic and up-to-date tasks.

Collaboration and Teamwork

Operational definitions of collaboration are provided in Table 2.6. Collaboration presents a different set of challenges for large-scale assessments. At the most basic, school level assessments are focused on getting measures of individual performance. Consequently, when faced with a collaborative task, the most important question is how to assign credit to each member of the group, as well as how to account for differences across groups that may bias a given student's performance. This becomes an even larger issue within international assessments where cultural boundaries are crossed. For example, ALL researched the potential for measuring teamwork. While the designers could generate teamwork tasks, at that time accounting for cultural differences became an insurmountable obstacle.

Several important research initiatives have worked on getting measures of individual performance that address key components of collaboration and measurement (Laurillard 2009). For example, Çakir et al. (2009) have shown how group participants, in order to collaborate effectively in group discourse on a topic like mathematical patterns, must organize their activities in ways that share the significance of their utterances, inscriptions, and behaviors. Their analysis reveals methods by which the group co-constructs meaningful inscriptions in the interaction spaces of the collaborative environment. The integration of graphical, narrative, and symbolic semiotic modalities facilitates joint problem solving. It allows group members to invoke and operate with multiple realizations of their mathematical artifacts, a characteristic of deep learning of mathematics. Other research shows how engaging in reflective activities in interaction, such as explaining, justifying, and evaluating problem solutions, collaboratively can potentially be productive for learning (Baker and Lund 1997). Several studies have also shown how taking part in collaborative inquiry toward advancing a shared knowledge object can serve as a means to facilitate the development of metaskills.

Table 2.6 Ways of working – collaboration, teamwork

Knowledge	Skills	Attitudes/values/ethics
Interact effectively with others • Know when it is appropriate to listen and when to speak *Work effectively in diverse teams* • Know and recognize the individual roles of a successful team and know own strengths and weaknesses, and recognizing and accepting them in others *Manage projects* • Know how to plan, set, and meet goals and to monitor and re-plan in the light of unforeseen developments	*Interact effectively with others* • Speak with clarity and awareness of audience and purpose. Listen with care, patience, and honesty • Conduct themselves in a respectable, professional manner *Work effectively in diverse teams* • Leverage social and cultural differences to create new ideas and increase both innovation and quality of work *Manage projects* • Prioritize, plan, and manage work to achieve the intended group result *Guide and lead others* • Use interpersonal and problem-solving skills to influence and guide others toward a goal • Leverage strengths of others to accomplish a common goal • Inspire others to reach their very best via example and selflessness • Demonstrate integrity and ethical behavior in using influence and power	*Interact effectively with others* • Know when it is appropriate to listen and when to speak • Conduct themselves in a respectable, professional manner *Work effectively in diverse teams* • Show respect for cultural differences and be prepared to work effectively with people from a range of social and cultural backgrounds • Respond open-mindedly to different ideas and values *Manage projects* • Persevere to achieve goals, even in the face of obstacles and competing pressures *Be responsible to others* • Act responsibly with the interests of the larger community in mind

Two further lines of research are pertinent to including collaborative work in large-scale assessments. The first line of research begins with the idea of a simulation where one respondent interacts with pre-programmed virtual partners. The drawback here is the current lack of theoretical understandings of how collaborators would interact in this environment. The second line of research is best exemplified by group tasks where evidence of interaction patterns and self-reflections are captured. Research into how to rate these interactions would lead to a rubric that might either be criterion-referenced or be normed according to country, nationality, socioeconomic status, or other differentiating group characteristics. In conjunction with the product scores, it would be possible to generate a collaboration scale on the basis of such research.

It has been observed that as employers, we most often base our staff recruitment decisions on formal, school, and college-based qualifications, using these as a measure of an applicant's potential to operate well within our organizations. However, we make decisions to fire people on the basis of their team-working skills, their collaborative styles, and their approach to work. These are the skills that matter most to us as employers, and it is in these areas that employers have for many years looked to occupational psychologists for support. There are a large number of psychological profiling measures, most of which seek to provide a prose summary of the interpersonal styles of working likely to be adopted by an individual. These profile measures attempt to score, for example, the extent to which an individual might seek help, might use discussion and dialogue to move matters forward, or might be an effective solver of open-ended and ill-defined problems. SHL provide assessments such as OPQ and 16PF, which are conducted online and are widely used by employers. The OPQ assessments seek to measure likely behaviors in three areas: Relationships with People, Thinking Style, and Feeling and Emotions. For example, in measuring Feeling and Emotions, OPQ gauges the extent to which an individual is *relaxed, worrying, tough minded, optimistic, trusting,* and *emotionally controlled.* Similarly, OPQ measures a dimension called Influence and gauges the extent to which an individual is *persuasive, controlling, outspoken,* and *independent minded.* These – and other measures, such as Belbin's team styles – provide considerable overlap with the skills domain that interests twenty-first century educators and could well provide useful examples of the ways in which it is possible to assess students' ways of working.

Tools for Working

The newest set of skills is combined in this grouping of tools for working. These skills, information literacy and ICT literacy, are the future and mark a major shift that is likely to be as important as the invention of the printing press. Friedman (2007) describes four stages in the growing importance of ICT. He identifies four "flatteners" that are making it possible for individuals to compete, connect, and collaborate in world markets:

- The introduction of personal computers that allowed anyone to author his/her own content in digital form that could then be manipulated and dispatched.
- The juxtaposition of the invention of the browser by Netscape that brought the internet to life resulting in the proliferation of websites and the overinvestment into fiber optic cable that has wired the world. NTT Japan has successfully tested a fiber optic cable that pushes 14 trillion bits per second that roughly equals 2,660 CDs or 20 million phone calls every second.
- The development of transmission protocols that made it possible for everyone's computer and software to be interoperable. Consequently, everyone could become a collaborator.

- The expansion of the transmission protocols so that individuals could easily upload as well as download. For example, when the world was round, individuals could download vast amounts of information in digital formats that they could easily access and manipulate. But, in the flat world, the key is the individual's ability to upload. This has given rise to open-source courseware, blogs, and Wikipedia, to name only a few examples.

To paint a picture of how important it is to be truly literate in the use of these tools, consider that it is estimated that a week's worth of the New York Times contains more information than a person was likely to come across in a lifetime in the eighteenth century. Moreover, it was estimated that four exabytes (4.0×10^{19}) of unique information was be generated in 2010 – more than that the previous 5,000 years put together. In light of this information explosion, the coming generations must have the skills to access and evaluate new information efficiently so they can effectively utilize all that is available and relevant to their tasks at hand. One of the ways that they will manage this information explosion is through skilled use of ICT. Even now the use of ICT is growing. It has been reported that there are 31 billion searches on Google every month, up from 2.7 billion in 2006. To use Google, one must effectively use the internet. To accommodate the use of the internet, we have seen an explosion in the number of internet devices. In 1984, the number was 1,000, by 1992 it was 1,000,000, and in 2008 it had reached 1,000,000,000.

Information Literacy

Information literacy includes research on sources, evidence, biases, etc. Operational definitions of information literacy are provided in Table 2.7. These are clearly increasingly important skills.

The future consequences of recent developments in our societies due to globalization, networking (Castells 1996), and the impact of ICT are spawning a set of new studies. Hull and Schultz (2002) and Burbules and Silberman-Keller (2006) are examples of how such developments change conceptions of formal and informal learning and what some term distributed or networked expertise (Hakkarainen et al. 2004). Measurement procedures or indicators are still not clear with regard to these more future-oriented skills. For example, the *ImpaCT2* concept mapping data from the UK strongly suggests that there is a mismatch between conventional national tests, which focus on pre-specified knowledge and concepts, and the wider range of knowledge that students are acquiring by carrying out new kinds of activities with ICT at home (Somekh and Mavers 2003). By using concept maps and children's drawings of computers in their everyday environments, the research generates strong indication of children's rich conceptualization of technology and its role in their world for purposes of communication, entertainment, or accessing information. It shows that most children acquire practical skills in using computers that are not part of the assessment processes that they meet in schools. Some research has shown that students who are active computer users consistently underperform on paper-based tests (Russell and Haney 2000).

Table 2.7 Tools for working – information literacy

Knowledge	Skills	Attitudes/values/ethics
Access and evaluate information • Access information efficiently (time) and effectively (sources) • Evaluate information critically and competently *Use and manage information* • Use information accurately and creatively for the issue or problem at hand • Manage the flow of information from a wide variety of sources • Apply a fundamental understanding of the ethical/legal issues surrounding the access and use of information • Basic understanding of the reliability and validity of the information available (accessibility/acceptability) and awareness of the need to respect ethical principles in the interactive use of IST *Apply technology effectively* • Use technology as a tool to research, organize, evaluate, and communicate information • Use digital technologies (computers, PDAs, media players, GPS, etc.), communication/networking tools, and social networks appropriately to access, manage, integrate, evaluate, and create information to successfully function in a knowledge economy	*Access and evaluate information* • Ability to search, collect, and process (create, organize, and distinguish relevant from irrelevant, subjective from objective, real from virtual) electronic information, data, and concepts and to use them in a systematic way *Use and manage information* • Ability to use appropriate aids, presentations, graphs, charts and maps to produce, present, or understand complex information • Ability to access and search a range of information media including the printed word, video, and websites and to use internet-based services such as discussion fora and email • Ability to use information to support critical thinking, creativity, and innovation in different contexts at home, leisure, and work • Ability to search, collect, and process written information, data, and concepts in order to use them in study and to organize knowledge in a systematic way; Ability to distinguish, in listening, speaking, reading, and writing, relevant from irrelevant information	*Access and evaluate information* • Propensity to use information to work autonomously and in teams; critical and reflective attitude in the assessment of available information *Use and manage information* • Positive attitude and sensitivity to safe and responsible use of the internet, including privacy issues and cultural differences • Interest in using information to broaden horizons by taking part in communities and networks for cultural, social and professional purposes

ICT Literacy

EU countries, both on a regional and national level, and other countries around the world are in the process of developing a framework and indicators to better grasp the impact of technology in education and what we should be looking for in

assessing students' learning using ICT. Frameworks are being developed in Norway (see http://europa.eu/rapid/pressReleasesAction.do?reference=IP/09/1244), Norway (see Erstad, 2006), and Australia (see Ainley et al. 2006). According to the Summit on Twenty-first Century Literacy in Berlin in 2002 (Clift 2002), new approaches stress the abilities to use information and knowledge that extend beyond the traditional base of reading, writing, and mathematics, which has been termed *digital literacy* or *ICT literacy*. Operational definitions of information literacy are provided in Table 2.8.

In 2001, the Educational Testing Service (ETS) in the US assembled a panel for the purpose of developing a workable framework for ICT literacy. The outcome was the report *Digital transformation: A framework for ICT literacy* (International ICT Literacy Panel 2002). Based on this framework, shown in Table 2.9, one can define ICT literacy as "the ability of individuals to use ICT appropriately to access, manage and evaluate information, develop new understandings, and communicate with others in order to participate effectively in society." (Ainley et al. 2005) Different indicators of digital/ICT literacy can be proposed (Erstad 2010).

In line with this perspective, some agencies have developed performance assessment tasks of "ICT Literacy", indicating that ICT is changing our view on what is being assessed and how tasks are developed using different digital tools. One example is the tasks developed by the International Society for Technology in Education (ISTE) called *National Educational Technology Standards* (http://www.iste.org/standards.aspx), which are designed to assess how skilful students, teachers, and administrators are in using ICT.

In 2000, England's Department for Education commissioned the development of an innovative test of 14-year-old students' ICT skills. David Blunkett, at the time Secretary of State for Education, described his vision for education and attainment in the twenty-first century. He spoke of raising expectations of student capabilities. He also announced the development of a new type of online test of ICT, which would assess the ICT skills students need in the twenty-first century. These developed assessments are outlined in Fig. 2.3.

Development activity for the 14-year-old's test of ICT began in 2001. The original planned date for full roll-out and implementation was May 2009. In the event – and for a whole range of reasons – the original vision for the ICT tests was never realized. The test activities that were developed have been redesigned as standalone skills assessments that teachers in accredited schools can download and use informally to support their teacher assessment.

In Australia, a tool has been developed with a sample of students from grade 6 and grade 10 to validate and refine a progress map that identifies a progression of ICT literacy. The ICT literacy construct is described using three "strands": working with information, creating and sharing information, and using ICT responsibly. Students carrying out authentic tasks in authentic contexts are seen as fundamental to the design of the Australian National ICT Literacy Assessment Instrument (Ainley et al. 2005). The instrument evaluates six key processes: accessing information (identifying information requirements and knowing how to find and retrieve information); managing information (organizing and storing information for retrieval and reuse); evaluating (reflecting on the processes used to design and construct

Table 2.8 Tools for working – ICT literacy

Knowledge	Skills	Attitudes/values/ethics
Access and evaluate information and communication technology • Understanding of the main computer applications, including word processing, spreadsheets, databases, information storage and management • Awareness of the opportunities given by the use of Internet and communication via electronic media (e-mail, videoconferencing, other network tools) and the differences between the real and virtual world *Analyze media* • Understand both how and why media messages are constructed, and for what purposes • Examine how individuals interpret messages differently, how values and points of view are included or excluded, and how media can influence beliefs and behaviors • Understand the ethical/legal issues surrounding the access and use of media *Create media products* • Understand and know how to utilize the most appropriate media creation tools, characteristics, and conventions • Understand and know how to effectively utilize the most appropriate expressions and interpretations in diverse, multicultural environments	*Access and evaluate information and communication technology* • Access ICT efficiently (time) and effectively (sources) • Evaluate information and ICT tools critically and competently *Use and manage information* • Use ICT accurately and creatively for the issue or problem at hand • Manage the flow of information from a wide variety of sources • Apply a fundamental understanding of the ethical/legal issues surrounding the access and use of ICT and media • Employ knowledge and skills in the application of ICT and media to communicate, interrogate, present, and model *Create media products* • Utilize the most appropriate media creation tools, characteristics and conventions, expressions, and interpretations in diverse, multicultural environments *Apply technology effectively* • Use technology as a tool to research, organize, evaluate, and communicate information • Use digital technologies (computers, PDAs, media players, GPS, etc.), communication/networking tools, and social networks appropriately to access, manage, integrate, evaluate, and create information to successfully function in a knowledge economy • Apply a fundamental understanding of the ethical/legal issues surrounding the access and use of information technologies	*Access and evaluate information and communication technology* • Be open to new ideas, information, tools, and ways of working but evaluate information critically and competently *Use and manage information* • Use information accurately and creatively for the issue or problem at hand respecting confidentiality, privacy, and intellectual rights • Manage the flow of information from a wide variety of sources with sensitivity and openness to cultural and social differences • Examine how individuals interpret messages differently, how values and points of view are included or excluded, and how media can influence beliefs and behaviors *Apply and employ technology with honesty and integrity* • Use technology as a tool to research, organize, evaluate, and communicate information accurately and honestly with respect for sources and audience • Apply a fundamental understanding of the ethical/legal issues surrounding the access and use of information technologies

Table 2.9 Elaboration of key concepts of ICT literacy based on ETS framework

Category	Skills
Basic	Be able to open software, sort out and save information on the computer and other simple skills using the computer and software
Download	Be able to download different types of information from the internet
Search	Know about and how to get access to information
Navigate	Be able to orient oneself in digital networks, learning strategies in using the internet
Classify	Be able to organize information according to a certain classification scheme or genre
Integrate	Be able to compare and put together different types of information related to multimodal texts
Evaluate	Be able to check and evaluate if one has got the information one seeks to get from searching the internet. Be able to judge the quality, relevance, objectivity and usefulness of the information one has found. Critical evaluation of sources
Communicate	Be able to communicate information and express oneself through different meditational means
Cooperate	Be able to take part in net-based interactions of learning and take advantage of digital technology to cooperate and take part in networks
Create	Be able to produce and create different forms of information as multimodal texts, make web pages and so forth. Be able to develop something new by using specific tools and software. Remixing different existing texts into something new

ICT solutions and judgments regarding the integrity, relevance, and usefulness of information); developing new understandings (creating information and knowledge by synthesizing, adapting, applying, designing, inventing, or authoring); communicating (exchanging information by sharing knowledge and creating information products to suit the audience, the context, and the medium); and using ICT appropriately (critical, reflective, and strategic ICT decisions and considering social, legal, and ethical issues) (Ainley et al. 2005). Preliminary results of the use of the instrument show highly reliable estimates of ICT ability.

There are also cases where an ICT assessment framework is linked to specific frameworks for subject domains in schools. Reporting on the initial outline of a U.S. project aiming at designing a Coordinated ICT Assessment Framework, Quellmalz and Kozma (2003) have developed a strategy to study ICT tools and skills as an integrated part of science and mathematics. The objective is to design innovative ICT performance assessments that could gather evidence of use of ICT strategies in science and mathematics.

Living in the World

Borrowing the title of Bob Dylan's song, to say that "the times they are a changin'" is a gross understatement when one considers how different living and working in

the world will soon be. For example, the U.S. Department of Labor estimated that today's learner will have between ten and fourteen jobs by age 38. This reflects rapidly growing job mobility, with one in four workers having been with their current employer for less than a year, and one in two has been there less than 5 years. One might ask where these people are going as manufacturing and service industries move to places where there are abundant sources of cheap but sufficiently educated labor supplies. Essentially, people must learn to live not only in their town or country but also in the world in its entirety. As more and more people individually move in the twenty-first century to compete, connect, and collaborate, it is even more important that they understand all the aspects of citizenship. It is not enough to assume that what goes on in your own country is how it is or should be all over the globe. Hence, we have identified and group Citizenship, Life and Career, and Personal and Social Responsibility together as twenty-first century skills.

Citizenship, Global and Local

Citizenship as an educational objective is not new and has been part of curricula, especially in social studies. A central focus has been on knowledge about democratic processes. Citizenship as a competence, however, has been growing in importance, and implies certain challenges in measurement. Operational definitions of citizenship are provided as shown in Table 2.10.

Honey led a worldwide investigation into the use of twenty-first century assessments which investigated the existence and quality of assessments in key areas, including global awareness, concluding that "no measures currently exist that address students' understanding of global and international issues." (Ripley 2007, p. 5)

One example of a large-scale assessment of citizenship skills is the International Civic Education Study conducted by the International Association for the Evaluation of Educational Achievement (IEA). This research tested and surveyed nationally representative samples consisting of 90,000 14 year-old students in 28 countries, and 50,000 17 to 19 year-old students in 16 countries throughout 1999 and 2000.

The content domains covered in the instrument were identified through national case studies during 1996–1997 and included democracy, national identity, social cohesion and diversity. The engagement of youth in civil society was also a focus. Torney-Purta et al. (2001) reported the findings from these studies in the following terms:

- Students in most countries have an understanding of fundamental democratic values and institutions – but depth of understanding is a problem.
- Young people agree that good citizenship includes the obligation to vote.
- Students with the most civic knowledge are most likely to be open to participate in civic activities.
- Schools that model democratic practice are most effective in promoting civic knowledge and engagement.

Table 2.10 Living in the world – citizenship, local and global

Knowledge	Skills	Attitudes/values/ethics
• Knowledge of civil rights and the constitution of the home country, the scope of its government • Understand the roles and responsibilities of institutions relevant to the policy-making process at local, regional, national, and international level • Knowledge of key figures in local and national governments; political parties and their policies • Understand concepts such as democracy, citizenship, and the international declarations expressing them • Knowledge of the main events, trends, and agents of change in national and world history • Knowledge of the movements of peoples and cultures over time around the world	• Participation in community/neighborhood activities as well as in decision making at national and international levels; voting in elections • Ability to display solidarity by showing an interest in and helping to solve problems affecting the local or wider community • Ability to interface effectively with institutions in the public domain • Ability to profit from the opportunities given by the home country and international programs	• Sense of belonging to one's locality, country, and (one's part of) the world • Willingness to participate in democratic decision making at all levels • Disposition to volunteer and to participate in civic activities and support for social diversity and social cohesion • Readiness to respect the values and privacy of others with a propensity to react against antisocial behavior • Acceptance of the concept of human rights and equality; acceptance of equality between men and women • Appreciation and understanding of differences between value systems of different religious or ethnic groups • Critical reception of information from mass media

- Aside from voting, students are skeptical about traditional forms of political engagement, but many are open to other types of involvement in civic life.
- Students are drawn to television as their source of news.
- Patterns of trust in government-related institutions vary widely among countries.
- Gender differences are minimal with regard to civic knowledge but substantial in some attitudes.
- Teachers recognize the importance of civic education in preparing young people for citizenship.

The main survey has been replicated as the International Civic and Citizenship Education Study in which data have been gathered in 2008 and 2009 and from which the international report was released in June 2010 (Schulz et al. 2010).

The developments of the internet and Web 2.0 technologies have implications for the conception of citizenship as a competence. Jenkins (2006) says these developments create a "participatory culture." This challenges, both locally and globally, the understanding of citizenship, empowerment, and engagement as educational priorities. At the moment, no measures exist which assess these skills in online environments, even though the research literature on "young citizens online" has been growing in recent years (Loader 2007).

One example of how these skills are made relevant in new ways is the Junior Summit online community. This consisted of 3,062 adolescents representing 139 countries. The online forum culminated in the election of 100 delegates. Results from one study indicate "young online leaders do not adhere to adult leadership styles of contributing many ideas, sticking to task, and using powerful language. On the contrary, while the young people elected as delegates do contribute more, their linguistic style is likely to keep the goals and needs of the group as central, by referring to the group rather than to themselves and by synthesizing the posts of others rather than solely contributing their own ideas. Furthermore, both boy and girl leaders follow this pattern of interpersonal language use. These results reassure us that young people can be civically engaged and community minded, while indicating that these concepts themselves may change through contact with the next generation" (Cassell et al. 2006). In this sense, it also relates to the German term "Bildung" as an expression of how we use knowledge to act on our community and the world around us, that is, what it means to be literate in a society, or what also might be described as cultural competence as part of broader personal and social responsibility.

Life and Career

The management of life and career is included among the skills needed for living in the world. There is a long tradition of measurement of occupational preferences as one component for career guidance but no strong basis for building measures of skill in managing life and career. Suggestions for building operational definitions of this skill are provided in Table 2.11.

Personal and Social Responsibility

The exercise of personal and social responsibility is also included among the skills needed for living in the world. There are aspects of this skill in collaboration and teamwork, which is among the skills included among ways of working. Personal and social responsibility is taken to include cultural awareness and cultural competence. There is not a body of measurement literature on which to draw, but the scope intended is set out in the operational definitions offered in Table 2.12.

Challenges

The foregoing discussions have laid out principles for the assessment of twenty-first century skills, proposed ten skills, and given a sense of what they are and what measurements related to them might be built upon. That being said, there is still a very long row to hoe, as it is not enough to keep perpetuating static tasks within the assessments. Rather, to reflect the need for imagination to compete, connect, and collaborate, it is essential that transformative assessments be created. This cannot begin to happen without addressing some very critical challenges.

Table 2.11 Living in the world – life and career

Knowledge	Skills	Attitudes/values/ethics
Adapt to change • Be aware that the twenty-first century is a period of changing priorities in employment, opportunity, and expectations • Understand diverse views and beliefs, particularly in multicultural environments *Manage goals and time* • Understand models for long-, medium-, and short-term planning and balance tactical (short-term) and strategic (long-term) goals *Be self-directed learners* • Identify and plan for personal and professional development over time and in response to change and opportunity *Manage projects* • Set and meet goals, even in the face of obstacles and competing pressures • Prioritize, plan, and manage work to achieve the intended result	*Adapt to change* • *Operate in varied roles, jobs responsibilities, schedules, and contexts* *Be flexible* • Incorporate feedback effectively • Negotiate and balance diverse views and beliefs to reach workable solutions *Manage goals and time* • Set goals with tangible and intangible success criteria • Balance tactical (short-term) and strategic (long-term) goals • Utilize time and manage workload efficiently *Work independently* • Monitor, define, prioritize, and complete tasks without direct oversight *Interact effectively with others* • Know when it is appropriate to listen and when to speak *Work effectively in diverse teams* • Leverage social and cultural differences to create new ideas and increase both innovation and quality of work *Manage projects* • Set and meet goals, prioritize, plan, and manage work to achieve the intended result even in the face of obstacles and competing pressures *Guide and lead others* • Use interpersonal and problem solving skills to influence and guide others toward a goal • Leverage strengths of others to accomplish a common goal • Inspire others to reach their very best via example and selflessness • Demonstrate integrity and ethical behavior in using influence and power	*Adapt to change* • Be prepared to adapt to varied responsibilities, schedules, and contexts; recognize and accept the strengths of others • See opportunity, ambiguity and changing priorities *Be flexible* • Incorporate feedback and deal effectively with praise, setbacks, and criticism • Be willing to negotiate and balance diverse views to reach workable solutions *Manage goals and time* • Accept uncertainty and responsibility and self manage *Be self-directed learners* • Go beyond basic mastery to expand one's own learning • Demonstrate initiative to advance to a professional level • Demonstrate commitment to learning as a lifelong process • Reflect critically on past experiences for progress *Work effectively in diverse teams* • Conduct self in a respectable, professional manner • Respect cultural differences, work effectively with people from varied backgrounds • Respond open-mindedly to different ideas and values *Produce results* • Demonstrate ability to: – Work positively and ethically – Manage time and projects effectively – Multi-task – Be reliable and punctual – Present oneself professionally and with proper etiquette – Collaborate and cooperate effectively with teams – Be accountable for results *Be responsible to others* • Act responsibly with the interests of the larger community in mind

Table 2.12 Living in the world – personal and social responsibility

Knowledge	Skills	Attitudes/values/ethics
• Knowledge of the codes of conduct and manners generally accepted or promoted in different societies • Awareness of concepts of individual, group, society, and culture and the historical evolution of these concepts • Knowledge of how to maintain good health, hygiene, and nutrition for oneself and one's family • Knowledge of the intercultural dimension in their own and other societies	• Ability to communicate constructively in different social situations (tolerating the views and behavior of others; awareness of individual and collective responsibility) • Ability to create confidence and empathy in other individuals • Ability to express one's frustration in a constructive way (control of aggression and violence or self-destructive patterns of behavior) • Ability to maintain a degree of separation between the professional and personal spheres of life and to resist the transfer of professional conflict into personal domains • Awareness and understanding of national cultural identity in interaction with the cultural identity of the rest of the world; ability to see and understand the different viewpoints caused by diversity and contribute one's own views constructively • Ability to negotiate	• Showing interest in and respect for others • Willingness to overcome stereotypes and prejudices • Disposition to compromise • Integrity • Assertiveness

This section summarizes key challenges to assessing twenty-first century skills in ways that truly probe the skills of students and provide actionable data to improve education and assessments.

Using Models of Skill Development Based on Cognitive Research

The knowledge about acquisition of twenty-first century skills and their development is very limited. The developers of assessments do not yet know how to create practical assessments using even this partial knowledge effectively (Bennett and Gitomer 2009).

Transforming Psychometrics to Deal with New Kinds of Assessments

Psychometric advances are needed to deal with a dynamic context and differentiated tasks, such as tasks embedded in simulations and using visualization that may yield a number of acceptable (and unanticipated) responses. While traditional assessments

are designed to yield one right or best response, transformative assessments should be able to account for divergent responses, while measuring student performance in such a way that reliability of measures is ensured.

Making Students' Thinking Visible

Assessments should reveal the kinds of conceptual strategies a student uses to solve a problem. This involves not only considering students' responses but also interpreting their behaviors that lead to these responses. Computers can log every keystroke made by a student and thus amass a huge amount of behavioral data. The challenge is to interpret the meaning of these data and link patterns of behavior to the quality of response. These associations could then illuminate students' thinking as they respond to various tasks.

That computers can score student responses to items effectively and efficiently is becoming a reality. This is certainly true of selected-response questions where there is a single right answer. It is also quite easy to apply partial credit models to selected-response items that have been designed to match theories of learning where not quite fully correct answers serve as the distracters. Constructed responses pose challenges for automated scoring.

The OECD's PIAAC provides a good example of movement forward in machine scoring of short constructed responses. Some of the assessment tasks in PIAAC were drawn from the International Adult Literacy Survey (IALS) and the ALL Survey where all answers were short constructed responses that needed to be coded by humans. By altering the response mode into either drop and drag or highlighting, the test developers converted the items into machine scorable items. In these examples, however, all the information necessary to answer these types of questions resides totally in the test stimuli. Although the respondent might have to connect information across parts of the test stimuli, creation of knowledge not already provided is not required.

Machine scoring of extended constructed responses is in its infancy. Models do exist in single languages and are based on the recognition of semantic networks within responses. In experimental situations, these machine-scoring models are not only as reliable as human scorers but often achieve higher levels of consistency than can be achieved across human raters (Ripley and Tafler 2009). Work has begun in earnest to expand these models to cross languages and may be available for international assessments in the foreseeable future (Ripley 2009).

Interpreting Assisted Performance

New scoring rules are needed to take into account prompting or scaffolding that may be necessary for some students. Ensuring accessibility for as many students as possible and customization of items for special needs students within the design of the assessment are critical.

Assessing Twenty-First Century Skills in Traditional Subjects

Where the aims and goals of twenty-first century learning are described in countries' frameworks, they are generally specified as being taught through, within and across the subjects. However, computers can facilitate the creation of micro-worlds for students to explore in order to discover hidden rules or relationships. Tools such as computer-based simulations can, in this way, give a more nuanced understanding of what students know and can do than traditional testing methods. New approaches stress the abilities to use information and knowledge that extend beyond the traditional base of reading, writing, and mathematics. However, research shows that students still tuned into the old test situation with correct answers rather than explanations and reasoning skills can have problems in adjusting their strategies and skills. Without highly valued assessments of twenty-first century aims or goals requiring their teaching, it is difficult to see when or how education systems will change significantly for the majority of learners.

Accounting for New Modes of Communication

To date, newer modes of communication have rarely been represented in large-scale assessments. There is a mismatch between the skills young people gain in their everyday cultures outside of schools and the instruction and assessment they meet in schools. Different skills such as creativity, problem solving, and critical thinking might be expressed in different ways using different modes and modalities, which ICT provides. In light of the developments described in the chapter, it is essential that the radical changes in communication, including visual ways of communicating and social networking, be represented in some of the tasks of twenty-first century large-scale assessments. The speed with which new technologies develop suggests that it might be better to assess whether students are capable of rapidly mastering a new tool or medium than whether they can use current technologies.

Including Collaboration and Teamwork

Traditional assessments are focused on measuring individual performance. Consequently, when faced with a collaborative task, the most important question is how to assign credit to each member of the group, as well as how to account for differences across groups that may bias a given student's performance. This issue arises whether students are asked to work in pre-assigned complementary roles or whether they are also being assessed on their skills in inventing ways to collaborate in an undefined situation. Questions on assigning individual performance as well as group ratings become even more salient for international assessments where cultural boundaries are crossed.

Including Local and Global Citizenship

The assessment of citizenship, empowerment, and engagement, both locally and globally, is underdeveloped. At this time, no measures exist that assess these skills in online environments, even though the research literature on "young citizens online" has been growing in recent years. For international assessments, cultural differences and sensitivities will add to the challenge of developing tasks valid across countries. Having students solve problems from multiple perspectives is one way to address the challenge of cultural differences.

Ensuring Validity and Accessibility

It is important to ensure validity of standards on which assessments are based; accessibility with respect to skills demands, content prerequisites, and familiarity with media or technology and an appropriate balance of content and intellectual demands of tasks.

These important attributes of any assessments will prove particularly challenging for the transformative assessments envisaged in this paper. Careful development and piloting of innovative tasks will be required, including scoring systems that ensure comparability of complex tasks. Fluidity studies with technology are important in devising tasks for which experience with technology does not predict performance. Also, complex tasks typically demand access to intellectual resources (e.g., a search engine). This needs to be factored into designing complex assessment tasks as envisaged for transformative assessments.

Considering Cost and Feasibility

Cost and feasibility are factors operating for any assessment but will be greatly exacerbated for the innovative and transformative assessments that are to address the kinds of twenty-first century skills discussed in this paper. For sophisticated online assessments, ensuring that schools have both the technical infrastructure needed and the controls for integrity of data collection is mandatory. These latter matters are considered in Chap. 4.

References

Ainley, J., Fraillon, J., & Freeman, C. (2005). *National Assessment Program: ICT literacy years 6 & 10 report.* Carlton South, Australia: The Ministerial Council on Education, Employment, Training and Youth Affairs (MCEETYA).

Ainley, J., Pratt, D., & Hansen, A. (2006). Connecting engagement and focus in pedagogic task design. *British Educational Research Journal, 32*(1), 23–38.

Anderson, R. (2009, April). A plea for '21st Century Skills' white paper to include social and civic values. Memorandum to Assessment and Teaching of 21st Century Skills Conference, San Diego, CA.

Baker, E. L. (2007). The end(s) of testing. *Educational Researcher, 36*(6), 309–317.

Baker, M. J., & Lund, K. (1997). Promoting reflective interactions in a computer-supported collaborative learning environment. *Journal of Computer Assisted Learning, 13*, 175–193.

Banaji, S., & Burn, A. (2007). *Rhetorics of creativity*. Commissioned by Creative Partnerships. Retrieved November 30, 2009 www.creative-partnerships.com/literaturereviews

Bell, A., Burkhardt, H., & Swan, M. (1992). Balanced assessment of mathematical performance. In R. Lesh & S. Lamon (Eds.), *Assessment of authentic performance in school mathematics*. Washington, DC: AAAS.

Bennett, R. E. (2002). Inexorable and inevitable: The continuing story of technology and assessment. *Journal of Technology, Learning, and Assessment, 1*(1), 14–15.

Bennett, R. E., & Gitomer, D. H. (2009). Transforming K-12 assessment. In C. Wyatt-Smith & J. Cumming (Eds.), *Assessment issues of the 21st Century*. New York: Springer Publishing Company.

Bennett, R. E., Jenkins, F., Persky, H., & Weiss, A. (2003). Assessing complex problem solving performances. *Assessment in Education: Principles, Policy & Practice, 10*, 347–360.

Black, P., McCormick, R., James, M., & Pedder, D. (2006). Learning how to learn and assessment for learning: A theoretical inquiry. *Research Papers in Education, 21*(2), 119–132.

Black, P., & Wiliam, D. (1998). Assessment and classroom learning. *Assessment in Education, 5*(1), 7–71.

Boeijen, G., & Uijlings, P. (2004, July). *Exams of tomorrow: Use of computers in Dutch national science exams.* Paper presented at the GIREP Conference, Teaching and learning physics in new contexts, Ostrava, Czech Republic.

Buckingham, D., & Willett, R. (Eds.). (2006). *Digital generations: Children, young people, and new media*. Mahwah: Lawrence Erlbaum.

Burbules, N. C., & Silberman-Keller, D. (2006). *Learning in places: The informal education reader*. New York: Peter Lang.

Çakir, M. P., Zemel, A., & Stahl, G. (2009). The joint organization of interaction within a multimodal CSCL medium. *International Journal of Computer-Supported Collaborative Learning, 4*(2), 115–149.

Cassell, J., Huffaker, D., Ferriman, K., & Tversky, D. (2006). The language of online leadership: Gender and youth engagement on the Internet. *Developmental Psychology, 42*(3), 436–449.

Castells, M. (1996). *The rise of the network society* (The information age: Economy, society and culture, Vol. 1). Cambridge: Blackwell.

Cheng, L., Watanabe, Y., & Curtis, A. (Eds.). (2004). *Washback in language testing: Research contexts and methods*. Mahwah: Lawrence Erlbaum Associates.

Clift, S. (2002). *21st literacy summit white paper*. Retrieved from www.mail-archive.com/do-wire@tc.umn.edu/msg00434.html

Deakin Crick, R. D., Broadfoot, P., & Claxton, G. (2004). Developing an effective lifelong learning inventory: The ELLI project. *Assessment in Education: Principles, Policy & Practice, 11*, 247–318.

Draper, S. W. (2009). Catalytic assessment: Understanding how MCQs and EVS can foster deep learning. *British Journal of Educational Technology, 40*(2), 285–293.

Ericsson, K. A. (2002). Attaining excellence through deliberate practice: Insights from the study of expert performance. In M. Ferrari (Ed.), *The pursuit of excellence through education* (pp. 21–55). Mahwah: Lawrence Erlbaum Associates.

Erstad, O. (2006). A new direction? Digital literacy, student participation and curriculum reform in Norway. *Education and Information Technologies, 11*(3–4), 415–429.

Erstad, O. (2008). Trajectories of remixing: Digital literacies, media production and schooling. In C. Lankshear & M. Knobel (Eds.), *Digital literacies: Concepts, policies and practices* (pp. 177–202). New York: Peter Lang.

Erstad, O. (2010). Conceptions of technology literacy and fluency. In *International encyclopedia of education* (3rd ed.). Oxford: Elsevier.

Facione, P.A. (1990). *Critical thinking: A statement of expert consensus for purposes of educational assessment and instruction* (The Delphi Report). Millbrae: California Academic Press.

Forster, M., & Masters, G. (2004). Bridging the conceptual gap between classroom assessment and system accountability. In M. Wilson (Ed.), *Towards coherence between classroom assessment and accountability: 103rd Yearbook of the National Society for the Study of Education*. Chicago: University of Chicago Press.

Friedman, T. (2007). *The world is flat.* New York: Farrar, Straus and Giroux.

Gardner, J. (Ed.). (2006). *Assessment & learning*. London: Sage Publications.

Gee, J. P. (2007). *What video games have to teach us about learning and literacy* (2nd ed.). New York: Palgrave Macmillan.

Gick, M., & Holyoak, K. (1983). Scheme induction and analogical transfer. *Cognitive Psychology, 15*(1), 1–38.

Gipps, C., & Stobart, G. (2003). Alternative assessment. In T. Kellaghan & D. Stufflebeam (Eds.), *International handbook of educational evaluation* (pp. 549–576). Dordrecht: Kluwer Academic Publishers.

Hakkarainen, K., Palonen, T., Paavola, S., & Lehtinen, E. (2004). *Communities of networked expertise: Professional and educational perspectives*. Amsterdam: Elsevier.

Harlen, W. (2006). The role of assessment in developing motivation for learning. In J. Gardner (Ed.), *Assessment & learning* (pp. 61–80). London: Sage Publications.

Harlen, W., & Deakin Crick, R. (2003). Testing and motivation for learning. *Assessment in Education: Principles, Policy & Practice, 10*, 169–208.

Herman, J. L. (2008). Accountability and assessment in the service of learning: Is public interest in K-12 education being served? In L. Shepard & K. Ryan (Eds.), *The future of testbased accountability*. New York: Taylor & Francis.

Herman, J. L., & Baker, E. L. (2005). Making benchmark testing work. *Educational Leadership, 63*(3), 48–55.

Herman, J. L., & Baker, E. L. (2009). Assessment policy: Making sense of the babel. In D. Plank, G. Sykes, & B. Schneider (Eds.), *AERA handbook on education policy*. Newbury Park: Sage Publications.

Hof, R. D. (2007, August 20). Facebook's new wrinkles: The 35-and-older crowd is discovering its potential as a business tool. *Business Week*. Retrieved from http://www.businessweek.com/magazine/content/07_34/b4047050.htm

Holyoak, K. J. (2005). Analogy. In K. J. Holyoak & R. G. Morrison (Eds.), *The Cambridge handbook of thinking and reasoning* (pp. 117–142). Cambridge: Cambridge University Press.

Hull, G., & Schultz, K. (2002). *School's out! Bridging out-of-school literacies with classroom practice*. New York: Teachers College Columbia University.

International ICT Literacy Panel. (2002). *Digital transformation: A framework for ICT literacy*. Princeton: Educational Testing Service.

Jenkins, H. (2006). *Convergence culture: Where old and new media collide*. New York: New York University Press.

Johnson, M., & Green, S. (2004). *Online assessment: The impact of mode on student performance.* Paper presented at the British Educational Research Association Annual Conference, Manchester, UK.

Koretz, D., Broadfoot, P., & Wolf, A. (Eds.). (1998). *Assessment in Education: Principles, policy & practice* (Special issue: Portfolios and records of achievement). London: Taylor & Francis.

Kozma, R. B. (Ed.). (2003). *Technology, innovation, and educational change: A global perspective*. Eugene: International Society for the Evaluation of Educational Achievement.

Laurillard, D. (2009). The pedagogical challenges to collaborative technologies. *International Journal of Computer-Supported Collaborative Learning, 4*(1), 5–20.

Lee, E. Y. C., Chan, C. K. K., & van Aalst, J. (2006). Students assessing their own collaborative knowledge building. *International Journal of Computer-Supported Collaborative Learning, 1*(1).

Lessig, L. (2008). *Remix: Making art and commerce thrive in the hybrid economy*. New York: Penguin Press.
Lin, S. S. J., Liu, E. Z. F., & Yuan, S. M. (2001). Web-based peer assessment: Feedback for students with various thinking styles. *Journal of Computer Assisted Learning, 17*, 420–432.
Loader, B. (Ed.). (2007). *Young citizens in the digital age: Political engagement, young people and new media*. London: Routledge.
Loveless, A. (2007). *Creativity, technology and learning.* (*Update.*) Retrieved November 30, 2009 http://www.futurelab.org.uk/resources/publications-reports-articles/literature-reviews/Literature-Review382
McFarlane, A. (2001). Perspectives on the relationships between ICT and assessment. *Journal of Computer Assisted Learning, 17*, 227–234.
McFarlane, A. (2003). Assessment for the digital age. *Assessment in Education: Principles, Policy & Practice, 10*, 261–266.
Mercer, N., & Littleton, K. (2007). *Dialogue and the development of children's thinking*. London: Routledge.
National Center on Education and the Economy. (1998). New standards: Performance standards and assessments for the schools. Retrieved at http://www.ncee.org/store/products/index.jsp?setProtocol=true&stSection=1
National Research Council (NRC). (1996). *National science education standards*. Washington, DC: National Academy Press.
No Child Left Behind Act of 2001, United States Public Law 107–110.
Nunes, C. A. A., Nunes, M. M. R., & Davis, C. (2003). Assessing the inaccessible: Metacognition and attitudes. *Assessment in Education: Principles, Policy & Practice, 10*, 375–388.
O'Neil, H. F., Chuang, S., & Chung, G. K. W. K. (2003). Issues in the computer-based assessment of collaborative problem solving. *Assessment in Education: Principles, Policy & Practice, 10*, 361–374.
OECD. (2005). *Formative assessment: Improving learning in secondary classrooms*. Paris: OECD Publishing.
Pellegrino, J. W., Chudowsky, N., & Glaser, R. (Eds.). (2001). *Knowing what students know*. Washington, DC: National Academy Press.
Poggio, J., Glasnapp, D. R., Yang, X., & Poggio, A. J. (2005). A comparative evaluation of score results from computerized and paper and pencil mathematics testing in a large scale state assessment program. *Journal of Technology, Learning, and Assessment, 3*(6). Available from. http://www.jtla.org, 4–30
Pommerich, M. (2004). Developing computerized versions of paper-and-pencil tests: Mode effects for passage-based tests. *Journal of Technology, Learning and Assessment, 2*(6).
Quellmalz, E. S., & Kozma, R. (2003). Designing assessments of learning with technology. *Assessment in Education: Principles, Policy & Practice, 10*, 389–408.
Quellmalz, E., Kreikemeier, P., DeBarger, A. H., & Haertel, G. (2007). A study of the alignment of the NAEP, TIMSS, and New Standards Science Assessments with the inquiry abilities in the National Science Education Standards. Presented at the Annual Meeting of the American Educational Research Association, April 9–13, Chicago, IL
Raikes, N., & Harding, R. (2003). The horseless carriage stage: Replacing conventional measures. *Assessment in Education: Principles, Policy & Practice, 10*, 267–278.
Ridgway, J., & McCusker, S. (2003). Using computers to assess new educational goals. *Assessment in Education: Principles, Policy & Practice, 10*(3), 309–328.
Ridgway, J., McCusker, S., & Pead, D. (2004). *Literature review of e-assessment (report 10)*. Bristol: Futurelab.
Ripley, M. (2007). *E-assessment: An update on research, policy and practice*. Bristol: Futurelab. Retrieved November 30, 2009 http://www.futurelab.org.uk/resources/publications-reports-articles/literature-reviews/Literature-Review204
Ripley, M. (2009). JISC case study: Automatic scoring of foreign language textual and spoken responses. Available at http://www.dur.ac.uk/smart.centre1/jiscdirectory/media/JISC%20Case%20Study%20-%20Languages%20-%20v2.0.pdf

Ripley, M., & Tafler, J. (2009). JISC case study: Short answer marking engines. Available at http://www.dur.ac.uk/smart.centre1/jiscdirectory/media/JISC%20Case%20Study%20-%20Short%20Text%20-%20v2.0.pdf

Rumpagaporn, M. W., & Darmawan, I.N. (2007). Student's critical thinking skills in a Thai ICT schools pilot project. *International Education Journal*, *8*(2), 125–132. Retrieved November 30, 2009 http://digital.library.adelaide.edu.au/dspace/handle/2440/44551

Russell, M. (1999). Testing on computers: A follow-up study comparing performance on computer and on paper. *Education Policy Analysis Archives, 7*(20). Retrieved from http://epaa.asu.edu/epaa/v7n20

Russell, M., & Haney, W. (2000). Bridging the gap between testing and technology in schools. *Education Policy Analysis Archives, 8*(19). Retrieved from http://epaa.asu.edu/epaa/v8n19.html

Russell, M., Goldberg, A., & O'Connor, K. (2003). Computer-based testing and validity: A look into the future. *Assessment in Education: Principles, Policy & Practice, 10*, 279–294.

Scardamalia, M., & Bereiter, C. (2006). Knowledge building: Theory, pedagogy and technology. In R. K. Sawyer (Ed.), *The Cambridge handbook of the learning sciences*. New York: Cambridge University Press.

Schulz, W., Ainley, J., Fraillon, J., Kerr, D., & Losito, B. (2010). *Initial Findings from the IEA International Civic and Citizenship Education Study*. Amsterdam: IEA.

Sefton-Green, J., & Sinker, R. (Eds.). (2000). *Evaluating creativity: Making and learning by young people*. London: Routledge.

Shepard, L. (2007). Formative assessment: Caveat emptor. In C. Dwyer (Ed.), *The future of assessment: Shaping teaching and learning* (pp. 279–304). Mahwah: Lawrence Erlbaum Associates.

Shepard, L., Hammerness, K., Darling-Hammond, D., & Rust, R. (2005). Assessment. In L. Darling-Hammond & J. Bransford (Eds.), *Preparing teachers for a changing world: What teachers should learn and be able to do*. Washington, DC: National Academy of Education.

Shephard, K. (2009). E is for exploration: Assessing hard-to-measure learning outcomes. *British Journal of Educational Technology, 40*(2), 386–398.

Somekh, B., & Mavers, D. (2003). Mapping learning potential: Students' conceptions of ICT in their world. *Assessment in Education: Principles, Policy & Practice, 10*, 409–420.

Sweller, J. (2003). Evolution of human cognitive architecture. In B. Ross (Ed.), *The psychology of learning and motivation* (Vol. 43, pp. 215–266). San Diego: Academic.

Thurstone, L. L. (1927). A law of comparative judgment. *Psychological Review,* (34) 273–286.

Torney-Purta, J., Lehmann, R., Oswald, H., & Schulz, W. (2001). *Citizenship and education in twenty-eight countries: Civic knowledge and engagement at age fourteen*. Amsterdam: IEA.

Voogt, J., & Pelgrum, W. J. (2003). ICT and the curriculum. In R. B. Kozma (Ed.), *Technology, innovation, and educational change: A global perspective* (pp. 81–124). Eugene: International Society for Technology in Education.

Wall, D. (2005). *The impact of high-stakes examinations on classroom teaching* (*Studies in Language Testing*, Vol. 22). Cambridge: Cambridge University Press.

Walton, S. (2005). *The eVIVA project: Using e-portfolios in the classroom*. BETT. Retrieved June 7, 2007, from www.qca.org.uk/downloads/10359_eviva_bett_2005.pdf

Wasson, B., Ludvigsen, S., & Hoppe, U. (Eds.). (2003). *Designing for change in networked learning environments: Proceedings of the International Conference on Computer Support for Collaborative Learning 2003* (Computer-Supported Collaborative Learning Series, Vol. 2). Dordrecht: Kluwer Academic Publishers.

Webb, N.L. (1999). *Alignment of science and mathematics standards and assessments in four states* (Research Monograph No. 18). Madison: National Institute for Science Education.

Wegerif, R., & Dawes, L. (2004). *Thinking and learning with ICT: Raising achievement in primary classrooms*. London: Routledge Falmer.

Whitelock, D., with contributions from Road, M., & Ripley, M. (2007). *Effective practice with e-Assessment*. The Joint Information Systems Committee (JISC), UK. Retrieved November 30, 2009 http://www.jisc.ac.uk/publications/documents/pub_eassesspracticeguide.aspx

Williams, J. B., & Wong, A. (2009). The efficacy of final examinations: A comparative study of closed-book, invigilated exams and open-book, open-web exams. *British Journal of Educational Technology, 40*(2), 227–236.

Wilson, M., & Sloane, K. (2000). From principles to practice: an embedded assessment system. *Applied Measurement in Education, 13*(2), 181–208.

Woodward, H., & Nanlohy, P. (2004). Digital portfolios in pre-service teacher education. *Assessment in Education: Principles, Policy & Practice, 11*, 167–178.