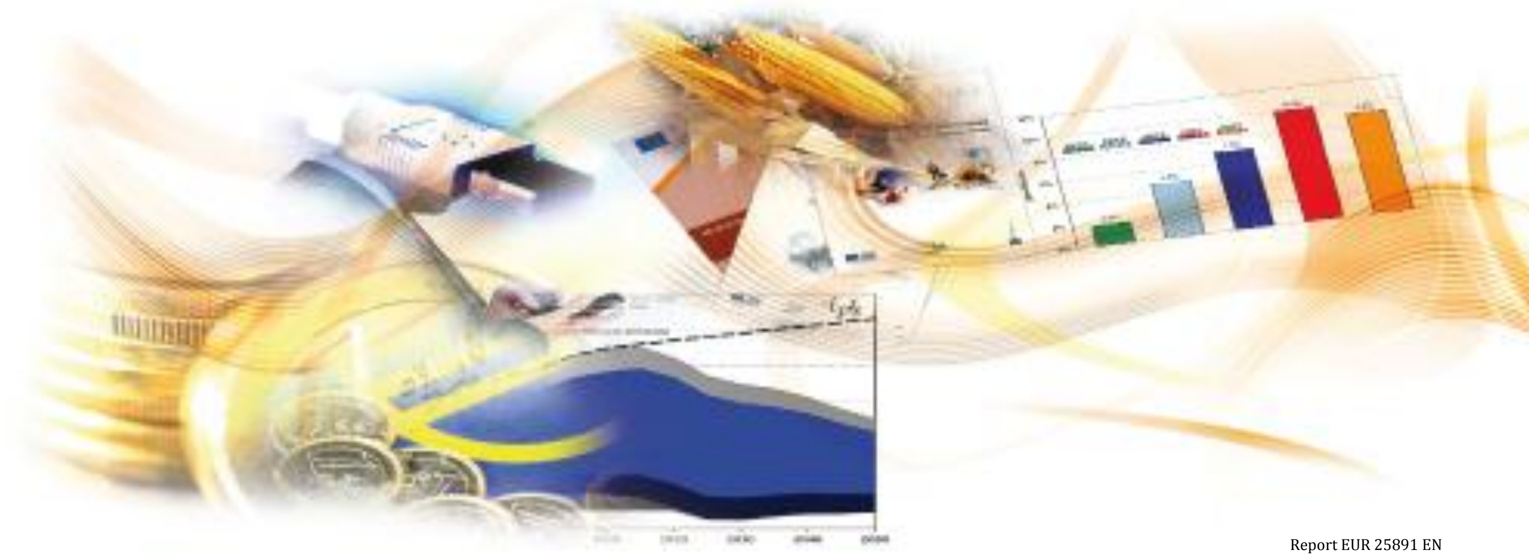
European Commission

JRC SCIENTIFIC AND POLICY REPORTS

The Use of ICT for the Assessment of Key Competences

Christine Redecker

2013



Report EUR 25891 EN

Joint Research Centre

European Commission
Joint Research Centre
Institute for Prospective Technological Studies

Contact information
Address: Edificio Expo. c/ Inca Garcilaso, 3. E-41092 Seville (Spain)
E-mail: jrc-ipts-secretariat@ec.europa.eu
Tel.: +34 954488318
Fax: +34 954488300

http://ipts.jrc.ec.europa.eu
http://www.jrc.ec.europa.eu

This publication is a Scientific and Policy Report by the Joint Research Centre of the European Commission.

JRC76971

EUR 25891 EN

ISBN978-92-79-29020-6 (pdf)

ISSN 1831-9424 (online)

doi:10.2791/87007

Luxembourg: Publications Office of the European Union, 2013

Printed in Spain

ACKNOWLEDGEMENTS

First and foremost I would like to thank my colleagues Yves Punie and Anusca Ferrari (DG JRC IPTS) as well as Tapio Säävalaa (DG EAC) for their valuable comments and feedback, both in shaping the project and in revising its subsequent drafts. I am furthermore deeply indebted to Riel Miller (UNESCO) and Øystein Johannessen (IT2ET Edu Lead) for their meticulous and critical reviews of a previous version of this report which have contributed to substantially improving its content, scope and quality.

I would also like to thank the members of the Thematic Working Group on the Assessment of Key Competences, led by DG EAC, for their valuable input and reflections. Finally, my thanks go to Patricia Farrer for editing and proofreading this report to prepare it for publication.

However, as ever, the views and conclusions expressed in the report, together with any errors or omissions, are the responsibility of the author.

EXECUTIVE SUMMARY

With the 2006 European Recommendation on Key Competences (Council of the European Union, 2006), all EU Member States have agreed on a framework of eight Key Competences for Lifelong Learning. Competences are defined as a combination of knowledge, skills and attitudes appropriate to the context and are seen as necessary by all citizens for personal fulfilment and development, active citizenship, social inclusion and employment. The eight Key Competences include: communication in the mother tongue, communication in foreign languages, mathematical competence and basic competences in science and technology, digital competence, learning to learn, social and civic competences, sense of initiative and entrepreneurship, and cultural awareness and expression.

European countries have made significant progress in incorporating these key competences into national curricula and other steering documents (Eurydice, 2012). However, one of the key challenges for education systems in many European Member States is the assessment of these competences (European Commission, 2012a).

Information and Communication Technologies (ICT) offer many opportunities for the provision of assessment formats which comprehensively capture all Key Competences, respecting the importance of the skills and attitudes dimensions and accounting for the less tangible themes underlying all Key Competences, such as critical thinking or creativity.

There is a vast range of formats and approaches which can foster different aspects of key competence development and can be used to address the specifics of each key competence in a targeted way. However, take-up and implementation in school education is still low. To seize the opportunities offered by ICT, targeted measures are needed to encourage the development, deployment and large-scale implementation of innovative assessment formats in school education.

The conceptual landscape

Currently, two conceptually different approaches to assessing Key Competences have been developing in parallel. On the one hand, Computer-Based Assessment (CBA) approaches have been employed for more than two decades and now go beyond simple multiple choice test formats (Generation 1&2 testing). With this new "Generation Re-Invention" or "transformative" testing, questions are increasingly integrated into more complex and authentic problem contexts, so that the full range of Key Competences can be assessed. Additionally, due to technological advances, a wider range of answer formats, including free text and speech, can be automatically scored. These question formats are already used to a certain extent in national and international large-scale tests and there are numerous interesting small-scale experiments, trials and pilots. However, in school education, the potential of transformative testing remains untapped.

On the other hand, technology-enhanced learning environments offer a promising avenue for embedded assessment of the more complex and behavioural dimensions of Key Competences, based on Learning Analytics. Many of the currently available technology-enhanced learning environments, tools and systems recreate learning situations which require complex thinking, problem-solving and collaboration strategies and thus allow for the development of generic skills. Some of these environments allow learners and teachers to assess performance, understand mistakes and learn from them. There is some evidence of how data on student engagement in these environments can be used as a basis for assessment and can also be directly fed back to students. However, on the whole, many of these programmes and environments are still experimental in scope and have not been mainstreamed in education and training. Embedded assessment has not yet matured, nor has it been widely or critically studied.

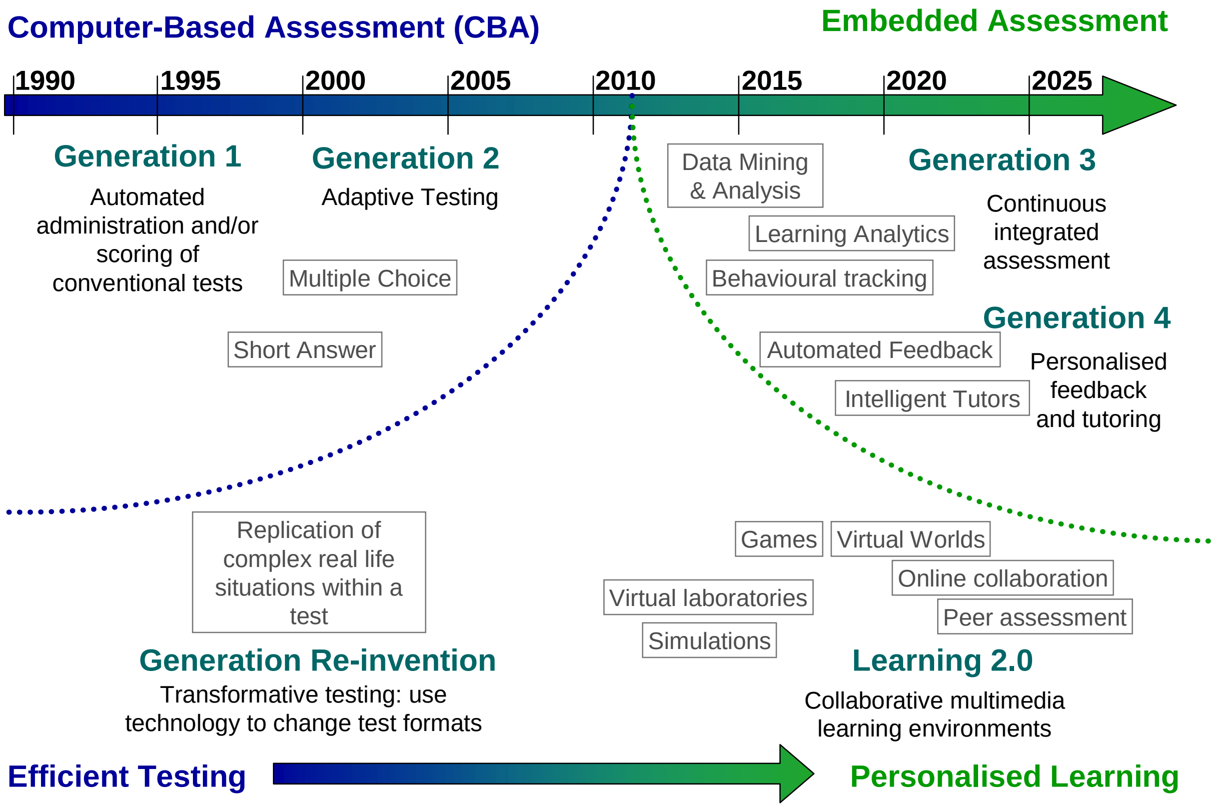


Figure 1: Overview of developments and trends in technology-enhanced assessment

Source: Elaborated by the author on the basis of Bunderson (1989), Martin (2008) and Bennett (2010)

Using ICT to support the assessment of Key Competences

Different Key Competences require different approaches to assessment. Thus, the strengths of different technology-enhanced assessment strategies depend on which competence and competence area is to be addressed. Figure 2 illustrates these focus areas. Currently, the potential of ICT in the comprehensive assessment of Key Competences, including less tangible and generic skills, remains largely untapped.

Current strategies for using ICT to foster competence-based assessment focus on Computer-Based Testing, online quizzes or simple games, and ePortfolios.

Computer-Based Testing is used widely and successfully for the summative and formative assessment of (basic) literacy and (advanced) reading skills and basic mathematical skills. Literacy and numeracy skills are being assessed in a range of national and international tests, which are, in many cases, electronic. Due to the nature of these competences, it has been possible to embed complex and authentic tasks in the multiple-choice format, so that mathematical competence can be comprehensively and reliably assessed by computer-based tests.

In general, however, computer-based tests tend to replicate traditional assessment formats, which focus on knowledge rather than skills and attitudes, and are not usually employed as a means of supporting more personalised, engaging, collaborative or authentic tasks. The advantage of computer-based tests over traditional assessment formats is that they provide instant and targeted feedback and can automatically adapt the difficulty of the test items to learners' different performance levels, to support formative assessment.

The internet is a vast resource for free and commercial **computer-based quizzes, games and tests** which can be used in the development and assessment of competences in literacy, reading and text comprehension and mathematics, in primary and secondary education. However, games, quizzes and engaging test formats are currently scattered, isolated, limited in scope, and ill-suited to comprehensive use in curricula and teaching.

ePortfolios are ideally suited to the assessment of collections of work produced by students and are thus particularly powerful tools for communication *in the mother tongue, communication in foreign languages* and *cultural awareness and expression*. ePortfolios are already widely used in European schools as a means of supporting the formative and summative assessment of students' creative productions. However, more innovative formats of cultural and artistic expression, such as blogs, wikis, tweets, audio and video recordings, etc., are seldom included. Educators often do not realize that ePortfolios can also be powerful tools for encouraging online collaboration and also self- and peer assessment, which contribute to and at the same time assess students' *learning to learn* skills.

	F = Formative Assessment **S = Summative Assessment**	**CBA**	**Quizzes / simple games**	**ePortfolios**	**Virtual worlds & games**	**Simulations**	**Intelligent Tutors**
1	Communication in the mother tongue	S	F	FS			
2	Communication in foreign languages	S		FS			
3a	Mathematical competence	S	F			FS	FS
3b	Basic competences in science and technology				FS	FS	
4	Digital competence					FS	
5	Learning to learn						
6	Social and civic competences				FS		
7	Sense of initiative and entrepreneurship				FS		
8	Cultural awareness and expression			FS	FS		

Figure 2: Overview of the potential of different ICT-based tools for the assessment of Key Competences

Promising trends for the future include the increased use of digital learning environments in education, the emergence of virtual environments which replicate authentic learning contexts and educational software which provides detailed and immediate feedback and allows teachers to adapt the learning process to each individual learner's pace and level. The key technology, which makes behaviour in these and other digital learning environments not only trackable, but also assessable, is Learning Analytics.

Technology-enhanced learning environments, which are often used in higher education and are starting to be deployed in school education as well, are used by some schools as a means of creating learning situations which require complex thinking, problem-solving and collaboration strategies. Some of these environments allow learners and teachers to assess performance, understand mistakes and learn from them. The use made of these tools depends highly on individual teachers' intentions.

Immersive environments and multiplayer games recreate learning situations which require complex thinking, problem-solving and collaboration strategies and thus allow the development of these skills, which are key components of all eight Key Competences. These environments replicate authentic contexts; encourage collaboration, empathy and negotiation; and reward strategic thinking, initiative and experimentation. For *competences in science*, in particular, **computer simulations and virtual laboratories** provide opportunities for students to develop and apply skills and knowledge in more realistic contexts and provide feedback in real time. **Practical tasks**, embedded in electronic environments, using mobile devices or online resources, are a further promising avenue for developing ICT-enabled assessment formats that better capture Key Competences, particularly *digital competences*. Since learners' behaviour in these electronic environments is tracked, their individual learning journeys – and also their possession of these skills – can, at least in principle, be automatically assessed.

Online simulations, virtual laboratories and games fostering competences in *science* are readily available, often for free. However, though they are being successfully used by many teachers and learners, these tools have not yet been mainstreamed. Environments that follow a games-approach are usually employed outside the curriculum, with no assessment of students' performance or learning gains.

Educational software solutions such as **intelligent tutoring systems** combine embedded (formative) assessment with instant feedback and targeted support. For example, these systems allow students to investigate mathematical concepts and problems in complex contexts, at their own pace. These systems include a series of tasks, which can be adapted in level of difficulty, and have helpful hints to encourage students to develop adequate solution strategies. Whereas these tools are popular in the US, they are not widely used in Europe.

In general, **Learning Analytics** is one of the most promising emerging technological trends for the comprehensive assessment of complex competences. Learning Analytics involves the interpretation of a wide range of data produced by and gathered on behalf of students in electronic environments in order to assess progress, and tailor education to individual students more effectively. Learning Analytics could allow assessment to be embedded in immersive environments, multiplayer games and computer simulations.

Challenges and Ways Ahead

On the whole, many of the more promising tools and environments for the assessment of Key Competences are still experimental in scope and have not become mainstream in education and training. In particular, learning analytics and embedded assessment, which are expected to become the most promising technological innovations for the assessment of Key Competences, have not yet matured and have not been widely or critically studied. Since technology is constantly evolving and many of the other more viable and interesting assessment formats for the assessment of Key Competences have only recently become available, it is not surprising (or worrying) that take-up in schools is slow.

Technological research and development should be focused on the most promising emerging techniques for comprehensive competence-based assessment, such as Learning Analytics and dynamic and interactive educational software for self-regulated learning.

Development, deployment and implementation of existing technological solutions should focus on increasing their scope, usability, variability and curricula-fit and also on the integration of complex and authentic assessment tasks and self- and peer-assessment options.

Pedagogical strategies that use ICT for the assessment of Key Competences should choose assessment formats that encourage alternative solutions and promote experimentation; promote self-regulated learning through self- and peer-assessment; and create learning contexts that allow learners to express themselves across a range of media and communication formats.

Policy Recommendations

To encourage the take-up of available tools and applications in schools, more policy support and guidance is needed for teachers, learners and parents.

In particular, the following policy options should be considered:

→ **Improve usability and curricula-fit**: Policy action is needed to support the development, use and accessibility of ICT environments and tools that take into account curricular needs and are better suited to use by teachers as part of their daily teaching practice.

→ **Teacher networking and knowledge exchange**: Many teachers are not aware of the possibilities that ICT offer to enhance assessment for learning. Teacher networks can facilitate knowledge exchange and learning and can contribute to upscaling and mainstreaming existing good practice.

→ **Research and development** should devote more attention to innovative learning and assessment environments, such as educational multiplayer games and simulations, and consider how learning analytics can meaningully be used to foster formative assessment.

→ **Encourage discussion and provide guidance:** A critical and open discourse among educators, researchers and policy makers is needed on the advantages and drawbacks of ICT-enhanced assessment strategies, in order to identify viable strategies that allow the comprehensive assessment of all Key Competences for Lifelong Learning.

TABLE OF CONTENTS

1. INTRODUCTION

We are currently living in an era of accelerated change as concerns not only technological developments, but also society on the whole. As a consequence, the skills and competences needed for work and life in the 21st century are continuously evolving. Policy is reacting towards these changes by calling for education to focus on the development of Key Competences for Lifelong Learning (Council of the European Union, 2006). The recent "Rethinking Education Strategy" (European Commission, 2012b) again emphasizes the need for the development of transversal skills and basic skills at all levels.

However, learning processes and goals can only change if assessment changes as well. This report aims to shed some light on how Information and Communication Technologies (ICT) can contribute to offering assessment formats that comprehensively capture students' competences and allow for skills and attitudes as they are displayed in authentic situations to be assessed, in a way to inform and assist both learners and teachers in improving each individual learner's competence development.

Policy Context

With the 2006 European Recommendation on Key Competences (Council of the European Union, 2006), all EU Member States have agreed on a framework of eight Key Competences for Lifelong Learning. Competences are defined as a combination of knowledge, skills and attitudes appropriate to the context and are foreseen as necessary by all citizens for personal fulfilment and development, active citizenship, social inclusion and employment. The eight Key Competences include: communication in the mother tongue, communication in foreign languages, mathematical competence and basic competences in science and technology, digital competence, learning to learn, social and civic competences, sense of initiative and entrepreneurship, cultural awareness and expression. The Key Competences Recommendation (2006) emphasizes, in particular, the transversal dimension of all eight Key Competences.

The "Education and Training 2020" (ET 2020) strategic framework for European cooperation in education and training (Council of the European Union, 2009) underlines that all citizens need to be able to acquire these Key Competencies. Acquiring Key Competences is a priority for European and Member States policies, as argued at European level in the Europe 2020 Strategy (European Commission, 2010), in particular the flagships "Digital Agenda", "New Skills and Jobs", "Youth on the move" and "Innovation Union".

The key role of assessment in the learning process and in the acquisition of Key Competences was first emphasised by the 2009 Joint Report of the Council and the Commission "Key competences for a Changing World" (European Commission, 2009). Based on Member States' reports, it shows that most countries have recently changed school curricula to incorporate at least some elements of the Key Competences, or even the entire European framework in some cases. Despite this trend, progress is still needed in teacher training, in the development of learning materials, and in the update of assessment methodologies. In this respect, the report concludes that "most current assessment methods have a strong emphasis on knowledge and recall and do not sufficiently capture the crucial skills and attitudes dimension of Key Competences. Also the assessment of transversal Key Competences and the assessment in the context of cross-curricular work appear inadequate."(European Commission, 2009) The report refers to the growing use of complementary methodologies, such as portfolios, peer assessment, and project work, and suggests that these should be examined and developed further.

The recent "Rethinking Education Strategy" (European Commission, 2012b) emphasizes the need to build skills for the 21st century and underlines the importance of transversal skills and basic skills. It also acknowledges that "What is assessed can often determine what is valued and what is taught. While many Member States have reformed curricula, it remains a challenge to modernise assessment to support learning." Both the Communication and its accompanying Staff Working Paper on the assessment of Key Competences point out that "the potential of new technologies to help find ways of assessing key competences needs to be fully explored" (European Commission, 2012b).

Technologies are increasingly adopted for learning outside formal educational environments (Ala-Mutka, 2010; Collins & Halverson, 2010a) and are steadily making their way into the classroom (Cachia & Ferrari, 2010). However, there is no clear evidence on how ICT is currently used to address the inadequacies currently encountered in the assessment of Key Competences.

Thus, there is a need to better understand how ICT for assessment can support modernising schools and education systems for providing future skills and Key Competences efficiently for all learners and on how the attitude and skills dimension, as well as the creative application of the Key Competences can adequately be supported through assessment.

Thus, the key questions for this study are:

1. In which ways ICT is currently used for assessment of knowledge, skills and attitudes as set out in the European Framework on Key Competences? What approaches have proved to be effective and what are the benefits and drawbacks?
2. What is the potential of ICT for assessing Key Competences? How can ICT be used to assess the skills and attitudes dimension of the Key Competences? What are the specific considerations that need to be taken into account as concerns teachers, learners and educational institutions and curricular aims?
3. What are the implications of the gap between the current use and the potential of ICT for policies related to the assessment and promotion of Key Competences?

Research Context

(NACCCE, 1999). Assessment is one of the most powerful influences on teaching and learning. It is an essential component of learning and teaching, as it allows the quality of both teaching and learning to be judged and improved (Ferrari, Cachia, & Punie, 2009). It often determines the priorities of education (NACCCE, 1999), it always influences practices and has backwash effects on learning (Ellis & Barrs, 2008). Moreover, changes in curricula and learning objectives are ineffective if assessment practices remain the same (Cachia, Ferrari, Ala-Mutka, & Punie, 2010), as learning and teaching tends to be modelled against the test rather than according to curriculum guidance or innovative best practice.

However, assessment as it is currently implemented in Europe tends to put too much emphasis on subject knowledge, and less on skills and attitudes, and to neglect altogether the increasingly important cross-curricular competences such as learning to learn or entrepreneurship (European Commission, 2012a). According to a recent JRC-IPTS study, collecting the opinions of almost 8,000 teachers from all over Europe, formal tests are still the predominant form of assessment in many schools, with 76% of respondents claiming to often or always assess their students this way (Cachia & Ferrari, 2010).

Assessment is usually been understood as having three purposes: diagnostic, formative and summative. Diagnostic assessment is used to analyse pupils' capabilities and aptitudes as a basis for planning; Formative assessment has the aim to gather evidence about the pupils' progress to influence teaching methods and priorities; and summative assessment is used to judge pupils' achievements at the end of a programme of work (NACCCE, 1999).

Information and Communication Technologies (ICT) are currently being exploited in different ways to support mainly summative, but also formative and diagnostic assessment. In particular, computer-based tests are currently being used widely and for a variety of educational purposes, especially in the US (Bennett, 2010; Bridgeman, 2009; Csapó, Ainley, Bennett, Latour, & Law, 2012) but increasingly also in Europe (Eggen & Straetmans, 2009; Moe, 2009; Wandall, 2009). However, computer-based tests are generally used as a means of increasing the efficiency and effectiveness of test administration; improving the validity and reliability of test scores; and making a greater range of test formats susceptible to automatic scoring. Thus, despite the variety of computer-enhanced test formats, eAssessment strategies tend to replicate traditional assessment approaches, based on the explicit testing of knowledge.

The European Framework on Key Competences poses three particular challenges to assessment in general, and to technology-enhanced assessment in particular: How can assessment capture and promote student's capacity to apply knowledge in *authentic situations* that require using a broad combination of Key Competences? How can assessment capture and foster *skills and attitudes*, such as critical thinking, creativity, initiative, problem solving, risk assessment, decision taking, and constructive management of feelings? How can assessment be meaningfully *integrated* into learning as to support formative assessment and self-regulated learning?

Outline of this report

This report addresses the key questions from two different angles, from the angle of research on technology-enhanced assessment strategies (chapter 2), and from the perspective of each of the eight Key Competences (chapter 3). Chapter 2 is dedicated to a description and discussion of common and emerging computer-enhanced assessment formats, with a focus on tools and environments that are currently used or could be used in the near future to support assessment in primary and secondary education. Chapter 3 discusses, for each of the eight European Key Competences, computer-aided assessment strategies which can contribute to enabling schools and teachers to move from knowledge-based to competence-based assessment.

Chapter 4 discusses challenges, obstacles, barriers and bottlenecks to the use of ICT for the assessment of Key Competences in the European Union and chapter 5 closes with the report conclusions and policy recommendations.

2. ICT-ENHANCED ASSESSMENT STRATEGIES

Over the last decade or two technologies have increasingly been used to support and shape assessment processes. At the end of the eighties, Bunderson, Inouye and Olsen (1989) published an influential article about the four generations of computerized educational measurement, foreseeing the following development (Martin, 2008):

Generation 1: Computerized testing: administering conventional tests by computer;

Generation 2: Computerized adaptive testing: tailoring the difficulty or contents of the next piece presented or an aspect of the timing of the next item on the basis of examinees' responses;

Generation 3: Continuous measurement: using calibrated measures embedded in a curriculum to continuously and unobtrusively estimate dynamic changes in the student's achievement trajectory and profile as a learner;

Generation 4: Intelligent measurement: producing intelligent scoring, interpretation of individual profiles, and advice to learners and teachers, by means of knowledge bases and inference procedures.

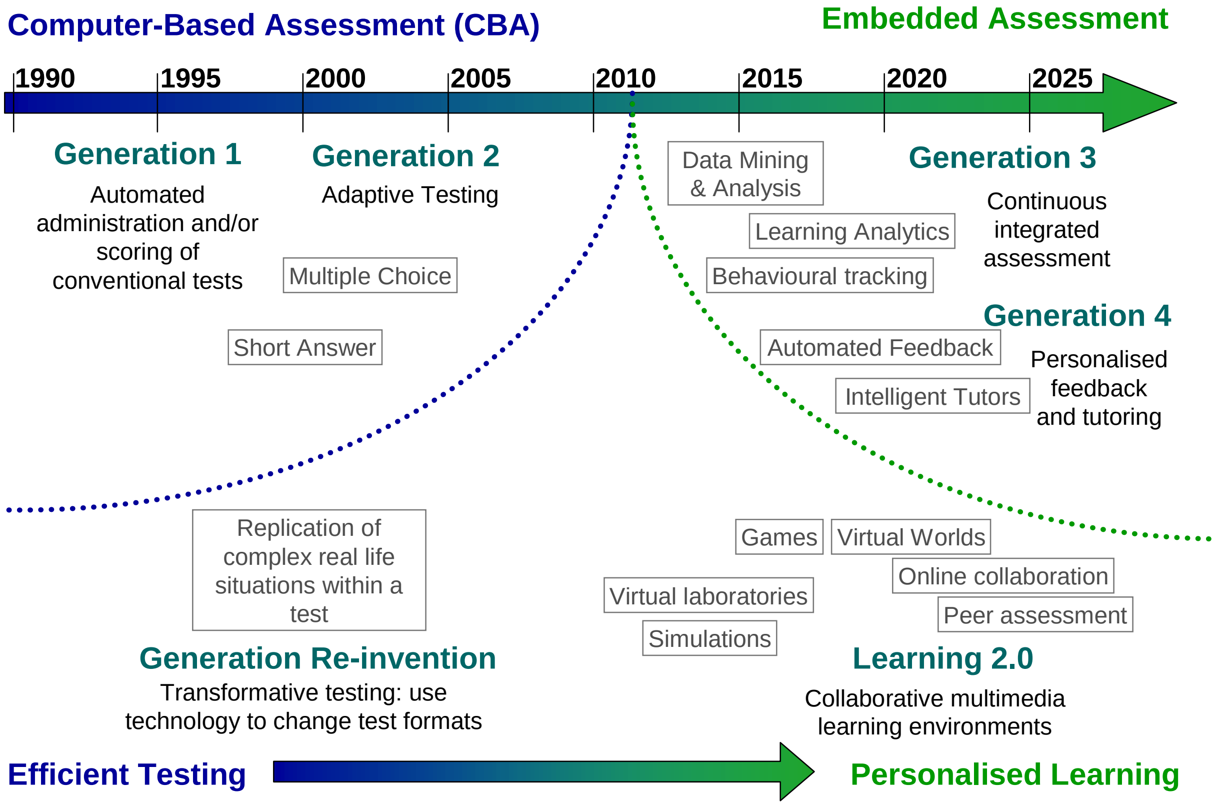


Figure 3: Overview of developments and trends in technology-enhanced assessment

Source: Elaborated by the author on the basis of Bunderson (1989), Martin (2008) and Bennett (2010)

Interestingly, these predictions are not far off the mark. The first two generations of eAssessment or Computer-Based-Assessment (CBA), which are more precisely referred to as Computer-Based-Testing (CBT), have by now become mainstream and are routinely employed in a range of large-scale tests (for more, see chapter 3). The main challenge currently lies in making the transition from the first two generations to the latter two. Thanks to current developments in data-mining and first developments and trials with intelligent electronic tutor systems, generations 3 and 4 could technologically become a reality within the next five years (L. Johnson, Smith, Willis, Levine, & Haywood, 2011). However, more time will be needed to conceptually make the leap between the era of computer-based testing and the era of technology-enhanced learning. While the first two generations of CBA centre on the notion of *testing* and on the use of computers to improve the efficiency of testing procedures, generation 3 and 4 do away with the notion of tests and seamlessly integrate holistic and personalised assessment into *learning*.

Thus, the vision underlying generations 3 and 4 is that computers will make explicit testing obsolete. Learners will be continuously monitored and guided by the electronic environment which they use for their learning activities, so that diagnostic, formative and summative assessment will become embedded in the learning process. Ultimately (generation 4), assessment will become inseparably intertwined with teaching and learning, with learning systems being able to provide instant and valid feedback and advise to learners and teachers concerning future learning strategies, based on the learners' individual learning needs and preferences, as displayed in his past and current learning activities.

In parallel to the conceptual shift emerging in the area of computer-based assessment, Western societies are currently realising that learning objectives need to be revised to more adequately reflect the skills and competences needed for life in the 21st century. The evolution of information and communication technologies (ICT) has not only changed assessment procedures, but has also, and more significantly, changed society on the whole and with it the skills and competences that students need to acquire to be successful in life.

Thus, while numeric, verbal, scientific and digital literacy are and will remain key building blocks for successful participation in society, problem-solving, reflection, creativity, critical thinking, learning to learn, risk-taking, collaboration, and entrepreneurship are increasingly becoming important in the future (Redecker, et al., 2010). The European Key Competence Recommendation (2006) responds to these changed competence needs – that in the research literature are usually alluded to as "21st century skills" (Binkley, et al., 2012) – by emphasizing them as "themes" that "play a role in all eight Key Competences", and include, in particular: "critical thinking, creativity, initiative, problem solving, risk assessment, decision taking, and constructive management of feelings" (Council of the European Union, 2006).

To foster and develop these important themes, Education and Training systems must change their curricula and their pedagogic practices. Most importantly, Education and Training in Europe has to reconsider assessment practice. Assessment strategies have to be developed that go beyond testing factual knowledge. These strategies must aim to include of the skills and attitudes dimension of Key Competences in the assessment process and allow for capturing the less tangible themes underlying all Key Competences.

It has been argued that, while ICT are to a large extend used to deliver traditional assessment formats more effectively and efficiently (the "migratory strategy"), ICT can also be used to modify and alter how competences are assessed (the "transformative strategy") and, in this respect, give rise to formats and environments which facilitate the assessment of skills that have been difficult to measure with traditional assessment formats following a test-approach (Binkley, et al., 2012; Ripley, 2009). Transformative or "Generation Reinvention" assessment strategies can support and trigger the assessment of the crosscutting themes that support all Key Competences, such as, in particular, problem solving, inquiry skills, organisation, collaboration and communication skills. By allowing for more complex cognitive strategies to be assessed while remaining grounded in the testing paradigm, these innovative testing formats could facilitate the paradigm shift between the first two and the last two generations of eAssessment, e.g. between explicit and implicit assessment.

Moreover, technology-enhanced learning environments offer a promising avenue for the assessment of the more complex and behavioural dimensions of the European Key Competences. Many of these technology-enhanced learning environments, tools and systems recreate learning situations which require complex thinking, problem-solving and collaboration strategies and thus allow for the development of these generic skills. Some of these environments allow learners and teachers to assess performance, understand mistakes and learn from them. There is some evidence on how data on student engagement in these environments can be used as a basis for assessment and can also be directly fed back to students. However, on the whole, many of these programs and environments are still experimental in scope and have not become mainstream use in education and training. Embedded assessment has not yet matured and has not widely or critically been studied.

Hence, for the current study, two research lines converge. On the one hand, the question arises how classical Computer-Based-Tests can support the assessment of Key Competences by allowing for the implementation of question and answer formats that more adequately capture Key Competences (sections 2.1 and 2.2). On the other hand, it is important to understand how newly emerging ICT-facilitated environments and tools, which are designed to foster the development of Key Competences and their underlying themes, can also – at least prospectively – support diagnostic, formative and summative assessment (sections 2.5 and 2.6). Additionally, there are assessment formats and strategies which are better suited than tests to capture the less tangible aspects of Key Competences, such as portfolio assessment (section 2.3) and peer assessment (section 2.4). The use of ICT can enhance the effectiveness and viability of these assessment strategies.

2.1 eAssessment

The terms e-Assessment, Computer-Based Assessment (CBA), Computer-Assisted/Aided Assessment (CAA), computerized testing and computer-administered testing, are often used interchangeably to denote an electronic assessment processes where ICT is used for the presentation of assessment activity and the recording of responses (JISC, 2006). Since these terms are usually used to refer to computer-based multiple-choice or short-answer response tests, the notion of Computer-Based Testing (CBT) more adequately reflects this assessment approach, in contrast to more integrated approaches, such as ePortfolio assessment. However, it has to be noted that lately more diverse question and answer formats have been integrated in what has been dubbed "generation re-invention" testing (see below, section 2.1.4), which shows that CBT can also be used in more innovative ways.

First and second generation tests, i.e. computer-based conventional tests and computer adaptive tests, have over the years led to a more effective and efficient delivery of traditional assessments (Martin, 2008). More recently, assessment tools have been expanded and enriched to better reflect the intended domains, to include more authentic tasks and to allow for the assessment of constructs that have either been difficult to assess or which have emerged as part of the information age (Pellegrino, 2010). As the measurement accuracy of all of these test approaches depends on the quality of the items it includes, item selection procedures – such as Item Response Theory or mathematical programming – play a central role in the assessment process (El-Alfy & Abdel-Aal, 2008).

Computer-based tests offer numerous advantages over paper-based tests. Advantages include the paperless test distribution and data collection, greater standardization of test administrations, monitoring of student motivation, obtaining machine-scorable responses for writing and speaking, providing standardized tools for examinees (e.g., calculators and dictionaries), and the opportunity for more interactive question types (Bridgeman, 2009).

Furthermore computer-based tests have a positive effect on students' motivation, concentration and performance; provide teachers with access to high quality materials; thanks to automatic scoring teachers can focus on analysis and interpretation of assessment results; and, more recently,. eAssessment applications are being develped which provide learners and teachers with detailed reports that describe strengths and weaknesses, thus supporting formative assessment (Ripley, 2009).

The growth of large-scale computer based tests has led to the development of pupil monitoring systems, such as the one developed by Cito in the Netherlands, which mirror the National test format and support the teacher in tailoring teaching to the individual level of a student and in comparing the progress and results of teaching with national standards (Glas & Geerlings, 2009).

2.1.1 Research on Compuer-Based Assessment (CBA)

Comparing Paper-Based and Computer-Based Tests

In 2005 the Education Testing Service (ETS) published the findings of a large scale comparison of paper-based and computer-delivered assessments (National Centre for Education Statistics, 2005). The empirical data were collected in 2001 and involved over 2,500 year 8 students who completed either mathematics or writing assessments. The traditional, paper-based tests were migrated to screen format, with little or no amendment made for the purpose of screen delivery. In mathematics the study found no significant differences between performance on paper and on screen, except for those students reporting at least one parent with a degree. These students performed better on paper. The study also found no significant differences in writing, expect for students from urban fringe/large town locations (Ripley, 2009). Another study comparing computer and paper-and-pencil versions of mathematics questions for end of key stage 2 and 3 assessments in England, also comes to the conclusion that in most cases the change to the different medium seems to make little difference. However, for some items the affordances of the computer profoundly affect how the question is attempted, and therefore what is being assessed when the item is used in a test (Threlfall, Pool, Homer, & Swinnerton, 2007).

Similarly, Hardré et al. (2007) compared paper-based, computer-based and web-based test administration and found that, overall, few differences in data quality were observed between the different administration conditions despite some evidence in favour of paper-based administration, in particular participants' preference for paper-based tests. However, in two nationally representative studies, one in mathematics and one in writing, both focused primarily on 8^{th} grade students, conducted by the US National Assessment of Educational Progress (NAEP) conducted two nationally representative studies, on in mathematics and one in writing, both focused primarily on 8^{th} grade students, the computer and paper tests did not appear to measure the same skills (Bennett, 2010; Horkay, Bennett, Allen, Kaplan, & Yan, 2006).

Reliability and Validity of Scores

In an experiment with fifth-grade and university students, Park (2010) compared short answer and multiple choice format. Consistent with previous research students attained higher score in the multiple choice format; the correlation between the short answer and the multiple choice format was high; and the item discrimination was higher in the short answer format.

Another study found that, multiple choice tests and oral examinations results were consistent when the multiple choice questions were paired in such a way as to detect and penalise guessing while both scores were significantly different from the un-paired multiple choice test result (Ventouras, Triantis, Tsiakas, & Stergiopoulos, 2011).

CBT for formative assessment

Investigating the use of an online assessment tool in a large, first-year business mathematics course at a major Australian university, Angus and Watson (2009) found that higher exposure to the online instrument robustly lead to higher student learning, all else being equal. Wilson et al. (2011) examined the effectiveness of voluntary computer-assisted multiple-choice practice tests on student performance in a large, first-year undergraduate geography course. They found students who use the computer-assisted practice quizzes earned significantly higher grades in midterm exams than those students who did not.

Future technological developments

The future of test construction may lie - at least for highly structured ability domains - in the field of computer based, automatic generation of psychological test items (Arendasy, Sommer, & Hergovich, 2007).

Various approaches allow for the selection of test items that most adequately respond to the course content and the learner's individual learning history and competence level (Barla, et al., 2010). Thus, tests can be created that can adapt the difficulty level of the questions to the individual learner (Di Bitonto, Laterza, Roselli, & Rossano, 2010). For example, Chatzopoulou, and Economides (2010) developed a Web-based adaptive testing system for assessing Greek high school students' programming knowledge. The system was shown to successfully predict the corresponding classification of students in Greek National Exams and to assist teachers and learners in identifying shortcomings.

Technological challenges

One of the major challenges for CBT is the selection of test items that reliably assess the examinees competence level. To improve item selection procedures for large item banks, heuristic search and machine learning approaches, including neural networks, are being developed, and abductive network modeling is investigated to automatically identify the most-informative subset of test items that can be used to effectively assess the examinees without seriously degrading accuracy (El-Alfy & Abdel-Aal, 2008).

Multidimensional adaptive testing (MAT) can increase measurement efficiency in educational and psychological testing, by reducing the number of presented items by about 30-50% compared to unidimensional adaptive testing and by about 70% compared to fixed item testing (Frey & Seitz, 2009).

Multidimensional item response theory (MIRT) is being developed as foundation for modeling performance in complex domains, taking into account multiple basic abilities simultaneously, and representing different mixtures of the abilities required for different test items. However, currently different MIRT models coexist which differ substantially as concerns educational assessment (Hartig & Höhler, 2009).

2.1.2 First Generation Computer Based Tests

For the first two generations of tests that rely essentially on the transposition of existing paper and pencil tests, the main advantages of the computer administration compared to paper and pencil relies in efficiency gains as concerns administration, data collection and processing, scoring and reporting (Pellegrino, 2010; Ripley, 2009).

The most notable distinction from a paper-based test is in the adaptive nature of delivery and scoring utilized by several of the programs. Critical issues identified by this first generation of tests are: infrastructure and comparability (Pellegrino, 2010).

These tests are currently already being administered on a large scaled for a variety of educational purposes, especially in the US (Csapó, et al., 2012), but increasingly also in Europe. In Norway, for example, national CBT are used in reading Norwegian, mathematics and reading English, and after initial trials, it was decided to opt for linear (1st generation) CBT, rather than adaptive tests (Moe, 2009).

Example 1: Diagnostic Assessment in Hungary in Mathematics, Reading and Science

The Centre for Research on Learning and Instruction at the University of Szeged in Hungary is developing an online assessment system with funding from the Social Renewal Operational Program and the European Union social funds. Through this project, the goal is to lay the foundation for a nationwide diagnostic assessment system for the grades 1 through 6.

The project involves: The development of assessment frameworks; exploration of the possibility of using diagnostic assessments in various fields; item construction and the development of item banks; the setting up and piloting of an online assessment system; extension of assessment to students with special education needs; preparation of teachers and educational experts for participation in a variety of assessment processes; secondary analysis of data collected in national and international assessment programs.

It is implemented in 7 work packages: 1. Developing Frameworks for Diagnostic Assessments in Grades 1-6 in the Domains of Mathematics, Reading and Science; 2. Exploring the possibility of expanding diagnostic assessment to include other cognitive and affective domains; 3. Development of item banks for diagnostic assessments; 4. Implementing technology-based assessment; 5. Development of assessment instruments for students with special educational needs; 6. In-service teacher training; 7. Secondary analysis of the results of large-scale national and international assessments.

The funding base for this project is approximately 2.5 million euro for the first phase of the project that lasts for 22 months. The funding for the second phase of the project, starting in the spring of 2012, is on the same level, according to the project management.The project will develop an item bank in nine dimensions (reading, mathematics and science in three domains) as well as a number of other minor domains. According to the project management, the project has a high degree of transferability and can regarded as close to the framework in the PISA assessment.

There are areas that the project would like to look into, but for capacity reasons can not cover. These areas include game-based learning, monitoring how children play games and assessing the efficiency of this kind of intervention. The project would also like to look into how technology can be used to capture emotions.

This project may be one of the largest studies on online assessment. According to the project, the project receives substantial funding from the European Commission. It covers both core subjects such as mathematics, science and reading as well as looking for the feasibility of expanding diagnostic assessment to include other cognitive and affective domains.

The proliferation of CBA especially in the US, has given rise to a number of (commercial and non-commercial) services, libraries and search engines that allow educators to identify, find and compare available CBT for their particular assessment purposes. ETS (http://www.ets.org/test_link/about), for example, has a searchable database of more than 25,000 tests and other measurements. Toolfind (http://www.toolfind.org/) is a directory designed to help professionals in youth-serving programs find measurement tools for elementary, middle and/or high school students and youth, parent, staff and teacher respondents.

2.1.3 Second Generation: Adaptive Tests

In adaptive testing, the examinee's skill level is re-estimated after each item is administered. The next item to be administered is then chosen based, in part, on that updated skill-level estimate, so that examinees are branched to easier or harder questions based on their performance on prior questions (Bennett, 2010; Bridgeman, 2009). The computerized adaptive testing has its main advantage over paper and pencil in a more efficient administration mode (less items and less testing time), while at the same time keeping measurement precision (Martin, 2008).

The most well-known type of algorithmic testing is CAT, which is a test where the algorithm is designed to provide an accurate point estimation of individual ability or achievement (N. A. Thompson & Weiss, 2009). A similar, but lesser-known, approach is computerized classification testing (CCT), also known as sequential testing, where the algorithm is designed to classify students. For example, students can be classified as pass/fail or into educational proficiency levels such as basic/proficient/advanced (N. A. Thompson & Weiss, 2009). The undesirable feature in CATs that item exposure varies greatly with item difficulty and that the most discriminating items are used at high frequencies have indeed been quite problematic for this type of testing in terms of test security (Martin, 2008).

CAT test use is very widespread, in particular in the US: The Measures of Academic Progress (MAP), for example, is a computer-adaptive test series for primary and secondary school assessments in reading, mathematics, language usage, and science used in over 3.400 school districts (cf. Bridgeman, 2009; Csapó, et al., 2012). Similarly, the Oregon Assessment of Knowledge and Skills (OAKS), is an adaptive test in grades 3-12 in reading, mathematics, science and social studies (Bennett, 2010; Csapó, et al., 2012). A series of adaptive tests are used as part of the admission process to higher education in the US, such as the GRE-CAT, a variety of graduate programs at the masters and doctoral levels, the GMAT for graduate management courses, and the TOEFL test for no-native speakers seeking university entrance in the US (cf. Bridgeman, 2009).

In a number of European countries, adaptive tests are used on a national level to assess a number of Key Competences at different levels: In the Netherlands, CAT are used to test the arithmetic skills of students entering primary school teacher training colleges (Eggen & Straetmans, 2009). In Denmark, CAT is used for nationwide obligatory tests of students in primary and lower secondary schooling in seven school subjects, as a means of school evaluation (Wandall, 2009).

2.1.4 Generation Reinvention: Transforming Assessment

Bennett (2010) differentiates between three generations of Computer-Based Assessment, broadly mapping onto the normal adoption phases for technology (substitution, innovation and transformation) (R. Meijer, 2008). The first two generations roughly coincide with generation 1 and 2 as described by Bunderson, Inouye and Olsen (1989), whereas his third category – Generation R (Reinvention) – refers to developments in (1st and 2nd generation) CBT that prepare for the transition towards 3rd and 4th generation computer-based assessment. Generation R tests use complex simulation, sampling of student performance repeatedly over time, integration of assessment with instruction, and the measurement of new skills in more sophisticated ways (Bennett, 2010).

Similarly Ripley (2009) discusses the "transformational" approach to CBA, whereby the test developer sets out to redefine assessment and testing approaches in order to lead educational change, in particular to allow for the assessment of 21st century skills, such as problem solving, communication, team working, creativity and innovation. A good example are computer based tests that measure process by tracking the activities students undertake on the computer while answering a question or performing a task, such as, for example in the ETS iSkills test. However, experience from this and other trials indicates that developing these tests is far from trivial (Lent, 2008). One of the greatest challenges for the developers of transformative assessments is to design new, robust, comprehensible and publicly acceptable means of scoring student's work (Ripley, 2009).

2.1.5 Automated Scoring

To make computer based assessment more efficient, especially when large group of learners concerned, automatised scoring programs are being developed, based on a number of different algorithms for automated language analysis. Automated Language Analysis refers to an electronic process by which candidates' responses to essaystyle questions (typed rather than handwritten) are analysed and marked electronically; words or short phrases can be marked by much simpler key word matching processes (JISC, 2006).

Automated scoring has the potential to dramatically reduce the time and costs associated with the assessment of complex skills such as writing, but its use must be validated against a variety of criteria for it to be accepted by test users and stakeholders (Weigle, 2010).

Automated scoring can be decomposed into isolating scorable components (feature extraction), judging those components (feature evaluation), and score generation (feature accumulation (Csapó, et al., 2012). Automated scoring programs for essay, for example, tend to use features that are easily computable, such as structure, word complexity, average word length etc. and combine them in ways that best predict the scores awarded by human judges under operational conditions, omitting features that cannot be easily extracted from an essay by machine (Ben-Simon & Bennett, 2007). Some e-assessment systems that automatically evaluate free-text students' answers have recently been extended to include automatically generated students' conceptual models (Pérez-Marín & Pascual-Nieto, 2010).

Assignments in programming languages, or other formal notations, can be automatically tested and evaluated (Amelung, Krieger, & Rösner, 2011). Automatic assessment of free-text responses is more difficult. However, automated scoring is already used for scoring essay-length responses, for example by the Intelligent Essay Assessor (Pearson Knowledge Technologies), Intellimetric (Vantage), Project Essay Grande (PEG) (Measurement, Inc.) and e-rater (ETS) (Bennett, 2010). MS Office Electronic Feedback Software, is an example of an assessment tool which automatically generates and emails MS Word processed reports to students, thus providing more detailed feedback[1] and allowing for plagiarism to be detected more easily[2].

---

1 Cf. (Denton, Madden, Roberts, & Rowe, 2008)

2 http://www.ljmu.ac.uk/ITHelp/software/feedback.asp.

Reliability and Validity of Automatic Scoring

Electronic scoring of essays closely mimics the results of human scoring, and the agreement of an electronic score with a human score is typically as high as the agreement between two humans, and sometimes even higher (Bennett, 2010; Bridgeman, 2009). Research indicates that automatic grading displays a high reliability, also for the assessment of open-ended questions (Wang, Chang, & Li, 2008). The comparison of human and automated scores on responses to TOEFL® iBT Independent writing tasks, for example, indicate moderate but consistent correlations between both human and e-rater scores and non-test indicators, providing criterion-related validity evidence for the use of e-rater along with human scores (Weigle, 2010). Also scoring programmes developped for secondary education show that machine-generated scores achieve high inter-rater reliability against human grading (Wang, et al., 2008).

However, automatic scoring programs usually use length and mechanics as dominant features to assess essay-length texts, with only minimal weight given to content or organisation and development (Ben-Simon & Bennett, 2007). Thus, a study comparing human and computer marking of approximately 600 essays produced by 11-year-olds in the UK showed good agreement between human and machine marking. Yet, discrepancies concerned essays marked higher by humans which exhibited more abstract qualities such as interest and relevance, while there was little, if any, difference on more mechanical factors such as paragraph demarcation (Hutchison, 2007).

Spoken responses can also be captured by the computer and automatically scored. Automated scoring for highly predictable speech, such as a one sentence answer to a simple question, correlates very highly with human ratings of speech quality, while for longer and more open-ended responses, automated speech scoring is not yet good enough for use in high stakes tests (Bridgeman, 2009).

For short-answer free-text responses of around a sentence in length automatic scoring has been shown to be at least as good as that of human markers (Butcher & Jordan, 2010). Additionally, programs are being developed which not only author and reliably mark short-answer free-text assessment tasks, but also give tailored and relatively detailed feedback on incorrect and incomplete responses, inviting examinees to repeat the task immediately so as to learn from the feedback provided (Jordan & Mitchell, 2009). Thus, the traditional concerns about the the value of automated scoring of text for formative assessment (Gipps, 2005) is being overcome, at least for some question and answer types.

Technological Developments

Currently, a lot of research is being invested in improving automatic scoring techniques for free text answers (Noorbehbahani & Kardan, 2011) and dedicated written text assignments, such as summary writing (He, Hui, & Quan, 2009). \

Also, for other types of more complex assignment task, automatic assessment is being developed. There are, for example, systems that are able to automatically assess programming assignments written in a variety of languages based on the structure of the source code and the correctness of the program's output (Blumenstein, Green, Fogelman, Nguyen, & Muthukkumarasamy, 2008). Similarly, programs are being developed which allow for the authoring and automatic checking of geometry exercises (Isotani & Brandão, 2008). Automatic scoring has furthermore already been tested for capturing the analyze the information searching abilities of individual students engaged in an problem-solving exercise, which in experimental settings showed a high correlation with human rating (Chiou, Hwang, & Tseng, 2009).

Even for creative problem-solving assignments in (high school) science education automated grading schemes for natural language responses are being developed which, in experimental trials, showed to be highly reliable (Wang, et al., 2008).

2.2 Computer-based assessment tools for schools

There are a large number of electronic tools and services on the market supporting various kinds of assessment activities. Some of these are offered as specific modules of learning management systems (LMS) that enable the management of (usually multiple-choice) items together with the administration and internet-based delivery of tests (e.g. Moodle, http://www.moodle.org). Furthermore, there are comprehensive assessment management systems with specific focus on summative and formative assessment, like Questionmark Perception (http://www.questionmark.com/uk/perception/index.aspx) and different more targeted and isolated authoring software tools which allow teachers to compile, administer and grade electronic tests, including, for example Hot Potatoes (http://hotpot.uvic.ca), which focuses on test administration and OpenSurveyPilot (http://www.opensurveypilot.org/) which is dedicated to data collection and presentation (Scheuermann & Pereira, 2008).

2.2.1 Assessment Systems

Perie et al. (2009) discuss commercial formative assessment systems (which they argue should more correctly be referred to as interim assessment systems) which use either web- or server-based item banks that teachers use to provide periodic checks of student understanding. However, these checks rarely provide diagnostic feedback that teachers and students can use to address immediate deficiencies.

Pellegrino (2010) provides a non-exhaustive list of such technology-based assessment resources. These systems often explicitly promote or provide formative interventions and very detailed diagnoses of student understanding, resulting in a much deeper assessment, but this depth often occurs at the expense of breadth. That is, these systems are typically not able to assess as much of a content area as the large-scale systems.

Name	Website
DataDirector	www.achievedata.com
NWEA Measures of Academic Progress (MAP) Tests	www.nwea.org
Odyssey	http://compasslearningodyssey.com
MyAccess, Algebra Readiness Diagnostic Testing Program, Student Progress Monitoring System	www.vantagelearning.com
Pearson Prosper Assessment System	http://formative.pearsonassessments.com/prosper/index.htm
Princeton Review Online	http://www.princetonreview.com/educators/instructional/assessment.asp
PLATO Assessment Solutions eduTest	http://www.plato.com/District-Solutions/Assessment-and-Datamanagement.aspx
Scantron Achievement Series	www.scantron.com
Harcourt Assessment Learnia	www.harcourtassessment.com
McGraw Hill Yearly Progress Pro	www.mhdigitallearning.com
Acuity	www.acuityforschool.com
DIBELS (Dynamic Indicators of Basic Early Literacy Skills)	http://dibels.uoregon.edu
Summary Street	http://www.pearsonkt.com

IMMEX (Interactive Multimedia Exercises)	www.immex.ucla.edu
DIAGNOSER	http://www.carnegielearning.com
Assistments	www.assistments.org

Table 1: Overview of technology-based assessment resources for schools
Source: Pellegrino (2010)

2.2.2 Learning Management Systems

The terms Virtual Learning Environment (VLE), Learning Management Systems (LMS) and Managed Learning Environment (MLE) refer to a set of interrelated learning and teaching tools for the management and facilitation of learning activities within an institution, along with the provision of content and resources required to help make the activities successful. Standard functions of VLEs are curriculum mapping, student tracking, communications tools, tutor and student support, assessment and learning delivery tools (JISC, 2006).

Learning management software packages such as Blackboard or Moodle enable teachers to carry out many learning management tasks more efficiently and effectively than more traditional face-to-face methods. A good learning management system can administer and collate needs analysis data from students; allow teachers to post course information, handouts, and other materials for students to download; enable students to submit assignments, and teachers to grade and return assignments electronically; document student achievement and archive learner portfolios containing samples of spoken and written language; administer, analyze, collate and store the results of classroom quizzes; administer, collate and present student evaluations of teachers.

Almost all learning management systems (LMSs) offer support for assessment, e.g., for the creation, execution, and evaluation of multiple choice tests (Amelung, et al., 2011). In some programs learners can select their own preferred learning styles and strategies and indicate the level of proficiency gain they would like to achieve by a certain date; the programme can recommend tasks and activities based on these choices and, if the system is linked to an online bank of materials, assemble a tailored set of learning resources for each individual learner (Nunan, 2010).

In online courses, where the content is tagged with lexical, grammatical, and functional information, the software can also perform and important diagnostic function. As students work through the materials, the programme can identify those aspects of the linguistic system where students are weak, and can suggest adjustments to the study plan to provide additional remedial practice in these areas. (Nunan, 2010)

Formative and summative assessment resources can also be integrated into Learning Management Systems. In Scotland, for example, COLA materials for colleges are a bank of more than 250 formative assessment resources, covering a wide range and level of subjects across the Scottish curriculum, which can be integrated in the college's Virtual Learning Environment.[3] Similarly, the SOLAR project makes over 600 e-Assessment resources available for colleges.[4]

2.2.3 Learner Response Systems (Clickers)

Learner response systems (LRSs) – sometimes referred to as electronic voting systems (EVS), clickers, or polling tools – allow students to simultaneously vote on a multiple choice question posed by the teacher (Looney, 2010; Means & Rochelle, 2010). Some devices also accept free text or numeric answers (Looney, 2010). Responses are aggregated and displayed on the teachers' computer in the form of bar charts or graphs (Costello, 2010).

3 http://www.rsc-sw-scotland.ac.uk/eAssessment/eAssessment.htm
4 http://www.sqa.org.uk/sqa/8165.html#center; http://www.rsc-sw-scotland.ac.uk/eAssessment/eAssessment.htm; http://www.sqasolar.org.uk/mini/27322.html

Research suggests that despite the limiting multiple-choice format, test questions can be designed in such a way as to trigger subsequent deep learning without direct teaching input (Draper, 2009). There is a broad range of practical uses for these systems, from testing the understanding of science concepts (from primary aged school children up to physics undergraduates), to undertaking evaluation of events as well as public participation in data collection for research on attitudes to careers (Moss & Crowley, 2011).

Evidence regarding the impact of polling technologies on formative assessment practice is still limited. As with all educational technology, whether learning benefits are achieved depends not on the technology but on whether an improved teaching method is introduced with it (Beatty & Gerace, 2009; Draper, 2009). Yarnall, Shechtman et al. (2006) found that polling devices tended to reinforce teachers' existing approaches to assessment and research by Beatty and Garace (2009) suggests that teachers' pedagogical skills, including skills for classroom management, are the single most important determinant of success.

2.2.4 Tools for Collaboration

A range of digital applications are available that enable interaction, collaboration and sharing between users and allow for collaboartive projects to be implemented in school education. In particular, social networking applications, wikis and blogs are increasingly being used in primary, secondary and higher education and have become standard options integrated in Learning management systems, such as Moodle or Elgg.

Social networking services can be broadly defined as internet- or mobile-device-based social spaces designed to facilitate communication, collaboration and content sharing across networks of contacts (Childnet International, 2008). They enable users to connect to friends and colleagues, to send mails and instant messages, to blog, to meet new people and to post personal information profiles, which may comprise blogs, photos, videos, images, audio content (OECD, 2007; Pascu, 2008). Prominent examples of social networking services include Facebook and MySpace (for social networking/socialising), LinkedIn (for professional networking), and Elgg (for knowledge sharing and learning).

"Weblogs" or "blogs", a term coined by Jorn Barger in 1997, are online public writing environments, which enable a single author or a group of authors to write and publicly display articles, called posts, which are listed in reversed chronological order (Anderson, 2007). Depending on the author's wishes, blogs can include visual, audio and video content, as well as features such as links to other blogs, information about the author, and comments from readers (Ellison & Wu, 2008; OECD, 2007). Children and young people are increasingly becoming authors of blogs (Owen, Grant, Sayers, & Facer, 2006) and a 2006 survey in the UK found that about half of the responding educational institutions reported using blogs (Open Source Software Watch, 2006). There are blog sites, like Edublogs, that offer free blogs aimed specifically for pupils and teachers.

Wikis. A wiki is a website that allows users to collaboratively add, remove and otherwise edit and change content, usually text (OECD, 2007). The most prominent example of a wiki is Wikipedia, a collaboratively created online encyclopaedia. Wikis have become very popular environments for collaborative projects in formal education and training. Research indicates that wikis can promote effective collaborative learning and confidence in formative self and peer assessment by facilitating rapid feedback, vicarious learning through observing others' contributions and easy navigation and tracking facilities (F. Su & Beaumont, 2010).

2.2.5 Educational Games

Game-based learning has grown in recent years as research continues to demonstrate its effectiveness for learning for students of all ages. Games for education span the range from single-player or small-group card and board games all the way to massively multiplayer online games and alternate reality games (L. Johnson, et al., 2011). Those at the first end of the spectrum are easy to integrate with coursework, and in many institutions they are already an option (L. Johnson, et al., 2011). For example, a range of software applications has been developed to foster emergent literacy in preschool children through word and alphabet games.

Research on the effectiveness of such products has produced mixed results and suggests that the child's entering skill level, the nature of the non-ICT-based instruction offered, and the phonetic regularity of the child's language may be factors in explaining why some students find and advantage of software use while others do not (Means & Rochelle, 2010).

2.2.6 ICT tools for formative assessment

In higher education, in particular, e-assessment approaches are currently being developed which allow for the automatic generation of formative assessments to support self-regulated learning (Al-Smadi & Guetl, 2011). Formative online quizzes, for example, have been shown to help students' awareness of their own ability, ideally leading to independent self study (Campbell & Gorra, 2009).

Assessment packages for Learning Management Systems such as Moodle are currently being developped, which allow for the integration of self-assessment, peer-assessment and summative assessment, based on the automatic analysis of learner data, and provides learners with recommendations for further training and resources (Florian & Fabregat, 2011; Florián G, Baldiris, Fabregat, & De La Hoz Manotas, 2010).

Considering that primary and secondary classrooms are increasingly characterized by the diversity of learner backgrounds and individual learning needs, these are promising developments for improving personalisation and supporting formative assessment. Self-assessment tools, embedded assessment based on data-mining techniques and intelligent tutoring systems which provide automated feedback can thus support teachers in diagnosing, assessing and responding to individual learners' needs.

2.3 ePortfolios

Electronic portfolios, or e-portfolios, are digital environments which allow learners to gather, and showcase digital artefacts as a proof of their learning process and achievements. (Büyükduman & Şirin, 2010; H. T. D. Huang & Hung, 2010; JISC, 2006). Thus, e-Portfolios are electronic collections of users' achievements which permit an assessment of their competences (Eurydice, 2011a). Since ePortfolios can contribute to fostering students' cooperation and collaboration in self-organized project-groups (Sporer, Steinle, & Metscher, 2010), they are perceived as a promising way of reconciling collaborative and constructive learning approaches with the assessment of individual progress and performance.

Although portfolio assessment is not new and has been used for some time without ICT, the use of digital tools has developed this type of assessment further by allowing for the integration of multimedia formats. In web-based instructional system assessment based on e-portfolios is currently the main method used to assess students' performance (Ni & Liu, 2009).

As a tool for formative assessment digital portfolios make it easier for teachers to keep track of documents, follow students' progress, and comment on students' assignments (Binkley, et al., 2012), while at the same time allowing for self- and peer-assessment to be documented. Portfolio assessment is, in general, perceived as a learner-empowering alternative to computer-based testing (Cummins & Davesne, 2009).

e-Portfolios are widely used all across Europe, in particular in higher education (Dysthe & Engelsen, 2011). Dysthe and colleagues (2007) analysed how portfolios were used and assessed in higher education in Norway. They found that, as a general tendency 'soft' disciplines had more reflection based and varied portfolio models and made more recourse to peer assessment than the 'hard' disciplines (maths, sciences and engineering).

In primary and secondary schools e-Portfolios are receiving increasing attention and importance across Europe. E-Portfolios have already been implemented in school education in Belgium, Austria, Portugal, Romania, UK and Turkey; while Bulgaria, Germany, France and Iceland are in the pilot phase and eight countries in the planning. In Portugal and the United Kingdom e-Portfolios are already available to students throughout their entire educational career and are assessed by awarding bodies in England, Wales and Northern Ireland. In contrast, Poland and Liechtenstein are focusing more on providing teachers with ICT tools to monitor pupil progress. (Eurydice, 2011a)

In most cases, e-portfolios are employed to record individual progress. They can, however, also be used to support student collaboration. Liu (2007), for example, set up a personal and group-based learning portfolio system, in which teachers allocate students in groups which work together on an assignment and evaluate other groups' work. This system can include multiple courses as well as records of each student's previous learning, thus enabling a more complete learning portfolio. In a study with university students Barbera (2009) interconnected students' e-portfolios in a unique netfolio such that each student assesses their peers' work and at the same time is being assessed. This process creates a chain of co-evaluators, facilitating a mutual and progressive improvement process. Similarly, Garrett and colleages (2009) linked (university) students' individual portfolios in a social network, encouraging them to share their work, with positive effect on student motivation and performance.

2.3.1 Research Findings

On the whole, the research literature on e-portfolios is rather sparse, and the claims made for the likely benefits of e-portfolios exceed the evidence available (Ridgway & McCusker, 2008). The efficacy of different forms of e-portfolios needs to be further studied and evaluated.

Benefits and drawbacks

Due to the vast variety of e-portfolios and the range of learning and assessment purposes, the success and effectiveness of e-portfolio assessment depends on many and diverse factors applying to each individual case. In a case of e-portfolios use in higher education, Segers and colleagues (2008) found, for example, that not the assessment design, but rather how it is implemented, and how students' learning is stimulated through the assessment, influences the quality of learning.

Üstünel and Deren (2010) investigated the use of e-portfolios among a group of primary school students, learning English in Turkey. Their results indicate that, while the e-portfolio work did not influence students' attitudes towards learning, it did improve their attitude towards exams. Romova and Andrew (2011) used portfolios collecting different successive drafts of written assignments in an academic writing university course. Their findings suggest that a multi-draft portfolio is an effective assessment tool, because it provides a feedback loop and enhances learners' understanding of writing as a recursive process.

Peacock and colleagues (2010) interviewed 23 tutors in a range of subject areas, from Scottish further and higher education on their experiences with e-portfolios. Tutors pointed out that e-portfolios could encourage personal development and a more reflective approach to studies; assist student transition; and, in some cases, support assessment. Concerns were raised, however, relating to moving paper-based assessed portfolios online, the legal issues of implementing an ePortfolio and the technical robustness and flexibility of systems.

e-Portfolios and performance

There is some evidence that portfolios increase student understanding and performance. Burks (2010) found that e-portfolios used in an undergraduate mathematics course lead to increased student performance. Interviewing (US) primary school students, their parents and teachers, McLeod and S. Vasinda (2009) found that all parties attributed subjective satisfaction to the portfolio process; that students developed deep-thinking skills and that teachers obtained valuable insights into students' thoughts. Similarly, Ocak. and Ulu (2009) investigated the opinion of more than 300 5- and 8-grade students, 37 teachers and 92 parents. Their findings indicate that all three groups positively agreed with using portfolio in learning and they all believed that the use of portfolio plays prominent roles in the assessment of students' progress.

Chang and Tseng (2009) compared the use of e-portfolio assessment among a group of juniour high school students with conventional assessment. Their experimental results indicate that e-portfolios have significant positive influence on students' performance, in particular as concerns reflection, self-assessment, continuous improvement, goal setting, problem solving, data gathering, work and peer interaction.

Kim and Olaciregui (2008) tested an e-portfolio system in a fifth-grade science class. The student-constructed science portfolio was a result of a collection of digital artefacts such as graphic images, instructional videos and textual files on terms and definitions relevant to the Earth's atmosphere. They found that the students who had followed the e-portfolio approach scored significantly higher than the control group, both in the information-processing performance test and in the 3-day delayed memory retention tests.

Marking and grading e-portfolios

E-portfolio assessment poses a number of difficuluties for assessors, in particular, the workload caused by the complexity of portfolios and a lack of commonly agreed appraisal criteria or standards (Tillema & Smith, 2007; Tisani, 2008).

There are attempts to use ICT to facilitate the (formative and summative) assessment load e-portfolios pose for assessors. Chen and Chen (2009) developed a mobile formative assessment tool using data mining, which helps teachers to precisely assess the learning performance of individual learner utilizing only the learning portfolios in a web-based learning environment. While experimental results are very promising, this approach clearly needs to be further developed.

Another strategy aimed at reducing the assessment load is the method of "comparative pairs" (Newhouse, 2011) or "Thurstone's graded pairs" (Ripley, 2009), in which assessors are presented with two students' portfolios at a time, and are asked to make a judgement about which one is better. Once the first judgement has been made, the marker is presented with a second pairing, then a third, and so on. No other scoring or grading is made. Newhouse (2011) compared the traditional analytical method of marking e-portfolios with the comparative pairs method in a study of 115 student portfolios and found that both methods of marking provided highly reliable scores, with the comparative pairs method being more reliable. Findings from the eSCAPE project similarly suggests that this method of scoring exhibits rates of reliability equal to, or slightly in excess of, the levels of reliability achieved on multiple-choice tests (Binkley, et al., 2012).

Example 2: e-scape portfolio assessment

E-scape is a UK research & development project for the creation of a system in which school learners use hand-held digital tools in the classroom to create real-time web-portfolios. There are two principal innovations in the e-scape system. First, the hand-held tools used by students are linked dynamically to their teachers' laptop – operating as a local server. This sends a series of tasks to the learners and 'hoovers-up' anything that they produce in response to them. The local server is enabled to upload – dynamically (in real time) - all the data from a class/group into a website where learners' web-portfolios emerge. In phase 3 (which ended in 2009) learners from 19 schools have created 350 design & technology, 60 science and 60 geography portfolios.

Second, a web-based assessment system has been founded on a 'Thurstone pairs' model of comparative assessment. The web-based portfolios can readily be distributed anywhere at anytime – enabling multiple judges to scrutinise the portfolios simultaneously. The judging in phase 3 involved 28 design & technology judges, 6 science and 6 geography judges. All the judging was completed on-line in a short timewindow and with extraordinarily high reliability (0.95).

An authoring tool allows teachers to design an activity – and modify it in the light of second thoughts or trial runs. It enables different sub-tasks to be selected for different learners – allowing teachers to personalise the activities to particular learners and their needs. Equally however for assessment activities, examination bodies can ensure that exactly the same activity is presented to all learners in all test schools.

Once the activity has been designed, it can be transferred into the exam management system (EMS) which teachers use to manage the activity in the classroom. At the start of the activity, the teacher activates sub-task 1 and this is sent to learners' devices. They work on it for the designated period (writing, drawing, taking photos etc), at which point the teacher's EMS 'hoovers-up' all their work back into the EMS and sends the second sub-task to learners. The closed network guarantees good data transmission between the teacher's and learners' devices. Throughout the activity, the EMS enables the teachers to check that all the learners' devices are connected and operating properly – and gives a simple visual check on the battery state of each device. At the end of the activity, the teacher uploads all the data from the EMS into the portfolio management system.

A "pairs engine" manages the grading process following the Thurstone method of graded pairs.

Source: http://www.gold.ac.uk/media/e-scape_phase3_report.pdf.

2.4 Peer assessment

Peer assessment is an educational arrangement where students judge a peer's product or performance quantitatively and/or qualitatively and which stimulates students to reflect, discuss and collaborate (Strijbos & Sluijsmans, 2010; Topping, 2009). Products to be assessed include writing, oral presentations, portfolios, test performance, or other skilled behaviors (Topping, 2009). Due to the increased use of peer assessment over the past decade, the term peer assessment now covers a multitude of sometimes incompatible practices (Gielen, Dochy, & Onghena, 2011). Nonetheless, research on peer assessment is currently still "in a stage of adolescence, grappling with the developmental tasks of identity formation and affiliation" (Kollar & Fischer, 2010). Empirical evidence for peer assessment effects on learning is scarce, mostly based on student self-reports or involving comparison of peers' and teachers' ratings or anecdotal evidence from case studies (Strijbos & Sluijsmans, 2010).

Although peer-assessment can support summative assessment (for examples see: Kaufman & Schunn, 2010), in the vast majority of cases, peer assessment is used for formative purposes, e.g. to encourage students to help each other plan their learning, identify their strengths and weaknesses, target areas for remedial action, and develop metacognitive and other personal and professional skills (Topping, 2009).

While peer assessment as a pedagogical strategy is not linked to the use of ICT, the use of electronic leaning environments and web-based interfaces lever the power of peer assessment in providing effective and timely feedback for more complex assignment tasks, which exert a high workload on teachers and assessors and can (not yet) adequately be responded to by computer programs (Paré & Joordens, 2008). The advent of dedicated web-based tools allows the frequent and efficient implementation of self and peer assessment activities even in large classes (Willey & Gardner, 2010). Wikis and other collaborative tools and environments Have been shown to enhance peer assessment effectiveness and quality (Xiao & Lucking, 2008).

2.4.1 Benefits of peer assessment

The literature suggests that peer assessment contributes to the development of student learning and promotes ownership of assessment processes (Bryant & Carless, 2010). It is considered a powerful technique to engage students in active learning, and to make them reflect about their own work (Dziedzic, Janissek, & Bender, 2008). Peer assessment has been shown to change the way students perceived their own abilities and their potential to make improvements in their work, encourage critical reflection; help develop skills of autonomous learning; and raise motivation and academic standards (McMahon, 2010). It has furthermore been argued that using a peer assessment strategy in the classroom could facilitate learners' critical thinking, meta-cognitive skills and deep-thinking (Hou, Chang, & Sung, 2007; Sitthiworachart & Joy, 2008). Peer reviewing can furthermore help build a stronger learning community (Søndergaard, 2009).

Moreover, peer feedback is often considered as a means of reducing teachers' assessment workload and, possibly, improving learning quality (Bouzidi & Jaillet, 2009; Paré & Joordens, 2008).

2.4.2 Effectiveness and validity of peer assessment

Findings on the effectiveness and validity of peer-assessment are mixed. Some research studies indicate that peer-assessment does not significantly enhance students' performance (Chang & Tseng, 2009) and that there is a lack of consistency across various student raters and with respect to teacher-assessment scores and end-of-course examination scores (Chang, Tseng, Chou, & Chen, 2011; C. h. Chen, 2010). These findings seem to imply that e-portfolio peer assessment is not a reliable indicator of performance.

However, other studies underline that peer assessment can be of equal reliability and validity to the assessment of a teacher: Tsivitanidou and colleagues (2011) studied the peer assessment skills of a group of 7th graders who were anonymously assigned to reciprocally assess their peers' science web-portfolios. Students were found to be able to define and use their own assessment criteria, whose overall validity and reliability, however, were found to be low. In an online peer assessment experiment in which 242 (university) students, enrolled in 3 different courses, took part, Bouzidi and Jaillet (2009) showed that peer assessment is equivalent to the assessment carried out by the professor, at least when simple calculations and basic scientific arguments are concerned.

A study by Liang and Tsai (2010) showed that while self-assessment scores were not consistent with the expert's scores, peer assessment scores demonstrated adequate validity with the expert's evaluation. The validity of the peer assessment increased with the number of peer assessment rounds. Similarly Saito (2008) found in two research trials that peer assessment is a robust system in which instruction on skill aspects may suffice to achieve a certain level of correlation with the teacher's judgement.

In a study among 240 undergraduate university students who gave and received peer feedback via a web-based learning environment, Paré and Joordens (2008) found that peers produced grades similar in level and rank order as those provided by expert graders, especially when students were made accountable for the grade awarded. Similarly, Sitthiworachart (2008) found that computer-mediated peer assessment to be an accurate assessment method in a programming course, a finding that Tseng and Tsai (2007) confirm for online peer assessment among 10th graders in a computer course.

2.4.3 Student perceptions

Findings on students' perceptions and attitudes towards peer assessment are equally divided. Several studies, mostly conducted in the context of higher education indicate that that students value the peer review activity, take peer reviews seriously and provide comprehensive and constructive reviews (Bauer, Figl, Derntl, Beran, & Kabicher, 2009; Bloxham & West, 2007). Additionally, students highly value the fact that using an online system for peer assessment allows their judgements to remain anonymous (D. Thompson & McGregor, 2009).

However, other studies suggest that students are concerned about the validity and objectiveness of their peers' judgements and question their peers' assessment skills (C. h. Chen, 2010; Kaufman & Schunn, 2010). Research indicates that students prefer written online reviews with the possibility of oral follow-up questions to reviewers (Bauer, et al., 2009). Tseng and Tsai (2007) found that different kinds of feedback had different effects: Reinforcing peer feedback and in some instances suggestive feedback was useful in helping students' improve their projects, while didactic and corrective feedback did not appear not to be favorable for improving students' work.

2.4.4 Performance gains

There are strong indications that peer assessment can contribute to enhancing students' understanding and performance. Gielen et al. (2010) found that for writing assignments of Grade 7 students in secondary education, receiving 'justified' comments in peer feedback improved performance, but this effect diminishes for students with better pre-test performance. A research study on online peer assessment with 10th graders in a computer course indicates that students significantly improved their projects based on the peer assessment activities (Tseng & Tsai, 2007). Kelly et al. (2010) found that a peer assessment, integrated in a structured, collaborative learning environment for undergraduate psychology students, facilitated active student engagement throughout the academic year, and was associated with improved marks in the final written exam.

Li et al. (2010) showed that the quality of peer-feedback received influences the quality of the final, revised product, when compared to the quality of the initial project. Online peer assessment of writing assignments can gradually improve the quality of student work with respect to coverage, richness and organization (Liang & Tsai, 2010). Similarly, Liang and Tsai (2010) found that in a university writing class, students significantly improved their writing skills as the peer assessment activity proceeded, in particular as concerns the coverage, richness and organization of their assignments.

Furthermore, peer and self assessment can contribute to the development of self-monitoring skills, which, in turn, can significantly improve students performance in the subject at hand (Cho, Cho, & Hacker, 2010).

2.4.5 Factors for success

The effectiveness of peer-assessment is contingent of a number of different factors. Cartney (2010) highlights the importance of appreciating the emotional as well as the cognitive aspects of peer learning. Bryant and Carless (2010) underline that assessment practices are deeply cultural and, in test-dominated settings, peer assessment may have most potential when explicit links are drawn with preparation for summative assessment. Moreover, research indicates that students who receive feedback from multiple peers improve their writing quality more than students receiving feedback from a single expert (Cho & MacArthur, 2010).

McMahon (2010) reports on a four-year project among undergraduate university students which succeeded in establishing a culture of providing immediate, reflective and useful peer feedback. The two key factors identified were the openly and exclusively formative character of peer assessment and the ownership of all data generated by the process in the student being assessed. While Saito (Saito, 2008)found that training only slightly improved peer assessment, an analysis of the literature, as conducted by van Zundert et al. (2010), reveals that training significantly enhances peer assessment quality and skills.

2.4.6 Efficiency gains

Peer assessment is also conceived as a means of increasing assessment efficiency for product and performance types, such as portfolios, essays or scientific writing exercises (Paré & Joordens, 2008). Lai (2010) compared the effectiveness of automated writing evaluation and of peer evaluation for a group of English as a foreign language learners in Taiwan and found that participants preferred peer feedback over automated feedback. More recent research suggests that peer review and automated grading evolve as complementary strategies. For example, when implemented in an electronic environment, the quality of the student's marking and commenting of their peers' work can be automatically graded (Davies, 2009).

Example 3: Personalisation by Pieces

Personalisation by Pieces (PbyP www.pbyp.co.uk) is an innovative method for assessing key competencies using ICT as the organizing platform. PbyP has concentrated on setting up continuous formative assessment systems that are based on online portfolios of evidence and uses peer assessment. These systems require the learners to actually use competencies 5-8 during the learning process as well as gathering evidence of progress made in these competencies. The process of this initiative is as follows:

Competences are broken down into single identifiable skills. Then 9 progressively more challenging steps are defined in a 'Skill ladder' ending in level 9 which is defined as the best example of this skill you have ever seen displayed by a professional adult. In total, 24 skills were defined that are the constituent skills of key competencies 5-8, and for each of those 9 levels of progression. These 'Ladders' were then incorporated into a website creating a matrix of 24 x 9 boxes. If a learner clicks on one of these boxes then they can view not only the statement describing it but, more importantly, hundreds of examples of that skill at that level in operation from evidence provided by thousands of children across different countries. Although the examples that the learners can see are chosen to illustrate one particular skill at one particular level, the majority of the examples are multi-disciplinary and integrate numerous skills, showing their interconnectivity as well as the diversity of contexts in which they apply. Learners can rate the examples they find most useful and the ideas for evidencing them that are most original. This results in a dynamic interpretation of contexts in which these competencies are used.

After viewing examples of work, learners can upload their own evidence of achievement in a particular skill at a particular level. Uploaded evidence goes to learners not in the same school who have already achieved the skill level in question at least twice and are therefore knowledgeable enough to give the learner some advice on how to improve as well as praise for what was best. This peer assessor can award a 'Pass'. If a learner receives two passes in a skill level then they become an assessor for this skill level themselves and will begin receiving uploaded evidence from learners in other schools. Students gain certificates as they progress and these have been authenticated by numerous assessors inside and outside the school.

The latest release of PbyP contains an action research tool for teachers so that they can agree on a ladder of progression as a staff and then try out ideas collaboratively for how they can embed competencies into the curriculum. The tool collects a snapshot of the evidence in learner portfolios before and after the teacher conducts their project. It also asks for feedback via learner questionnaire. All the analysis data collected before and after the teacher project is analysed on the site to let the teacher know if the lesson they gave actually did generate progress in key competencies or not. The teacher can then choose to submit their work as an exemplar for other teachers or not. If a teacher submits their work as an example, the analysis data goes with it so that other teachers can search for projects that the learners reported as having been useful rather than just rely upon the teacher's own perception. They can also see the impact in terms of evidence uploaded and peer assessed as a result of the project.

2.5 Embedded Assessment

2.5.1 Learning Analytics

Learning analytics – expected to be mature by 2015/16 – refers to the interpretation of a wide range of data produced by and gathered on behalf of students concerning their engagement, performance, and progress, in order to assess academic progress, predict future performance, and revise curricula and teaching strategies (L. Johnson, et al., 2011). Similarly, embedded assessment refers to a situation in which students engage in learning activities and in performative tasks as part of the normal pattern of learning and instruction, while an assessment system draws conclusions about their competencies based on what they do (Ridgway & McCusker, 2008). For example, spyware would be loaded (with user knowledge) onto users' computers, and would track the patterns of activity to investigate learners' internet (re-)search strategies. Data analysis could be used to provide feedback to the user to improve their search strategies, and to identify areas for their future development. (Ridgway & McCusker, 2008)

Data-mining techniques are already used to evaluate university students' activity patterns in Virtual Learning Environments for diagnostic purposes. Analytical data mining can, for example identify students who are at risk of dropping out or underperforming[5]; generate diagnostic and performance reports[6]; assess interaction patterns between students on collaborative tasks[7] and visualise collaborative knowledge work[8]. Researchers are aware of the fact that, although it is possible in theory to capture every observable action in the "click stream" of interactive behaviours in an electronic environment, there is a need to establish which of those actions are cognitively informative and to understand which aspects of the data provide the greatest informational value (ETS, 2012).

2.5.2 Intelligent Tutoring Systems and Automated Feedback

Research indicates that the timing of the feedback is critical. The closer the feedback is to the actual performance, the more powerful is its impact on subsequent performance as well as learner motivation (Nunan, 2010). Online instructional programs can provide instant feedback, either directly, through automated feedback, or through electronic tutors which are used as scaffold the learning process needs (Looney, 2010). Tuomi (2006) has coined the term "pedagogical veils" to refer to applications such as electronic tutors that guide learners and scaffold their learning process.

Most programs can not only tell students which answers are correct and which wrong, but also provide qualitative information on why particular responses are incorrect (Nunan, 2010). Although in some cases this feedback may be fairly generic, some programmes search for patterns in student work to better target feedback and to then adjust the level of difficulty in subsequent exercises according to needs (Looney, 2010). Some programs such as AutoTutor are designed to promote, among other, self-regulated learning (Graesser, 2009) and meta-cognition (Sullins, Jeon, D'Mello, & Graesser, 2009).

Huang et al. and Wang (2011; 2010), for example, developed an intelligent argumentation assessment system for elementary school pupils based on machine learning techniques. The system analyses the structure of students' scientific arguments posted on a moodle discussion board and issues feedback in case of bias. In a first trial the system was shown to be effective in

---

5 For example: the Signals system at Purdue University, http://www.itap.purdue.edu/tlt/signals/; the Academic Early Alert and Retention System at Northerm Arizona University, http://www4.nau.edu/ua/GPS/student/.

6 Cf. http://www.socrato.com/.

7 Cf. http://research.uow.edu.au/learningnetworks/ seeing/snapp/index.html.

8 http://emergingmediainitiative.com/project/ learning-analytics/

classifying and improving student's argumentation level and assisting them in learning the core concepts taught at a natural science course on the elementary school level.

Currently, intelligent tutoring systems such as AutoTutor[9] and GnuTutor[10] (cf. Olney, 2009) are being developed which teach students by holding a conversation in natural language. There are versions of AutoTutor that guide interactive simulation in 3D micro-worlds, that detect and produce emotions, and that are embedded in games (Graesser, 2009). In its latest version, AutoTutor has been enabled to detect learners' boredom, confusion, and frustration by monitoring conversational cues, gross body language, and facial features (D'Mello, Craig, Fike, & Graesser, 2009; D'Mello, Dowell, & Graesser, 2009). GnuTutor, a simplified open source variety of AutoTutor intends to create a freely available, open source ITS platform that can be used by schools and researchers (Olney, 2009).

In school education, intelligent tutors and educational software tools with feedback functions are mainly used in mathematics instruction. In the US, Intelligent Tutoring Systems (ITSs) are being widely used in the US. The most popular system "Cognitive Tutors" is used by half a million students in around 2600 US middle and high schools (Ritter, Towle, Murray, Hausmann, & Connelly, 2010). In Norway, the Kikora Software for mathematics has been rather successful.

Example 4: The Norwegian Kikora Software for Mathematical Competence

Kikora is both a Norwegian software company as well as a piece of educational software for mathematics. Kikora, the company, was founded in 2005, and the development of the software was facilitated through a development grant from the Norwegian Ministry of Education and Research in 2006-2007.

The core of the Kikora software is that it contains a feedback function that gives the learner line by line feedback in all calculations. The software also contains a reporting tool that enables the teachers to track the progress and critical learning areas of the students, which in turn facilitates mapping and planning of 1 to 1 activities based on the individual student profiles. The program also contains an item bank with approximately 3000 tasks for students in lower and upper secondary education. The selection of the items pays attention to the need for items that will support the progress of learners with mathematics challenges as well as learners who perform well in mathematics. The content is strictly anchored in the requirements of the mathematics curriculum in Norwegian education.

Seen from a national/Norwegian point of view, the relative success of Kikora is also a reminder that high quality educational software is not always readily available. The bi-annual Monitors of ICT use in Norwegian schools clearly indicate this. Kikora's success is linked to the fact that the software can give learners line-by-line instant feedback, thus combining assessment with tutoring. According to the company, this approach to feedback loops and the reporting system for teachers for group or class monitoring stimulate deeper learning among the students. Examples such as the Kikora software may act as an illustrative example of how educational software can strengthen formative assessment in mathematics education, which is a critical issue given the motivational issues learners especially in lower secondary education are struggling with.

---

9 http://www.autotutor.org/.

10 http://gnututor.com/.

2.6 Augmented reality, virtual worlds, immersive games

Immersive environments and games are specifically suitable for acquiring 21st century skills such as problem-solving, collaboration and inquiry, because they are based on the fact that the knowledge that needs to be acquired is not presented explicitly to learners but must be inferred from the environment (de Jong, 2010). By making the context of learning similar to the contexts within which students will apply their learning, educators expect to promote inquiry skills while also making learning activities more motivating and increasing the likelihood that acquired skills and dispositions will transfer to real-world situations (Means & Rochelle, 2010). In particular, immersive game-based learning environments provide a powerful new form of curriculum for teaching and learning science, leading to significantly better learning results than traditional learning approaches (Barab, et al., 2009).

Assessment can be integrated in the learning process withing virtual environments and games. Within a virtual world like Second Life, for example, educational material can be delivered and assessment can be administered, using a mixture of multiple choice and environment interaction questions, while respecting and encouraging an exploratory attitude to learning (Bloomfield & Livingstone, 2009).

2.6.1 Augmented Reality

Augmented reality refers to the layering of information over a view or representation of the normal world, offering users the ability to access place-based information in ways that are compellingly intuitive. Augmented reality is considered a promising technology for enhancing education, because it can be used for visual and highly interactive forms of learning; it responds to user input; and allows for dynamic processes, extensive datasets, and objects too large or too small to be manipulated to be brought into a student's personal space at a scale and in a form easy to understand and work with (L. Johnson, et al., 2011). Augmented reality can be combined with games designed and used with handheld devices to promote environmental and science education in secondary schools (Klopfer & Squire, 2008; Squire & Klopfer, 2007).

Johnson et al. (2011) give a number of examples for enhancing school education by using augmented reality: In chemistry, for example, using handheld devices, students can explore a physical space to uncover clues and receive data related to a simulated environmental disaster detailed in a game-based scenario using AR simulations. In geography, students can study an augmented globe in a textbook, and gain both a better representation of the cartographic information and greater options for interaction and comprehension. In history, visiting actual locations tagged with information, students view images and information from the past in situ, enhancing their comprehension.

Museums, in particular, are exploiting the potential of augmented reality to make history and art more accessible and interesting for a young audience. London's Natural History Museum, for example, gives visitors handheld screens featuring an interactive video that allows users to learn about the evolution of dinosaurs, which are seen in the video moving around the actual space of the museum. The Museum of London, follows a different approach and released a free iPhone application called StreetMuseum that uses GPS positioning and geo-tagging to allow users as they travel around the city of London to view information and 3D historical images overlaid on contemporary buildings and sites. Similarly, the iTacitus project in allowed users to visit historical locations, such as the Coliseum, with their mobile device, and witness an event from the past, and, in Australia the Powerhouse Museum developed an augmented reality application that allows visitors to use their mobile phones to see Sydney, Australia, as it appeared one hundred years ago[11].

[11] http://www.powerhousemuseum.com/layar/.

In primary and secondary Education and Training, augmented reality is usually integrated in a game environment. In the Radford Outdoor Augmented Reality (ROAR) Project[12], for example, AR is used to help teach primary and secondary students more about Native American history and teamwork through a game called Buffalo Hunt.

2.6.2 Virtual Worlds

Virtual environments, like Second Life[13], or similar online 3D virtual worlds, such as Active Worlds[14], Entropia Universe[15], and Dotsoul Cyberpark[16] provide users with a online game-like 3D digital environment to which users subscribe (OECD, 2007). The user is represented by an avatar, i.e. the interactive representation of a human figure in a three-dimensional interactive graphical environment (De Freitas, 2007).Users can build, display, and store virtual creations, as well as host events and businesses or real university courses (OECD, 2007).

In March, 2007, more than 250 universities, 2500 educators and the New Media Consortium, with over 225 member universities, museums and research centres, had a presence in Second Life (Calongne, 2007). A survey among 209 educators using Second Life, conducted by the New Media Consoritum (NMC) in early 2007, indicates the manifold uses of 3D environments for educational purposes (NMC, 2007): 60% of educators took (43%) or are planning to take (17%) a class in Second Life; 58 % taught (29%) or are planning to teach (28%) a class in Second Life. Learning/teaching related activities include: supervising class projects and/or activities; conducting research in SL; class meetings; virtual office hours; mentoring student research projects; student services and support activities . Asked about the potential of Second Life for education, a majority of respondents see a significant or high potential for role-playing (94%), simulation and scenario activities (87%), artistic expression (86%), group work, collaboration and meetings (78%), distance learning programs (74%), team building (73%), conducting training (71%), professional development (68%), and teaching full courses (60%).

2.6.3 Immersive Games

Since 2003, when James Gee began to describe the impact of game play on cognitive development, research and interest in the potential of gaming on learning has exploded, as has the diversity of games themselves, with the emergence of serious games as a genre, the proliferation of gaming platforms, and the evolution of games on mobile devices (L. Johnson, et al., 2011).

Multiplayer online games are one of the most powerful forms of modern gaming allowing as they do the possibility of reliving situations and conflicts in different settings and conditions in groups (De Freitas, 2007). According to IDATE, more than 100 MMORPGs exist today worldwide.[17] Playing games online is attracting a quarter of the total worldwide Internet population; in Europe one in five web users plays online games (Pascu, 2008).

The use of online games for collaborative game play in leisure time contexts (e.g. Everquest and World of Warcraft) has increased dramatically over the last five to ten years with the growth of usage of the internet. Currently, there are over 4 million users of Everquest worldwide, 6 million users of World of Warcraft and over 7 million registered users for America's Army (De Freitas, 2007). The average online gamer visits a gaming site 9 times a month.[18] More than 10 million people are reported to have played MMOs worldwide in 2006 and the number is doubling every

12 http://gameslab.radford.edu/ROAR.html.
13 http://secondlife.com/.
14 http://www.activeworlds.com/.
15 http://www.entropiauniverse.com/.
16 http://www.dotsoul.net/.
17 IDATE DigiWorld 2007
18 comSCore July 2007 http://www.comscore.com/press/release.asp?press=1521

year.[19] As of July 2006, there are over thirteen million active subscriptions to MMOG worldwide.[20] More than a third of US adult Internet users play online games on a weekly basis, compared with 29% who watch short online videos and 19% who visit social networking sites with the same frequency (Pascu, 2008).

Serious games include popular commercial games such as Quest Atlantis, Civilization III, The Triple A Game Show, Revolution, Mad City Mystery (de Jong, 2010). In the best-selling game Civilization, for example, players have the opportunity to relive the development of global social and economic history (Collins & Halverson, 2010b). The creation of massively multiplayer online (MMO) games especially designed for learning is projected to become a reality in about five years' time (L. Johnson, et al., 2011).

2.6.4 Simulations

Simulations present students with an authentic scenario that, ideally, requires them to behave as they would in the real world, testing their theories and applying their knowledge to complete complex tasks. Data gathered in simulation-based assessments can measure not just the correctness of one final answer, but multiple aspects of the student's ability to apply skills to solve problems; for example, efficient use of information and tools, systems thinking, as well as accuracy in decisions made at each step (ETS, 2012).

Currently virtual laboratories and simulations are primarily used (if at all) as learning, rather than testing environments. As such they provide an important tool for a more realistic and problem-oriented way of assessing scientific competence. Recent advances in simulation, data collection, and data analysis are making it possible to gain insights into learners' thinking from their progress through the tasks in a simulation and thus use behaviour deisplayed in a virtual environment as a basis for formative and summative assessment. By using simulations and virtual laboratories the whole scientific process of conducting and evaluating experiments can be integrated in the assessment task. Students' skills and attitudes, e.g. their diligence in planning and conducting experiments and generating data, as well as their interpretation, analysis and critical skills, can become part of the assessment process. While these skills can also be assessed using real-life experiments as assessment tasks, virtual environments increase the variety and scope of experiments that can realistically be conducted by students themselves, with limited time and resources.

Research on the use of games, virtual laboratories and simulations in school education is still in its beginning. Evidence on whether these tools increase performance on traditional tests is divided. In a two year study with fifth grade students on the WISE learning environment, which supports students in collecting, synthesizing and analyzing information from a variety of authentic sources and simulations, in this on plant growth and development, it was shown that, overall, students made significant gains in understanding standards-based science concepts (Williams & Linn, 2002). Yet, a study on "Astronomy Village", a software program designed to engage secondary science students in authentic and inquiry-based learning over core topics in astronomy, indicates that students' on a standards-oriented test did not improve. However, students' performance did improved with respect to the (inquiry) curriculum-oriented exam (Taasoobshirazi, Zuiker, Anderson, & Hickey, 2006).

These examples illustrate that ICT provides powerful tools for the assessment of more complex and applied skills, such as scientific inquiry, analysis, interpretation and reflection. While this potential has not yet been exploited fully for assessment purposes in schools, experimentation in the school context indicates that simulations, virtual laboratories and multiplayer games, can promote effective learning in real-life contexts and support more adequate and applied assessment strategies, in particular as concerns formative assessment.

---

[19] IDATE Digiworld 2007.

[20] www.mmogchart.com.

Example 5: EcoMUVE & EcoMobile

EcoMUVE and EcoMOBILE are two curriculum research projects at Harvard university to teach ecological concepts using an immersive virtual environment (EcoMUVE) and mobile technologies for real-life tasks (EcoMOBILE).

In EcoMUVE, the scientific investigations occur within a Multi-User Virtual Environments (MUVE). which recreates authentic ecological settings within which students explore and collect information. Students work individually at their computers and collaborate in teams within the virtual world. Student's avatar works collaboratively with the avatars of others and computerized agents to figure out, e.g., what has caused fish in a local pond to suddenly die. EcoMUVE includes two ecosystems science curricular modules that together take approximately ten 50-minute class periods to implement. These include two MUVEs for teaching various aspects of ecosystems science, with full technical documentation, ancillary materials, and teacher guide and training.

EcoMOBILE (Ecosystems Mobile Outdoor Blended Immersive Learning Environment) is an extension of the EcoMUVE curriculum. Students take handheld digital devices into a real field location where "hotspots" bring up visualizations, video, 3D models, and multiple choice or open-ended questions.

The mobile devices are used to access and collect information and clues; capture pictures, video, or voice recordings as evidence in solving an environmental mystery; and access special features through an Augmented Reality (AR) interface,, which provides students with information that would not otherwise be apparent in the natural environment. Furthermore, students use environmental probes that allow collection of real-time data similar to the kinds of data ecosystems scientists study. These probes will allow students to collect some of the same data (dissolved oxygen concentrations, temperature, turbidity, and pH) that they collected in the virtual environment.

Source: http://ecomuve.gse.harvard.edu/index.html.; http://ecomobile.gse.harvard.edu/; (ETS, 2012).

3. THE USE OF ICT FOR KEY COMPETENCE ASSESSMENT IN SCHOOLS

The current use of ICT for assessment in and for school education

Since the publication of the European Key Competences Recommendation (2006), almost all European countries have included these EU Key Competences in their steering documents (Eurydice, 2011a; Halász & Michel, 2011), albeit with differences in the way in which these are defined and "unpacked" (cf. Pepper, 2011). Many European countries recommend using ICT to teach these competences, without, however, indicating how (Eurydice, 2011a). As far as the use of ICT for assessment is concerned, there is a variety of ways in which approaches to assessment are recommended across Europe. Eight countries, in different areas of Europe, recommend the use of ICT in pupil assessment. Estonia, Austria, the United Kingdom and Norway recommend the use of ICT as an information source for use in traditional tests. Denmark (for primary education), Spain, Austria and Norway have central recommendations for on-screen testing, while four countries have them for interactive testing. Denmark (for primary education), Austria and Norway also recommend the use of interactive testing (Eurydice, 2011a).

According to data collected from the different European education ministries by Eurydice (2009), ICT is used in national assessment only in in the Netherlands, Denmark and Norway. While in Denmark more innovative, adaptive computer-based methods are employed, the Netherlands and Norway employ traditional forms of computer-based tests, which replicate tests previously (or alternatively) taken in paper-and-pencil format. Besides the possible use of computers for national written examinations in both the Netherlands and Norway, pupils in the latter may also use ICT in oral exams during a preparation period or for a presentation of the exam (Eurydice, 2009).

Compared to the level of ICT use on national high stakes tests, ICT enhanced assessment strategies in primary and secondary schools are far less advanced. Again, the Netherlands are an exception. Already in the 1990s the National Institute for Educational Measurement in the Netherlands developed a pupil monitoring system, accompanied by a software programme, for primary education to assess pupils' learning progress on a longitudinal basis (Vlug, 1997). Subsequently and as a consequence of national high-stakes ICT-based testing computer-based monitoring and assessment systems for schools have been developed and offered to public schools[21].

However, in general, computer programmes supporting assessment in primary and secondary educational institutions are not used on a widespread basis and there is a lack of evidence on actual classroom and school use of ICT for assessment purposes. There are some experimental trials that allow some insights on the benefits and drawbacks of different assessment strategies using ICT. Furthermore, a number of open-source and commercial products designed for use by educators allow some conclusions on how ICT are currently used in primary and secondary education for formative and summative assessment purposes.

There is, however, no reliable or comparable data that would allow a more comprehensive evaluation of the of ICT for assessment purposes in schools in Europe. Furthermore, most conventional e-assessment systems are not suited for examining analytic, creative and constructive skills (Usener, Gruttmann, Majchrzak, & Kuchen, 2010). Thus, far more research and experimentation is needed to conclude on the benefits and drawbacks of using ICT for the assessment of kex competences. The following sub-chapters should therefore be understood as a collection of ideas of how ICT could be used to support the formative and summative assessment of the skills and attitudes dimensions, in particular, of Key Competences.

[21] Cf. for example http://www.leeds.ac.uk/educol/documents/000000853.htm for primary schools in the Netherlands.

Different Key Competences – different assessment tools

	F = Formative Assessment S = Summative Assessment	CBA	Quizzes / simple games	ePortfolios	Virtual worlds & games	Simulations	Intelligent Tutors
1	Communication in the mother tongue	S	F	FS			
2	Communication in foreign languages	S		FS			
3a	Mathematical competence	S	F			FS	FS
3b	Basic competences in science and technology				FS	FS	
4	Digital competence					FS	
5	Learning to learn						
6	Social and civic competences				FS		
7	Sense of initiative and entrepreneurship				FS		
8	Cultural awareness and expression			FS	FS		

Figure 4: Overview of the potential of different ICT-based tools for the assessment of Key Competences

Looking at the different ICT-based assessment approaches as these are described in the previous chapter and considering the specificities of each of the eight – or in this respect: nine – Key Competences, some focus areas, as illustrated in Figure 2, emerge that indicate the particular potential of ICT in enhancing competence-based assessment strategies.

Computer-based testing: In particular, for the *reading and text comprehension* as well as *mathematical competence*, computer-based tests can provide contexts that are rich enough to comprehensively assess these competences. The advantages of computer-based tests over traditional assessment formats lie in the provision of instant and targeted feedback and in the possibility to automatically adapt the difficulty of the test items to learners' different performance levels.

ePortfolios are ideally suited as environments scaffolding the assessment of collections of works produced by students and are thus particularly powerful tools for communication *in the mother tongue, communication in foreign languages* and *cultural awareness and expression*. ePortfolios can furthermore support online collaborations, self- and peer assessment, which contribute to fostering and at the same time assessing students' *learning to learn* skills, as they promote self-regulated learning and improve motivation, engagement and participation.

Immersive environments and multiplayer games recreate learning situations which require complex thinking, problem-solving and collaboration strategies and thus allow for the development of these skills, which are key components of all eight Key Competences. These environments replicate authentic contexts; encourage collaboration, empathy and negotiation; and reward strategic thinking, initiative and experimentation. They are thus specifically suitable for *competences in science and technology*, for *social and civic competences* and the development of *sense of initiative and entrepreneurship*. Since learners' behaviour in these electronic environments is tracked, their individual learning journey – and with it these skills – can, at least in principle, be automatically assessed.

For *competences in science*, in particular, **computer simulations and virtual laboratories** provide opportunities for students to develop and apply skills and knowledge in more realistic contexts and provide feedback in real time. Practical tasks, embedded in electronic environments, using mobile devices or making available online resources, are a further promising avenue for developing ICT enabled assessment formats that better capture Key Competences, in particular as concerns *digital competence*.

In general, one of the most promising emerging technological trends for the comprehensive assessment of complex competences are **Learning Analytics,** i.e. the interpretation of a wide range of data produced by and gathered on behalf of students in electronic environments in order to assess progress, and tailor education to individual students more effectively. Learning Analytics are currently still in an experimental and development phase, but could become a reality within the next five years (L. Johnson, et al., 2011). Learning Analytics could allow assessment to be embedded in immersive environments, multiplayer games and computer simulations, thus levering the potential of these tools in assessing Key Competences as these are applied in real-life contexts.

Educational software solutions such as **intelligent tutoring systems** take this idea one step further by offering embedded assessment with instant feedback and targeted support. In particular for *mathematical competence* these systems allow students to investigate mathematical concepts and problems in complex contexts, at their own pace, through a series of tasks adapted in level of difficulty and with the help of hints that encourage them to develop adequate solution strategies.

In the following sections, the particular strengths and weaknesses of different ICT enhanced assessment formats are presented, illustrated and discussed. For each Key Competence, a brief analysis of the current use of ICT for assessment, the particular strengths of ICT, current bottlenecks and possible policy interventions is provided.

Example 6: innovating national assessment: US K-12 State Tests for 2014-15

In the US a new series of national tests has been commissioned, which will be rolled out in the school year 2014/15. A total of 45 states plus the District of Columbia are participating in two assessment consortia which were awarded $360 million in "Race to the Top Assessment grants" in 2010 to design, develop and pilot test new systems of summative assessments in English language arts (ELA) and mathematics for students in Grades 3-12 to replace existing state tests.

The new, fully digital assessments, which include both interim assessments and end-of-year summative assessments, are being designed to measure mastery of the Common Core State Standards. In English language these standards place strong emphasis on students' ability to read complex texts, conduct electronic searches, evaluatesources, draw evidence from them, and craft well-supportedwritten arguments. The mathematics the number of topics to be taught in a given school year is reduced to promote greater depth of understanding and mastery of core skills, particularly in the early grades. The objective is to measure individual growth as well as proficiency; assess hard-to-measure skills such as critical thinking and the application of skills to solve complex problems.

Since test developers believe that "much of what is new, different, and important in these standards cannot be adequately assessed by conventional methods," the new series of tests will allow for a richer range of items and tasks, and for the capture of student responses during activities involving research, design, and problem solving. To do this, a range of technology-enhanced assessment strategies are being combined.

In English language the assessment environment will include popup features for support (e.g. definitions for potentially unfamiliar words); automated essay-scoring programs; audio, video and search features;

hyperlink environments to simulate Internet resources; electric organisers (e.g. diagrams, tables, charts, Powerpoint); like/dislike voting features to engage students and tease out their opinions.

In mathematics, real life tasks are being replicated in an electronic environment, within an activity that seeks to model good instruction and yields useful diagnostic information. In a sample task, for example, the assessment simulates the experience of making punch and employs an interactive "sweetness meter" computer tool. Students are asked "to reveal their strategies as they work through the task" so that interventions can be made before they move to the next section of the task" and the assessment can gauge where their knowledge falls in the progression of levels of understanding.

Source: http://www.k12center.org/rsc/pdf/a-sea-change-in-assessment-letter-size.pdf.

3.1 Communication in the mother tongue

Box 1: Definition of "Communication in the mother tongue"

Communication in the mother tongue is the ability to express and interpret concepts, thoughts, feelings, facts and opinions in both oral and written form (listening, speaking, reading and writing), and to interact linguistically in an appropriate and creative way in a full range of societal and cultural contexts; in education and training, work, home and leisure.

Source: Recommendation of Key Competences for Lifelong Learning (Council of the European Union, 2006)

Computer use in the language of instruction and in foreign languages is low in most European countries. According to PISA 2009 data around 80 % of European students reported never using computers in neither of the two subject areas; only in Denmark, the Netherlands, Sweden, Liechtenstein, Norway and Turkey around 40 % or more students reported using computers in language of instruction classes on a weekly basis for up to 60 minutes or even more (Eurydice, 2011a).

This finding is all the more disappointing as ICT can contribute to increasing engagement and motivation in reading and writing and are better suited to address changing reading and writing patterns in a digital society.

A recent survey conducted in the United Kingdom (Clark, Osborne and Dugdale, 2009) shows that technology-based materials are the most frequently read, with nearly two-thirds of children and adolescents reading websites every week and half the sample also reading blogs, networking websites, and e-mails every week (Clark, Osborne, & Dugdale, 2009). Apparentlz, especially for adolescent learners, digital texts present a more natural form of reading and writing than written texts.

Considering that girls on average outperform boys in reading, at least on traditional, paper-based tests (Eurydice, 2011b), computerized tests can be a way of addressing boys' reading preferences and a way of increasing their motivation for reading. A set of three studies comparing computerised with paper-and-pencil literacy tests found that while there were no significant gender differences in the computerised versions of the tests, girls performed significantly higher than boys on the paper versions of the spelling modules (Horne, 2007).

Particularly in primary education where literacy levels show a great diversity, ICT can help tailor learning activities to individual learners needs and offer them interesting and engaging reading activities that blend play and assessment. Furthermore, computer-based testing can assist teachers in diagnosing pre-schoolers' speech and language disorders. The computer system can thus partially compensate the lack of experienced clinicians in the school settings and furthermore offers recommendatios for intervention and remediation (Toki & Pange, 2010).

Computer-Based Testing is used widely and successfully for the summative and formative assessment of (basic) literacy and (advanced) reading skills. In a recent study (Sainsbury & Benton, 2011) a pair of electronic tests assessing early reading was administered to over 1000 pupils aged 5-7 from 26 schools to support formative assessment. Test items based on a range of distinct skills, including phonological segmentation, rhyming and word recognition, and were administered on screen, with visual and aural prompts. An automated marking and analysis system provided teachers with formative 'profiles' for each student, together with indicators for the next steps in teaching.

Furthermore, **computer-based quizzes, games and tests** can contribute to formative assessment of student litercy by offering more detailed information on each learners' competence profile.

Games in particular are suited for teaching and assessing reading skills. In most countries a vast variety of free and commercial computer games are available which train reading skills at primary school level. In the English context, for example, numerous webpages combine educational materials with little games supporting early literacy[22].

Example 7: Sample game item from http://www.letters-and-sounds.com: Players score on identifying meaningful and meaningless words

Source: http://www.phonicsplay.co.uk/PicnicOnPluto.html.

These games offer primary pupils and teachers and interesting and engaging alternative to traditional formative assessment formates; enable differentiation and personalisation by allowing each learner to play at their particular level of competence; and provide instant feedback in a motivating way.

Similarly, to support individual tutoring for children's initial reading, an automated reading tutor for initial reading in Dutch has been developed which automatically assesses a child's reading level, provides oral feedback to a child at the phoneme, syllable or word level, and tracks where a child is reading, for automated screen advancement or for direct feedback to the child (Duchateau, et al., 2009).

Especially in the US, but increasingly also in Europe, computerized tests are used for diagnosing and summatively assessing literacy skills of (pre-)Kindergarten and primary school children. The perceived advantages of using computer-based test for Kindergarten and Key-stage 1 children include: allowing whole classes to be tested together without additional personnel; assessing a large number of content areas in reduced time and with fewer questions than a standard paper and

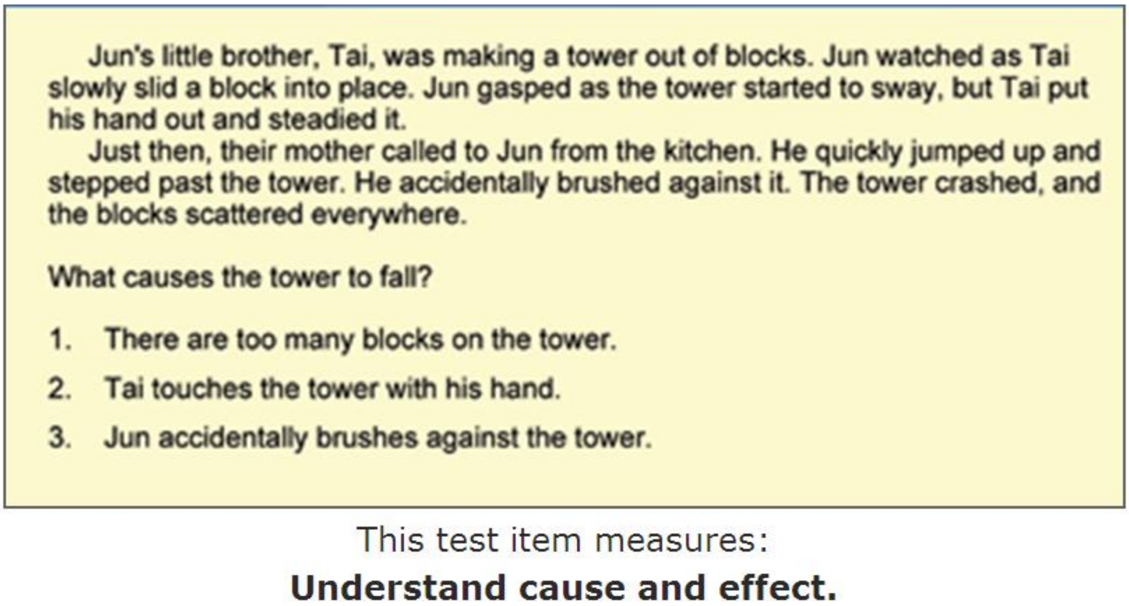
---

[22] Cf. http://learnenglishkids.britishcouncil.org/en/; http://www.letters-and-sounds.com; http://www.phonicsplay.co.uk/freeIndex.htm; http://www.familylearning.org.uk/phonics_games.html; http://www.bbc.co.uk/schools/ks1bitesize/literacy/.

pencil test; immediate and accurate score reports; engaging students with animations during the test (Shamir, Johnson, & Brown, 2009).

The Dynamic Indicators of Basic Early Literacy Skills (DIBELS)[23], for example, are a set of procedures and measures for assessing the acquisition of early literacy skills from kindergarten through sixth grade. They are designed to be short (one minute) fluency measures used to regularly monitor the development of early literacy and early reading skills and are being used widely in the US. DIBELS are not conceived of as summative assessment tools, but as a means to diagnose individual students tutoring needs with a view to allocating resources and planning student support.

Another example, the Waterford Assessment of Core Skills (WACS)[24] is a computerized adaptive test of early literacy for students in Kindergarten through 2nd grade, which includes assessments in letter recognition, letter sound and initial sound recognition, blending, segmenting, reading real and non-words, reading comprehension, listening comprehension, and vocabulary (Shamir, et al., 2009). Similarly, STAR Reading and STAR Early Literacy are standardized, computer-adaptive assessments for use in primary and secondary education in the US, which aim to provide information to teachers about students' competences compared to national norms, as a means of predicting achievement on other standardized (high-stakes) tests. The programme automatically scores and generates reports.[25]


Jun's little brother, Tai, was making a tower out of blocks. Jun watched as Tai slowly slid a block into place. Jun gasped as the tower started to sway, but Tai put his hand out and steadied it.

Just then, their mother called to Jun from the kitchen. He quickly jumped up and stepped past the tower. He accidentally brushed against it. The tower crashed, and the blocks scattered everywhere.

What causes the tower to fall?

1. There are too many blocks on the tower.
2. Tai touches the tower with his hand.
3. Jun accidentally brushes against the tower.

This test item measures:

Understand cause and effect.

Example 8: Sample test question from STAR reading test

Source http://www.renlearn.com/sr/overview/sample.aspx.

These examples illustrate how ICT can be used to make formative and summative assessment for reading more efficient and effective. There are other examples where ICT is used to provide a reading curriculum in a digital learning environment, where the integrated formative assessment and feedback provided is conceived as a means to personalising learning and addressing different learners' reading needs. SuccessMaker's Reader's Workshop[26] and Accelerated Reader[27] are two of the many commercial reading software products for primary education that are very popular in the US. These tools provide ICT-based instruction with animations and game-like scenarios. Assessments are embedded within the programmes; feedback is automatically generated and

23 https://dibels.uoregon.edu/; http://dibels.uoregon.edu/; https://dibels.uoregon.edu/docs/dibelsinfo.pdf; https://dibels.uoregon.edu/dibelsinfo.php; https://dibels.uoregon.edu/techreports/.

24 https://wacs.waterford.org/wacs/home.htm.

25 http://en.wikipedia.org/wiki/STAR_(software).

26 http://www.successmaker.com/Courses/c_awc_rw.html.

27 http://www.renlearn.com/ar/.

instantly provided. Learning can be customized for remedial, developmental, and accelerated students and each lesson can be adapted to each student's demonstrated strengths and weaknesses. These programmes have been evaluated as having positive impacts on learning (Looney, 2010).

Another, European, example is iSTART (interactive strategy training for active reading and thinking). iSTART is a Web-based tutoring programme that uses animated agents to teach reading strategies to young, adolescent (grade 8-12) and collage-aged students. Animated agents are used to teach comprehension strategies such as paraphrasing, predicting. As the learner progresses through the modules, he or she creates self-explanations that are evaluated by the agent (Means & Rochelle, 2010). iSTART has been shown to improve the quality of students' self-explanation that, in turn, were reflected in improved comprehension scores (DeFrance, Khasnabis, & Palincsar, 2010). In a long term experiment with 389 students an improvement in performance was found for all students, indicating that students' self-explanation abilities improve and converge as a function of practice (Jackson, Boonthum, & McNamara, 2010). However, some research findings indicate that the programme is most beneficial for students with the least knowledge about the subject domain, as well as for students who are less strategic in their approach to text (McNamara, et al., 2009; McNamara, Levinstein, & Boonthum, 2004). Other research findings suggest that different readers improve at different levels of comprehension (Magliano, et al., 2005).

There are also e-Assessment strategies for **writing**. Computerized Adaptive Tests for the evaluation of written English knowledge are continuously being improved to further increase measurement accuracy (Barrada, Olea, Ponsoda, & Abad, 2006). Automatic scoring mechanisms for written text are being further developed and refined, although difficulties persist. As mentioned above, while computer scores are highly correlated and consistent with human scores, there are concerns about the adequacy of the criteria used in automatic scoring programmes. Automated scoring for essay length responses is therefore unlikely to become common practice in primary and secondary schools in the near future.

However, with the increasing importance of digital texts, the range of writing genres and styles increases, and schools should take into account that blog posts, wiki entries and even tweets can and should become part of the curriculum for communication in the mother tongue and should therefore also be considered for assessment. Similarly, with the increasing ease of recording spoken language, the assessment of oral communication skills will be facilitated.

ePortfolios can be a means of gathering a rich variety of written and oral communication exercises and assignments and making them available for peer-, self- and teacher assessment. Furthermore, since there is evidence that students who successfully develop self-monitoring skills are more likely to improve their writing skills (Cho, et al., 2010), self-assessment tools can be a promising avenue for developing and improving writing skills.

In short: ICT for the assessment of communication in the mother tongue

The particular strenghth of ICT for the assessment of this key competence lies in:

- Assessment tasks in the form of quizzes and games that support formative assessment, provide instant feedback, and encourage self-regulated learning.

Currently:

- Both, computer-based tests and educational games, are available and being used, at least in some countries, particularly at primary school level as a means of formatively assessing literacy, reading and comprehension skills.
- Computer-based assessment formats are being employed in some countries for national tests, mainly for evaluating schools and curricula.

- ePortfolios are being employed as a means of supporting the assessment of written and oral expression, both for formative and summative purposes.

However:

- Games, quizzes and engaging test formats are currently scattered, isolated, limited in scope, and ill-fitted to comprehensively support curricula and teaching.
- Computer-based tests tend to replicate traditional assessment formats and are not employed as a means of supporting more personalised, engaging, collaborative or authentic tasks.
- More innovative formats for written and oral expression, such as blogs, wikis, tweets, audio and video recordings, etc., are seldom included in ePortfolios.

To seize the opportunities offered by ICT:

- Comprehensive environments, linking games, quizzes and tests to learning content, should be developed which offer teachers a variety of versatile, adaptive and engaging assessment tools that are fitted to curricula and focus on assessing skills and attitudes, based on a range of authentic text genres and communication contexts.
- Digital communication formats, which are becoming increasingly important and are often a more natural way of communication for young people, need to become and integral part of curricula and assessment.

Policy support is needed to:

- Encourage and support the development of more comprehensive, versatile and adequate ICT-based assessment formats for reading and comprehension skills that allow teachers to more closely monitor learning progress and to more timely and effectively to react towards individual learners' strengths and weaknesses, which are at the same time enjoyable and engaging for learners.
- Re-focus curricula on newly emerging communication patterns, such as digital text formats and audio and video communication, and supply teachers and learners with guidance and support on how digital artefacts collected in ePortfolios can and should be assessed, including self- and peer assessment.

3.2 Communication in foreign languages

Box 2: Definition of "Communication in foreign languages"

Communication in foreign languages broadly shares the main skill dimensions of communication in the mother tongue: it is based on the ability to understand, express and interpret concepts, thoughts, feelings, facts and opinions in both oral and written form (listening, speaking, reading and writing) in an appropriate range of societal and cultural contexts (in education and training, work, home and leisure) according to one's wants or needs. Communication in foreign languages also calls for skills such as mediation and intercultural understanding. An individual's level of proficiency will vary between the four dimensions (listening, speaking, reading and writing) and between the different languages, and according to that individual's social and cultural background, environment, needs and/or interests.

Source: Recommendation of Key Competences for Lifelong Learning (Council of the European Union, 2006)

In most European countries only 20-40% of students use computers in foreign language instruction; only in Denmark and Norway around 60% and 50% of students report using computers in foreign language lessons (Eurydice, 2011a).

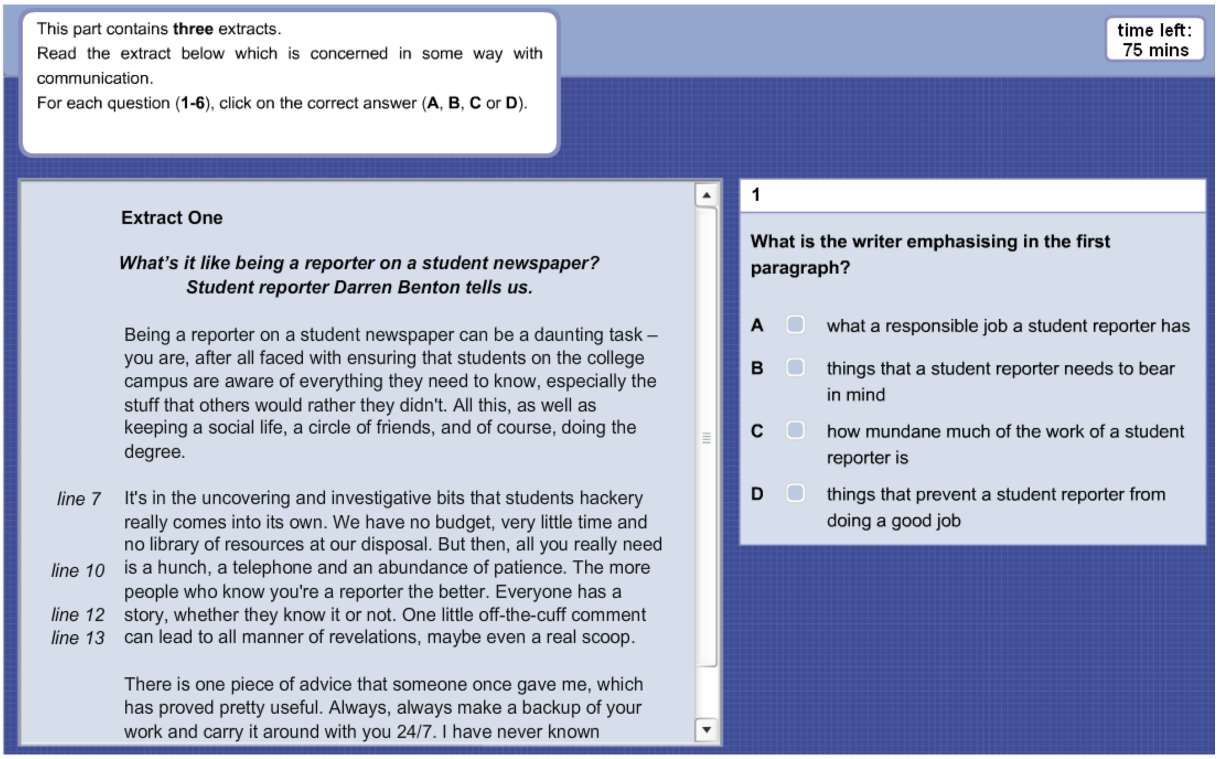
Yet computer-enhanced learning and teaching has been shown a powerful tool in foreign language education, in particular, as it allows for more personalised tuition and allows for the smooth integration of different communication contexts. For example, a three-year case study with elementary school children in Greece investigated the degree to which comprehension skills, such as perception, information retrieval and concentration, are improved by using a computer-assisted instructional tool to deliver a typical introductory course of the French language. A comparison of the performance of the experimental group with the control group over three subsequent years indicates that the use of the computer had a significant positive influence on the abilities of perception and information retrieval (Tzortzidou & Hassapis, 2001).

Assessment methods which best suit language learning in online environments include personalization services for adaptive educational hypermedia and online portfolios to measure performance based on collections of student-created work (Agudo, Rico, Edwards, & Sánchez, 2009). e-gramm (cf. Looney, 2010; Sanz, 2008) is an example illustrating how ICT can be used effectively in assessing and providing feedback on written tasks in foreign language instruction. E-gramm is an ICT-based programme developed at the Universidad Nacional de Educación a Distancia in Spain for learners of English as a Foreign Language, which provides detailed feedback on written compositions. The developers analysed common mistakes across hundreds of student compositions, encoded them, and developed feedback to address these kinds of mistakes and to support learners in modifying their own writing. By 2008, the programme was able to detect 60 to 70% of common mistakes made by Spanish mother tongue students. The developers believe that students get much more feedback than they would from teachers, who have little time to provide detailed comments or suggestions.

In school settings, ePortfolios and self-assessment tools have been used successfully as a means to support formative and summative assessment and foster students' self-regulated learning. Furthermore, developments in speech recognition can contribute to providing electronic test formats for oral language skills, thus facilitating the assessment of these skills.

CBA

Online computer-based language tests are often used for selection and placement purposes in higher education, complementing or supplementing self-assessed foreign language proficiency (Meurant, 2009). Computerization of foreign language reading tests has been of interest among language assessment researchers for the past 15 years, but few empirical studies have evaluated the equivalence of the construct being measured in computerized and conventional foreign language reading tests and the generalizability of computerized reading test results to other reading conditions (Sawaki, 2001).



Example 9: Sample test question in reading, Cambridge Certificate in Advanced English
Source: http://olpt.s3.amazonaws.com/online-practice-cae/index.html.

Currently, research focuses on the development of fully automatic tests of spoken language ability, which are already available and have been shown to provide valid scores for communication skills in a second language (Bernstein, van Moere, & Cheng, 2010). The TOEFL iBT, for example, includes a speaking section in which the examinee responds via microphone to audio prompts delivered through the headset. Questions ask the examinee to do such things as listen to a short lecture and summarize it; or read a brief passage, listen to a short lecture and answer a question calling for integration of information from the two sources (Bennett, 2010).

So-called 'facility-in-L2' tests, a family of automated spoken language tests in Spanish, Dutch, Arabic, and English, have been shown to adequately measure receptive and productive language ability as test-takers engage in a succession of tasks with meaningful language. Automated assessment of non-native speech is not yet advanced enough to compete with human assessment. However, advances are being made, with speech assessment programmes that show moderately high correlations with pronounciation scores(L. Chen, Evanini, & Sun, 2010). In general, research indicates that scores from the automated tests are strongly correlated with the scores from oral proficiency interviews (Bernstein, et al., 2010).

For school education, however, these developments are of minor interest, as they might eventually contribute to making testing more efficient, but will have little effect on developing learners' intercultural communication skills.

On the whole, computer-based testing as such will not revolutionize foreign language learning and teaching as it replicates traditional assessment procedures and is usually employed to increase testing efficiency. However, if used formatively, computer-based assessment in the form of online quizzes can help teachers in diversifying their teaching strategies and offering more personalised learning opportunities, by supplying learners with what they perceive as games, in which they can train different solution strategies, get immediate feedback and assistance, and monitor their own progress.

Self-assessment

In general, students' perception of their foreign language skill is not very accurate or reliable. A study among USA university students, for example, indicates that self-assessment of foreign language reading ability is not an accurate predictor of test performance and actual reading skills for advanced learners (Brantmeier, 2006).

However, as an instructional tool for formative assesment promoting self-regulated learning, continuous or recurring self-assessment exercises can enable primary and secondary students to better understand their achievements and learning needs and subsequently improve self-confidence and performance. A study among 254 6th grade students in South Korea studying English as a foreign language who were asked to perform self-assessments on a regular basis for a semester during their English classes, shows that students improved their ability to self-assess their performance over time. There is furthermore some positive effects of self-assessment on the students' English performance as well as their confidence in learning English, though the effect sizes were rather small (Butler & Lee, 2010).

ePortfolios

The European Language Portfolio (ELP)[28] and its American adaptations, LinguaFolio and the Global Language Portfolio, are tools to be used with the Common European Framework of Reference for languages and the American national standards. The ELP is divided in three parts, consisting of a language passport, a language biography, and a dossier, builds on earlier research on portfolios and second language assessment (Cummins & Davesne, 2009). The objective of the ELP is to stimulate reflective learning in which goal setting and self-assessment play a central role (Little, 2009). The European Language Portfolio has, for example, been integrated into the Diversificación Curricular English syllabus at a Secondary School in Asturias (Spain), where it has been shown to be a useful tool to promote self-assessment and self-reflection in students with learning difficulties (Alonso, 2011).

Portfolio assessment in general is perceived as a useful tool in language teaching at primary and secondary school level, as it can help students understand their own progress in the foreign language and build up self-confidence. A study investigating the impact of portfolio assessment on foreign language writing skills of iranian students of English as a foreign language found that students in the portfolio assessment group outperformed the students in the control group in their overall writing ability and in the sub-skills of focus, elaboration, organization and vocabulary. The findings suggest that portfolio assessment empowers students' learning of English writing, hence emphasizing the formative potential of portfolio assessment in foreign language classes (Ghoorchaei, Tavakoli, & Ansari, 2010).

Immersive environments and games

In the future, immersive environments replicating authentic communication context could also be used in school education as a means of more effectively embedding formative assessment, feedback and guidance in the learning process. Alelo[29], for example, is a commercial provider of crash courses on basic communicative skills in foreign languages and cultures, employing immersive simulations of real-life social communication that are based on interactive 3D video games involving spoken dialogs and cultural protocols conducted with "socially intelligent virtual humans. Alelo utilizes game design techniques to promote learning, e.g., by providing learners with missions to achieve, supporting fluid game-play in the form of simulated conversations with non-player characters, and continual feedback on learner performance within a game scenario context. Corrective feedback is embedded in the game in a fluent way. Artificial intelligence (AI) plays

---

28 http://www.coe.int/t/dg4/portfolio/default.asp?l=e&m=/main_pages/welcome.html; www.coe.int/portfolio.

29 www.alelo.com.

multiple roles in this learning environment: to process the learner's speech, to interpret and evaluate learner actions, to control the response of non-player characters, to generate hints, to assess the trainee's mastery of the skills, and to assist in the generation and validation of lesson content (W. L. Johnson, 2010).

Example 10: Alelo
Source: www.alelo.com.

In short: ICT for the assessment of communication in foreign languages

The particular strenghth of ICT for the assessment of this key competence lies in:

- Automatic scoring techniques for written text which allow learners to immediately detect and correct the most salient errors and thus also reduces teachers' working load; ideally combined with tutoring software which explains the underlying concepts and offers targeted training exercises.
- ePortfolios to showcase work done across a variety of communication formats.

Currently:

- As with communication in the mother tongue, computer-based tests and educational games are available and being used to support language learning, although to a somewhat lesser extent.
- ePortfolios are being employed widely as a means of supporting the assessment of written and oral expression, both for formative and summative purposes.

However:

- Automated scoring and correction techniques targeted at language learning have not reached a level of sophistication that would allow their more widespread deployment –

despite the availability and relatively high reliability of automatic scoring for texts written in a foreign language.

- In general, experimentation with ICT as a means of supporting the competence-based assessment of foreign language learning is less pronounced when compared to other Key Competences. The potential of ICT in shortening feedback loops and in combining assessment and targeted training is under-researched.

To seize the opportunities offered by ICT:

- More research and experimentation is needed to eventually develop coherent and comprehensive training programmes that offer teachers and learners targeted, versatile, adaptive and engaging training and assessment tools that provie instant feedback and promote self-regulated learning.

Policy support is needed to:

- Raise awareness for the specific potential of ICT in fostering language learning and set incentives for more targeted research in this area.

3.3 Mathematical competence

Box 3: Definition of "Mathematical competence"

Mathematical competence is the ability to develop and apply mathematical thinking in order to solve a range of problems in everyday situations. Building on a sound mastery of numeracy, the emphasis is on process andactivity, as well as knowledge. Mathematical competence involves, to different degrees, the ability and willingness to use mathematical modes of thought (logical and spatial thinking) and presentation (formulas,models, constructs, graphs, charts).

Source: Recommendation of Key Competences for Lifelong Learning (Council of the European Union, 2006)

Technology has become essential to the practice of mathematics and has subsequently changed focus and strategies for teaching and learning mathematics, favouring more creative, collaborative and constructive learning approaches. Research indicates, for example, that classrooms with graphing calculator use tend to foster a more constructive climate, with more conjecturing, more frequent use of multiple solutions, and higher levels of discourse than those classrooms with infrequent calculator use (Nathan, 2010).

When technologies are used in mathematics education it is important to assess how activities will support the development of mathematical understanding and the technical expertise that students will need (Forster, 2006). Research indicates, for example that using graphing technology for non-routine activities, such as mathematical discovery and complex problem solving, tend to support increased conceptual understanding and higher achievement, while use of technology for routine calculations does not (Guerrero, Walker, & Dugdale, 2004).

However, in general, technology used in the mathematics classroom seems to contribute to fostering innovative and learner-centred pedagogies. An analysis of the SITES 2006 data suggests that mathematics teachers in countries with a relative high percentage of ICT, compared to those in countries with low ICT use, tend to apply a learner-centered approach in their educational practice, have a focus on life long learning competencies, and are encouraged by their school leaders in using ICT and in applying new ways of teaching and learning (Pelgrum & Voogt, 2009).

One of the reasons for the positive effect of technology on changing learning and teaching strategies in mathematics might be that from the very beginning the use of technology in mathematics instruction has been conceived as a means of re-focusing learning objectives, moving from purely operational skills and routine tasks (which are delegated to a technological device) towards applied mathematics and more complex higher-order skills, with a focus on (data) analysis, interpretation, reflexion, problem-solving and transfer skills. Consequently many ICT enhanced programmes, learning environments and applications are designed to foster these more generic skills. Research points towards the viability of this strategy, by confirming that complementing mathematics instrution in schools with ICT applications designed to foster problem-solving and transfer skills, does in fact raise student performance, especially for students which are already relatively high in intelligence and mathematical ability (J. Meijer & Riemersma, 2002).

As far as assessment is concerned, **ePortfolios** play a minor role in mathematics instrution, although it has been shown that they can lead to increased student performance in mathematics, by fostering ownership and self-regulated learning (Burks, 2010).

Games are playing an important role, and, by fostering self-assessment, support formative assessment, but are currently not used for summative or diagnostic purposes. Puppetman, for example, is a rational number addition video game which allows teachers to examine process data from game play for formative assessment (e.g. time spent on each level, strategies used). The advantage of a game such as Puppetman is that, although it can be used for formative assessment, students do not perceive the game as a test (Vendlinski, Delacruz, Buschang, Chung, & Baker, 2010).

The internet is a vast resource for free and commercial numeracy and mathematics games, both for primary and secondary school mathematics, but often with a focus on basic operations and concepts[30].

Catch the Fly

Mr. Frog needs you to give him the x, y coordinates for a fly so he can dine!

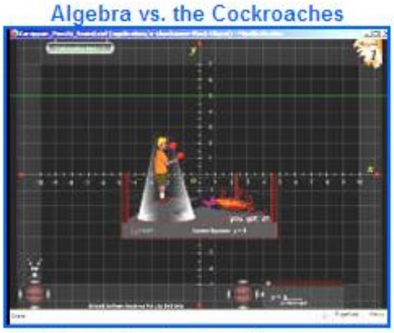
(Middle School and up)

Number Cop

Pick numbers by their divisibility, and don't crash into any of them!

(Middle School and up)

Algebra vs. the Cockroaches



It's you against the cockroaches crawling on your graph paper!

(High School and up)

No-sound version - faster download

Factortris

Remember Tetris? Here it is with a new dimension!

(Middle School and up)

No-sound version - faster download

Example 11: Mathematical games for secondary school education

Source: http://hotmath.com/games.html.

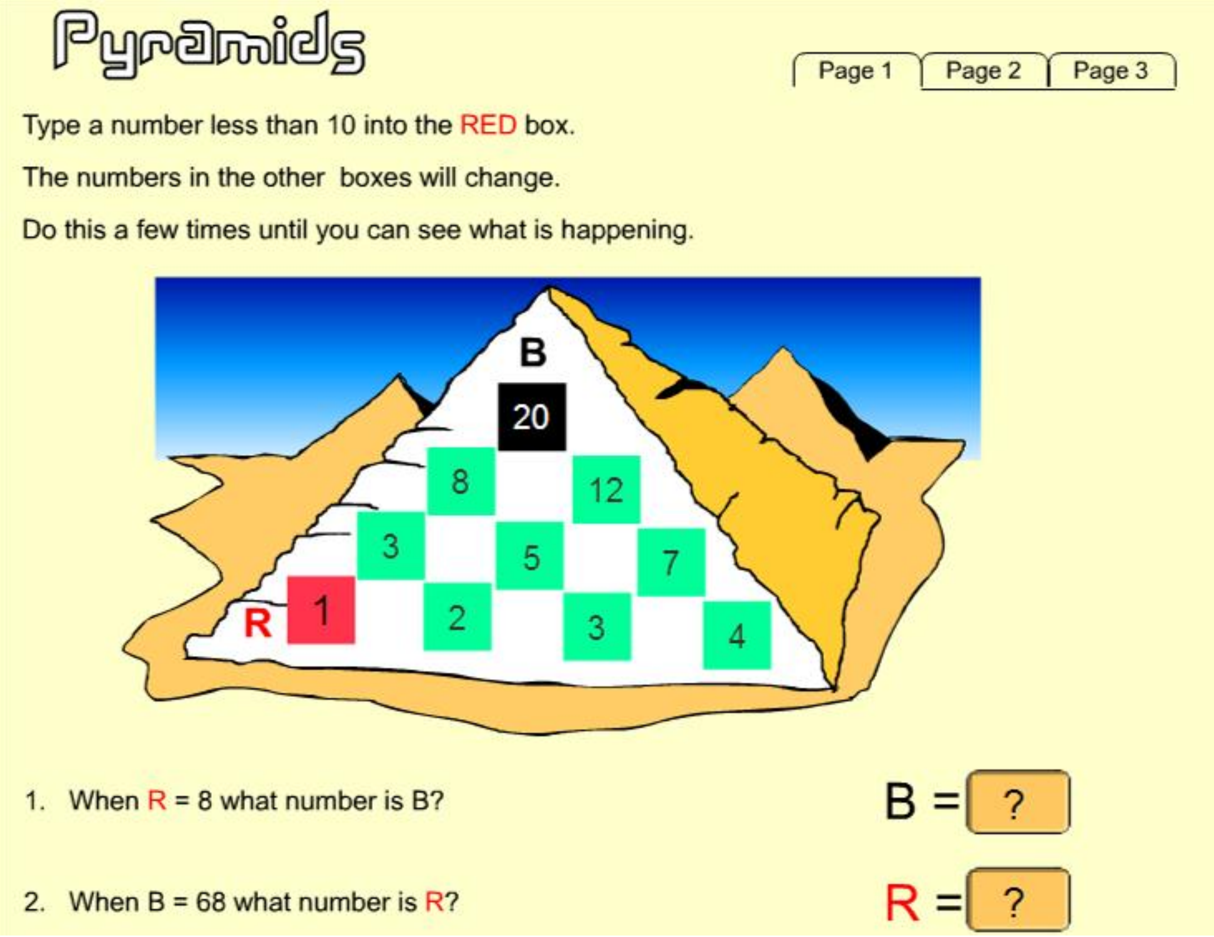
However, the most commonly used computer-based tool for formative and summative assessment in mathematics are **computer-based tests,** which are increasingly integrated in intelligent tutoring environments to link assessment with targeted and tailormade tutoring. Since mathematical answers are more susceptible to automatic scoring, even when alternative solutions are possible, it has been possible to develop computer-based test environments which allow for diverse, innovative and engaging tasks to be implemented, which embed feedback and tutoring options guiding students in the development of their solution strategies.

With the development of "Generation Re-Invention" testing formats, it has become possible to design mathematical test items in a way to better assess problem-solving and transfer skills. World Class Arena[31], for example, is an international initiative designed to identify and assess gifted and talented students around the world. It was devised by the British government Department for Education and Skills (DfES). The tests are designed to assess higher order thinking skills in mathematics and problem solving for students aged 9-14. Each test requires students to apply creative thinking and logic to respond to problems and clearly communicate their thought

[30] For example: http://www.bbc.co.uk/schools/ks1bitesize/numeracy/; http://mathplayground.com/games.html; http://www.coolmath-games.com/; http://www.primarygames.com/math.htm; http://hotmath.com/games.html.

[31] http://www.worldclassarena.org/en/home/home.htm.

processes. Since the first test session in 2001, over 18,000 students in over 25 different countries worldwide have taken the tests (Binkley, et al., 2012; Ripley, 2009).



Example 12: Mathematical problem solving task for 13 year olds, in a CBA format
Source: http://www.worldclassarena.org/v5/flash/13_year_old/pyramids.swf.

Just like literacy, numeracy skills are being assessed in a range of national and international tests, which, in many cases are being conducted electronically. The US easyCBM™ assessment system, for example, provides school districts, administrators, and teachers with a full suite of assessment and reporting options in reading and mathematics[32].

As with games, the internet is a rich resource for electronic tests and quizzes for school mathematics[33]. Furthermore, a variety of (commercial) software solutions, such as, for example, Accelerated Maths[34], are available that allow teachers to create math assignments tailored to each student's current level; automatically score all math practice, including assignments and tests; provide ongoing feedback on students' daily practice; and differentiate math instruction, addressing each student's individual needs. Similarly, STACK[35] provides mathematical questions, which are answered using a simple linear syntax. The submitted answer is assessed and immediate formative feedback is provided. Multiple attempts at a question are usually encouraged and randomly generated questions are offered to allow students to practice.

Thus, computer-based tests can be powerful tools supporting formative and low-stakes summative assessment. In this sense, computer-based assessment formats with integrated feedback are

[32] http://easycbm.com/.

[33] For an overview of tests and quizzes available in the UK see: http://www.assessmentfocus.com/k12-math.php.

[34] http://www.renlearn.com/am/overview/.

[35] http://www.stack.bham.ac.uk/.

particularly useful for students who need extra support and attention and allow teachers to offer each student support at their individual level of competence. Two studies with special-needs pupils (aged 8-12) suggests, that an ICT-based assessment format, in which test items were enriched with an optional auxiliary tool that students could use for solving the problems, can reveal weak pupils' learning potential and improve their strategy use (Peltenburg, Van Den Heuvel-Panhuizen, & Doig, 2009; Peltenburg, van den Heuvel-Panhuizen, & Robitzsch, 2010).

Research in higher education confirms the usefulness of CBA for formative assessment. Research on the regular use of an online assessment instrument within a university mathematics course has, for example, been shown to lead to higher student learning, all else being equal (Angus & Watson, 2009). Additionally, tests delivered in interactive, immersive environments have the advantage of providing teachers with diagnostic tools that help them to tailor instruction to the needs of students at different achievement levels (Bottge, Rueda, Kwon, Grant, & LaRoque, 2009). Furthermore, making online assessments accessible to students, for self-assessment purposes, can enable them to learn independently, reduce anxiety, and improve self-efficacy (Morris & Dowdall, 2011).

In higher education, computer aided assessment packages such as AiM, STACK and MapleTA are often integrated into computer algebra systems such as Maple, which are being used widely in higher education mathematic courses. Maple additionally allows assignments and examinations to be completed electronically within Maple or downloaded, and student work to be compiled in Maple and uploaded. Assignments and tests can be completely or partially be marked automatically in Maple, allowing for different task formats to be integrated, and a marking report is generated by the programme and returned to the student (Blyth & Labovic, 2009).

For primary and secondary schools, however, **intelligent tutoring systems** are more promising electronic environments schaffolding the learning process in mathematics. The advantage of these systems is that they can provide immediate constructive feedback (Ljungdahl & Prescott, 2009), offer help which is tailored to the individual problem-solving approach and adapt the level of difficulty of the tasks administered to the individual learners' progress and learning needs.

Intelligent Tutoring Systems (ITSs) are being widely used in the US, where the most popular system "Cognitive Tutors" provides differentiated instruction which encourages problem-solving behaviour to half a million students in around 2600 US middle and high schools (Ritter, et al., 2010). The programme selects mathematical problems for each student adapted in level of difficulty. Correct solution strategies are annotated with hints, which allow students to access instruction that is directly relevant to the problem they are working on and the strategy they are following within that problem.[36] In Norway, the Kikora Software for mathematics (cf. **Example 4**, page 33) is used widely. Kikora offers detailed automatic feedback and allows teachers to track progress.

Research indicates that students who used Cognitive Tutor significantly outscored their peers on national exams, an effect that was especially noticable for students with limited English proficiency or special learning needs (Ritter, Anderson, Koedinger, & Corbett, 2007). Research on the implementation of a web based intelligent tutoring system "eFit" for mathematics at lower secondary schools in Germany confirms this finding, by showing that children using the tutoring system significantly improved their arithmetic performance over a period of 9 months (Graff, Mayer, & Lebens, 2008).

Thus, ICT has led and will further contribute to changing the shape, scope and focus of mathematics instruction and assessment, favouring more advanced, complex and applied competences. The greatest advantage and benefit of using electronic environments in mathematics instruction lies in supporting effective formative assessment, by providing instant feedback to students and allowing them to proceed at their own pace, through a series of tasks adapted in level of difficulty and with the help of hints that aim to develop adequate solution strategies.

---

[36] http://carnegielearning.com/static/web_docs/2010_Cognitive_Tutor_Effectiveness.pdf.

In short: ICT for the assessment of mathematical competence

The particular strenghth of ICT for the assessment of this key competence lies in:

- Quizzes, games and educational software which provide instant feedback to students and allow them to proceed at their own pace, through a series of tasks adapted in level of difficulty and with the help of hints that aim to develop adequate solution strategies.

Currently:

- Computer-based testing in mathematics is used widely, on national, inter-national and school level. Due to the nature of mathematical inquiry, it has been possible to embed complex and authentic tasks so that mathematical competence can comprehensively and reliably be assessed through computer-based tests.
- The internet is a rich resource for electronic tests, quizzes and games for school mathematics which support formative assessment.
- A variety of (commercial) software solutions are available that allow teachers to create assignments tailored to each student's current level; automatically score all practice exercises, assignments and tests; provide ongoing feedback on students' daily practice; and differentiate math instruction, addressing each student's individual needs.

However:

- Games, quizzes and online tools encouraging mathematical inquiry are currently scattered across the internet, isolated and limited in scope, and ill-fitted to comprehensively support curricula and teaching.
- In Europe, intelligent tutoring systems and other environments supporting the comprehensive and personalised assessment of mathematical competence are scarcely used.

To seize the opportunities offered by ICT:

- Comprehensive environments, linking games, quizzes and tests to learning content, should be developed which offer teachers a variety of versatile, adaptive and engaging assessment tools that are fitted to curricula and focus on assessing skills and attitudes, based on a range of authentic tasks and problem contexts.
- Existing educational software with feedback and tutoring functions should be adapted to curricula in all European countries and further improved by providing more authentic and complex tasks that capture the full scope of mathematical competence.

Policy support is needed to:

- Encourage take-up and harmonisation of existing tools and environments.
- Encourage and enable teachers to make use of existing assessment tools, to critically examine their value and to adapt them to their learners' needs.

3.4 Basic competences in science and technology

Box 4: Definition of "Basic competences in science and technology"

Competence in science refers to the ability and willingness to use the body of knowledge and methodologyemployed to explain the natural world, in order to identify questions and to draw evidence-based conclusions.Competence in technology is viewed as the application of that knowledge and methodology in response to perceived human wants or needs. Competence in science and technology involves an understanding of the changes caused by human activity and responsibility as an individual citizen.

Source: Recommendation of Key Competences for Lifelong Learning (Council of the European Union, 2006)

There is a general consensus that students' ability to construct evidence-based explanations in classrooms through scientific inquiry is critical to successful science education (C. J. Huang, et al., 2011). Research suggests that inquiry-based, learner-centered learning experience in science is associated with long-term improvements in learning (Derting & Ebert-May, 2010).

Based on experimental evidence, it is expected that inquiry-oriented science curricula can substantially be enhanced by the use of engaging ICT applications with integrated assessment, and is suggested that a large-scale effort to do so might have a lasting impact on science education (Taasoobshirazi, et al., 2006). It has been shown, for example, that interactive visualizations combined with online inquiry and embedded assessments can deepen student understanding of complex ideas in science (Linn, Lee, Tinker, Husic, & Chiu, 2006).

Conventional computer-based tests can, of course, be used to assess basic competences in science, as illustrated by the 2006 PISA Computer-Based Assessment of Student Skills in Science[37]. By now, automated grading schemes for the assessment of creative problem-solving in science education have been developed and trialed in the context of secondary Earth science education, where the machine-generated scores achieved high inter-rater reliability against human grading (Wang, et al., 2008).

Similarly, **mobile technologies** can be used, both for the learning and the assessment process. In a research project in Singapore mobile technologies were used during the 2009 school year to deliver the primary (grade 3) science curriculum. It was found that the experimental class performed better than other classes as measured by traditional assessments in the science subject. Additionally, students were found to learn science in personal, deep and engaging ways and developed a positive attitudes towards mobile learning (Looi, et al., 2011).

Scientific inquiry capabilities can also be assessed by evaluating students' scientific inquiry **portfolios** in actual hands-on experiments (J. M. Su, Lin, Tseng, & Lu, 2011). Scientific e-portfolios can be supplemented with electronic tools which automatically assess and diagnose students' scientific inquiry abilities and generate personalized diagnostic reports, which diagnose learning problems and provide corresponding reasons and remedial suggestions based on teacher-defined assessment knowledge (J. M. Su, et al., 2011).

Computer simulations, scientific games and virtual laboratories

For science education, the most powerful and targeted technology-based learning and assessment environments are provided by online simulations and virtual laboratories, supplemented by the remote access to tools and sensors such as atomic force microscopes or telescopes. These ICT tools can make scientific phenomena accessible that are too dangerous or too expensive to include in science classes at the primary or secondary level (Delgado & Krajcik, 2010). Computer simulations, scientific games and virtual laboratories provide opportunities for students to develop and apply skills and knowledge in more realistic contexts and provide feedback in real time. Simulations may

[37] http://www.oecd.org/document/9/0,3746,en_32252351_32236191_45938505_1_1_1_1,00.html.

involve mini-laboratory investigations, or "predict-observe-explain" demonstrations. The programmes usually provide opportunities for students to reflect on their own actions and patterns they may detect in the responses provided by the simulation programmes (Looney, 2010).

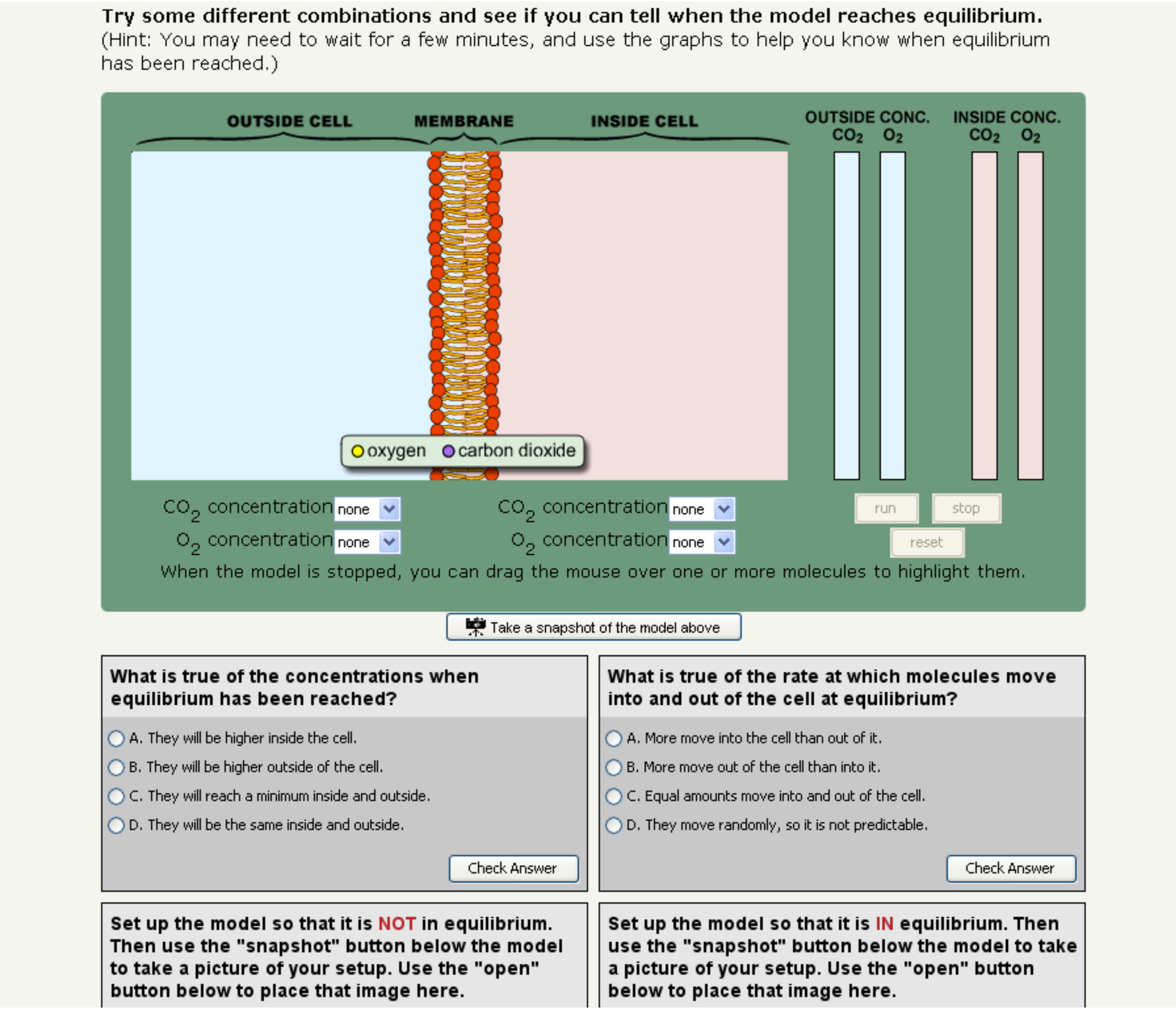
Category	**Description**	**Examples**
Real-time data acquisition and graphing	Probes for pH, temperature, voltages, pressure, dissolved oxygen level, motion etc. that connect to a computer, graphing calculator or handheld device	Verneier probes for use with TI 83 plus calculator-based laboratory; Pasco PASPORT and Xplorer handheld probes
Simulations	Computer-generated versions of real-world objects and phenomena that often allow the user to control settings	Virtual frog dissection kit (http://froggy.lbl.gov/virtual/) Chemical kinetics simulation (http://www.chem.uci.edu/undergrad/applets/sim/simulation.htm); Molecular Workbench (http://workbench.concord.org)
Virtual laboratories	Environments allowing students to design and carry out experiments	www.chem.ox.ac.uk/vrchemestry http://learn.arc.nasa.gov/vlab/features.html
Remote access to instruments	Students interact with and manipulate instruments that are not available at their school, over the internet	University of Carolina atomic force microscope; weather data
Inquiry-structuring software and sites	Software or websites that help teachers and learners plan and manage long, complex investigations	WISE (wise.berkeley.edu) Ideakeeper (???)
Learning environments	We-based environments integrating curriculum, multimedia, simulations , support for learners, embedded assessment	WISE (wise.berkeley.edu) BioLogica (http.//biologica.concord.org)

Table 2: Overview of Computer Environments and Tools Supporting Scientific Inquiry
Source: Delgado & Krajcik (2010)

There are a number of examples in which ICT is used to make professional scientific data and empirical observations accessible to pupils. The Kids as Global Scientists and World Watcher curricular projects access the actual data of professional scientists, using the Internet and custom software (Delgado & Krajcik, 2010). The Microelectronics WebLab (recently re-named: iLab)[38] at MIT allows students to do actual (not simulated) laboratory research on state-of-the art equipment through the Internet. A study of WebLab indicates that WebLab allows undergraduates to learn at their own pace and on their own schedules; enables them to use different processes of learning (intuitive, visual, abstract); and gives them an opportunity to link individual and collaborative effort in creative combinations. (Fischer, Mitchell, & Del Alamo, 2007). In other projects, such as Project Feeder-Watch, BioKIDS, and GLOBE, students generate genuine scientific data themselves. Remote access to sophisticated tools such as telescopes and microscopes allows students to use the instruments of practicing scientists. All of these factors increase the relevance and authenticity of science learning for students (Delgado & Krajcik, 2010).

[38] http://ilab.mit.edu/iLabServiceBroker/.

Dynamic websites, such as Web of Inquiry[39], allow students carry out scientific inquiry projects to develop and test their theories; learn scientific language, tools, and practices of investigation; engage in self assessment; and provide feedback to peers (Herrenkohl, Tasker, & White, 2011). The PhET Interactive Simulations Project at the University of Colorado Boulder provides more than 115 free, research-proven, interactive simulations for science, technology and mathematis education (http://phet.colorado.edu).



Example 13: Molecular Workbench: Example "Diffusion"
Source: http://mw.concord.org/modeler/index.html.

Some science-learning environments can have embedded formative assessments that teachers can access immediately in order to gauge the effectiveness of their instruction and modify their plans accordingly (Delgado & Krajcik, 2010). Simulations such as provided by Molecular Workbench[40], for example, make visible phenomena that are too small or too fast to observe, such as chemical reactions or gas at the molecular level. Molecular Workbench provides visual, interactive computational experiments for teaching and learning science, which can be customized and adapted by the teacher. Embedded assessments allow teachers to generate real-time reports which provide a complete view of student learning progression that teachers can track.

Similarly, BioLogica[41], a software tool for teaching high school genetics, embeds curriculum and assessment functions within a computer-based manipulable model, which focuses on independent

---

[39] http://www.webofinquiry.org.
[40] http://mw.concord.org/modeler/index.html.
[41] http://biologica.concord.org/.

inquiry and scientific reasoning. In 2000/01 BioLogica was piloted in ten high schools with nearly twenty teachers and approximately 700 students.

Example 14: SimScientist

Source: http://simscientists.org.

The SimScientists[42] research program explicitly focuses on exploiting the potential of simulations as (learning and) assessment tools. The prototype assessments address key topics in middle and high school science: forces and motion (physical science) and ecosystems (life science). The assessments developed engage students in problem-based investigations designed to test deep science knowledge and extended inquiry skills. Ongoing follow-up projects (Calipers II) focus on creating a new generation of simulation-based, curriculum-embedded formative assessments with immediate, individualized coaching for students. Students and teachers receive diagnostic reports and teachers use supporting reflection activities to improve student learning during instruction.

As part of the Advanced Technologies for Learning in Authentic Settings (ATLAS)[43] research project a number of computer-based learning environments have been designed, with a focus on science education. BioWorld, for example, is a computer-based learning environment that provides a realistic environment for students to learn about diseases through solving specific patient cases. BioWorld provides a hospital simulation where students learn diagnostic reasoning by visiting patients, interpreting patient symptoms, conducting diagnostic tests, and collecting appropriate information in the library. Solving a patient case in BioWorld not only consists of submitting the right diagnosis, but also requires the student to select and organize evidence that supports and justifies decisions made throughout the case resolution process.

There are also a number of commercial games, such as Spore, which are not explicitly meant for the education market, but embody scientific principles and can therefore be used to teach school science in an engaging way. Furthermore, a variety of educational games for science education

[42] http://simscientists.org.
[43] http://www.mcgill.ca/atlas-lab/.

have been developed. ARIES (Acquiring Research Investigative and Evaluative Skills) is a computerized educational game in which players attempt to stop extraterrestrials from implicitly stunting scientific progress on Earth by publishing bad research in a variety of fields. Players progress through three modules: 1) read and be tested on an on-line science text, 2) evaluate potentially flawed research articles, and 3) learn question-asking skills. ARIES incorporates multiple learning principles, such as testing effects, generation effects, and formative feedback. (Wallace, et al., 2009)

Example 15: Quest Atlantis
Source: http://atlantis.crlt.indiana.edu/.

Another example is Quest Atlantis,[44] an international learning and teaching project that uses a 3D multi-user environment to immerse children, ages 9-16, in educational tasks. It combines strategies used in the commercial gaming environment with lessons from educational research on learning and motivation. To successfully solve the problems encountered in the games, students need to demonstrate causal reasoning skills, subject knowledge in physics and chemistry, and be able to understand how systems work at both macro and micro level.

River City is an environment in which use their knowledge of biology along with the results of tests conducted online with equipment such as virtual microscopes to investigate the mechanisms through which a disease is spreading in a simulated 18th century city. Students collaborate to write up their research findings as a report to River City's mayor. Prompts and expert input, which are gradually faded as students acquire stronger inquiry skills, guide the learning process. Using data-

[44] http://atlantis.crlt.indiana.edu/. (Cf. S. Barab, Thomas, Dodge, Carteaux, & Tuzun, 2005; S. A. Barab, Gresalfi, Dodge, & Ingram-Goble, 2010; S. A. Barab, Sadler, Heiselt, Hickey, & Zuiker, 2007; Dodge, et al., 2008; Hickey, Ingram-Goble, & Jameson, 2009; Lim, Nonis, & Hedberg, 2006; Peppler & Solomou, 2011; Thomas, Barab, & Tuzun, 2009; Warren, Stein, Dondlinger, & Barab, 2009; Zheng, Young, Wagner, & Brewer, 2009)

mining allows teachers to assess individual student's inquiry skills and document gains in students' engagement, learning and self-efficacy (Dede, 2010; Means & Rochelle, 2010).

Similarly, the Virtual Performance Assessment project[45] relies on simulated, game-like environments to assess students' ability to perform scientific inquiry in order to solve a problem. The assessment is embedded in authentic settings, which allow better observation and measurement of complex cognition and inquiry processes (Binkley, et al., 2012).

A bit different in approach, Outbreak@The Institute is played across a university campus, on handheld computers (PDAs). Players' PDAs are connected wirelessly to a network, which provides them with positioning information. Players take on the roles of doctors, medical technicians, and public health experts to contain a disease outbreak. Players can interact with virtual characters and employ virtual diagnostic tests and medicines. They are challenged to identify the source and prevent the spread of an infectious disease that can spread among real and/or virtual characters according to an underlying model. (Rosenbaum, Klopfer, & Perry, 2007)

While online simulations, virtual laboratories and games are readily available, often for free, and are being successfully used by many teachers and learners, use of these tools has not yet become mainstreamed. On the contrary, as a recent study by Eurydice (2011) finds, primary and secondary students rarely use computers for conducting experiments of simulations of natural phenomena in science lessons.

Compared to other Key Competences, virtual learning environments are relatively advanced for science as a school subject, possibly because scientific experiments have always been considered an essential element of scientific discovery and instruction, although their use within the science classroom has always been limited by resources and practical constraints, rendering valid empirical experiments difficult and time-consuming. Thus, the perceived need for simulations might be lower for other Key Competences, such as social and civic competences or cultural expression and awareness. Yet, looking at the examples discussed in this chapter, it is obvious that the assessment of these Key Competences could equally benefit from a virtual environment approach using simulations and games.

In short: ICT for the assessment of competences in science and technology

The particular strenghth of ICT for the assessment of this key competence lies in:

- Simulations, virtual laboratories and multiplayer games, which embed learning and assessment in authentic real-life contexts, thus supporting the development and assessment of scientific inquiry, analysis, interpretation and reflection.

Currently:

- Compared to other Key Competences, virtual learning environments are relatively advanced for science as a school subject. A great number and variety of simulations and virtual laboratory environments supporting scientific inquiry are available. Some of these include or embed formative assessment.
- Educational games are available and being used at school level to develop and (in some cases also) assess scientific inquiry skills and social competences in authentic contexts.

However:

- While online simulations, virtual laboratories and games are readily available, often for free, and are being successfully used by many teachers and learners, use of these tools has

45 http://vpa.gse.harvard.edu/a-case-study-of-the-virtual-performance-assessment-project/.

not yet become mainstreamed.

- Only few environments for scientific inquiry include or embed assessment.

To seize the opportunities offered by ICT:

- Existing environments for scientific inquiry and investigation should be further harmonised with curricula.
- Assessment should become integrated into virtual environments for scientific inquiry in such a way that the whole process of scientific investigation and the respective student's strategic approach is reflected.

Policy support is needed to:

- Encourage take-up and harmonisation of existing tools and environments.
- Encourage and enable teachers to make use of existing environments, to critically examine their value and to adapt them to their learners' needs.
- Set incentives for the development and deployment of assessment modules for environments for scientific inquiry which comprehensively assess skills and attitudes.

3.5 Digital competence

Box 5: Definition of "Digital competence"

Digital competence involves the confident and critical use of Information Society Technology (IST) for work, leisure and communication. It is underpinned by basic skills in ICT: the use of computers to retrieve, assess, store, produce, present and exchange information, and to communicate and participate in collaborative networks via the Internet.

Source: Recommendation of Key Competences for Lifelong Learning (Council of the European Union, 2006)

Digital Competence comprises a range of competences that are very different in nature, ranging from purely operational, technical skills, to collaborative and critical skills. Different Digital Competence frameworks, curricula and assessment approaches highlight on different aspects and work on a common European framework outlining the different dimensions and sub-competences has only just started (cf. Ferrari, 2012).

The very nature of digital competence invites for technology-based assessment formats. However, many of the most currently used assessment tools for digital competence employ a knowledge-based, traditional multiple choice format.

The European Computer Driving Licence (ECDL)[46], internationally known as ICDL, is one of the leading authorities of computer skills certification programmes. Certification is based on a traditional multiple-choice test, which is administered electronically in a certified test centre. Similarly, the Internet and Computing Core Certification (IC³®)[47], uses knowledge-based multiple-choice test format to assess and certify work-related ICT skills.

Against this background, it should be noted that, also for the assessment of digital competence more interesting and interactive assessment formats are possible and have been shown to be viable. A research study on **online peer assessment** with 10th graders in a computer course indicates that students significantly improved their projects based on the peer assessment activities and that peer assessment scores were highly correlated with those marked by the experts (Tseng & Tsai, 2007).

However, **practical tasks replicating real-life ICT use** are a more promising approach for adequately capturing the full scope of digital competence. As a study among 200 undergraduate students shows, who were asked to complete an **online scenario based ICT assessment** in which they were required to actually perform tasks associated with accessing, evaluating, integrating and communicating information, the majority of students performed below their own perceived level (Hilberg & Meiselwitz, 2008).

The Key stage 3 ICT tests (UK) are an example of an assessment format which requires (14 year old) students to use multiple ICT tools in concert (word processor, browser, spreadsheet) much in the same way these are used in real work and academic environments (Bennett, 2010). The project led to the development of extended, authentic tasks assigned to students completing tests of ICT skills in a virtual desktop environment (Ripley, 2009). Similarly, the iSkills[48] assessment intends to measure students' critical thinking and problem-solving skills in a digital environment. In a one-hour exam real-time, scenario-based tasks are presented that measure an individual's ability to navigate, critically evaluate and understand the wealth of information available through digital technology. The programme provides individual and group data for use in student evaluation and placement.

The national ICT skills assessment programme in Australia (cf. MCEECDYA, 2008) is designed as an authentic performance assessment, mirroring students' typical 'real world' use of ICT. In the 2005 and 2008 rounds of the assessment, students completed tasks on computers using software that

46 http://www.ecdl.org/programmes/index.jsp.

47 http://www.certiport.com/Portal/desktopdefault.aspx?tabid=229&roleid=102.

48 http://www.ets.org/iskills/.

included a seamless combination of simulated and live applications. Some tasks were automatically scored and others were stored and marked by human assessors. The tasks (items) were grouped in thematically linked modules each of which followed a narrative sequence covering a range of school-based and out-of-school based themes. Each module typically involved students collecting and appraising information as well as synthesising and reframing the information.

eVIVA[49] was an Ultralab research project (2002-2004) funded by the UK Qualifications and Curriculum Authority (QCA). eVIVA uses mobile phones, voice recognition technology and the internet to support formative and summative assessment. The teacher made a holistic assessment of the pupil's ICT capabilities based on the milestones and work submitted in the e-portfolio, student reflections or annotations (on their own and fellow students' work), the recorded eVIVA answers submitted via telephone and any written answers attached to the questions, and classroom observations (Binkley, et al., 2012).

These examples illustrate that the full scope of digital competence as a creative, collaborative and critical competence, can best be assessed if the test situation replicates every day ICT use and involves several tools, devices and programmes used in applied situations, in which several resources and solution strategies must be combined. The case of Norway illustrates that it is possible to implement ICT-based assessment formats that endeavour to capture the full scope of Digital Competence on a national or regional level.

Example 16: Testing digital competence among Norwegian schoolchildren

The Norwegian Centre for ICT in Education has been piloting tests for digital literacy in Oslo and Bergen, the two largest cities in Norway. The Oslo test has been conducted annually since 2008. Today, learners in the 4th grade and in the 12th grade (upper secondary education) take this test. In each grade, there are about 10 000 students. The tests assess the following competencies: basic digital skills, digital communication, digital information processing, digital production and critical thinking. The different elements of the test are weighed differently.

The Bergen test has been conducted annually since 2009. Today, learners in 5th grade take the test. Annually, approximately 5 000 students participate in the test. The Bergen test follows the same structure as the Oslo test. The development of a diagnostic test in digital skills started in late 2011, and the test is supposed to be have been finalized in March 2012. This test will be an optional test, but experiences from similar tests indicate that most students (about 60 000 per cohort) will take these tests. The diagnostic test is based on a newly developed framework developed by the Ministry of Education and Research and the Norwegian Directorate for Education and Training. The framework defines four "digital areas": To acquire and process digital information, produce and process digital information, critical thinking and digital communication.

The tests will typically comprise 50-70 questions distributed across the four areas. They are developed and implemented on different platforms with different functionalities. This means that test types, test operation and other aspects related to the tests vary. According to the Norwegian Centre for ICT in Education, there is a need to elaborate the test platform in order to capture the digital practice among learners, e.g. through simulation and interactive test items.

This test has been piloted at 30 schools with 800 students. The test items will be anchored in the national curriculum and related to a number of competence goals across various subject curricula. The choice of the 4th grade for the test is related to the importance of mapping the knowledge level, strengths and weakness of students, a type of early intervention principle. The 4th grade is also a relevant mapping level, because the learners will according to the curriculum start using digital tools in the 2nd grade.

The development and availability of test items is governed by a new framework for the basic skills in the Norwegian curriculum being implemented in 2012. The test items cover a spectrum of test tasks, however, according to the informant at the National Centre for ICT in Education, there is a lack of test items that require simulations and test items anchored in real world problems and challenges.

---

[49] http://www.ofqual.gov.uk/files/Eviva_Final_Report_2004-feb2008_copy.pdf.

In short: ICT for the assessment of digital competence

The particular strenghth of ICT for the assessment of this key competence lies in:

- Practical tasks replicating real-life ICT use, either embedded and documented in a virtual environment or undertaken in a real life (collaborative) context with the aid of technological devices and programmes.

Currently:

- If a technological approach is chosen for the assessment of digital competence, most commonly computer-based tests are employed. In particular, certification schemes, such as the European Computer Driving License, resort to multiple choice tests.
- There are a number of national trials and pilots for school education, in Europe and beyond, in which digital competence is being assessed through an authentic task which requires a number of digital devices and tools to be used.

However:

- Computer-based tests used in certification schemes for digital competence tend to focus on knowledge rather than (practical) skills and attitudes.
- More complex and authentic practical contexts assessing the applied use of digital tools have not yet become mainstream use. Deployment is currently hindered by the complexity of the assessment context to be used, which, in case of a virtual environment, requires substantial investments in setting up the corresponding virtual environment, and for real life tasks requires examiners to consider a high number of interactions and observations to be considered, increasing their work load.

To seize the opportunities offered by ICT policy support is needed to:

- Encourage further experimentation and research in the use of authentic tasks as a means of assessing digital competence, in view of generating good practice and mainstreaming existing efforts.
- Encourage and enable teachers and examiners to develop viable strategies for assessing digital competence in real life contexts.
- Firmly embed the need to assess digital competences including also other key elements such a creativity, collaboration, and critical competence, in practical contexts in curricula.
    - Develop a common understanding across all Member States of the key elements and building blocks of the concept of digital competence and how this competence can and should be assessed.

3.6 Learning to learn

Box 6: Definition of "Learning to Learn"

'Learning to learn' is the ability to pursue and persist in learning, to organise one's own learning, including through effective management of time and information, both individually and in groups. This competence includes awareness of one's learning process and needs, identifying available opportunities, and the ability to overcome obstacles in order to learn successfully. This competence means gaining, processing and assimilating new knowledge and skills as well as seeking and making use of guidance. Learning to learn engages learners to build on prior learning and life experiences in order to use and apply knowledge and skills in a variety of contexts: at home, at work, in education and training. Motivation and confidence are crucial to an individual's competence.

Source: Recommendation of Key Competences for Lifelong Learning (Council of the European Union, 2006)

There is no evidence on the (de facto or potential) use of ICT for the assessment of Learning to Learn. However, one of the strengths of ICT is to actively involve the learner in the learner process. ICT can thus be used to support self-regulated learning processes, through **self-assessment** and peer assessment in particular, which, in turn, contribute to developing students' learning to learn abilities.

There is a general awareness among school educators that ICT can help students to self-assess by providing them with immediate feedback on their performance and by allowing the sharing of information. In a number of European countries, ICT are already being used to support self-assessment. Liechtenstein, for example uses ICT tools for self-assessment in secondary education; Bulgaria, Lithuania and Iceland have pilot projects, while France, Malta and Slovenia plan the use of self-assessment (Eurydice, 2011a).

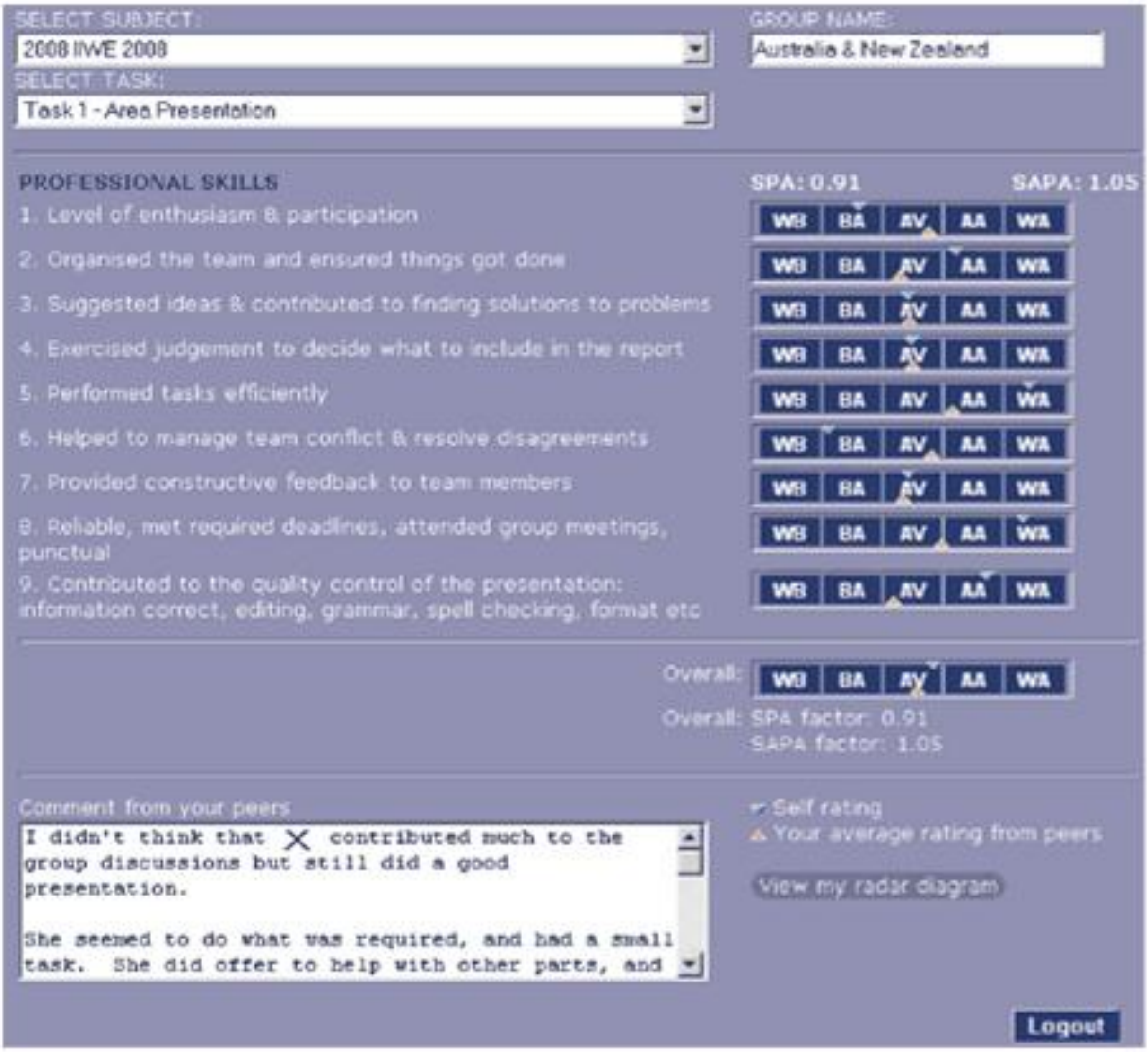
Self-assessment is an important means for encouraging self-awareness, self directed learning and fostering motivation. Additionally, self assessment also provides teachers a view of what students need or their perception of needs (Mortazavi, 2010). A study in higher education suggests that self-assessment exercises provide students the opportunity to reflect on the course and their performance, help them monitor their own progress, motivate them to do well in the course, and provide them the opportunity to give feedback to the instructor (Walser, 2009). Research shows that adolescents' self-concept significantly influences achievement motivation, motivation for creativity and partially also on academic achievement (Ďuricová, 2009). Self-assessment can enable adolescent students to better understand their particular strengths and learning needs and to pro-actively engage in learning.

ICT can assist in implementing self-assessment strategies which promote the active development of learning to learn skills in several ways.

Online quizzes as a form of individual self-assessment can help students' awareness of their own ability, ideally leading to independent self study (Campbell & Gorra, 2009). Such web-based systems for testing and self-assessment can also be implemented with open technologies (Caric, Tuba, & Moisil, 2010). A study on the use of electronic self-assessment tools among univaersity students indicates that those students that used interactive self-assessment tools performed better than those that did not (Ibabe & Jauregizar, 2010).

Mobile assessment systems that have recently become available foster more flexible assessment arrangements, more efficient use of time, and more opportunities for student reflection on learning and assessment (C. h. Chen, 2010). Mobile applications supporting self-assessment have been shown to improves the achievement of high school and university students, especially amongst younger learners, with a relatively low impact on teaching activities and methodology (De-Marcos, et al., 2010).

Using self- and peer assessment in conjunction with **collaborative** peer learning activities improves student engagement and can lead to self-reflection and the experimentation with different learning strategies. The regular use of self and peer assessment in different contexts can promote effective peer learning, increase engagement and encourage students to learn (Willey & Gardner, 2010).



Example 17: SparkPlus peer assessment tool

Source: http://spark.uts.edu.au/.

SparkPlus[50], for example, is a web-based self and peer assessment kit. It enables students to confidentially rate their own and their peers' contributions to a team task or individual submissions and to improve their judgement through benchmarking exercises (Willey & Gardner, 2010).

PeerWise[51], for example, is a collaborative web-based system that engages students in the creation of a test bank of multiple-choice questions. Thus students encouraged not only to test their knowledge, but also to think of and create relevant questions for a given subject matter, thus actively developing their learning strategies and correcting one another on test questions that might be formulated erroneously. Studies involving large university courses in New Zealand and the US indicate that using PeerWise improves exam performance (Denny, Hanks, & Simon, 2010; Denny, Luxton-Reilly, & Hamer, 2008).

[50] http://spark.uts.edu.au/.

[51] http://peerwise.cs.auckland.ac.nz/.

Simliarly, StudySieve[52] is an online tool in which students create and share assessment questions with their peers. The submitted questions can be answered, and all answers are visible to the entire community. Both the questions and answers are evaluated by students in a double-blind peer review process (Luxton-Reilly, Plimmer, & Sheehan, 2010).

ePortfolios can be used to foster reflection, self-assessment, continuous improvement, goal setting, problem solving, data gathering and peer interaction, and thus improve (secondary school) students' performance (Chang & Tseng, 2009). A study on the use of the European Language Portfolio (ELP) in English for Specific Purposes (ESP) classes in the higher education indicates that self-confidence, self-reflection and self- assessment are improved (Hismanoglu & Hismanoglu, 2010). Similarly, a study on the use of ePortfolios by grade 8 students taking senior high school computer courses revealed no significant effect on student achievement, but a statistically positive effect on self-perceived learning performance (Chang & Tseng, 2011).

As with peer assessment, research findings on the coherence of self-assessment and teacher assessment on Portfolio performance are mixed. A comparative analysis of teacher-assessment, student self-assessment and peer-assessment in a Web-based portfolio assessment environment for high school students revealed significant differences, but indicated that self-assessment was highly consistent with teacher-assessment and with end-of-course examinations (Chang, Tseng, & Lou). Other studies indicate that self-assessment scores are not consistent with the expert's scores, while peer assessment scores demonstrate adequate validity (Liang & Tsai, 2010). For example, the study on the use of ePortfolios by grade 8 students in a computer course quoted above, revealed significant differences in teacher-assessment and self-assessment (Chang & Tseng, 2011). However, on the whole, research findings indicate that self-assessment and self-reflection enable learners to better understand their strengths, weaknesses and learning leads and to proactively engage in learning activities that improve their performance.

In short: ICT for the assessment of "Learning to Learn"

The particular strenghth of ICT for the assessment of this key competence lies in:

- Providing environments and tools, such as ePortfolios, quizzes, and tutoring systems, which support self- and peer-assessment and thus contribute to fostering self-regulated learning.

Currently:

- Learning to learn is a transversal competence that is not usually explicitly assessed. Consequently, there are no computer-based assessment tools being developed or used to assess this competence.
- ePortfolios are often used as a means of encouraging learners to reflect upon their learning process, to self-assess their performance, provide feedback to peers and react upon peer-assessment received. In this sense, one can say that ePortfolios are being used as a means of fostering and assessing learning to learn.

To seize the opportunities offered by ICT:

- Other computer-based assessment tools and environments, such as intelligent tutoring systems, virtual laboratories and game environments, should better integrate self- and peer-assessment options to encourage learners to reflect upon their learning process and develop viable learning strategies.

Policy support is needed to:

[52] http://www.cs.auckland.ac.nz/courses/compsci111s1c/studysieve-web-documentation.pdf.

- Entice a reflection on the meaning and importance of Learning to Learn to encourage teachers, learners and software developers to adequately reflect this vital competence in their assessment strategies.
- Encourage research and development on assessment tools and environments that foster learning to learn by supplying collaborative assessment contexts and encouraging self-assessment and reflection.

3.7 Social and civic competences

Box 7: Definition of "Social and civic competences"
"[Social and civic competences] include personal, interpersonal and intercultural competence and cover all forms of behaviour that equip individuals to participate in an effective and constructive way in social and working life, and particularly in increasingly diverse societies, and to resolve conflict where necessary. Civic competence equips individuals to fully participate in civic life, based on knowledge of social and political concepts and structures and a commitment to active and democratic participation.
Source: Recommendation of Key Competences for Lifelong Learning (Council of the European Union, 2006)

There is limited evidence on the use of ICT to assess social and civic competences. However, looking at examples from the research literature, three potentially promising strategies for assessing this key competence emerge, in order of increasing importance: Psychometric tests, complex real-life problems which are to be solved using online resources, and game-environments, with virtually replicate real-life social and global challenges, which the player is asked to solve.

Psychometric tests aim to assess a person's behavioural and personality profile, usually in view of a specific job profile. Popular psychometric tests, often used in workplaces are the Myers-Briggs test, the Hermann Brain Dominance Instrument (HBDI), and the DISC Assessment. Many psychometric tools address inter-personal and communication skills. 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 (Binkley, et al., 2012).

Currently, psychometric tests are pre-dominantly used in workplaces and for recruitment. The Occupational Personality Questionnaire (OPQ), a commercial test used in the US, for example, seeks to measure likely occupational behaviours in the areas of relationships with people, thinking style, feeling and emotions.[53] In Europe, for example the European Commission's European Personnel Selection Office (EPSO) employs a "situational judgment test" for pre-selecting potential candidates for recruitment to the European Commission, which aims to assess candidates' behaviour in typical work situations, assessing, among others, skills such as communicating, learning and development, working with others, and leadership[54].

The wide use of psychometric assessments for recruitment has motivated significant research into the development of psychometric testing procedures which can provide accurate and efficient estimates of the parameters of interest (Remus & Collins, 2008). However, the very nature of these tests, designed to assess a person's invariable personality traits, makes their use in Education and Training questionable, since here the student's performance is considered a dynamic process which is (ideally) characterized by continuous competence improvement and refinement.

In Education and Training, psychometric tests are currently used, mainly for diagnostic purposes. Toolfind[55] is an example of a database with 46 test tools for teachers and educators which assess competences in eleven outcomes areas including problem-solving, positive behaviour (self-control, cooperation, conduct in school, responsibility), leadership, learning orientation (motivation, persistence, study habits), and academic skills. These tools could be a starting point for developing and critically discussing whether and how social competences can and should be assessed in such a format.

Practical tasks, embedded in online environments or making available online resources, are a more promising avenue for investigation, as these allow more complex social behaviour to be observed and assessed. Unfortunately, there are not many examples illustrating this approach. The College Work and Readiness Assessment (CWRA) is such a test. It was introduced in St. Andrew's

[53] http://www.shl.com/WhatWeDo/SHLReports/default.aspx.
[54] http://europa.eu/epso/doc/selection_procedure_en.pdf.
[55] http://www.toolfind.org/

School in Delaware to test students' readiness for college and work, and it quickly spread to other schools across the US It consists of a single 90-minute task that students must accomplish by using a library of on-line documents, from one-page newspaper editorials to 20 page research reports. Students must address real-world dilemmas (like helping a town reduce pollution), making judgments that have economic, social, and environmental implications, and articulate a solution in writing.

Online games take this idea one step further, by transposing the complete social and political context of the task at hand into a virtual environment and asking the student to solve the problem posed and to adequately react towards changing circumstances. There are a number of examples illustrating this approach, although none of these are currently used as formal assessment tasks.

Example 18: Global Conflicts
Source: http://www.globalconflicts.eu/.

Radford Outdoor Augmented Reality (ROAR)[56] is an augmented reality game which uses Augmented Reality to help teach K-12 students more about Native American history and teamwork through a game called Buffalo Hunt. Global Conflicts[57] is an educational game designed to help teach concepts in citizenship, geography, and media. Developed by Serious Games International, it has detailed lesson plans and assignments for students.

Mass Extinction[58] is a game on climate change which took place in the spring of 2011 and was developed by MIT's Education Arcade. PeaceMaker Game[59] is designed to teach concepts in diplomacy and foreign relations. The game allows the player to take on the role of either the Israeli Prime Minister or Palestinian President, trying to find peaceful resolutions to conflicts before the term of office expires.

Most of these games are educational in approach and aim to raise awareness for certain social and global conflicts and challenges. Most of them allow students and teachers to assess how well the problem at hand was solved, and in some cases alternative solution strategies are offered and

56 http://gameslab.radford.edu/ROAR.html.
57 http://www.globalconflicts.eu/.
58 http://shass.mit.edu/research/cms_game.
59 http://www.peacemakergame.com/game.php.

discussed. They therefore, arguably, do support formative and even summative assessment. However, as concerns the games' set-up, this is a side-effect, rather than an objective.

In short: ICT for the assessment of social and civic competences

The particular strenghth of ICT for the assessment of this key competence lies in:

- Multiplayer games, which embed learning and assessment in authentic real-life contexts, which replicates social and political conflicts, thus supporting the development and assessment of social and civic competences.

Currently:

- ICT are not commonly used for the assessment of social and civic competences.
- Multiplayer games presenting a social or political conflict in a 3D virtual environment and requiring players to collaboratively solve the conflict are being experimented with in educational contexts.

However:

- Games are not considered a serious environment for learning. If they are used at all, they are employed outside the curriculum and without assessing students' performance or learning gains.

To seize the opportunities offered by ICT:

- A greater number and variety of educational games supporting the development of social and civic competences is needed.
- Educational games supporting the development of social and civic skills should be more directly linked to curricula and support the core learning objectives.
- Educational games should include features allowing for the formative and/or summative assessment of student's performance, based on peer- and self-assessment as well as on the actual behaviour as recorded by the electronic environment.

Policy support is needed to:

- Encourage the development and dissemination of educational games for the development and assessment of social and civic competences.

3.8 Sense of Initiative and Entrepreneurship

Box 8: Definition of "Sense of initiative and entrepreneurship"

Sense of initiative and entrepreneurship refers to an individual's ability to turn ideas into action. It includes creativity, innovation and risk-taking, as well as the ability to plan and manage projects in order to achieve objectives. This supports individuals, not only in their everyday lives at home and in society, but also in the workplace in being aware of the context of their work and being able to seize opportunities, and is a foundation for more specific skills and knowledge needed by those establishing or contributing to social or commercial activity. This should include awareness of ethical values and promote good governance.

Source: Recommendation of Key Competences for Lifelong Learning (Council of the European Union, 2006)

As with "Learning to Learn" evidence on how to foster and assess sense of initiatve and entrepreneurship is scarce. Being a transversal skill, there is a lack of evidence and insight on how this key competence is, can and should be integrated in curricula and taught in schools. The situation is not much better in cases where "enterprise education" has been introduced in schools as a dedicated subject as there is a concern that the "delivery" of enterprise education takes place in ways which are not "enterprising" forms of learning, and that changes to definitions, frameworks and pedagogy are needed to clarify its future educational role (Draycott & Rae, 2011).

There are, of course, computer-based tests, such as the GMAT test, which regulate entrance into business study programmes. However, these tests focus on candidates' abstract, numerical and verbal reasoning skills, rather than their entrepreneurial spirit.

Example 19: CityOne and Innov8

Innov8 and CityOne are simulation games, developed by IBM to teach business students and those working in businesses and municipalities to effectively manage complexity. Both games provide continuous feedback.

INNOV8, the IBM Business Process Management (BPM) 3-D simulation game, intends to give IT and business players a better understanding of how effective BPM impacts an entire business ecosystem. Players see how practical process improvements can help meet profitability, customer satisfaction and environmental goals while addressing real problems faced by municipalities and businesses, as concerns traffic, customer service and supply chain management. The game features a fictional call center agency, who has a process model that is functioning sub-optimally. As the protagonist Logan, the learner must discover the current model, find out why it is under-performing and then optimize it to meet the demands of the market.

According to ICM: "Over 1000 schools worldwide have downloaded the game and more than 100 universities worldwide have built custom curriculum using our serious game to help students learn about business process management and SOA. One study found that a great lecture can improve learning by 17% but serious games can improve learning by 108%."

In CityOne, the player's mission is to solve real-world business, environmental and logistical problems. The objective is to understand how technology can revolutionize these industries, explore ways to accelerate process change, integrate with trading partners, and control costs with a flexible IT infrastructure. The key learning objective is to understand how Business Process Management, Collaborative Technologies, and Service Oriented Architecture enable industry solutions that help organizations and industries adapt to new demands and build a sustainable advantage.

Source: http://www-01.ibm.com/software/solutions/soa/innov8/full.html; http://www-01.ibm.com/software/solutions/soa/innov8/cityone/

In some instances **games** have been shown to be useful tools for fostering initiative and entrepreneurship. A sports business professor at the University of Oregon, for example, has taken a commercial game, Madden NFL, and used one of its modes for developing football franchises to help teach students about marketing and business decisions.[60] IBM has developed games for business students to better understand the complexity of business decisions and the learn to optimize business processes (cf. Example 19).

Apart from strategy games, which could, at least in principle, be used as a tool to teach assess initiative and entrepreneurship, computer-based **psychometric tests** are another set of tools that can assess behavioural and attitudinal personal characteristic, such as, among others, initiative and entrepreneurship.

As with other social competences, skills related to sense of initiative and entrepreneurship, such as leadership, organising, managing and planning, risk-taking, pro-active attitude, etc., are already being tested in psychometric tests that are currently being used to assess a candidate's suitability for a certain job profile. For instance, the Myers-Briggs Type Indicator, distinguishes between extraversion and introversion (Binkley, et al., 2012); the EPSO situational judgement test assesses skills which contribute or are related to entrepreneurship, such as prioritising and organising, resilience and leadership.

For example, the "Enterprize™ Questionnaire"[61] aimed at identifying innovative people is a commercial psychometric test offered to companies to predicts a person's capability to contribute towards innovation within their organisation. On a five-point Likert scale the test assesses an individual's attributes in: Innovation, creativity, and imagination; Opportunistic behaviour and initiative within the workplace; Commitment and the desire to prove one's self; Risk tolerance and risk management; Leadership and the ability to inspire others . The entrepreneur questionnaire takes approximately 20 minutes to complete, after which a comprehensive personalised report is available.

Entrepreneurship

Please read each statement and choose to what extent you agree or disagree with each statement on a 5 point scale or 'Don't understand the question':

	STRONGLY AGREE	AGREE	NEITHER AGREE OR DISAGREE	DISAGREE	STRONGLY DISAGREE	DON'T UNDERSTAND THE QUESTION
1. Just because they built a business, entrepreneurs cannot remain independent	○	○	○	○	○	○
2. An organisation can't afford to have people running around with different ideas - they need to be controlled by management	○	○	○	○	○	○
3. Innovation should be part of corporate strategy	○	○	○	○	○	○
4. I like to encourage others to have new ideas	○	○	○	○	○	○
5. A structured group approach is very helpful in testing new ideas	○	○	○	○	○	○
6. Once a business exceeds a certain size, it is best if the business founder hands it over to professional managers	○	○	○	○	○	○
7. "Ideas people" should be taken off new projects if they cannot demonstrate quick success	○	○	○	◉	○	○
8. Being a management consultant is a lonely, tough life	○	○	○	○	○	○

Example 20: Enterprise Questionnaire
Source: http://careeringahead.com.au/entre.asp.

60 http://it.uoregon.edu/itconnections/playing-for-agood-grade.
61 http://www.psychpress.com/psychometric/talent-psychometric-testing.asp?entrepreneurship.

Obviously, such instruments will need to undergo profound changes to be used with a younger student population. Also, the concepts of initiative and entrepreneurship would have to be clearer defined to be susceptible to assessment in such a format. A critical discussion and sound research of this test approach might even reveal that it is less suited to reliably capture the core of this key competence. Psychometric tests which are currently used for diagnostic purposes within Education and Training may serve as a starting point for developing and critically discussing how initiative and entrepreneurship can possibly be assessed, with or without the use of ICT.

In short: ICT for the assessment of "Sense of Initiative and Entrepreneurship"

The particular strenghth of ICT for the assessment of this key competence lies in:

- Providing environments and tools, such as multiplayer games, in which entrepreneurial skills can be developed and assessed.

Currently:

- Sense of initiative and entrepreneurship is a transversal competence that is not usually explicitly assessed. Consequently, there are no computer-based assessment tools being developed or used to assess this competence.
- Existing psychometric tests, which assess entrepreneurial skills in a multiple-choice format, are too static to serve the purpose of assessing this competence in a school context.
- A range of commercial multiplayer games encourages initiative and entrepreneurship. However, these games are not developed for or used in educational contexts.

To seize the opportunities offered by ICT:

- Educational games should be developed, on the basis of existing multiplayer games, which encourage the entrepreneurial spirit and decision making and, through self-reflection and peer-feedback allows for the assessment of this competence.

Policy support is needed to:

- Encourage the adaptation and development of multiplayer games that foster and assess sense of initiative and entrepreneurship and can be embedded in a school context.

3.9 Cultural awareness and expression

Box 9: Definition of "Cultural awareness and expression"

Appreciation of the importance of the creative expression of ideas, experiences and emotions in a range of media,including music, performing arts, literature, and the visual arts. […]Cultural knowledge includes an awareness of local, national and European cultural heritage and their place in the world. It covers a basic knowledge of major cultural works, including popular contemporary culture. It is essential tounderstand the cultural and linguistic diversity in Europe and other regions of the world, the need to preserve it and the importance of aesthetic factors in daily life.

Source: Recommendation of Key Competences for Lifelong Learning (Council of the European Union, 2006)

There is a general consensus that interactive media tools can contribute to realising discovery-based activities, cooperative and collaborative learning strategies, and new forms of artistic expression (Cuthbertson, et al., 2007). However, arts student-oriented computer education is a difficult task for most education institutions and research on pedagogical approaches for using ICT in arts education is scarce (Cao & Wang, 2009).

Thus, while new media give rise to new forms of artistic expression previously unaccessible to students, such as video and web-content production, there is a lack of evidence on how ICT can be used to support the assessment of these and other forms of artistic expression. Similarly, while ICT makes it easier for students to obtain historic and cultural information, more insight is needed on how a critical, reflective and active use of this information can be promoted and eventually also be assessed.

An obvious and very suitable assessment format for cultural and artistic expression are **ePortfolios**. In the ePortfolio, students can present and annotate the artefacts and insights they have generated or collected; comment, assess and conclude on their learning process; and showcase their work, to their teachers and peers or to external examiners. ePortfolios can also be used to schaffold exam-like assessment situations, replacing traditional, papar-based assessment formats with exam formats that are more suited to foster creative and artistic production, as the e-scape case illustrates (cf. Example 2, p. 27). In the e-scape project, a 6 hour collaborative design workshop replaced England's school examinations for 16 year-old students in Design and Technology. Students work individually, but within a group context, to build their design solution. Students are given a number of staged assessment instructions and information via a personal, portable device (PDA). The handheld device also acts as the tool to capture assessment evidence – via video, camera, voice, sketchpad and keyboard. During the 6 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 multi-media portfolio which is loaded to a secure website. The worked is scored using Thurstone's graded pairs. The project involved eleven schools across England and resulted in 250 performance portfolios. The reliability of the assessment method was very high (Binkley, et al., 2012; Ripley, 2009).

Games are a another promising avenue for combining teaching and assessing cultural awareness and expression while involving students actively in the development and critical use of their knowledge. Melody Mixer is a game developed at the University of Wisconsin-Madison that teaches music students how to read and compose music. It encourages students to experiment with sound and composition to better learn how pieces are constructed (L. Johnson, et al., 2011). Gamestar Mechanic[62] is a learning environment that encourages students to design their own video games, practicing systems thinking, problem solving, critical thinking, iterative design, creativity, collaboration.

62 www.gamestarmechanic.com.

Ghosts of a Chance[63] is a game which allowed visitors to the Smithsonian American Art Museum a chance to decipher codes, follow treasure maps, send text messages, and uncover hidden objects in this multimedia scavenger hunt. Its follow-up, "Pheon"[64] is a variation on the classic game Capture the Flag, in which players compete in obtaining the game's virtual talisman, the pheon, to restore balance to a virtual world called Terra Tectus. Players first determine their alliance to one of two groups and then complete various missions focused on the museum's art collections, exhibitions, and programs to earn points and propel the game. Pheon can be played on-site or online and like other mission-based games, revolves around the completion of tasks, the making of objects, discovery, and documentation.

Example 21: Pheon

Source: http://pheon.org; accessed via Facebook.

Far more research and experimentation is needed to better understand how ICT can foster the assessment of cultural awareness and expression, in particular new media are very powerful tools in promoting a creative and productive interaction with cultural heritage and foster students' artistic expression by allowing for a far greater variety and mix of genres and formats than traditional media.

63 http://www.ghostsofachance.com/.

64 http://americanart.si.edu/multimedia/games/pheon.cfm; http://pheon.org/.

In short: ICT for the assessment of cultural awareness and expression

The particular strenghth of ICT for the assessment of this key competence lies in:

- ePortfolios which allow students to produce, collect and reflect upon cultural artefacts in a range of different formats.
- Virtual reality environments which replicate (and annotate) historic or contemporary reality.

Currently:

- ePortfolios are already being used widely as a means of showcasing cultural artefacts produced by students.
- Augmented reality environments, such as virtual museums or cities, exist which allow students to explore other cultures and civilisations.
- There is some experimentation with virtual games that encourage learners to engage with cultural artefacts and/or historic events in an engaging way.

However:

- Virtual cities and museums as well as online games are not primarily conceived as educational tools for schools and therefore do not consider assessment options.
- The fact that ePortfolios are used for the assessment of cultural artefacts produced by students does not always lead to a greater range of modes of cultural expression being exploited by learners and teachers. Not always are more innovative formats for cultural expression, such as videos, photos, recordings, digitally produced and/or manipulated images and audio file, etc. included.

To seize the opportunities offered by ICT:

- Tools should be developed that make learners' engagement with virtual reality environments assessable.
- More research is needed on how to improve the valdity and effectiveness of portfolio grading, which is often perceived by examiners as a tedious process susceptible to selective and subjective judgements.

Policy support is needed to:

- Provide guidance and support on how artefacts collected in ePortfolios can and should be assessed, including self- and peer assessment.

4. CHALLENGES

This report has highlighted a number of promising avenues for the use of ICT in supporting the assessment of Key Competences. However, it has also become clear, that, in most cases, these innovative strategies do not reflect every day classroom practice. On the one hand, it has to be noted that technology is still advancing at accelerating speeds, while assessment in education institutions underlies rigorous and strict curricular constraints, which cannot keep pace with these developments. It is therefore not surprising that mainstream assessment practice is only slowly realising the potential benefits of ICT. Rather, it is encouraging that on small-scale and grass-root level, ICT are being embraced by teachers and school leaders as a means of supporting more competence-based assessment, as the wide-spread use of ePortfolios, peer-assessment and self-assessment illustrates.

On the other hand, even for technology-based assessment formats that have been available for quite some time and are widely used for assessment purposes outside primary and secondary education institutions, such as Computer-Based Tests, take-up in European schools has been slow. As concerns test formats that replicate traditional tests in an electronic environment the general scepticism with which ICT is greeted might even be a blessing. As the introduction of ICT in higher education shows, too much technological enthusiasm often distracts from asking crucial questions about the pedagogical approach and usefulness of different ICT-enabled environments and test formats.

Thus, one major concern against the introduction of ICT for assessment in schools, namely that the assessment format determines and limits what is assessed and that more complex and applied competences cannot be assessed by a rather reduced multiple-choice format, deserves careful consideration. Especially when moving from knowledge-based to competence-based assessment, more traditional forms of e-Assessment may not be adequate tools for assessment. Care must be taken that the ICT tools, applications and environments selected are suitable to adequately and comprehensively assess the competences at stake.

However, the introduction of ICT in schools has in the majority of cases discussed above been carefully reflected against pedagogical criteria and objectives. As the examples highlight, ICT is in many cases conceived as a means of promoting more innovative and engaging learning and assessment formats that more adequately address and measure students' needs with a view to offering each one of them the support they need. Thus the question does arise: What hinders the more widespread use of ICT for assessment?

Access to satisfactory ICT infrastructure has in the past been one of the most important factors contributing to the effective use of information technologies in all subjects and for all students. While some infrastructure problems persist, which are delaying the integration of new technologies into teaching and learning, by now ICT is generally and readily available in European schools (Eurydice, 2011a). Thus, infrastructure problems that in the early days of ICT use in European schools posed a major obstacle to wider deployment are currently being overcome. Moreover, many European countries have launched so-called "1-to-1 learning" initiatives that promote each students' access to their own mobile personal computing device. While internet connections are not always available to adequately support 1-to-1 learning strategies, infrastructure constraints, on the whole, play a minor role in hindering ICT take-up for assessment.

However, a lack of educational software and support staff poses a major obstacle to the effective and efficient use of ICT in schools in Europe (Eurydice, 2011a). Not all countries offer a high number and variety of educational tools supporting assessment. The educational applications, games and learning environments that are available for free or at a low cost are often still limited in their functionalities and distributed on different platforms, in different formats, requiring teachers to invest time and effort to identify adequate tools for each specific learning objective. Often, loading the programme, understanding its functionalities and navigating within the environment require more effort from the student than actually answering the questions posed, thus distracting from the main task at hand.

More funding, policy support and guidance are needed to provide viable learning environments that answer in a more holistic manner to educators' needs. Take-up also lags behind, because the development of technological environments that support the assessment of more complex competences is an expensive and time-consuming process in which software developers and educators need to collaborate.

One further obstacle prevailing still today is the time involved, in particular on the part of the teacher, in integrating computer-based assessment in teaching and learning. Teachers and learners need more support in acquiring the necessary skills to effectively use the full range of ICT tools available to support the assessment of Key Competences. More senior teachers, in particular, need additional support and encouragement to engage with technologies and realise their benefit for their teaching.

As has become clear above, the main objective of the use of ICT for the assessment of Key Competences cannot and should not be to replace direct human interaction with computer mediated communication. One key concern that might hinder the take-up of promising ICT-strategies for assessment is the fear that virtual environments could replace real life interaction and that schools might lose sight of their ultimate objective, namely successfully integrating young people in the real world and enabling them to play an active and productive role as European citizens in the 21st century. In fact, especially for younger students, the direct interaction with teachers and peers is very important and teachers should be advised to ensure that ICT is used and conceived of as a tool for promoting these relationships, rather than substituting them. Thus, in primary and secondary education, face-to-face instruction guided by teachers and embedded in a social environment of peers is to be preferred over distance education.

However, ICT assessment approaches in school education usually take these observations into account and are, in their majority, designed to support and further enhance the relationship to teachers and peers, by allowing teachers to better monitor student achievement and progress and thus provide more adequate and targeted feedback and by supporting collaborative work. Teachers, students and learners need to be better informed and advised on the benefits, drawbacks and dangers involved, to better understand how ICT can contribute to improving and enhancing school education.

The main challenge to the use of ICT for the assessment of Key Competences is the lack of software solutions – electronic tools, programs or environments – that are comprehensive and versatile enough to support curricular based classroom teaching and learning. Furthermore, many of the more promising programs and environments for the assessment of Key Competences are still experimental in scope and have not become mainstream in education and training. In particular, learning analytics and embedded assessment, which are expected to become the most promising technological innovations for the assessment of Key Competences, have not yet matured and have not widely or critically been studied.

A further threat to seizing the potential of ICT for assessment is the fact that currently deployment and research are distributed unequally across the different Key Competences. While there are a range and number of ICT-based learning and assessment environments available for basic literacy, numeracy and scientific investigation skills, there is a lack of tools and also a lack of research efforts on other Key Competences.

For example, there is a need for more research on strategies – whether ICT-based or not – for the assessment of *social and civic competences*, *learning and learn* and *sense of initiative and entrepreneurship*. *Digital competence*, on the contrary, is being assessed by several national and international certification schemes and standards. However, these assessment schemes employ computer-based tests which test candidates' knowledge rather than their practical skills. In other areas, such as in *communication in foreign languages*, research efforts are directed at the automatic scoring of spoken language to increase the efficiency of the testing process, whereas the potential of automatic correction of written text, which can support (self-)assessment and foster self-regulated learning is not being given the same attention.

5. CONCLUSIONS AND POLICY RECOMMENDATIONS

Information and Communication Technologies (ICT) offer many opportunities for supporting assessment formats that can capture complex competences otherwise difficult to assess. There is a vast range of formats and approaches which could foster different aspects of key competence development and could be used to address the specifics of each key competence in a targeted way. However, take-up and implementation in school education is still low. To seize the opportunities offered by ICT, targeted measures are needed to encourage the development, deployment and large-scale implementation of innovative assessment formats in school education.

The state of the art

Computer-Based Testing is used widely and successfully for the summative and formative assessment of basic literacy and advanced reading skills and basic mathematical skills. In particular, literacy and numeracy skills are being assessed in a range of national and international tests, which are in many cases electronic. Due to the nature of these competences, it has been possible to embed complex and authentic tasks in the multiple-choice format, so that mathematical competence can be comprehensively and reliably assessed by computer-based tests.

In general, however, computer-based tests tend to replicate traditional assessment formats as they focus on knowledge, rather than skills and attitudes. Furthermore, they are not usually employed as a means of supporting more personalised, engaging, collaborative or authentic tasks. The advantage of computer-based tests over traditional assessment formats is that they provide instant and targeted feedback and can automatically adapt the difficulty of the test items to learners' different performance levels, to support formative assessment. Their potential, however, is currently untapped.

The internet is a vast resource for free and commercial **computer-based quizzes, games and tests** which can be used for the development and assessment of competences in literacy, reading and text comprehension and mathematics, in primary and secondary education. However, games, quizzes and engaging test formats are currently scattered, isolated, limited in scope, and ill-suited to comprehensive use in curricula and teaching.

ePortfolios are ideally suited to the assessment of collections of work produced by students and are thus particularly powerful tools for communication *in the mother tongue, communication in foreign languages* and *cultural awareness and expression*. ePortfolios are already being used widely in European schools for the formative and summative assessment of students' creative productions. However, more innovative formats of cultural and artistic expression, such as blogs, wikis, tweets, audio and video recordings, etc., are seldom included. Educators often do not realize that ePortfolios can also be powerful tools for fostering online collaboration and self- and peer assessment, which contribute to fostering and at the same time assessing students' *learning to learn* skills.

Technology-enhanced learning environments are often used in higher education and are starting to be deployed in school education as well. They are used by some schools as a means of creating learning situations which require complex thinking, problem-solving and collaboration strategies. Some of these environments allow learners and teachers to assess performance, understand mistakes and learn from them. The use made of these tools depends highly on individual teachers' intentions.

Immersive environments and multiplayer games recreate learning situations which require complex thinking, problem-solving and collaboration strategies. They also encourage the development of these skills, which are key components of all eight Key Competences. These environments replicate authentic contexts; encourage collaboration, empathy and negotiation; and reward strategic thinking, initiative and experimentation. They are thus specifically suitable for *competences in science and technology*, for *social and civic competences* and the development of *sense of initiative and entrepreneurship*. Since learners' behaviour in these electronic environments is tracked, their individual learning journeys – and their possession of these skills – can, at least in principle, be automatically assessed.

Online simulations, virtual laboratories and games fostering competences in *science* are readily available, often for free. Though they are being successfully used by many teachers and learners, this use has not yet been mainstreamed. Only a few environments for scientific inquiry include or embed assessment. Multiplayer games, presenting natural catastrophes or a social or political conflict in a 3D virtual environment and requiring players to collaboratively investigate and solve the problem, are being experimented with in educational contexts as a means of fostering scientific inquiry skills and social and civic competences. However, these games are usually employed outside the curriculum, with no assessment of students' performance or learning gains.

Educational software solutions such as **intelligent tutoring systems** take this idea one step further by offering embedded assessment with instant feedback and targeted support. In particular for *mathematical competence* these systems allow students to investigate mathematical concepts and problems in complex contexts. Students complete a series of tasks adapted in level of difficulty at their own pace, with helpful hints that encourage them to develop adequate solution strategies. Whereas these tools are popular in the US, they are not widely used in Europe.

Learning Analytics is one of the most promising emerging technological trends for the comprehensive assessment of complex competences. Learning Analytics interprets a wide range of data produced by and gathered on behalf of students in electronic environments in order to assess progress, and tailor education to individual students more effectively. Learning Analytics could allow assessment to be embedded in immersive environments, multiplayer games and computer simulations thus leveraging the potential of these tools in providing complex and authentic contexts in which Key Competences can be acquired and displayed.

Ways Ahead

There are a number of different challenges to be addressed, if we are to reap the benefits of ICT for the assessment of Key Competences:

Technological research and development should be focused on the most promising emerging techniques supporting comprehensive competence-based assessment. These are:

- Learning analytics which will enable educators to embed assessment into engaging virtual environments, such as multiplayer games, online simulations and virtual laboratories;
- Educational software which continually assesses the learning process and provides targeted feedback. This could be used to personalise the learning process to learners' individual needs and strengths and encourage self-regulated learning.

Development, deployment and implementation of existing technological solutions should focus on:

- Increasing scope, variability and curricula fit of existing tools. Improve their usability by allowing teachers and learners to adapt assessment tasks to their needs;
- Developing complex and authentic assessment tasks;
- Implementing self- and peer-assessment options.

Pedagogical strategies including the use of ICT for the assessment of Key Competences should:

- Choose assessment formats that encourage alternative solutions and promote experimentation;
- Promote self-regulated learning through self- and peer-assessment;

- Create learning contexts that allow learners to express themselves across a range of media and communication formats, experiment with different search and research strategies, and use ICT-based assessment as a means of making these more innovative learning strategies more readily assessable.

Policy Recommendations

To encourage the take-up of available tools and applications in schools, more policy support and guidance is needed for teachers, learners and parents.

In particular, the following policy options should be considered:

→ **Encourage the development of ICT environments and tools that holistically support curricula**. While there are a vast number and variety of ICT tools that support learning and assessment, most of them are limited in scope and do not necessarily support the learning progression foreseen in curricula. Policy action is needed to support ICT environments and tools that take into account curricular needs and are better targeted at deployment by teachers.

→ **Encourage the development of ICT environments and tools that allow teachers to quickly, easily and flexibly create customized electronic learning and assessment environments**. Open source tools that can be adapted by teachers to fit their teaching style and their learners' needs should be better promoted. Teachers should be involved in the development of these tools and encouraged to further develop, expand, modify and amend these themselves.

→ **Encourage teachers to network and exchange good practice**. Many of the ICT-enhanced assessment practices within schools are promoted by a small number of teachers who enthusiastically and critically engage with ICT for assessment. To upscale and mainstream and also to establish good practice, it is necessary to better support these teachers, encourage them to exchange their experiences and establish good practice.

→ **Set incentives for research and development** of promising technologies for the assessment of Key Competences, in particular as regards: the use of authentic tasks; the development and dissemination of educational multiplayer games; the development of automatic assessment and correction tools for written text; and the use of learning analytics to enable assessment to be embedded in virtual learning environments and games.

→ **Encourage discussion and offer guidance on viable ICT-enhanced assessment strategies**. While deployment of ICT in schools is lagging behind, given the vast range and variety of ICT strategies supporting assessment, a critical discourse on the advantages and drawbacks of ICT should be started among educators and policy makers. This could lead to the development of recommendations for the take-up of ICT in the comprehensive assessment of Key Competences.

6. BIBLIOGRAPHY

Agudo, J. E., Rico, M., Edwards, P., & Sánchez, H. (2009). *Personalization in hypermedia language assessment*, Lisboa.

Al-Smadi, M., & Guetl, C. (2011). *Supporting self-regulated learners with formative assessments using automatically created QTI-questions*, Amman.

Ala-Mutka, K. (2010). *Learning in Informal Online Networks and Communities* (No. 24149 EN): European Commission-Joint Research Centre-Institute for Prospective Technological Studies, Seville.

Alonso, A. C. (2011). The Portfolio as a resource for reflection and self-evalaution with students having learning difficulties. *El Portafolio como recurso para la reflexión y la autoevaluación en alumnos con dificultades de aprendizaje*(16), 137-153.

Amelung, M., Krieger, K., & Rösner, D. (2011). E-assessment as a service. *IEEE Transactions on Learning Technologies, 4*(2), 162-174.

Anderson, P. (2007). What is Web 2.0? Ideas, technologies and implications for education. *JISC Technology and Standards Watch, Feb. 2007*. Retrieved from http://www.jisc.ac.uk/media/documents/techwatch/tsw0701b.pdf.

Angus, S. D., & Watson, J. (2009). Does regular online testing enhance student learning in the numerical sciences? Robust evidence from a large data set. *British Journal of Educational Technology, 40*(2), 255-272.

Arendasy, M. E., Sommer, M., & Hergovich, A. (2007). Psychometric technology: Automatic dual-component item generation exemplified using a new item type for the measurement of arithmetic fluency. *Psychometrische Technologie Automatische Zwei-Komponenten-Itemgenerierung am Beispiel Eines Neuen Aufgabentyps zur Messung der Numerischen Flexibilität, 53*(3), 119-130.

Barab, S. A., Scott, B., Siyahhan, S., Goldstone, R., Ingram-Goble, A., Zuiker, S. J., et al. (2009). Transformational play as a curricular scaffold: Using videogames to support science education. *Journal of Science Education and Technology, 18*(4), 305-320.

Barbera, E. (2009). Mutual feedback in e-portfolio assessment: An approach to the netfolio system. *British Journal of Educational Technology, 40*(2), 342-357.

Barla, M., Bieliková, M., Ezzeddinne, A. B., Kramár, T., Šimko, M., & Vozár, O. (2010). On the impact of adaptive test question selection for learning efficiency. *Computers and Education, 55*(2), 846-857.

Barrada, J. R., Olea, J., Ponsoda, V., & Abad, F. J. (2006). Item selection rules in a Computerized Adaptive Test for the assessment of written English. *Estrategias de selección de ítems en un test adaptative informatizado para la evaluaion de inglés escrito, 18*(4), 828-834.

Bauer, C., Figl, K., Derntl, M., Beran, P. P., & Kabicher, S. (2009). *The student view on online peer reviews*, Paris.

Beatty, I. D., & Gerace, W. J. (2009). Technology-enhanced formative assessment: A research-based pedagogy for teaching science with classroom response technology. *Journal of Science Education and Technology, 18*(2), 146-162.

Ben-Simon, A., & Bennett, R. E. (2007). Toward a more substantively meaningful automated essay scoring. *Journal of of Technology, Learning and Assessment, 6*(1),

Bennett, R. E. (2010). Technology for Large-Scale Assessment. In P. Peterson, E. Baker & B. McGaw (Eds.), *International Encyclopedia of Education* (3rd ed., Vol. 8, pp. 48-55). Oxford: Elsevier.

Bernstein, J., van Moere, A., & Cheng, J. (2010). Validating automated speaking tests. *Language Testing, 27*(3), 355-377.

Binkley, M., Erstad, O., Herman, J., Raizen, S., Ripley, M., & Rumble, M. (2012). Defining 21st Century Skills. In P. Griffin, B. McGaw & E. Care (Eds.), *Assessment and Teaching of 21st Century Skills* (pp. 17-66). Dordrecht, Heidelberg, London, New York: Springer.

Bloomfield, P. R., & Livingstone, D. (2009). Multi-modal learning and assessment in Second Life with quizHUD. *Conference in Games and Virtual Worlds for Serious Applications*, 217-218.

Bloxham, S., & West, A. (2007). Learning to write in higher education: Students' perceptions of an intervention in developing understanding of assessment criteria. *Teaching in Higher Education, 12*(1), 77-89.

Blumenstein, M., Green, S., Fogelman, S., Nguyen, A., & Muthukkumarasamy, V. (2008). Performance analysis of GAME: A generic automated marking environment. *Computers and Education, 50*(4), 1203-1216.

Blyth, B., & Labovic, A. (2009). Using Maple to implement eLearning integrated with computer aided assessment. *International Journal of Mathematical Education in Science and Technology, 40*(7), 975-988.

Bottge, B. A., Rueda, E., Kwon, J. M., Grant, T., & LaRoque, P. (2009). Assessing and tracking students' problem solving performances in anchored learning environments. *Educational Technology Research and Development, 57*(4), 529-552.

Bouzidi, L., & Jaillet, A. (2009). Can online peer assessment be trusted? *Educational Technology and Society*, *12*(4), 257-268.

Brantmeier, C. (2006). Advanced L2 learners and reading placement: Self-assessment, CBT, and subsequent performance. *System, 34*(1), 15-35.

Bridgeman, B. (2009). Experiences from Large-Scale Computer-Based Testing in the USA. In F. Scheuermann & J. Björnsson (Eds.), *The Transition to Computer-Based Assessment*. Luxembourg: Office for Official Publications of the European Communities.

Bryant, D. A., & Carless, D. R. (2010). Peer assessment in a test-dominated setting: Empowering, boring or facilitating examination preparation? *Educational Research for Policy and Practice*, *9*(1), 3-15.

Bunderson, V. C., Inouye, D. K., & Olsen, J. B. (1989). The four generations of computerized educational measurement. In R. L. Linn (Ed.), *Educational measurement* (Third ed., pp. 367-407). New York: Macmillan.

Burks, R. (2010). The student mathematics portfolio: Value added to student preparation? *PRIMUS, 20*(5), 453-472.

Butcher, P. G., & Jordan, S. E. (2010). A comparison of human and computer marking of short free-text student responses. *Computers and Education, 55*(2), 489-499.

Butler, Y. G., & Lee, J. (2010). The effects of self-assessment among young learners of English. *Language Testing, 27*(1), 5-31.

Büyükduman, I., & Şirin, S. (2010). *Learning portfolio (LP) to enhance constructivism and student autonomy*, Famagusta.

Cachia, R., & Ferrari, A. (2010). *Creativity in Schools: A Survey of Teachers in Europe*. Seville: European Commission - Joint Research Centre -Institute for Prospective Technological Studies.

Cachia, R., Ferrari, A., Ala-Mutka, K., & Punie, Y. (2010). *Creative Learning and Innovative Teaching: Final Report on the Study on Creativity and Innovation in Education in EU Member States* (No. EUR24675): JRC-IPTS.

Calongne, C. (2007). A View from Second Life's Trenches: Are You a Pioneer or a Settler? *Proceedings of the NMC Summer Conference, 2007, 111-119*. Retrieved from http://www.nmc.org/publications/2007-conference-proceedings

Campbell, J., & Gorra, A. (2009). *Using formative e-assessment to support students' self-awareness of their abilities*, Algarve.
Cao, X. M., & Wang, X. (2009). *A computer- assisted assessment and diagnosis system for arts students-oriented computer education*, Singapore.
Caric, M., Tuba, M., & Moisil, I. (2010). *Web-based testing and self-assessment system implemented with open technologies*, Corfu Island.
Cartney, P. (2010). Exploring the use of peer assessment as a vehicle for closing thegap between feedback given and feedback used. *Assessment and Evaluation in Higher Education, 35*(5), 551-564.
Chang, C. C., & Tseng, K. H. (2009). Use and performances of Web-based portfolio assessment. *British Journal of Educational Technology, 40*(2), 358-370.
Chang, C. C., & Tseng, K. H. (2011). Using a web-based portfolio assessment system to elevate project-based learning performances. *Interactive Learning Environments, 19*(3), 211-230.
Chang, C. C., Tseng, K. H., Chou, P. N., & Chen, Y. H. (2011). Reliability and validity of Web-based portfolio peer assessment: A case study for a senior high school's students taking computer course. *Computers and Education, 57*(1), 1306-1316.
Chang, C. C., Tseng, K. H., & Lou, S. J. (2011). A comparative analysis of the consistency and difference among teacher-assessment, student self-assessment and peer-assessment in a Web-based portfolio assessment environment for high school students. *Computers and Education, 58*(1), 303-320.
Chatzopoulou, D. I., & Economides, A. A. (2010). Adaptive assessment of student's knowledge in programming courses. *Journal of Computer Assisted Learning, 26*(4), 258-269.
Chen, C. h. (2010). The implementation and evaluation of a mobile self- and peer-assessment system. *Computers and Education, 55*(1), 229-236.
Chen, C. M., & Chen, M. C. (2009). Mobile formative assessment tool based on data mining techniques for supporting web-based learning. *Computers and Education, 52*(1), 256-273.
Chen, L., Evanini, K., & Sun, X. (2010). *Assessment of non-native speech using vowel space characteristics*, Berkeley, CA.
Childnet International. (2008). Young People and Social Networking Services: A Childnet International Research Report. Retrieved from http://www.digizen.org/downloads/fullReport.pdf
Chiou, C. K., Hwang, G. J., & Tseng, J. C. R. (2009). An auto-scoring mechanism for evaluating problem-solving ability in a web-based learning environment. *Computers and Education, 53*(2), 261-272.
Cho, K., Cho, M. H., & Hacker, D. J. (2010). Self-monitoring support for learning to write. *Interactive Learning Environments, 18*(2), 101-113.
Cho, K., & MacArthur, C. (2010). Student revision with peer and expert reviewing. *Learning and Instruction, 20*(4), 328-338.
Clark, C., Osborne, S., & Dugdale, G. (2009). *Reaching out with role models: Role models and young people's reading*: London: National Literacy Trust.
Collins, A., & Halverson, R. (2010a). The second educational revolution: rethinking education in the age of technology. *Journal of computer assisted learning, 26*(1), 18-27.
Collins, A., & Halverson, R. (2010b). Technology Supports for Lifelong Learning. In P. Peterson, E. Baker & B. McGaw (Eds.), *International Encyclopedia of Education* (3rd ed., Vol. 8, pp. 184-188). Oxford: Elsevier.

Costello, P. (2010). A cost-effective classroom response system. *British Journal of Educational Technology, 41*(6).
Council of the European Union. (2006). Recommendation of the European Parliament and the Council of 18 December 2006 on key competences for lifelong learning. (2006/962/EC). Official Journal of the European Union, L394/10.
Council of the European Union. (2009). Council conclusions of 12 May 2009 on a strategic framework for European cooperation in education and training ('ET 2020'). OJ C119, 28.5.2009.
Csapó, B., Ainley, J., Bennett, R., Latour, T., & Law, N. (2012). Technological Issues for Computer-Based Assessment. In P. Griffin, B. McGaw & E. Care (Eds.), *Assessment and Teaching of 21st Century Skills* (pp. 143-230). Dordrecht, Heidelberg, London, New York: Springer.
Cummins, P. W., & Davesne, C. (2009). Using electronic portfolios for second language assessment. *Modern Language Journal*, *93*(SUPPL. 1), 848-867.
Cuthbertson, A., Hatton, S., Minyard, G., Piver, H., Todd, C., & Birchfield, D. (2007). *Mediated education in a creative arts context: Research and practice at Whittier Elementary School*, Aalborg.
D'Mello, S., Craig, S., Fike, K., & Graesser, A. (2009). Responding to learners' cognitive-affective states with supportive and shakeup dialogues. *Lecture Notes in Computer Science 5612*, 595-604.
D'Mello, S., Dowell, N., & Graesser, A. (2009). Cohesion relationships in tutorial dialogue as predictors of affective states. *Frontiers in Artificial Intelligence and Applications, 200*(1), 9-16.
Davies, P. (2009). Review and reward within the computerised peer-assessment of essays. *Assessment and Evaluation in Higher Education, 34*(3), 321-333.
De-Marcos, L., Hilera, J. R., Barchino, R., Jiménez, L., Martínez, J. J., Gutiérrez, J. A., et al. (2010). An experiment for improving students performance in secondary and tertiary education by means of m-learning auto-assessment. *Computers and Education, 55*(3), 1069-1079.
De Freitas, S. (2007). Learning in Immersive Worlds. A review of game-based learning. *JISC e-Learning Programme*. Retrieved from http://www.jisc.ac.uk/media/documents/programmes/elearninginnovation/gamingreport_v3.pdf
de Jong, T. (2010). Technology Supports for Acquiring Inquiry Skills. In P. Peterson, E. Baker & B. McGaw (Eds.), *International Encyclopedia of Education* (3rd ed., Vol. 8, pp. 167-171). Oxford: Elsevier.
Dede, C. (2010). Technological Support for Acquiring Twenty-First -Century Skills. In P. Peterson, E. Baker & B. McGaw (Eds.), *International Encyclopedia of Education* (3rd ed., Vol. 8, pp. 158-166). Oxford: Elsevier.
DeFrance, N., Khasnabis, D., & Palincsar, A. S. (2010). Reading and Technology. In P. Peterson, E. Baker & B. McGaw (Eds.), *International Encyclopedia of Education* (3rd ed., Vol. 8, pp. 150-157). Oxford: Elsevier.
Delgado, C., & Krajcik, J. (2010). Technology and Learning - Supports for Subject Matter Learning. In P. Peterson, E. Baker & B. McGaw (Eds.), *International Encyclopedia of Education* (3rd ed., Vol. 8, pp. 197-203). Oxford: Elsevier.
Denny, P., Hanks, B., & Simon, B. (2010). *PeerWise: Replication study of a student-collaborative self-testing web service in a U.S. setting*, Milwaukee, WI.
Denny, P., Luxton-Reilly, A., & Hamer, J. (2008). *Student use of the PeerWise system*, Madrid.

Denton, P., Madden, J., Roberts, M., & Rowe, P. (2008). Students' response to traditional and computer-assisted formative feedback: A comparative case study. *British Journal of Educational Technology, 39*(3), 486-500.
Derting, T. L., & Ebert-May, D. (2010). Learner-centered inquiry in undergraduate biology: Positive relationships with long-term student achievement. *CBE Life Sciences Education, 9*(4), 462-472.
Di Bitonto, P., Laterza, M., Roselli, T., & Rossano, V. (2010). *An adaptive test for learning objects: Item calibration*, Oak Brook, IL.
Draper, S. W. (2009). Catalytic assessment: Understanding how MCQs and EVS can foster deep learning. *British Journal of Educational Technology, 40*(2), 285-293.
Draycott, M., & Rae, D. (2011). Enterprise education in schools and the role of competency frameworks. *International Journal of Entrepreneurial Behaviour and Research, 17*(2), 127-145.
Duchateau, J., Kong, Y. O., Cleuren, L., Latacz, L., Roelens, J., Samir, A., et al. (2009). Developing a reading tutor: Design and evaluation of dedicated speech recognition and synthesis modules. *Speech Communication, 51*(10), 985-994.
Ďuricová, L. (2009). Self-concept of university students and their motivation. *New Educational Review, 17*(1).
Dysthe, O., & Engelsen, K. S. (2011). Portfolio practices in higher education in Norway in an international perspective: Macro-, meso- and micro-level influences. *Assessment and Evaluation in Higher Education, 36*(1), 63-79.
Dysthe, O., Engelsen, K. S., & Lima, I. (2007). Variations in portfolio assessment in higher education: Discussion of quality issues based on a Norwegian survey across institutions and disciplines. *Assessing Writing, 12*(2), 129-148.
Dziedzic, M., Janissek, P. R., & Bender, A. P. (2008). *Assessment by peers - An effective learning technique*, Saratoga Springs, NY.
Eggen, T. J. H. M., & Straetmans, G. J. J. M. (2009). Computerized Adaptive Testing of Arithmetic at the Entrance of Primary School Teacher Training College. In F. Scheuermann & J. Björnsson (Eds.), *The Transition to Computer-Based Assessment*. Luxembourg: Office for Official Publications of the European Communities.
El-Alfy, E. S. M., & Abdel-Aal, R. E. (2008). Construction and analysis of educational tests using abductive machine learning. *Computers and Education, 51*(1), 1-16.
Ellis, S., & Barrs, M. (2008). The Assessment of Creative Learning. In J. Sefton-Green (Ed.), *Creative Learning* (pp. 73-89). London: Creative Partnerships.
Ellison, N., & Wu, Y. (2008). Blogging in the Classroom: A Preliminary Exploration of Student Attitudes and Impact on Comprehension. *Journal of Educational Multimedia and Hypermedia, 17*, 99-122.
ETS. (2012). *Sea Change in Assessment: How Technology is Transforming K-12 Testing*: http://www.k12center.org/rsc/pdf/a-sea-change-in-assessment-letter-size.pdf.
European Commission. (2009). Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions: “Key Competences for a changing world: Progress towards the Lisbon Objectives in Education and Training: Analysis of implementation at the European and National Levels. COM(2009) 640.
European Commission. (2010). Commission Communication "Europe 2020 – A strategy for smart, sustainable and inclusive growth". COM (2010) 2020.
European Commission. (2012a). Assessment of Key Competences in initial education and training: Policy Guidance *Staff Working Document. Accompanying the Communication from the Commission on Rethinking Education: Investing in skills for better socio-economic outcomes*. SWD(2012) 371 final.

European Commission. (2012b). *Rethinking Education:Investing in skills for better socio-economic outcomes*.

Eurydice. (2009). *Teaching Reading in Europe: Context, Policies and Practices*: Education, Audiovisual and Culture Executive Agency.

Eurydice. (2011a). *Key Data on Learning and Innovation through ICT at School in Europe 2011*: Education, Audiovisual and Culture Executive Agency.

Eurydice. (2011b). *Teaching Reading in Europe: Context, Policies and Practices*: Education, Audiovisual and Culture Executive Agency.

Eurydice. (2012). *Developing Key Competences at School in Europe: Challenges and Opportunities for Policy*: Education, Audiovisual and Culture Executive Agency.

Ferrari, A. (2012). *Digital Competence in practice: An analysis of frameworks*. Seville: JRC-IPTS.

Ferrari, A., Cachia, R., & Punie, Y. (2009). *Innovation and Creativity in Education and Training in the EU Member States: Fostering Creative Learning and Supporting Innovative Teaching. Literature review on Innovation and Creativity in E&T in the EU Member States (ICEAC)*: JRC-IPTS.

Fischer, J., Mitchell, R., & Del Alamo, J. (2007). Inquiry-learning with WebLab: Undergraduate attitudes and experiences. *Journal of Science Education and Technology, 16*(4), 337-348.

Florian, B., & Fabregat, R. (2011), *14th International Conference on Human-Computer Interaction, HCI International 2011: Vol. 173 CCIS* (pp. 138-142). Orlando, FL.

Florián G, B. E., Baldiris, S. M., Fabregat, R., & De La Hoz Manotas, A. (2010). *A set of software tools to build an author assessment package on Moodle: Implementing the AEEA proposal*, Sousse.

Forster, P. A. (2006). Assessing technology-based approaches for teaching and learning mathematics. *International Journal of Mathematical Education in Science and Technology, 37*(2), 145-164.

Frey, A., & Seitz, N. N. (2009). Multidimensional adaptive testing in educational and psychological measurement: Current state and future challenges. *Studies in Educational Evaluation, 35*(2-3), 89-94.

Garrett, N., Thoms, B., Alrushiedat, N., & Ryan, T. (2009). Social ePortfolios as the new course management system. *On the Horizon, 17*(3), 197-207.

Ghoorchaei, B., Tavakoli, M., & Ansari, D. N. (2010). The impact of Portfolio assessment on Iranian EFL students' essay writing: A process-oriented approach. *GEMA Online Journal of Language Studies, 10*(3), 35-51.

Gielen, S., Dochy, F., & Onghena, P. (2011). An inventory of peer assessment diversity. *Assessment and Evaluation in Higher Education, 36*(2), 137-155.

Gielen, S., Peeters, E., Dochy, F., Onghena, P., & Struyven, K. (2010). Improving the effectiveness of peer feedback for learning. *Learning and Instruction, 20*(4), 304-315.

Gipps, C. V. (2005). What is the role for ICT-based assessment in universities? *Studies in Higher Education, 30*(2), 171-180.

Glas, C. A. W., & Geerlings, H. (2009). Psychometric aspects of pupil monitoring systems. *Studies in Educational Evaluation, 35*(2-3), 83-88.

Graesser, A. (2009). Autotutor and the world of pedagogical agents: Intelligent tutoring systems with natural language dialogue. *22nd International Florida Artificial Intelligence Research Society Conference, FLAIRS-22*, 3.

Graff, M., Mayer, P., & Lebens, M. (2008). Evaluating a web based intelligent tutoring system for mathematics at German lower secondary schools. *Education and Information Technologies, 13*(3), 221-230.

Guerrero, S., Walker, N., & Dugdale, S. (2004). Technology in Support of Middle Grade Mathematics: What Have We Learned? *Journal of Computers in Mathematics and Science Teaching, 23*(1), 5-20.

Halász, G., & Michel, A. (2011). Key Competences in Europe: interpretation, policy formulation and implementation. *European Journal of Education, 46*(3), 289-306.

Hardré, P. L., Crowson, H. M., Xie, K., & Ly, C. (2007). Testing differential effects of computer-based, web-based and paper-based administration of questionnaire research instruments. *British Journal of Educational Technology, 38*(1), 5-22.

Hartig, J., & Höhler, J. (2009). Multidimensional IRT models for the assessment of competencies. *Studies in Educational Evaluation, 35*(2-3), 57-63.

He, Y., Hui, S. C., & Quan, T. T. (2009). Automatic summary assessment for intelligent tutoring systems. *Computers and Education, 53*(3), 890-899.

Herrenkohl, L. R., Tasker, T., & White, B. (2011). Pedagogical practices to support classroom cultures of scientific inquiry. *Cognition and Instruction, 29*(1), 1-44.

Hilberg, J. S., & Meiselwitz, G. (2008). *Undergraduate fluency with information and communication technology: Perceptions and reality*, Cincinnati, OH.

Hismanoglu, M., & Hismanoglu, S. (2010). The european language portfolio in ESP classes: A case study of learner reflection and self-assesment. *European Journal of Social Sciences, 12*(4), 671-684.

Horkay, N., Bennett, R. E., Allen, N., Kaplan, B., & Yan, F. (2006). Does it Matter if I Take My Writing Test on Computer? An Empirical Study of Mode Effects in NAEP. *Journal of Technology, Learning, and Assessment, 5*(2),

Horne, J. (2007). Gender differences in computerised and conventional educational tests. *Journal of Computer Assisted Learning, 23*(1), 47-55.

Hou, H. T., Chang, K. E., & Sung, Y. T. (2007). An analysis of peer assessment online discussions within a course that uses project-based learning. *Interactive Learning Environments, 15*(3), 237-251.

Huang, C. J., Wang, Y. W., Huang, T. H., Chen, Y. C., Chen, H. M., & Chang, S. C. (2011). Performance evaluation of an online argumentation learning assistance agent. *Computers and Education, 57*(1), 1270-1280.

Huang, C. J., Wang, Y. W., Huang, T. H., Liao, J. J., Chen, C. H., Weng, C. H., et al. (2010). *Implementation and performance evaluation of an intelligent online argumentation assessment system*, Wuhan.

Huang, H. T. D., & Hung, S. T. A. (2010). Implementing electronic speaking portfolios: Perceptions of EFL students: Colloquium. *British Journal of Educational Technology, 41*(5).

Hutchison, D. (2007). An evaluation of computerised essay marking for national curriculum assessment in the UK for 11-year-olds. *British Journal of Educational Technology, 38*(6), 977-989.

Ibabe, I., & Jauregizar, J. (2010). Online self-assessment with feedback and metacognitive knowledge. *Higher Education, 59*(2), 243-258.

Isotani, S., & Brandão, L. d. O. (2008). An algorithm for automatic checking of exercises in a dynamic geometry system: iGeom. *Computers and Education, 51*(3), 1283-1303.

Jackson, G. T., Boonthum, C., & McNamara, D. S. (2010), *10th International Conference on Intelligent Tutoring Systems, ITS 2010: Vol. 6095 LNCS* (pp. 349-351). Pittsburgh, PA.

JISC. (2006). e-Assessment Glossary.

Johnson, L., Smith, R., Willis, H., Levine, A., & Haywood, K. (2011). *The 2011 Horizon Report*. Austin, Texas: The New Media Consortium.

Johnson, W. L. (2010). Serious Use of a Serious Game for Language Learning. *International Journal of Artificial Intelligence in Education, 20*(2).

Jordan, S., & Mitchell, T. (2009). e-Assessment for learning? The potential of short-answer free-text questions with tailored feedback. *British Journal of Educational Technology, 40*(2), 371-385.

Kaufman, J. H., & Schunn, C. D. (2010). Students' perceptions about peer assessment for writing: their origin and impact on revision work. *Instructional Science*, 1-20.

Kelly, D., Baxter, J. S., & Anderson, A. (2010). Engaging first-year students through online collaborative assessments. *Journal of Computer Assisted Learning, 26*(6), 535-548.

Kim, P., & Olaciregui, C. (2008). The effects of a concept map-based information display in an electronic portfolio system on information processing and retention in a fifth-grade science class covering the Earth's atmosphere. *British Journal of Educational Technology, 39*(4), 700-714.

Klopfer, E., & Squire, K. (2008). Environmental detectives-the development of an augmented reality platform for environmental simulations. *Educational Technology Research and Development, 56*(2), 203-228.

Kollar, I., & Fischer, F. (2010). Peer assessment as collaborative learning: A cognitive perspective. *Learning and Instruction, 20*(4), 344-348.

Lai, Y. H. (2010). Which do students prefer to evaluate their essays: Peers or computer program. *British Journal of Educational Technology, 41*(3), 432-454.

Lent, G. v. (2008). Important Considerations in e-Assessment: An Educational Measurement Perspective on Identifying Items for an European Research Agenda. In F. Scheuermann & A. G. Pereira (Eds.), *Towards a Research Agenda on Computer-Based Assessment. Challenges and needs for European Educational Measurement*. Luxembourg: Office for Official Publications of the European Communities.

Li, L., Liu, X., & Steckelberg, A. L. (2010). Assessor or assessee: How student learning improves by giving and receiving peer feedback. *British Journal of Educational Technology, 41*(3), 525-536.

Liang, J. C., & Tsai, C. C. (2010). Learning through science writing via online peer assessment in a college biology course. *Internet and Higher Education, 13*(4), 242-247.

Linn, M. C., Lee, H. S., Tinker, R., Husic, F., & Chiu, J. L. (2006). Teaching and assessing knowledge integration in science. *Science, 313*(5790), 1049-1050.

Little, D. (2009). Language learner autonomy and the European Language Portfolio: Two L2 English examples. *Language Teaching, 42*(3), 222-233.

Liu, E. Z. F. (2007). Developing a personal and group-based learning portfolio system. *British Journal of Educational Technology, 38*(6), 1117-1121.

Ljungdahl, L., & Prescott, A. (2009). Teachers' use of diagnostic testing to enhance students' literacy and numeracy learning. *International Journal of Learning, 16*(2), 461-476.

Looi, C. K., Zhang, B., Chen, W., Seow, P., Chia, G., Norris, C., et al. (2011). 1:1 mobile inquiry learning experience for primary science students: A study of learning effectiveness. *Journal of Computer Assisted Learning, 27*(3), 269-287.

Looney, J. (2010). Making it Happen: Formative Assessment and Educational Technologies. *Promethean Thinking Deeper Research Papers, 1*(3).

Luxton-Reilly, A., Plimmer, B., & Sheehan, R. (2010). *Studysieve - A tool that supports constructive evaluation for free-response questions*, Auckland.

Magliano, J. P., Todaro, S., Millis, K., Wiemer-Hastings, K., Kim, H. J., & McNamara, D. S. (2005). Changes in reading strategies as a function of reading training: A comparison of live and computerized training. *Journal of Educational Computing Research, 32*(2), 185-208.

Martin, R. (2008). New possibilities and challenges for assessment through the use of technology. In F. Scheuermann & A. G. Pereira (Eds.), *Towards a Research Agenda on Computer-Based Assessment*. Luxembourg: Office for Official Publications of the European Communities.

MCEECDYA. (2008). *National Assessment Program. ICT Literacy Years 6 and 10 Report 2008*. Australia: Ministerial Council for Education, Early Childhood Development and Youth Affairs (MCEECDYA).

McLeod, J. K., & Vasinda, S. (2009). Electronic portfolios: Perspectives of students, teachers and parents. *Education and Information Technologies, 14*(1), 29-38.

McMahon, T. (2010). Peer feedback in an undergraduate programme: Using action research to overcome students' reluctance to criticise. *Educational Action Research, 18*(2), 273-287.

Means, B., & Rochelle, J. (2010). An Overview of Technology and Learning. In P. Peterson, E. Baker & B. McGaw (Eds.), *International Encyclopedia of Education* (3rd ed., Vol. 8, pp. 1-10). Oxford: Elsevier.

Meijer, J., & Riemersma, F. (2002). Teaching and testing mathematical problem solving by offering optional assistance. *Instructional Science, 30*(3), 187-220.

Meijer, R. (2008). Stimulating innovative item use in assessment. In F. Scheuermann & A. G. Pereira (Eds.), *Towards a Research Agenda on Computer-Based Assessment. Challenges and needs for European Educational Measurement*. Luxembourg: Office for Official Publications of the European Communities.

Meurant, R. C. (2009) & D. Slezak, W. I. Grosky, N. Pissinou, T. K. Shih, T. H. Kim & B. H. Kang (Vol. Ed.)*: Vol. 60* (pp. 84-91).

Moe, E. (2009). Introducing Large-scale Computerised Assessment Lessons Learned and Future Challenges. In F. Scheuermann & J. Björnsson (Eds.), *The Transition to Computer-Based Assessment*. Luxembourg: Office for Official Publications of the European Communities.

Morris, P., & Dowdall, S. (2011). *Computer-aided self-assessment and independent learning in higher education*, Noordwijkerhout.

Mortazavi, B. (2010). *Self assessment surveillance using e-Portfolio*, Tehran.

Moss, K., & Crowley, M. (2011). Effective learning in science: The use of personal response systems with a wide range of audiences. *Computers and Education, 56*(1), 36-43.

NACCCE. (1999). *All Our Futures: Creativity, Culture and Education*.

Nathan, N. J. (2010). Technology Supports for Acquiring Mathematics. In P. Peterson, E. Baker & B. McGaw (Eds.), *International Encyclopedia of Education* (3rd ed., Vol. 8, pp. 172-183). Oxford: Elsevier.

Newhouse, C. P. (2011). Using IT to assess IT: Towards greater authenticity in summative performance assessment. *Computers and Education, 56*(2), 388-402.

Ni, H., & Liu, F. (2009). *Quantitative assessment of students' learning activity based on students' electronic portfolio of web-based instruction system*, Nanjing.

NMC. (2007). The Spring, 2007 Survey: Educators in Second Life. Retrieved from http://www.nmc.org/pdf/2007-sl-survey-summary.pdf

Noorbehbahani, F., & Kardan, A. A. (2011). The automatic assessment of free text answers using a modified BLEU algorithm. *Computers and Education, 56*(2), 337-345.

Nunan, D. (2010). Technology Supports for Second Language Learning, *International Encyclopedia of Education* (3rd ed., Vol. 8, pp. 204-209). Oxford: Elsevier.

Ocak, G., & Ulu, M. (2009). The views of students, teachers and parents and the use of portfolio at the primary level. *Procedia - Social and Behavioral Sciences, 1*(1), 28-36.

OECD. (2007). Participative Web and User-created Content. Web 2.0, Wikis and Social Networking. . from http://213.253.134.43/oecd/pdfs/browseit/9307031E.PDF

Olney, A. M. (2009). GnuTutor: An open source intelligent tutoring system based on AutoTutor. *Cognitive and Metacognitive Educational Systems: Papers from the AAAI Fall Symposium, FS-09-02*, 70-75.

Owen, M., Grant, L., Sayers, S., & Facer, K. (2006). Social software and learning. Futurelab Opening Education Reports, 2006.

Paré, D. E., & Joordens, S. (2008). Peering into large lectures: Examining peer and expert mark agreement using peerScholar, an online peer assessment tool. *Journal of Computer Assisted Learning, 24*(6), 526-540.

Park, J. (2010). Constructive multiple-choice testing system. *British Journal of Educational Technology, 41*(6), 1054-1064.

Pascu, C. (2008). An Empirical Analysis of the Creation, Use and Adoption of Social Computing Applications: JRC Scientific and Technical Reports. EUR 23415 EN.

Peacock, S., Gordon, L., Murray, S., Morss, K., & Dunlop, G. (2010). Tutor response to implementing an ePortfolio to support learning and personal development in further and higher education institutions in Scotland. *British Journal of Educational Technology, 41*(5), 827-851.

Pelgrum, W. J., & Voogt, J. (2009). School and teacher factors associated with frequency of ICT use by mathematics teachers: Country comparisons. *Education and Information Technologies, 14*(4), 293-308.

Pellegrino, J. W. (2010). Technology and Learning - Assessment. In P. Peterson, E. Baker & B. McGaw (Eds.), *International Encyclopedia of Education* (3rd ed., Vol. 8, pp. 42-47). Oxford: Elsevier.

Peltenburg, M., Van Den Heuvel-Panhuizen, M., & Doig, B. (2009). Mathematical power of special-needs pupils: An ICT-based dynamic assessment format to reveal weak pupils' learning potential. *British Journal of Educational Technology, 40*(2), 273-284.

Peltenburg, M., van den Heuvel-Panhuizen, M., & Robitzsch, A. (2010). ICT-based dynamic assessment to reveal special education students' potential in mathematics. *Research Papers in Education, 25*(3), 319-334.

Pepper, D. (2011). Assessing Key Competences across the Curriculum — and Europe. *European Journal of Education, 46*(3), 335-353.

Pérez-Marín, D., & Pascual-Nieto, I. (2010). Showing automatically generated students' conceptual models to students and teachers. *International Journal of Artificial Intelligence in Education, 20*(1), 47-72.

Redecker, C., Leis, M., Leendertse, M., Punie, Y., Gijsbers, G., Kirschner, P., et al. (2010). *The Future of Learning: Preparing for Change*. Seville: JRC-IPTS.

Remus, J. J., & Collins, L. M. (2008). Comparison of adaptive psychometric procedures motivated by the Theory of Optimal Experiments: Simulated and experimental results. *Journal of the Acoustical Society of America, 123*(1), 315-326.

Ridgway, J., & McCusker, S. (2008). Challenges for Research in e-Assessment. In F. Scheuermann & A. G. Pereira (Eds.), *Towards a Research Agenda on Computer-Based Assessment*. Luxembourg: Office for Official Publications of the European Communities.

Ripley, M. (2009). Transformational Computer-based Testing. In F. Scheuermann & J. Björnsson (Eds.), *The Transition to Computer-Based Assessment*. Luxembourg: Office for Official Publications of the European Communities.

Ritter, S., Anderson, J. R., Koedinger, K. R., & Corbett, A. (2007). Cognitive tutor: Applied research in mathematics education. *Psychonomic Bulletin and Review, 14*(2), 249-255.

Ritter, S., Towle, B., Murray, R. C., Hausmann, R. G. M., & Connelly, J. (2010), *10th International Conference on Intelligent Tutoring Systems, ITS 2010: Vol. 6095 LNCS* (pp. 452). Pittsburgh, PA.

Rosenbaum, E., Klopfer, E., & Perry, J. (2007). On location learning: Authentic applied science with networked augmented realities. *Journal of Science Education and Technology, 16*(1), 31-45.

Sainsbury, M., & Benton, T. (2011). Designing a formative e-assessment: Latent class analysis of early reading skills. *British Journal of Educational Technology, 42*(3), 500-514.

Saito, H. (2008). EFL classroom peer assessment: Training effects on rating and commenting. *Language Testing, 25*(4), 553-581.

Sanz, F. (2008). Spanish-language Literature Review. In J. Looney (Ed.), *Teaching, Learning and Assessment for Adults: Improving Foundation*

Skills. Paris: OECD.

Sawaki, Y. (2001). Comparability of conventional and computerized tests of reading in a second language. *Language Learning and Technology, 5*(2), 38-59.

Scheuermann, F., & Pereira, A. G. (2008). What software do we need? Identifying quality criteria for assessing language skills at a comparative level. In F. Scheuermann & A. G. Pereira (Eds.), *Towards a Research Agenda on Computer-Based Assessment. Challenges and needs for European Educational Measurement*. Luxembourg: Office for Official Publications of the European Communities.

Segers, M., Gijbels, D., & Thurlings, M. (2008). The relationship between students' perceptions of portfolio assessment practice and their approaches to learning. *Educational Studies, 34*(1), 35-44.

Shamir, H., Johnson, E. P., & Brown, K. (2009). *Waterford assessment of core skills: A Computerized adaptive reading test for pre-K through 2nd grade*, Lisboa.

Sitthiworachart, J., & Joy, M. (2008). Computer support of effective peer assessment in an undergraduate programming class. *Journal of Computer Assisted Learning, 24*(3), 217-231.

Søndergaard, H. (2009). *Learning from and with peers: The different roles of student peer reviewing*, Paris.

Sporer, T., Steinle, M., & Metscher, J. (2010), *5th European Conference on Technology Enhanced Learning, EC-TEL 2010: Vol. 6383 LNCS* (pp. 584-589). Barcelona.

Squire, K., & Klopfer, E. (2007). Augmented reality simulations on handheld computers. *Journal of the Learning Sciences, 16*(3), 371-413.

Strijbos, J. W., & Sluijsmans, D. (2010). Unravelling peer assessment: Methodological, functional, and conceptual developments. *Learning and Instruction, 20*(4), 265-269.

Su, F., & Beaumont, C. (2010). Evaluating the use of a wiki for collaborative learning. *Innovations in Education and Teaching International, 47*(4), 417-431.

Su, J. M., Lin, H. Y., Tseng, S. S., & Lu, C. J. (2011). Opass: An online portfolio assessment and diagnosis scheme to support web-based scientific inquiry experiments. *Turkish Online Journal of Educational Technology, 10*(2), 151-173.

Sullins, J., Jeon, M., D'Mello, S., & Graesser, A. C. (2009). The relationship between modality and metacognition while interacting with autotutor. *Frontiers in Artificial Intelligence and Applications 200*(1), 674-676.

Taasoobshirazi, G., Zuiker, S. J., Anderson, K. T., & Hickey, D. T. (2006). Enhancing inquiry, understanding, and achievement in an astronomy multimedia learning environment. *Journal of Science Education and Technology, 15*(5-6), 383-395.

Thompson, D., & McGregor, I. (2009). Online self- and peer assessment for groupwork. *Education and Training, 51*(5), 434-447.

Thompson, N. A., & Weiss, D. J. (2009). Computerized and Adaptive Testing in Educational Assessment In F. Scheuermann & J. Björnsson (Eds.), *The Transition to Computer-*

Based Assessment. Luxembourg: Office for Official Publications of the European Communities.

Threlfall, J., Pool, P., Homer, M., & Swinnerton, B. (2007). Implicit aspects of paper and pencil mathematics assessment that come to light through the use of the computer. *Educational Studies in Mathematics, 66*(3), 335-348.

Tillema, H., & Smith, K. (2007). Portfolio appraisal: In search of criteria. *Teaching and Teacher Education, 23*(4), 442-456.

Tisani, N. (2008). Challenges in producing a portfolio for assessment: In search of underpinning educational theories. *Teaching in Higher Education, 13*(5), 549-557.

Toki, E. I., & Pange, J. (2010). *The design of an expert system for the e-assessment and treatment plan of preschooler's speech and language disorders*, Cairo.

Topping, K. J. (2009). Peer assessment. *Theory into Practice, 48*(1), 20-27.

Tseng, S. C., & Tsai, C. C. (2007). On-line peer assessment and the role of the peer feedback: A study of high school computer course. *Computers and Education, 49*(4), 1161-1174.

Tsivitanidou, O. E., Zacharia, Z. C., & Hovardas, T. (2011). Investigating secondary school students' unmediated peer assessment skills. *Learning and Instruction, 21*(4), 506-519.

Tuomi, I. (2006). The Future of Learning in the Knowledge Society: Disruptive Changes for Europe by 2020 (eds.). In Y. Punie, M. Cabrera, M. Bogdanowicz, D. Zinnbauer & E. Navajas (Eds.), *The Future of ICT and Learning in the Knowledge Society* (pp. 47-85).

Tzortzidou, S., & Hassapis, G. (2001). Assessment of the reading skill improvement in the computer-assisted teaching of a foreign language. *Education and Information Technologies, 6*(3), 177-191.

Usener, C. A., Gruttmann, S., Majchrzak, T. A., & Kuchen, H. (2010). *Computer-supported assessment of software verification proofs: Towards high-quality e-assessments in computer science education*, Chongqing.

Üstünel, E., & Deren, E. (2010). *The effects of e-portfolio based assessment on students' perceptions of educational environment*, Istanbul.

Vendlinski, T. P., Delacruz, G. C., Buschang, R. E., Chung, G. K. W. K., & Baker, E. L. (2010). Developing High-Quality Assessments That Align with Instructional Video Games. Retrieved from http://www.cse.ucla.edu/products/reports/R774.pdf

Ventouras, E., Triantis, D., Tsiakas, P., & Stergiopoulos, C. (2011). Comparison of oral examination and electronic examination using paired multiple-choice questions. *Computers and Education, 56*(3), 616-624.

Vlug, K. F. M. (1997). Because every pupil counts: The success of the pupil monitoring system in the Netherlands. *Education and Information Technologies, 2*(4), 287-306.

Wallace, P., Graesser, A., Millis, K., Halpern, D., Cai, Z., Britt, M. A., et al. (2009). Operation ARIES!: A computerized game for teaching scientific inquiry. *Frontiers in Artificial Intelligence and Applications, 200*(1), 602-604.

Walser, T. M. (2009). An action research study of student self-assessment in higher education. *Innovative Higher Education, 34*(5), 299-306.

Wandall, J. (2009). National Tests in Denmark – CAT as a Pedagogic Tool. In F. Scheuermann & J. Björnsson (Eds.), *The Transition to Computer-Based Assessment*. Luxembourg: Office for Official Publications of the European Communities.

Wang, H. C., Chang, C. Y., & Li, T. Y. (2008). Assessing creative problem-solving with automated text grading. *Computers and Education, 51*(4), 1450-1466.

Weigle, S. C. (2010). Validation of automated scores of TOEFL iBT tasks against non-test indicators of writing ability. *Language Testing, 27*(3), 335-353.

Willey, K., & Gardner, A. (2010). Investigating the capacity of self and peer assessment activities to engage students and promote learning. *European Journal of Engineering Education, 35*(4), 429-443.

Williams, M., & Linn, M. C. (2002). WISE inquiry in fifth grade biology. *Research in Science Education*, *32*(4), 415-436.
Wilson, K., Boyd, C., Chen, L., & Jamal, S. (2011). Improving student performance in a first-year geography course: Examining the importance of computer-assisted formative assessment. *Computers and Education*, *57*(2), 1493-1500.
Xiao, Y., & Lucking, R. (2008). The impact of two types of peer assessment on students' performance and satisfaction within a Wiki environment. *Internet and Higher Education*, *11*(3-4), 186-193.
Yarnall, L., Shechtman, N., & Penuel, W. R. (2006). Using handheld computers to support improved classroom assessment in science: Results from a field trial. *Journal of Science Education and Technology*, *15*(2), 142-158.

European Commission
EUR25891– Joint Research Centre – Institute for Prospective Technological Studies

Title: The Use of ICT for the Assessment of Key Competences

Author: Christine Redecker

Luxembourg: Publications Office of the European Union

2013 – 98 pp. – 21.0 x 29.7 cm

EUR – Scientific and Technical Research series –ISSN 1831-9424 (online)

ISBN 978-92-79-29020-6 (print)

doi:10.2791/87007

Abstract

This report assesses current trends in the area of ICT for learning and assessment in view of their value for supporting the assessment of Key Competences. Based on an extensive review of the literature, it provides an overview of current ICT-enabled assessment practices, with a particular focus on more recent developments that support the holistic assessment of Key Competences for Lifelong Learning in Europe. The report presents a number of relevant cases, discusses the potential of emerging technologies, and addresses innovation and policy issues for eAssessment. It considers both summative and formative assessment and considers how ICT can lever the potential of more innovative assessment formats, such as peer-assessment and portfolio assessment and how more recent technological developments, such as Learning Analytics, could, in the future, foster assessment for learning. Reflecting on the use of the different ICT tools and services for each of the eight different Key Competences for Lifelong Learning it derives policy options for further exploiting the potential of ICT for competence-based assessment.

LF-NA-25891-EN-N

As the Commission's in-house science service, the Joint Research Centre's mission is to provide EU policies with independent, evidence-based scientific and technical support throughout the whole policy cycle.

Working in close cooperation with policy Directorates-General, the JRC addresses key societal challenges while stimulating innovation through developing new standards, methods and tools, and sharing and transferring its know-how to the Member States and international community.

Key policy areas include: environment and climate change; energy and transport; agriculture and food security; health and consumer protection; information society and digital agenda; safety and security including nuclear; all supported through a cross-cutting and multi-disciplinary approach.

Publications Office

ISBN 978-92-79-29020-6

9 789279 290206