JRC TECHNICAL REPORTS

Evidence of innovative assessment: Literature review and case studies

Kapsalis, G., Ferrari, A.,
Punie, Y., Conrads, J.,
Collado, A., Hotulainen, R.,
Rämä, I., Nyman, L.,
Oinas, S., Ilsley, P.

2019

Joint
Research
Centre

EUR 29882 EN

Contact information
Name: Yves Punie
Address: Calle Inca Garcilaso 3, 41092 Spain
Email: Yves.PUNIE@ec.europa.eu
Tel.: +34 9544-88229

EU Science Hub
https://ec.europa.eu/jrc

JRC118113

EUR 29882 EN

PDF ISBN 978-92-76-12116-9 ISSN 1831-9424 doi:10.2760/552774

Luxembourg: Publications Office of the European Union, 2019

How to cite this report: Kapsalis, G., Ferrari, A., Punie, Y, Conrads, J., Collado, A., Hotulainen, R., Rämä, I., Nyman, L., Oinas, S., & Ilsley, P., Evidence of innovative assessment: Literature review & case studies, EUR 29882 EN, Publications Office of the European Union, Luxembourg, 2019, ISBN 978-92-76-12116-9, doi:10.2760/552774, JRC118113.

Foreword

Assessment is a central feature of education. Innovating assessment can drive change towards modernising our education systems, affecting how teaching and learning takes place in an increasingly digital society and economy. The study provides empirical evidence on innovative assessment, looking at its effectiveness and how it occurs in practice, through a number of case studies. The report highlights challenges and success factors for adopting innovative practices in assessment and identifies a series of key recommendations to advance innovative assessment.

The report is done by the JRC, supported by an external consortium, on behalf of, and in collaboration with the Directorate-General for Education, Youth, Sport and Culture.

The report is part of the JRC research on "Learning and Skills for the Digital Era" which has undertaken, since 2005, around 30 major studies on these issues, resulting in more than 120 different publications. Recent work has focused on the development of digital competence frameworks for citizens (DigComp), educators (DigCompEdu), educational organisations (DigCompOrg) and consumers (DigCompConsumers). A framework for opening up higher education institutions (OpenEdu) was also published in 2016, along with a competence framework for entrepreneurship (EntreComp). Some of these frameworks are accompanied by self-reflection instruments, such as SELFIE, focussed on digital capacity building of schools.

In 2019, JRC starting working, on behalf of DG EAC, on a new competence framework for Personal, Social and Learning to Learn competences (LifEComp). Background research for the new LifEComp framework was published early October 2019.

In addition, a series of 4 reports were published on innovating Continuous Professional Development, in school education and higher education, as well as a methodological guide on conducting evaluations of the provision of open digital textbooks. In addition, practical guidelines on open education for academics were released. Past research has been undertaken on Learning Analytics, MOOCs (MOOCKnowledge, MOOCs4inclusion), Computational thinking (Computhink) and policies for the integration and innovative use of digital technologies in education (DigEduPol), and the potential of blockchain in education.

More information on all our studies can be found on the JRC Science hub: https://ec.europa.eu/jrc/en/research-topic/learning-and-skills.

Abstract

This report presents the outcomes and analyses of the study *Evidence of Innovative Assessment*. It provides an overview of innovative (digital and non-digital) assessment approaches and evidence on how these have been implemented to various settings.

The first part describes the rational of the study, defines innovative assessment and gathers evidence on the effectiveness of a variety of assessment practices such as self- and peer-assessment, open badges, simulation and learning analytics.

The second part presents eight case studies that have integrated innovative assessment approaches from a range of different contexts (formal, non-formal learning, employment, elderly care), covering different age groups, assessment purposes and implementation strategies. Through cross comparisons, the report identifies the challenges and success factors and the replicability of these cases. The report ends with recommendations for research, educational policy and practice.

Contents

Acknowledgements

The authors extends a special acknowledgement and thanks to those who have helped to dwell deeper into understanding the various aspects of assessment as well those who provided the information for the case studies.

Special thanks to Christine Redecker who initially started the project of *Evidence of Innovative Assessment*.

Authors

Georgios Kapsalis, Anusca Ferrari, Yves Punie, Johannes Conrads, Antonio Collado, Risto Hotulainen, Irene Rämä, Laura Nyman, Sanna Oinas, Paul Ilsley

Co-author for the *Learning analytics* part: Riina Vuorikari.

Executive summary

The New skills agenda for Europe emphasises the need to invest in education and lifelong learning and maintains that citizens need to be equipped with a wider range of competences, which are required for personal fulfilment and social inclusion (European Parliament & Council of the European Union, 2006: 13). Decisions on assessment have an impact on how competences are taught. As competence-based education emphasizes strategic thinking and decision making, there is a need to develop new forms of assessment through which learners have an active role, become aware of their learning processes and needs, and develop a sense of responsibility for their learning (European Commission, 2018). Furthermore, assessment should be sustainable and prepare students for their future learning needs. It is therefore necessary to look for innovative and efficient assessment practices, which comply with the new educational priorities, preparing learners to operate effectively in the longer term as citizens of the world.

Innovative assessment refers to a form of assessment that is novel, flexible and adaptable to different contexts and approaches, which integrate a variety of methods and techniques while giving enough space to learners to engage in reflective practices and actively accommodate their learning needs. Digital technology can increase the validity and reliability of assessment practices, yet decisions should be based on pedagogy. If digital tools are used, they should be tailored to the needs of educators and learners. Moreover, innovative assessment should encourage deep learning through the development of certain disposition for practice such as reflection and evaluation, developing assessment criteria, engage in peer-assessment, working with peers and use of transversal skills.

Key findings

The study provides empirical evidence on innovative assessment practices from two sources. First, it reviews research outcomes on the effectiveness of a variety of innovative assessment practices. Second, it describes eight case studies that have been implemented within Europe and highlights the challenges and success factors of such schemes.

Given that learners are at the center of assessment, self- and peer-assessment are essential components of innovative assessment and valuable means of empowerment. Through self- and peer-assessment, learners gain ownership of their learning, their motivation and engagement is increased and self-sustained learning is fostered. There is sound evidence on the effectiveness of self- and peer-assessment on learners' behavior and learning outcomes and a variety of digital tools can facilitate such processes. Most importantly, self-reflection is the core element of any assessment procedure and developing such a competence is of utmost importance.

In addition, digital approaches such as open badges, simulation and learning analytics have been reviewed. Open Badges provide a comprehensive account of ones learning accomplishments (e.g. competences, achievements) throughout life. Simulations can assist learning in real like environments and transferring of knowledge in the workplace while learning analytics can provide valuable information and support learning through real time personalised feedback and visualisation of progress. However, there is a considerable lack of evidence in relation to the effectiveness of open badges and learning analytics in accommodating users' needs.

The case studies provide valuable insights on the challenges and success factors of innovative assessment projects. Among the success factors are the cooperation of teachers and exchange of good practices, the active role of the learners during self/peer-assessment and simulation, effective technology integration in the classroom, the commitment of the involved participants, a supportive school leadership and high quality professional training for teachers. Barriers include the need for time and energy for the

participants to adapt and be familiarized with the new practices and the lack of confidence in their digital skills. Finally, the obstacles that reduce the potential replicability are mostly related to teacher training.

Key recommendations

The recommendations concern all educational levels and call for action of the following:

Educational policy and governance practice

- Develop a framework of comprehensive assessment with an emphasis on formative assessment.
- Alignment of all sectors for the development of clear goals and reference points for innovative assessment practices.
- Constant evaluation and improvement of assessment systems.
- Increase awareness of stakeholders on how innovative tools can motivate and recognize authentic learning.

Institutional leadership and governance practice

- Comprehensive guidance on how to blend/integrate diverse assessment methods.
- Implementation of bottom-up approaches.
- Description of performance criteria in relation to national curriculum goals.

Collaboration and networking

- Collaboration among various stakeholders in designing assessment tools.
- Exchange of experiences and good practices through networks of instructors and learners.

Teaching and learning practices

- Self- and peer-assessment should be encouraged through the whole learning process.
- Learners should be given time and space to practice self-reflection.
- Digital assessment practices should be adapted to learners' progress and individual needs.

Capacity building

- Professional development of educators/instructors should become a top priority.
- Teachers' initial training should facilitate innovative assessment practises.

Infrastructure

- Development of authentic learning environments and open source digital tools adaptable to educators' and learners' needs.
- Provision of an open access platform with e-assessment tools and evidence about their effectiveness.

Research

- More research on the effectiveness of innovative assessment practices.
- Collaboration among universities, professional organisations and institutions for the effective development of innovative assessment tools (e.g. digital badges, simulations).

1. Introduction

1.1 Background and context

Assessment has influence and impact on curriculum, pedagogies, learners' performance and ultimately to education as a whole. It has been a central feature of education for centuries, sparking fierce debates on its role, relevance and delivery modes. Assessment has a backwash effect on teaching and learning: it dictates what learners should know and be able to do at the end of a unit of study. It therefore affects the ways in which educators teach, and the content they cover during lessons. It allows monitoring and eventually improving the quality of both teaching and learning.

Assessment exerts its effects on educators and learners alike in all educational settings (formal, non-formal, informal, life-long). This report focuses on scoping out innovative forms of assessment that could better accommodate the needs and priorities of those involved in education.

The term 'innovative assessment' should be understood in this report as an umbrella term under which we collect assessment methods and practices that, while being efficient at capturing complex learning processes and outcomes, provide 'new and better' solutions and approaches. Despite a specific interest in identifying how digital technologies can foster innovative practices, the scope of the study considers all forms and modes of assessment that facilitate innovation in teaching and learning.

The key questions for this study are:

- What constitutes 'innovative assessment'?
- How can innovative assessment practices be successfully implemented?
- How can policy and practice foster innovative assessment that addresses learners' needs?

1.2 Structure and contents of the report

The report is organized as follows:

- Chapter 1 gives an overview of the report, its scope and its methodology.
- Chapter 2 discusses the rational of the study, describes the meaning and functions of assessment and explains why innovative assessment is important.
- Chapter 3 provides a definition of innovative assessment and suggests how it can be successfully implemented.
- Chapter 4 synthesises empirical evidence on a variety of innovative assessment practices (self- and peer-assessment, open badges, simulations, learning analytics).
- Chapter 5 presents eight case studies of innovative assessment schemes from various EU countries, makes cross comparisons and outlines the insights gained from practitioners.
- Chapter 6 summarises the key conclusion and makes a set of recommendations for educational policy and practice.
- Finally, Annex provides an overview of digital assessment tools.

1.3 Scope of the report

This project aims to bring evidence to the debate about the assessment methods that are already in place as well as to contribute to the discussion on how to innovate assessment.

In particular, the main objectives of the current study are:

- Develop a concept of innovative assessment.
- Describe how innovative assessment practices can be established.
- Provide empirical evidence on innovative assessment practices, which enhance learners' achievements and support the development of key competences.
- Propose policy recommendations to further implement innovative assessment.

1.4 Reports' methodology

For the purposes of this report, extended desk research has been performed to identify relevant literature. The selection and analysis of the available documents (e.g. articles, books, reports) was performed with the purpose to acquire a deeper understanding of the field and provide a definition of what constitutes innovative assessment. Five innovative assessment approaches are described in more detail and evidence of their effectiveness is provided based on recent research, metanalyses and best evidence syntheses published in scientific journals.

The case studies were selected with a view to capturing a variety of the best possible innovative practices currently existing in Europe. Eight innovative assessment practices from formal education and business were obtained. For each examined practice, desk research was carried out along with several rounds of enquiries and a 45-minute phone interview with a leading staff member/representative from the implementing authority. A short case study report (i.e. 5-7 pages) was produced for each. The reports provide a concise description of the design and implementation of the innovative practices, identify the challenges, the success factors and their potential in terms of replicability.

2. Rationale of the study

2.1 Meaning of assessment

Assessment of student learning is defined as *"the systematic collection of information about student learning, using the time, knowledge, expertise and resources available, in order to inform decisions about how to improve learning"* (Walvoord, 2004 in Zacharis, 2010: 61). Through assessment, educators develop an understanding on what learners have learned, how effectively they accomplished the assigned tasks as well as the efficiency of materials, methods and techniques applied upon students' learning (Zacharis, 2010).

Nevertheless, assessment should not be narrowed on how well a system or a learner performs but, based upon pedagogy, to decide where to go next and pave the path towards that direction (Hattie, 2012; Thummaphan, 2017). Furthermore, it supports his/her personal development as, *"recognition of one's accomplishments is key to developing the identity of the individual"* (Jones, Hope & Adams, 2018: 430).

Assessment comprises formative, diagnostic and summative. Formative assessment is the process of gathering information about students' progress and making interpretations with the purpose to modify the teaching - learning processes according to learners' needs (Black & Wiliam, 1998). Diagnostic assessment examines learners' prior knowledge and identifies misconceptions, which causes problems in learning with the purpose to acknowledge the nature of their difficulties (Fuchs, Fuchs, Hosp & Hamlett, 2003 in Dega, 2019: 14) and is considered as part of formative assessment. Summative assessment judges what the learner has learned at a particular time (e.g. completion of a unit or course) in relation to some goals or standards (Siarova, Sternadel & Mašidlauskaitė, 2017).

Crossouard (2011) proposes an integrated approach of assessment blending summative and formative approaches while having learners assume an active role in monitoring their learning process and reflecting on it. Such an approach, he claims, brings together the advantages of both models with a positive impact on learners' development.

2.2 Arguments for innovation in assessment

Assessment often determines the priorities of education. It highlights what is mostly valued and the way it is implemented shapes learning processes and how education is organised. Most importantly, it provides a vision for the kind of education that a society is envisaging.

The reflection paper on the social dimension of Europe (European Commission, 2017) illustrates the radical and profound changes Europe confronted from 1990 onwards which transformed its economy and society. The New skills agenda for Europe emphasises the need to invest in education and lifelong learning and posits that citizens need to be equipped with a wider range of competences. *"Competencies are defined as a combination of knowledge, skills and attitudes appropriate to the context"* which a person needs for personal fulfilment and social inclusion (European Parliament & Council of the European Union, 2006: 13). The updated version of the reference framework includes eight set of competences (European Commission, 2018) (see Figure 1) which is a combination of traditional subjects and competences related to lifelong learning, personal development, successful integration and democratic participation within a radically changing society.

In addition, the rapid growth of digital tools used by learners demands a change of shift in education. **Today people should not only be able to elicit information but also to evaluate, analyse, create and apply those skills to address new problems and find new solutions, collaborate effectively and communicate persuasively.** People are expected to be ever alert and ready to make complex decisions in situations with access to unlimited information and to adjust their actions and attitudes according to possible risks and unforeseen problems. They must also learn quickly to use newer and better technological devices and programmes with ever-shorter life spans while leaning on their problem-solving skills (Halpern, 2008).

Decisions on assessment will have an impact on how competences are taught. As competence-based education emphasizes strategical thinking and decision making there is a **need to develop new forms of assessment through which learners have an active role, become aware of their learning processes and needs, and develop a sense of responsibility for their learning** (European Commission, 2018). Boud (2010 in Boud & Soler, 2016) maintains that assessment should be sustainable and should focus on two key elements: foster the desire for self-reflection and form dispositions for practice to prepare students for their future learning needs. It has therefore, become a necessity to look for innovative and efficient assessment practices which comply with the new educational priorities, preparing learners to operate effectively in the longer term as citizens of the world.

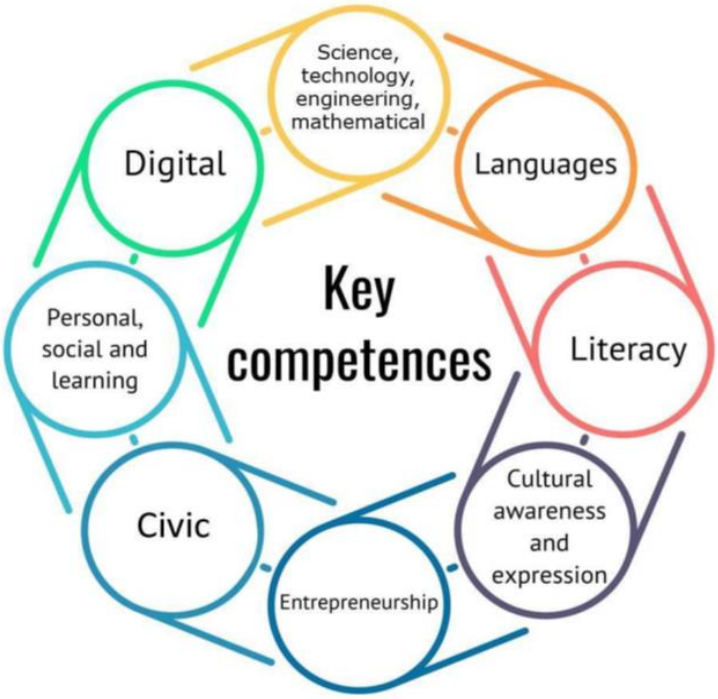

Figure 1. The European reference framework of key competences (in European Commission, 2018: 40).

3. Innovative assessment

3.1 Defining innovative assessment

The definition of innovative assessment varies according to the scope and the context of its implementation. At a basic level, it refers to assessment methods and techniques that are original and novel to a specific context (Mowl, 2006).

Some researchers suggest that innovative assessment is a flexible form of assessment, as opposed to that undertaken from everyone at the same time and place (Béguin, 2016). It refers to procedures that can determine what learners have already comprehended and provide evidence about their performance other than end-exams (Perrotta, 2014; Gozuyesil & Tanriseven, 2017). Through flexible assessment, educators decide about students' learning and the support they need (Eggen, 2016; Gozuyesil & Tanriseven, 2017), they differentiate teaching and engage them in customized assessment experiences (Chiappe, Pinto & Arias, 2016). Harris and Bell (1990 in Mowl, 2006) also claim that innovative assessment relies upon a different use of available resources, as a genuine attempt to improve classroom procedures through a variety of practices (Mowl, 2006). **Therefore, innovative assessment comes as a combination of various methods and techniques (new ones, contemporary and 'old-fashioned') united to improve the quality of students' learning.**

Innovative assessment has also been considered as any form of assessment that attempts to enhance students' learning through different approaches and address learners' needs in a more efficient way (Vincent-Lancrin, Kärkkäinen, Pfotenhauer, Atkinson, Jacotin & Rimini, 2014). It can also include a variety of practices that respond to learners' diverse expectations (Boud and Associates, 2010). These changes do not mean they have not been applied in the past, but rather fit better to the circumstances of the learning context (Zacharis, 2010).

Moreover, innovation in assessment is bound to context and may differ from one discipline to another (Hounsell et al., 2007). Furthermore, the innovator's stance shapes the approach at hand, since *"an innovation in one situation may be something already established elsewhere, but its importance ... is that initiative takers and participants see it as an innovation in their circumstances"* (Hannan & Silver, 2000 in Zacharis, 2010: 63). On the other hand, innovative assessment does not last for long as it gradually fades (Hounsell et al., 2007) thus someone should constantly 'invest' on being innovative.

On the other hand, assessment can take various forms (e.g. formative, summative, self-assessment, peer-assessment) that should not be seen as competing (Lau, 2016) but rather complementing each other (Boud, 1995; Siarova et al., 2017). The implementation of a variety of assessment methods and techniques can be perceived as a *"complex system of communication"* (Knight, 2002: 285) that can grasp and describe more accurately the diverse processes taking place in a classroom. Clearly, each method maintains certain advantages yet, *"there is not one "right" way to build an innovative assessment system"* (Marion, Lyons & Pace, 2017: 5). Rather the synergy of various types of assessment can support learners and cater their needs more effectively (Siarova et al., 2017). Consequently, the **combination of approaches** along with adaptation into context and learners' needs holds promise than the exclusive use of one method. Such an approach offers greater flexibility to policy makers and practitioners for establishing a learner-centred assessment system.

Innovative assessment is also closely related to classroom processes, aiming to improve instruction practices and enhance learners' engagement (Marion et al., 2017). In this respect, students are key-classroom-players and an innovative attempt should focus to *"the redistribution of educational power when assessment becomes not just something which is 'done to' learners but also 'done with' and 'done by' learners"* (Heron, 1981 in Mowl, 2006: 2). Despite being time consuming, educators need to allocate space for learners' self- and peer-assessment. It should also acknowledge learner's individual assessment experiences and priorities, which may include personal learning styles and different levels of understanding (Bevitt, 2015). Moreover, learners need to make decisions based on criteria not merely prescribed by the educator. Rather, they can adapt a set of innovative to them- approaches compatible to their learning needs and style and exercise their self-regulating skills (Valdivia, 2009 in Chiappe et al., 2016). Therefore, **innovative assessment should assist, encourage and inspire learners to assume a more active role and participate in the design and implementation of an innovative assessment approach**.

Moreover, OECD (2016: 32) suggested that although *"...innovation in education is not synonymous with the introduction of digital technology, [it] should include the smart implementation and use of technology in a way that leverages their potential for better teaching and learning practices"*. Thus, new methods that ensue from the use of digital technologies could assist the measurement of complex skills and processes through easy to access and free applicable tools, open to collaboration, adjustable to context and able to meet the diverse learning needs (Thummaphan, 2017; Chiappe et al., 2016).

Digital technologies can improve the validity and reliability of assessment practices (Oldfield, Broadfoot, Sutherland & Timmis, n.d.). They can as well assist in combining formative and summative techniques (e.g. e-portfolios, learning analytics) and formats (e.g. self-, peer- or group assessment). Digital tools can moreover monitor students' learning in real time (OECD, 2016; Brauer & Siklander, 2017) with increased personalization, flexibility and relevance to learners' interests and needs (Hofer, Duggan & Moellendick, 2018). An integrated approach can easily trace individual or group progress, aggregate evidence, comprehensively accumulate data, provide immediate feedback and make assessment more meaningful and authentic. Such a paradigm does not provide a glimpse of what has been mastered but rather gives a more complete account of learners' performance. On the other hand,

we should be aware that pedagogy should come first since *"technology alone cannot transform assessment practices"* (Broadfoot, Timmis, Payton, Oldfield & Sutherland, 2012/2013: 3) and therefore, digital tools should be tailored to the needs of educators and learners.

Most importantly, innovative assessment needs to be **sustainable**. It should not only focus on the current needs of the learners but also prepare them for their future needs. Boud (Boud, 2007: 19 in Boud & Soler, 2016) proposes that both educators and learners should engage in informed judgement, which refers to *"the capacity to evaluate evidence, appraise situations and circumstances astutely, to draw sound conclusions and act in accordance with this analysis"*. **The ability to reflect and self-regulate are at the heart of such processes.** An innovative approach should encourage self-reflection and the development of certain dispositions for practice (e.g. working with peers, develop assessment criteria, engage in peer-assessment). Such processes will assist learners to become deep rather than surface learners and make them capable of pursuing self-criticism and evaluation.

To sum up, innovative assessment refers to approaches that are novel, flexible enough and adaptable to context yet customised to better understand learning processes. They should capitalise on the benefits of the available assessment methods, techniques and adopt an integrative approach, while giving enough space to learners to engage in reflective practices and actively accommodate their learning needs.

3.2 Implementation of innovative assessment

The introduction of an innovation needs the coordination of all aspects of the educational system. In particular, policies, curricula, teaching methods and learning outcomes should be aligned and all implementation levels should be perceived as interrelated in order to address learners' needs (Rust, 2007). Yet, challenges lie among micro-, meso-, macro-levels (micro-level: classroom; meso-level: curricula; macro-level: education policy). They should be perceived as interrelated since *"it is difficult to speak of both the successes and the challenges that occur at one level without acknowledging their impact on the other two levels"* (Lock, Kim, Koh & Wilcox, 2018: 12). Thus, the impact of all levels has to be acknowledged and its successful implementation requires constant steering, support and negotiation among partners (Siarova et al., 2017).

The introduction of an innovative system may be hindered by internal and external factors. At the school level, external barriers may be the poor network infrastructure, the lack of adequate resources (hardware & software) and in-school technical support (Lucas, 2018). The structure and the culture of an organisation may also become a demotivating factor since restrictions or lack of an incentive system hinder individuals from taking the risk to innovate (Tierney, 2014). Furthermore, tight timetables and time investment for mastery of digital tools at school discourage the implementation of technology-based activities (Lucas, 2018).

At policy level, heavily prescribed curriculums and compulsory high-stake testing (e.g. national exams) may leave very little space for innovative practices. The lack of appropriate top-down support does not facilitate the establishment of flexible and innovative assessment practices (Brečko, Kampylis & Punie, 2014; Shimasaki, 2015 in Lucas, 2018). Apart from adequate financing, an implementation plan for mainstreaming and up-scaling is also needed. The sustainability of an escalated innovation depends on a monitoring working group, which coordinates stakeholders (Balanskat et al., 2013 in Lucas, 2018; Veiga Ávila et al., 2017). Most importantly, the added value of educational innovation has to be clearly communicated and supported by policy measures. Many educational authorities invest on infrastructure rather on educators' initial and continuous training which impedes their professional development (Broadfoot et al., 2012/2013; Lopes, 2010 in Lucas, 2018).

Internal barriers include teachers' attitudes and beliefs of the positive effect of innovative assessment on students' learning, their knowledge and skills of various assessment tools and how they can successfully be integrated within teaching-learning processes (Lucas, 2018). In particular, educators' value beliefs strongly predicted the quantity and quality of technology integration. Evidence also indicates that according to their value beliefs they place different weight to external barriers. Interestingly, this affects the way educators perceive access constraints, for example, those with high value belief on technology try to overcome such barriers through alternative solutions (Vongkulluksn, Xie & Bowman, 2018; Ertmer, Ottenbreitl-Leftwich, Sadik, Sendurur & Sendurur, 2012). Therefore, educators' professional development programs should foster positive value beliefs towards innovative assessment. Educators need to become aware on how innovative approaches can pedagogically promote different assessment models.

However, the process of balancing various assessment methods and techniques is a challenging task. Educators have to take decisions: which methods/techniques to select and how to blend them. This requires a sound knowledge of assessment practices and how they affect learning and an understanding of the particular purposes and circumstances at a given time (Hounsell, Xu & Tai, 2007). Therefore, **the effectiveness of innovative assessment depends on the efficiency of educators' decision-making.** Hence, teachers' continuing professional development is the cornerstone of a successful implementation (The Assessment Reform Group, 1999).

Policy makers should provide schools with autonomy and educators should be given enough space for experimentation and support if change towards innovative assessment is to be achieved (Lucas, 2018).

4. Innovative assessment approaches

Given that learners have to be the central axis of any system of assessment, self- and peer-assessment hold a prominent position. This section presents evidence on the impact of self-and peer-assessment on teaching learning processes and how they can assist the development of key competences. In addition, three digital assessment tools were selected which are context based, meaningful and provide a comprehensive view of authentic learning: digital badges, simulations and learning analytics.

4.1 Self-assessment

Self-assessment refers to the ability of learners to evaluate the process of their learning as well as the quality of their completed tasks. It is considered as an integrated part of self-regulated learning since the learner is engaged in monitoring and evaluating of both learning process and outcomes. During self-assessment, the learner has usually to evaluate his/her learning against some performance criteria (Brown & Harris, 2014).

However, **self-assessment is most effective when the learner engages in critical reflection** that may lead to significant insights and enhances self-understanding. During self-reflection, learners have to be able to examine their thoughts and emotions, to question their assumptions and the way they perceive and interpret events while taking into consideration external factors (Desjarlais & Smith, 2011; Melrose, 2017). Hence, through critical reflection the learner has to change thinking and consider new ideas. This, in turn may prompt incidental learning (unexpected learning which is not related to predetermined goals) and open their thinking beyond the boundaries of a particular discipline or a learning event (Melrose, 2017).

The process of self-assessment requires from the learner three steps of action. First, to be aware of the gap between his actual level of competence and the expected one. Second, to know what alternatives actions exist and choose the appropriate one. Thirdly, to execute the action and accomplish the expected level of competence. Self-assessment is an essential part of formative assessment, as the learner has to take the steps and close the gap between the current level and the expected one (Sadler, 1989).

Forms of self-assessment include assessment without given criteria (subjective self-rating), rubrics which include criteria and performance standards, scripts where learners have to answer questions which focus on the task process (Panadero, Brown & Strijbos, 2016) and reflective journals (Melrose, 2017). Other formats are checklists of tasks or process, smiley face, ordinal ranks where learners judge the quality of their work (self-rating) and self-marking using objective scoring guides (Brown & Harris, 2014). Self-marking and subjective self-rating are considered to be less powerful as they do not require complex evaluations and they are considered as a shallow learning approach to the task (Boud & Falchikov, 1989; Brown & Harris, 2014). However, **recent outcomes suggest that positive outcomes depend on the combination of different self-assessment interventions and the way they are implemented rather than specific formats**. For example, students who used rubrics to assess their drafts but who also generated criteria from a model paper produced more effective writing compared to those who used only rubrics (Andrade, Wang, Du & Akawi, 2009; Andrade, Du & Mycek, 2010).

The ability of the learner to identify what went right and what went wrong to amend behavior is a key requirement of self-assessment. Such inferences need to be realistic in the sense that their descriptions are accurate (Panadero et al., 2016). The level of accuracy can be determined by the alignment of self-judgment on actual task performance or against the judgment of an expert (Topping, 2003). If learners are able to accurately detect what is right and wrong in their work and why they end in such an outcome, it is easier for them to find pathways towards self-improvement (Boud & Falchikov, 1989).

However, outcomes from various professional contexts consistently illustrate that when people assess themselves they make significant errors. They overestimate their expertise, skills and character traits. Evidence suggest that learners do not possess all the information required to reach accurate self-assessment and even if they do have such information, they might give little weight or not even take them into account (Dunning, Heath & Suls, 2004). Hence, they overrate the performance, the skills being mastered and they often identify their weakness inaccurately (Regehr & Eva, 2006; Colthart et al., 2008). However, even if assessment is inaccurate, learner's engagement in such a process is worthwhile as the aim of self-assessment is not to provide an accountable monitoring of the level reached by the learner, but rather to **stimulate reflection on their own learning**, making their learning visible to themselves and reflect on how to improve their perform.

Many studies have shown that self-assessment skills can be taught and that learners' engagement in developing assessment criteria has positive effect on the accuracy of their self-judgements (Brown & Harris, 2014), while clear, concrete and well understood criteria enhances the accuracy of self-assessment (Panadero & Romero, 2014). On the other hand, some learners prefer to rely on educators' assessment than assessing themselves hence they need to be persuaded about the benefits of assessing the quality of their work (Panadero, Brown & Courtney, 2014). Learners may also not wish to disclose their self-assessments in particular if they are negative (Cowie, 2009 in Panadero et al., 2016). Nevertheless, learners need to become aware of the discrepancy between their own evaluation and an external one for improvement to occur. Initially, a learning conversation with the provision of quality feedback can assist learners in such a process without harming their self-esteem. Most importantly, a relationship of trust and safety is required for honest disclosure to occur (Alonso-Tapia & Pardo, 2006).

The level of learners' expertise and prior knowledge has considerable impact on the quality of their self-assessment. Novice learners do not have knowledge of the standards for quality work, they have difficulties to change their actions and to evaluate their products and less able students tend to overestimate the quality of their work (Boud & Falchikov, 1989). In contrast, those who underestimate their performance are usually very competent. It has been suggested that expertise increases a person's awareness of how much more there is to learn which influences their self-judgment (Dunning, Heath & Suls, 2004).

Cognitive psychologists suggest that the lack of prior knowledge (both declarative and procedural) poses high cognitive load, the capacity of working memory is limited and the process of self-assessment becomes challenging for the learner. If self-assessment repeatedly confirms poor performance, then the learner will not be motivated to allocate effort into learning a task. **In view of the fact that self-efficacy is a key component for self-regulated learning educators' scaffolding of learning tasks and of self-assessment is essential** (Kostons, van Gog & Paas, 2009).

Models of self-regulated learning suggest a cyclical process consisting of three phases: a preparatory (forethought), a performance and an appraisal phase (self-reflection) (Panadero, Jonsson & Botella, 2017). Empirical evidence illustrates the significant impact of self-assessment on all three phases with the strongest effect on the forethought and self-reflection phase (Panadero & Romero, 2014). Hence, learners are to be engaged in self-assessment even before they perform their task, as this will assist task monitoring and evaluation.

Moreover, research suggest that when learners focus on some aspects of their work according to their developmental characteristics and the nature of the task, learning improves and they become able to generate assessment criteria against which to evaluate work (Zimmerman, 1989). Finally, assessment should not be limited to externally devised objectives and goals but should also be based on personal goals and expectancies. Apart from deeper learning, reflection can prompt the development of competencies which are not explicitly taught but which enrich the professional identity of the learner and therefore satisfy personal needs (Bourke, 2014).

4.1.1 Technology-supported self-assessment

Apart from traditional modes, self-assessment is delivered through computer and mobile devices. Students consider computer/mobile assessment as an interesting task and perceive themselves as able to perform well (Chua, 2012; Chua & Don, 2013). Mobile devices are more appealing to young learners and combined to their self-efficacy beliefs they show greater motivation and engagement, leading to better learning achievements (Hwang & Chang, 2011). Overall, research points out that learners who perceive themselves as efficient users of digital tools are more motivated and achieve better learning outcomes (Castillo-Merino & Serradell-Lopez, 2014).

Research on the impact of assessment modes during self-assessment confirms the above data. High school students engaged in self-assessment in physics through computer/mobile devices showed increased motivation compared to paper based. Apart from increased motivation, they also demonstrated higher learning performance. Most importantly, low achievers had the highest gains compared to medium and high achieving learners (Nikou & Economides, 2016). Furthermore, online self-assessment predicts exam results even when class attendance was taken into account (Buchanan, 2001) and improves final exams pass rates (Ćukušić, Garaća & Jadrić, 2014). Similarly, learners who engaged in computer self-assessment achieved 10% better exam results compared to those who did not (Wilson, Boyd, Chen & Jamal, 2011).

Conclusively, young learners feel more confident with digital self-assessment, tasks become more appealing and therefore, demonstrate greater engagement. In addition, learning achievement is higher for those who engage in computer self-assessment.

4.1.2 Evidence on self-assessment

A systematic review of studies has demonstrated the **positive impact** of self-assessment on learning and achievement across a range of grades and subjects (Brown & Harris, 2014) and on student's self-regulated strategies and self-efficacy (Panadero et al., 2017). **It has been suggested that engagement in self-assessment enhances deeper learning and therefore learners have better performance.** Subsequently, this generates feelings of worth, a perception of improved capability that increases their self-efficacy. Further evidence points out that self-efficacy was one of the constructs with the strongest effect on learning for adults along with goal level, persistence and effort (Sitzmann & Ely, 2011).

Moreover, self-assessment empowers learners as they gain ownership of their learning and that increases further the use of self-regulatory strategies (Black & William, 1998 in Siarova, Sternadel & Mašidlauskaitė, 2017; Taras, 2010; Tan, 2012 in Panadero et al., 2016). Empirical evidence also suggests that self-assessment is related to increased student motivation, engagement, behavior and quality of student-teacher relationship (Griffiths & Davies, 1993; Schunk, 1996; Olina & Sullivan 2002; Munns & Woodward, 2006; Glaser et al., 2010 in Panadero et al., 2016). Furthermore, when self-assessment was accompanied by external feedback the correlation between self-assessment and learning was much stronger (Sitzmann, Ely, Brown & Bauer, 2017).

Learners' outcomes differ according to their ability although evidence is contradictory. Low performing students have larger learning gains (Sadler & Good, 2006) while other researchers suggest that average students who are more accurate in their self-assessment benefit the most (Boud, Lawson & Thompson (2013).

Conclusively, self-assessment is an essential component of innovative assessment not only for improving performance but mostly as a valuable means for learner's empowerment and **self-sustained** learning.

4.2 Peer-assessment

Peer-assessment refers to ***"a reciprocal process whereby students' produce feedback reviews on the work of peers and receive feedback from peers on their own work"*** (Nicol, Thomson & Breslin, 2014: 102). Peer-assessment can be formative or summative, quantitative (providing grades) or qualitative (providing extended verbal feedback) and a variety of products can be peer-assessed such as written assignments, presentations, portfolios, oral statements, scientific problems etc. (Topping, 2017).

Although there is considerable research on scores and grades awarded by peers for summative purposes (Topping, 1998), the current section will focus on the formative type of peer-assessment which includes qualitative feedback. Such assessment requires three skills. First the ability of defining assessment criteria in relation to the work to be assessed, second to identify its strengths and weaknesses and third to suggest areas for future learning (Sluijsmans, Brand-Gruwel, van Merrienboer & Martens, 2004).

Research has shown that both receiving review from peers and producing reviews for peers improve learners' performance in writing assignments (Cho & MacArthur, 2011; Cho & Cho, 2011). Receiving reviews, learners become aware of the different ways readers may interpret their work which may

confirm, supplement or conflict the learners' evaluation of their own work. Such feedback highlights discrepancies and assists them to focus on aspects that need improvement. On the other hand, producing reviews learners have the opportunity to critically think, to apply criteria and engage in reflection. Nicol and colleagues (2014: 116) have suggested that during the reviewing process learners evaluate peer work *"against an internal representation of their own work"*. Apart from the external criteria provided by the educator, learners use implicit criteria deriving from their own experience when completing similar with their peers' assignment. When students have to review a number of peer-work they come in front to a greater range of possibilities compared to the alternatives offered by one person even if that one is an expert. In turn, the learner may generate richer criteria but most importantly, the experience of applying such criteria in practice has shown to **facilitate internalisation and transfer of learning** (Price & O'Donovan, 2006; Nicol et al., 2014).

Learners find it easier to analyse others' work compared to their own because they can adopt a distanced perspective. Furthermore, by reviewing a variety of examples they gradually become aware of the desired performance (Black, Harrison, Lee, Marshall & Wiliam, 2003). Reinholz (2016) suggest that through peer-assessment learners **develop objective lenses**, which they can later apply to their own work. They also have to explain their own reasoning that promotes self-awareness. Therefore, apart from the development of communication skills and conceptual understanding, peer-assessment supports the development of self-assessment (Black et al., 2003). Finally, the act of critical appraisal will assist them in their future careers as they will have to appraise and comment on the work or performance of others as well as enhance their ability to produce quality work and therefore, prompt the development of their professional skills (Topping, 2017).

Overall, students have positive attitudes towards peer-feedback and they positively appraise the received feedback (Rotsaert, Panadero, Schellens & Raes, 2018). Students as assessors use a more comprehensible language and their comments are more likely adjusted to the learners' level of understanding as both assessors and assessees face similar challenges (Panadero, Jonsson & Alqassab, 2018). Although the accuracy of peer-assessment varies, empirical evidence suggests that it is beneficial for learning (see section 4.2.2). Peers are not considered experts and their feedback induce uncertainty that stimulates learners to explore further confirmation and/or perform self-corrections, thus, a deeper understanding of the subject is achieved. In contrast, educators' feedback is not questioned since learners' regard them as 'knowledge authority' (Yang, Badger & Zhen, 2006).

Providing feedback is not an easy task and educators' support is needed for quality improvement. Research has shown that feedback is considerably improved when guiding questions are given to learners. They provide a more balanced account of negative and positive judgements and more elaborations of what the assesse should improve and why (Gielen & De Wever, 2012, 2015). In addition to guiding questions, the use of rubric and continuous practice, assisted students to make more sound evaluations and develop further their expertise (Rotsaert et al., 2018). Most importantly, generating criteria through discussion with students increases significantly the reliability of peer-assessment (Falchikov & Goldfinch, 2000). Interestingly, students' perceptions of their improvement in peer-assessment comply with their actual progress in producing quality feedback over time (Rotsaert et al., 2018).

A number of interpersonal factors influence the quality of peer-assessment. Overall, students tend to overscore their peers in some cases with the purpose to enhance relationship with them. They also do not feel comfortable marking friends, fearing that they may be rejected. Thus, anonymity can alleviate peer pressure and establish conditions of psychological safety. Students may also not trust their peers as

assessors and consider such feedback as less valid. Constant practice and discussion about the assessment criteria establish transparency and increase trust to the whole process (Panadero, 2016).

Conclusively, peer-assessment becomes a constructive task through which the learner has to receive and give feedback, provide informed judgements, extract meaning and implement suggestions for improvement, yet practice and constant monitoring from educators is needed, if optimal learning outcomes are to be achieved. Most importantly, it **fosters a sense of shared purpose and responsibility for learning** which empowers learners and prepares them for their future learning needs.

4.2.1 Technology-supported peer-assessment

Recent technological advances have developed a number of online peer-assessment tools (Babik, Gehringer, Kidd, Pramudianto & Tinapple, 2016). PeerScholar is an example of such a tool used in various studies, which consists of four phases. Learners in the first phase (create) submit their assignment according to educators' instruction and rubric guidelines. In the second phase (assess), students have to anonymously assess a number of peer assignments; in the third phase (reflect/revise) they had to reflect on peer-feedback and then proceed with revisions of their work. Educators in the final phase (evaluate) have to evaluate the revised assignment and the peer-assessment processes. Various research studies have shown student's positive attitudes towards such assessment (Davies, 2004; Paré & Joordens, 2008; Collimore, Paré & Joordens, 2015).

Students who participated in peerScholar report that they liked the anonymity and reading the opinions of their peers while they acknowledged that peer-feedback helped them to improve their work (Collimore, Paré & Joordens, 2015). Most importantly, peer-assessment assisted the development of **critical thinking** skills as students had to examine their peers' assignment, point out the strengths and weakness of their work, justify their comments and make suggestions for improvement (Pare & Joordens, 2009, 2008). This in turn, influenced their own work as they became more competent in applying assessment criteria (Li, Liu & Steckelberg, 2010) and develop self-regulatory strategies (Gikandi & Morrow, 2016). Interestingly, peer- and teacher marking were almost similar while evidence indicate that 5-6 peer assessors are the optimal for a valid outcome (Paré & Joordens, 2009).

Peer-assessment through digitized learning formats holds **significant advantages**. First, with online tools learners can give and receive feedback immediately to peers and educators alike, a vast improvement over classroom based communication practices. Second, when needed, online tools offer anonymity for students receiving and giving feedback. By removing the risk of exposure in front of peers, students experience less pressure and fear of disapproval, they tend to be more positive and their satisfaction and compliance towards peer-assessment is increased (Vanderhoven, Raes, Motrieux, Rotsaert & Schellens, 2015; Güler, 2016). Research points out that students who gave feedback in a web forum anonymously provided five time more critical feedback compared to those whose identity was known highlighting how interpersonal variables impact the quality of peer-assessment (Howard, Barrett & Frick, 2010).

4.2.2 Evidence on peer-assessment

Systematic review of studies points out the positive effects of peer-assessment on learners' achievement (van Gennip, Segers & Tillema, 2009). In relation to the quality of peers' feedback, the use of justifications significantly improved performance but the effect diminished for students with high

performance (Gielen, Peeters, Dochy, Onghena & Struyven, 2010). There is also evidence that peer-assessment has a positive impact on learners' motivation (Hsia, Huang & Hwang, 2016; Lai & Hwang, 2015), creativity (Hwang, Hung & Chen, 2014), self-regulation skills (Gikandi & Morrow, 2016), self-efficacy (Hsia et al., 2016), critical thinking (Harrison, O'Hara & McNamara, 2015; Lai & Hwang, 2015; Nicol et al., 2014), problem-solving skills (Hwang et al., 2014; Moore & Teather, 2013) and overall enhancement of student learning and performance (Hsia et al., 2016; Hwang et al., 2014; Kablan, 2014; Mulder, Baik, Naylor & Pearce 2014).

Nevertheless, the effectiveness of peer-assessment depends on several aspects. In particular, learners should have opportunities to give and receive peer-feedback more than once in a particular task, to discuss about their given and received feedback (Gikandi & Morrow, 2016; Reinholz, 2016) and direct attention to the learning task, task processing strategies and self-regulation strategies instead of the 'self' (Hattie & Timperley, 2007).

4.3 Digital badges

Digital badges are digital visual rewards for non-tangible accomplished tasks, competencies, providing an account of one's life-long learning[1] trajectory. They may refer to either autonomous or prescribed learning pathways and are awarded by groups, institutions or organizations (Frederiksen, 2013; Gibson, Ostashewski, Flintoff, Grant & Knight, 2013; Anderson & Staub, 2015; O'Byrne, Schenke, Willis & Hickey, 2015; Liyanagunawardena, Scalzavara & Williams, 2017; Carey & Stefaniak, 2018; Hofer et al., 2018). They are available online, contain metadata[2] (e.g. information about the issuer, evaluation criteria, process and result of the accomplishment) that validate acquired skills (Gibson et al., 2013; Anderson & Staub, 2015; Devedžić & Jovanović, 2015; Ellis, Nunn & Avella, 2016; Eaglen Bertrando, 2017) and acknowledge prior learning (Lius, 2016). Some badges are credentials of learning within a close system (e.g. Duolingo[3] for foreign languages) yet, most of them are open and their metadata can be transferred into other systems (Farmer & West, 2016).

They can also assist the creation of learners' e-Portfolio and present their profile of interests, accomplishments, competences and experiences (Lius, 2016; Cheng, Watson & Newby, 2018) that can be transparent through metadata (Gibson et al., 2013; Iwata, Clayton & Saravan, 2013; Casilli & Hickey, 2016) and systematically evaluated (Iwata, Telloyan, Murphy, Wang & Clayton, 2013). Thus, they act as a way to document life-long learning from which both learners and educators benefit (Ellis et al., 2016; Hamson-Utley & Heyman, 2016).

Digital badges communicate information about learners' achievements from formal, non-formal and informal education (Jovanovic & Devedzic, 2014). There is a growing demand for soft accreditation as learners seek acknowledgement of their informal learning activities (Law & Law, 2014). Digital badges offer such credentials and are already used in applications for recruitment, for pre-professional practice, human resource training programmes, informal out of school learning and for many other purposes (Gibson et al., 2013).

[1] Learning that occurs throughout life in formal, non-formal and informal settings (Schuetze, 2007).
[2] *"Data that describes other data"* (https://whatis.techtarget.com/definition/metadata).
[3] https://www.duolingo.com/

4.3.1 Benefits of digital badges

There is evidence that digital badges can overcome the assessment challenges of traditional courses as they can recognize diverse learning trajectories and competencies that previously were not acknowledged, such as **21st century skills** and social skills (Abramovich, 2016; Farmer & West, 2016). They appear as a response to the revolution of the e-world, shifting achievement measurement from exams to personalised accomplishments (O'Byrne et al., 2015). Moreover, badging can bridge formal and informal learning as it can strengthen the learning outcomes from traditional degree programs (Carey & Stefaniak, 2018).

Digital badges encourage learners to personalise performance by planning in advance, even select content and criteria that are relevant to their preferences and needs (Farmer & West, 2016). Learners can develop their own learning path and accomplish a task in small fractions (granular learning) following the pace that suits them (Brauer & Siklander, 2017; Eaglen Bertrando, 2017; Carey & Stefaniak, 2018). On the other hand, educators provide scaffolding, guidance, support and encourage peer- and self-assessment (Jovanovic & Devedzic, 2014; Anderson & Staub, 2015; Devedžić & Jovanović, 2015). In this way, learners can self-regulate their professional development. There are successful examples of collaboration between universities and professional organisations that have developed badging programs. In United States, the National Science Teachers Association has collaborated with NASA and Penn State University and developed 63 professional development activities for educators. They were free to select activities and create their own learning journey and even decide about their level of achievement (high achievement: badge award; low achievement: stamp award) (Farmer & West, 2016). Given that both educators and learners had choices, the design of the program encouraged the development of autonomy and self-direction.

Digital badges can increase the expectations for success (Abramovich, Schunn & Higashi, 2013) -as they reward not only accomplishment but also engagement (Jovanovic & Devedzic, 2014; Carey & Stefaniak, 2018; Garnett & Button, 2018) through affective, behavioural and cognitive indicators (Hatzipanagos & Code, 2016) and thereby act as intrinsic and extrinsic motivators (Abramovich et al., 2013; Gibson et al., 2013; Ellis et al., 2016; Eaglen Bertrando, 2017; Cheng et al., 2018; Motheeram, Herselman & Botha, 2018).

Apart from giving incentives for their engagement (Gibson et al., 2013; Motheeram et al., 2018), digital badges maintain formative feedback (Abramovich et al., 2013) as learners may examine the metadata and reflect upon their intrapersonal and interpersonal development (Anderson & Staub, 2015). By facilitating metacognitive processes, it has been suggested that digital badges have the potential to *"improve learners' self-efficacy"* (Cheng et al., 2018: 193) and thereby lead to increased achievement (Jones et al., 2018).

One of the biggest challenges in the quality of badging systems is the lack of rigor. There is a flood of lightweight badges, which are awarded without the use of certain criteria. Their extended use *"may lead to increased use of narrow assessment formats … characterized as conformative*[4] *and deformative*[5] *assessment"* (Casilli & Hickey, 2016: 124). Due to their widespread use, such badges carry no weight and their potential for being credentials of authentic learning is lost. Lightweight badges can have positive effect if they reflect components of learning over time and hence, a collection of such badges illuminates the interests and profile of the learner (West & Randall, 2016). Yet, surveys point out that learners seek to heavyweight badges that carry social capital as evidence of performance that is valued by the

[4] Assessment that *"encourages instrumentalism"* and *"criteria compliance"* (Torrance, 2007: 282) rather than learning, while educators control learners' assessment *"by means of detailed instructions and a standardised way of assessing"* (Tolgfors & Öhman, 2015: 158).

[5] Formative assessment that does not necessarily have a positive impact on students' learning (Torrance, 2012).

professional community (Law & Law, 2014). West and Randall (2016) argue that emphasis should be given to the criteria required for someone to earn a badge and to the establishment of valid and reliable assessment practices yet further research is needed into the principles which will ensure quality of badges and their effectiveness on learning outcomes. Moreover, decision makers in business and governments are still unaware of the potential of badges to motivate and enhance learning and therefore dissemination of their benefits within society is needed (Farmer & West, 2016).

4.3.2 Evidence on effectiveness

Research points out that digital badges have a positive impact on learners' participation. Learners in MOOC **courses with a badge system participated five times more** (voting, posing questions and responding to questions) **compared to courses without a badge system** (Anderson, Hutterlocher, Kleinberg & Leskovec, 2013). In another study, students who had access to a badging system were significantly more engaged with the online learning tool (peer wise) and answered more questions compared to those who did not have access to badges. Yet, there was no effect on the number of learners' questions (Denny, 2013).

Chou and He (2017) investigated the impact of badges on class participation and interaction (original postings, comments or replies) according to delivery format (face to face and online university courses) and to pedagogical orientation (technology-activity based and read-write-reflect-comment course design). Learners' overall participation was higher in courses with a badge system compared to those without one. However, a significant effect was found only for comments and replies and not for learners' original posts (assignment submission). Interestingly, a badging effect was found only on the read-write-reflect-comment online course. Therefore, the nature of activities within online courses is a factor that also needs to be taken into account. Badges seem also to be less effective on interactive courses, which focus on hands on activities and projects.

Further evidence suggests that the impact of badges varies according to learners' age and their learning ability. Abramovich, Schunn and Higashi (2013) applied an artificial intelligent tutoring program with badges in order to improve middle school students' Math ability. The system issued badges for participating to the system and mastery of skills. Participatory badges assisted low performing students to stay engaged to the task. Yet, the more badges low performing students' earned, the less concerned they were about their performance, while the motivation of high performing ones was not affected. Similarly, middle school students with low mastery orientation seem to benefit from badges when participating in a Geometry game. Interestingly, the group with no badges outperformed the group with badges. Yet, the outcome was mitigated by mastery orientation (high correlation with post-test performance) indicating that badges may hinder students' performance of high mastery orientation (Biles, Plass & Homer, 2014). Another study has also investigated the effectiveness of badging platforms on learners' behaviour and performance (61 students' 13-14 years old applied ClassDojo and class badges in Geometry) and they were evaluated for their interaction with others, their effort to do an activity, their participation and the effect of class badges on their cognitive skills. The study used observations, semi-structured interviews and questionnaires. Learners with best averages scores in various participatory indicators had more badges, demonstrating an increased engagement in Geometry tasks (da Rocha Seixas, Gomes & de Melo Filho, 2016).

Further evidence also comes from research on **gamification**, which refers to online learning materials/services that integrate gaming design features such as digital badges. There are examples of online services, which unlock badges and positively affect usage activity. Experimental research in a peer-

to-peer trading system (Sharetribe) was applied with and without a badging system in a Finnish University over two years. The system provided badges for general user activity, carpooling, borrowing, offering help and for giving a free item on Christmas. They were awarded at three levels (bronze, silver, gold) according to usage frequency and users could view their badge in their profile. These findings suggest that badges had a positive effect on the number of page views, the number of comments, transactions and trade proposals (Hamari, 2017). In another experimental study, university students who followed traditional exercises were compared with students who had gamified experiences (earned badges) in ICT course. Students with gamified experiences had higher motivation, performed significantly better in practical assignments while students with the traditional exercises participated more (higher attendance/exercises completion) and achieved higher scores in written examination (Domínguez et al., 2013).

Moreover, **students respond differently according to the type of badge**. University students who completed interactive automatically assessed exercises about data and algorithms were awarded three types of badges (time management, carefulness and learning) which however did not affect their final grade. Time management badges had strong effect on Computer Science major students while carefulness on minor course students. A significant difference between treatment and control group appeared for one learning badge (C2) which required from students to redo a complete round of exercises (Hakulinen, Auvinen, Korhonen, 2013).

Conclusively, research indicates a **badging effect on participation and interaction on online environments and on motivation** yet, more evidence is needed on their effect on learning and performance while taking into consideration learners' age, ability, individual characteristics, pedagogical features and learning environments. Future research can elucidate how individual badges encourage desired learning behaviours.

4.4 Simulations

Simulations create scenarios-based environment that imitate the real world. They are **dynamic tools** where learners can apply their knowledge, practice skills, adopt various roles and experiment with different strategies in a safe environment. Most importantly learners can observe the outcomes of their actions, thereby assume responsibility of their decisions (Vlachopoulos & Makri, 2017). Simulations are also integrated in many games. Simulation games/scenarios are greatly used in Health Sciences, Biology and Business Marketing and are considered as ideal instruments for situated learning and transferring of knowledge in the workplace (Lukosch, Kurapati, Groen, & Verbraeck, 2016).

4.4.1 Benefits of simulations

Real time feedback in simulation games reduces stress and uncertainty as participants have a clear view of the objectives of the game (Nkhoma et al., 2014). Furthermore, learners' engagement and motivation are enhanced in simulated learning environments due to feelings of satisfaction that derives from peer-learning and cooperation with other participants. When instruction focus on strategies, which they can use in the virtual environment, learners' interest and engagement is augmented (Chang, Peng & Chao, 2010). The provision of visual feedback and the opportunities for manipulation during simulation training encourages learners to examine their assumptions and action while solving the task, thus enhancing reflection in action (Söderström, Häll, Nilsson, Ahlqvist, 2014/2015).

Furthermore, investigation of students' mistakes has shown that simulations are **ideal for training in**

decision making within complex and dynamic situations (Pasin & Giroux, 2011; Lin & Tu, 2012). The most effective instructional design features for simulation-based education are: variation in task complexity, opportunities for repetitive practice, practice over a period of time, learners' cognitive engagement (through task variation, intentional task sequencing, feedback, multiple repetitions), the use of multiple learning strategies, training tailored to individual learning needs, mastery learning of a clearly-defined standard of performance, provision of feedback during or after the simulation activity, longer time in practice and variation in the clinical context (Cook et al., 2013).

The role of the instructor is important, as s/he has to emphasize the learning goals, facilitate and support learners when new information and high order skills are involved (Kovalik & Kuo, 2012; Wouters & van Oostendorp, 2013). In particular, s/he has to prompt students to formulate hypothesis, describe observations, provide explanation and interpret the context to construct knowledge and deepen their understanding (Hämäläinen & Oksanen, 2014).

Research has also highlighted the benefits of debriefing which constitutes an essential component of simulation-based education (Tannenbaum & Cerasoli, 2012/2013). **Debriefing** refers to a discussion between two or more individuals where aspects of performance are analysed with the aim to gain insight that impacts professional practice. It is a form of formative assessment as the new insights are co-created by the instructor and the learner during discussion and aims to improve learners' current performance through constructive feedback (Eppich & Cheng, 2015). There are various models of debriefing, which usually use methods such as self-assessment, focused facilitation, directive assessment or a combination of them. In self-assessment the learner has to identify what went well, what problems occurred and suggest solutions to remedy them. During focused facilitation, the learner has to focus on performance deficits, discuss the reasons of their appearance and identify solutions. The discussion may take place between the learner and an expert but peers may also participate. During directive assessment, the instructor provides feedback in a didactic manner. S/he clarifies important learning points and provides information when knowledge gaps or performance deficits are identified (Cheng et al., 2015).

However, the management of debriefing is not easy and training is needed to secure its optimal use. Simulation educators consider debriefing as an overwhelming process as they want to avoid being harsh on the learner. Eppich and Cheng (2015) have developed a model for promoting excellence and reflective learning in simulation (PEALS). The particular model adopts a blended approach while ensuring that learning is active, collaborative, self-directed and learner centred. It has been developed for Health Care simulation, yet it can be useful across professions, disciplines and for different debriefing environments and its implementation consists of four phases: reaction, description, analysis and summary phase (see Figure 2).

In the first phase, learners are encouraged to express initial thoughts and their feelings. In the description phase, a summary of key events is provided which assist all group members to focus on main issues. In the analysis phase the educator has to decide about the ideal strategy while taking into account a number of variables such as the performance domain addressed (cognitive, behavioural, technical), if the rationale of performance deficit is evident, the time available (long, short, moderate) and the level of learners' expertise. A clear view of the above can guide decision making of the appropriate debriefing method (learner self-assessment, focused facilitation or directive feedback). In the summary phase, learners express what messages have been taken, the enablers and barriers they anticipate in their practice or/and the educator may provide a review of the main points (Eppich & Cheng, 2015). Thus, apart from performance improvement, learners who participate in debriefings are trained to prioritise topics, to transit from one topic to another, to redirect discussion and to deal with difficult situations. The

quality of debriefing can be assessed with various tools [e.g. Debriefing Assessment for Simulation in Health Care (DASH), Objective Structured Assessment of Debriefing (OSAD)] which have good reliability and validity measures and can be used for self- and peer-assessment as well as for formative and/or summative purposes (Cheng et al., 2015).

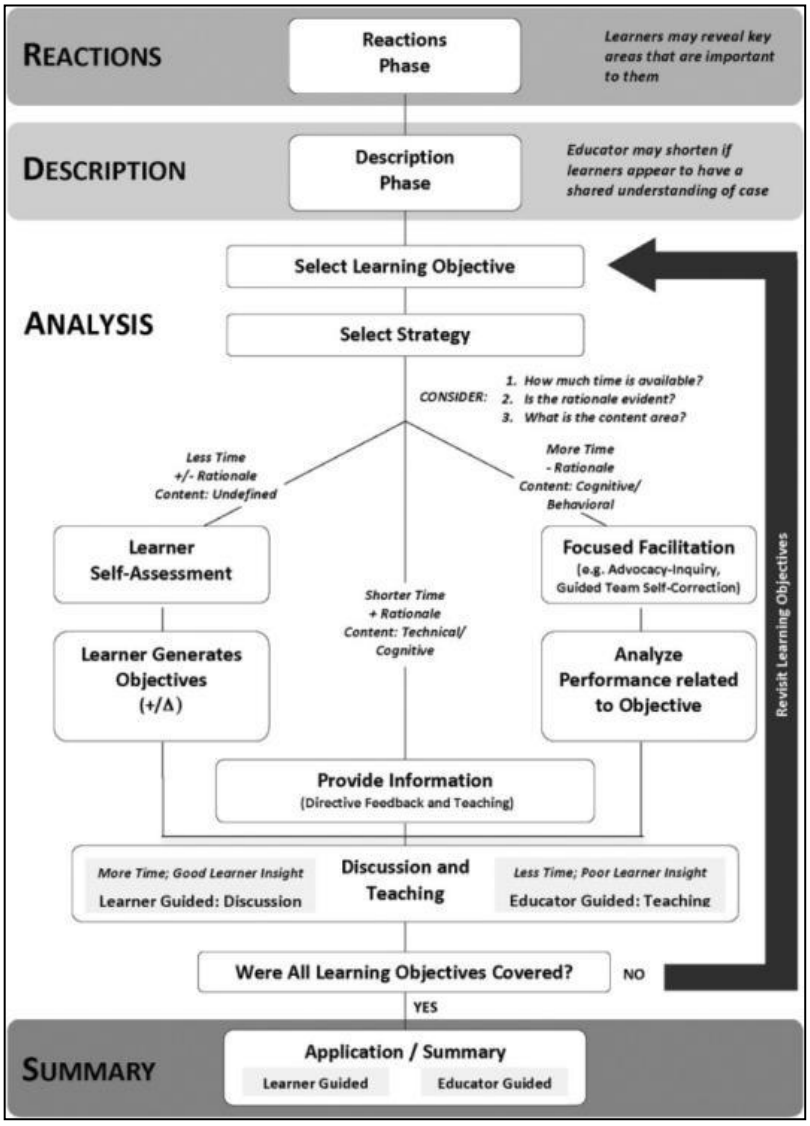

Figure 2. Phases of the debriefing process (in Eppich & Cheng, 2015: 110).

To sum up, simulations are ideal for situated learning as learners can practice skills and apply prior knowledge in authentic-like situations, which they will encounter in their professional life and thereby, they are considered as innovative assessment tools of great value. In addition, features integrated within simulation such as real time feedback, the quality of feedback by the instructor and learners' self-reflection of their simulated experiences during debriefing, are essential components that accentuate the effectiveness of simulation on learners' professional development.

4.4.2 Evidence on effectiveness

Simulation based assessment is widely applied in **Health Sciences** since it permits the testing of learners' performance without patient safety concerns. Overall, large positive effects appear on technology enhanced simulation compared to no intervention (Vogel et al., 2006; Cook et al., 2011; McGaghie, Issenberg, Cohen, Barsuk & Wayne, 2011; Ilgen, Sherbino & Cook, 2013; Cheng et al., 2014; Kim, Park & Shin, 2016) and small to medium effects compared to other instructional methods (Cook et al., 2012; Ilgen et al.,2013). On the other hand, meta-analyses of evidence demonstrates a large effect on learners' knowledge, skills and behaviour, a medium effect for time behaviour (how long it takes for task completion) and a small for patients' outcomes (Cook et al., 2011; McGaghie et al., 2011; Brydges, Hatala, Zendejas, Erwin & Cook, 2015).

Further evidence from all education sectors (K-12 and undergraduate) revealed a significant effect of simulations on learning outcomes even though the effect size was smaller compared to educational games (Merchant, Goetz, Cifuentes, Keeney-Kennicutt & Davis, 2014). In Compulsory Education, simulation has mostly been applied in Science Education and positive results have been reported on students' motivation and attitude, time on task, comprehension of lab task and the acquisition of practical laboratory skills (Rutten, van Joolingen & van der Veen, 2012).

Various studies suggest that simulation assist self-assessment (Arias Aranda, Haro Domiguez & Romerosa Martinez, 2010), **higher order thinking**[6] (Crocco, Offenholley & Hernandez, 2016) and the development of complex cognitive skills (Siewiorek, Saarinen, Lainema & Lehtinen, 2012) which facilitate deep learning. When simulated scenarios incorporate problem solving and reflective practices, **metacognitive thinking** is considerably enhanced (Hou, 2015). Multi role simulations where students have to develop arguments, make judgements and evaluate situations also assist the development of critical thinking and self-awareness (Silvia, 2012). In addition, when the learner has control on the level of difficulty and receives feedback after the simulation, **self-efficacy and transfer of learning** is significantly improved (Gegenfurtner, Quesada-Pallarès & Knogler, 2014).

Students outperformed when simulation was used in combination with other instructional practices than if used as a stand-alone practice (Sitzmann, 2011; Merchant et al., 2014). The quality of feedback also moderated learning outcomes. For declarative tasks, elaborative explanation was more effective than visual cues. For procedural tasks, knowledge of correct response feedback was more effective thus, knowing the correct answer learners could navigate their actions in the simulated environment more effectively (Merchant et al., 2014).

The highest effects on cognitive and affective outcomes also appear for high fidelity simulations, while medium fidelity simulations had highest effect on psychomotor outcomes (Kim et al., 2016). Therefore, the degree of realism needed may depend on the learning task and its context with considerable variations according to education outcomes (Ilgen et al., 2013). Interestingly, simulation studies that assessed learners' knowledge level were more effective compared to studies, which assessed their skill level suggesting that skill acquisition is more demanding and extended practice is needed. Furthermore, simulation studies use more researcher-developed instruments than standardized ones. Notably, such studies yield better learning outcomes raising concerns about the reliability and validity of such instruments (Merchant et al., 2014).

Protocols of good practice suggest that a session should consist of pre-briefing followed by simulation and ending with a debriefing discussion (Decker et al., 2013). Debriefing is a mean of formative

[6] Higher order thinking skills refers to *"critical, logical, reflective, metacognitive, and creative thinking … all intellectual tasks which call for more than information retrieval … [as well as] transfer and problem solving"* (Husamah, Fatmawati & Setyawan, 2018: 252-253).

assessment since it intends to shape future actions by lessons learned through reflective practices usually with the assistance of a facilitator. Findings indicate that debriefs improve individual and team performances by 20%-25% and such effects were similar across simulated and real settings in medical and non-medical organizations (Tannebaum & Cerasoli, 2012/2013). Moreover, there are various models of debriefing (Sawyer, Eppich, Brett-Fleegler, Grant & Cheng, 2016) yet, empirical evidence in support of a specific debriefing method is limited (Cheng et al., 2014; Raemer et al., 2011) although any debriefing method can be effective if used appropriately by well-trained facilitators (Sawyer et al., 2016).

Furthermore, there is no conclusive evidence on the effectiveness of key features during simulation debriefing (e.g. content, timing, use of video, educator presence, duration). Post event debriefing is the most commonly used key feature and research outcomes point out significant performance improvement in various contexts (Cheng et al., 2014; Raemer et al., 2011; Gegenfurtner et al., 2014) while within event debriefing seems beneficial for technical skills and mastery learning goals (Eppich, Hunt, Duval-Arnould, Sidall & Cheng, 2015; Van Heukelom, Begaz & Treat, 2010). On the other hand, the use of video during debriefing does not seem to have a significant effect on learners' outcomes (Cheng et al., 2014).

To conclude, evidence from systematic reviews and meta-analyses confirm that simulations lead to improved affective, cognitive and behavioural outcomes with debriefing being an essential quality component. However, the relative merits of different debriefing methods according to the context and topic of instruction still remain unknown. Finally, the effectiveness of simulation implementation is influenced by the quality of facilitators' feedback, design features and the level of learners' engagement, variables which need to be taken into account during assessment implementation.

4.5 Learning analytics

Learning analytics is a field of research that has developed over the last decade and continues to grow quickly. Though practical applications are beginning to emerge, the technology is still not widely used in educational settings. According to the Society of Learning Analytics Research, learning analytics involves the *"measurement, collection, analysis and reporting of data about learners and their contexts for purposes of understanding and optimizing learning and the environment in which it occurs"* (Banff, 2011 in Ferguson et al., 2016: 12). Learning analytics has its roots in many fields of educational and technical research, including assessment, personal learning and social learning, in business intelligence and data mining. It draws on theory and methodologies from disciplines such as statistics, artificial intelligence and computer science (Dawson, Gašević, Siemens & Joksimovic, 2014).

A review from 2016 gathered evidence of implementations of learning analytics in educational contexts focusing on the use and the processes of implementing learning analytics in any tier of education (Ferguson et al., 2016). Although the review was not exhaustive, it illustrates well the kind of practical applications of learning analytics that are already possible today. It concluded that although implementations across Europe are promising, the field is currently fragmented in terms of the use of learning analytics to improve -and innovate- education and there is a wide gap between the potential roles for learning analytics and what has been put into practice by ICT/learning technology vendors, developers and researchers.

As a general premise, learning analytic tools can aggregate data and generate information about learners' behavior and activities (e.g. learners' learning records, strategies applied, learning content accessed, questions-answers posed, learners' engagement with the online system). Based on such information, the system can provide **intelligent feedback** to both educators and learners (Li & Chen, 2013) which can be presented in various formats, including **information visualization** within

dashboards[7]. As a filed, learning analytics has mainly focused on areas such as performance prediction, detection of at-risk students, data visualization, intelligent feedback, course recommendation, learners' skill estimation and detection of their behavioral patterns, planning and scheduling, analysis of social networks and concept maps' development (Sin & Muthu, 2015).

Systematic reviews point out that most research focus on student prediction and technical aspects of data mining (Papamitsiou & Economides, 2014; Sin & Muthu, 2015). Moreover, even if the learning analytics tools *seemed* to be focusing on visualising learner engagement and activity providing early alerts, these data visualisations were not necessarily 'actionable' in the way that learning analytics should eventually lead to a targeted pedagogical intervention. In other words, they do not reveal what actions should be taken to improve learning and teaching. In addition, efforts focus less on innovative pedagogical processes and practices or on helping educational organisations to fully embrace the digital era (Ferguson et al., 2016).

Another issue with current tools is finding evidence for their formal validation (e.g. whether the tools fulfil their intended purpose, such as having a positive impact on learning, encouraging more efficient learning or more effective learning). The issue is partly related to the timeframe; very little hard evidence is currently available that is based on anything other than short-term studies. Some positive work in the field of learning analytics is cited in the 'LACE Evidence Hub'[8], which gathers research evidence on tools improving learning outcomes, for example, supporting teachers.

4.5.1 Evidence of the use and effectiveness

A number of learning analytics tools take advantage of innovative pedagogy and theoretical approaches to teaching and learning. Some of them have been described in two reports along with information about the context being applied (Ferguson et al., 2016; Steiner, Kickmeier-Rust & Türker, 2014). Some tools and impact studies are highlighted below:

Improving students' learning habits: CLARA is a tool that aims to make students aware of their learning dispositions (the habits of minds they bring to their learning). The survey tool platform generates a 'learning power' profile visualisation for each student and interventions based on these learning profiles. In addition, students receive coaching and mentoring from trained peers and staff. The tool is developed by the University of Technology, Sydney and a case study is provided in Ferguson and colleagues (2016: 121).

Helping students to reflect: Open Essayist is a tool that provides automated feedback to learners on draft essays in order to support their reflection and development. It presents a computer-based analysis of the most important sections and key words in a draft so that learners can compare those to what they intended to convey and adjust their writing in the light of that comparison (more information can be found in Ferguson et al., 2016: 64).

Yet, evidence on the impact of learning analytics on learning is still sparse. One of the few large-scale studies taking place at Purdue University, Indiana has shown that real time feedback had a positive effect on learners' grades and retention behavior (Arnold & Pistilli, 2012). Learners received feedback about their progress before the end of the course in their personal system. Their progress is represented with traffic lights (red indicate that they will fail, yellow most likely to fail and green indicates that progress is satisfactory). Notably, students who received a red or yellow signal were alerted and the majority of them had considerable improvement. Only 10.6% from the initial red group remained at the same level

[7] Dashboards are frames divided into sections, which can present concurrently different type of information.
[8] http://evidence.laceproject.eu/

(Arnold, 2010). In another small-scale study, adaptive selection of text questions tailored to the individual needs of learners, improved testing outcomes for below average students (Barla et al., 2010). The investigation of the effectiveness of digital assessment tools on learning outcomes is still very limited, further empirical work is needed in this area (Dawson et al., 2014).

5. Case studies of innovative assessment

5.1 Case studies

This chapter presents eight case studies that were selected with a view to capture a variety of innovative assessment practices from formal education and business currently existing in Europe (Table 1).

Section 5.1. provides an overview of each case study. Although, it does not comprehensively describe all aspects of the studies, yet it focuses on those parts most relevant to the objectives of the study. Section 5.2 makes cross-comparisons on a variety of dimensions, while section 5.3 provides a summary of the insights gained from the implementation of the eight innovative practices.

Table 1. Case studies examined in this report.

Country	Case study
10 EU countries	Assessment of Transversal Skills (ATS2020)
Sweden	e-Assessment of Prior Learning in Swedish elderly care (e-APL)
United Kingdom	Objective Structured Clinical Examination (OSCE)
Finland	Finnish Matriculation Examination (FME)
Hungary	Multipoly Next
France	Neo Alta
Slovakia	Teach for Slovakia (Teach4SK)
Spain	Assessment based on an Online Collaborative Project (AOCP)

5.1.1 Assessment of Transversal Skills (ATS2020)

The ATS2020 project developed a comprehensive learning model, which was piloted in lower secondary schools of 10 EU countries. Co-funded by the European Commission, the model is focused on supporting the development and assessment of transversal skills. Project partners include Ministries of Education, National Assessment Bodies and Teacher Education Institutions.

Fact box	
Education level(s)	Upper Primary Education; Lower Secondary Education
Target group(s)	Students
Specific subject(s) addressed	Transversal/soft skills
Transversal skills addressed	Digital competence
Assessment objectives	Assess transversal or soft skills; Improve student engagement & learning; Improve feedback to students
Assessment methods	Teachers' formative & summative assessment; Students' peer- & self-assessment
Assessment format	Collaboration; e-Portfolio assessment

Rationale and assessment objectives

Putting the assessment of transversal skills in the spotlight

ATS2020 is an innovative policy experimentation project co-funded by the European Commission. The innovation focus of the project lies in the provision of a tested, hands-on learning methodology based on an e-Portfolio process for the acquisition and assessment of students' transversal skills along with guidelines, assessment techniques and supporting scaffolding tools developed for teachers and students. Targeting upper primary and lower secondary education classes, ATS2020 links age-suitable transversal competences to the diverse national curricula of European countries on the basis of specific learning objectives.

ATS2020 provides a framework for the enhancement of students' indispensable transversal skills and new approaches and innovative tools to teachers for the development and assessment of these skills.

ATS2020
Assessment of Transversal Skills

Source: http://www.ats2020.eu/

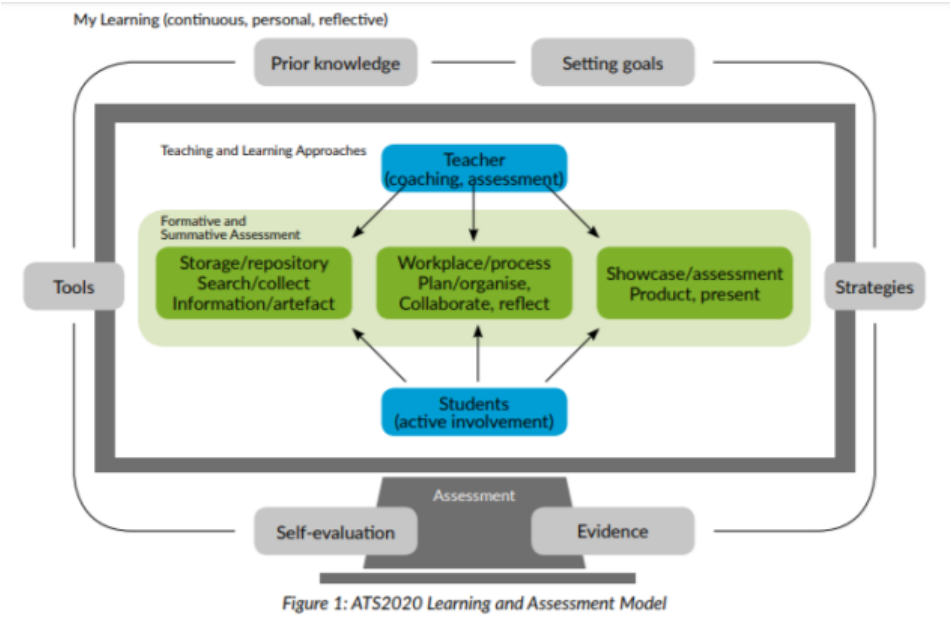

Figure 3. The ATS2020 learning and assessment model.[9]

The project is comprised of 17 partners from 11 EU countries. The ATS2020 project is a follow-up project of EUFOLIO which aimed to develop digital skills through an e-Portfolio approach, a learning path that students follow to develop transversal skills.

Initially, the project developed the ATS2020 Learning and Assessment Model (see Figure 3) building on existing learning models. The model aims to elaborate learning as a process and product making use of a web of activities leading to learning outcomes showcased in the form of e-Portfolios.

The ATS2020 model was piloted in 2016-2017 by means of a quantitative and qualitative evaluation methodology. The pilot was implemented in 10 of the partner countries and involved a total of 224 schools, 747 teachers and 11,891 students.

Key features

The project developed a comprehensive conceptual framework, the ATS2020 learning and assessment model. The model is founded on the premise that skills cannot be assessed at one point alone instead assessment should be part of a continuous process.

In a nutshell, the learning model includes the following key features:

- A learning design process for the development of transversal skills within the curriculum.
- Incorporation of digital technologies using common learning environments, engaging in digital learning activities and using digital resources and tools.
- A digital journal -the so-called “My Learning Journal”- maintained by students for their learning.
- E-Portfolios created and owned by students and function as a repository, workspace and showcase. In addition, they contain the "My Learning Journal" used by students to organise and create their own learning plan. The e-Portfolios are subject to assessment in each learning cycle.

[9] Source: http://www.ats2020.eu/what-is-ats2020

- Formative assessment tools are used to assess students' skills via teacher assessment, peer- and self-assessment. The scaffolding tools can be created by students and further developed by teachers.

The learning model is implemented through learning cycles, usually corresponding to the different subject units. Each teacher is expected to implement about 3-5 learning cycles in a school year, depending on the subject curriculum. Teachers are supported in designing and developing their own learning scenarios and supporting material or in adapting existing ones from the ATS2020 repository.

Design and Technology: The technical process / Transport	
I. GENERAL DESCRIPTION	
Country	Belgium
Author/Editor	Tom Van den Broeck
School	Het Laerhof
Subject	**DESIGN AND TECHNOLOGY**
LD Title	**The technical process / Transport**
Year group	Secondary Education – Second Year (13-14 years)
Duration	10 periods of 50 minutes
Short description	Students are introduced to the technical process. They run true successive steps of the technical process to realize a simple technical system (problem definition, design phase, implementation, initial starting and evaluation). They do research of the different components and subsystems in a technical system. They make choices involving technical systems and think about criteria such as environment, society, price, quality, safety and efficiency. They brainstorm about the problem solution, they search and collect modes of transport, research the evolution of modes of transport, transmissions, materials, switch leg. In addition, they present the evolution of modes of transport, research and discuss the impact of different modes of transport on people, planet and profit, search job vacancies and research of employment. The final product is building a race car which is sufficient enough to transport a marble.
Hosted (URL)	https://resources.ats2020.eu/resource-details/LEDE/technology-the-technical-process

Figure 4. Learning scenario/design example.[10]

In the learning scenarios/designs (see Figure 4) teachers connect the national educational curricula with the ATS2020 transversal skills formulated as learning objectives (see section *Transversal skills focus*). Supporting tools include a visualised learning design methodology, step-by-step guidance with templates defining the tools, activities, learning goals, required competences and final learning outputs at macro-, meso- and micro-level.

Transversal skills focus

The assessment of transversal skills is the key theme of the project. As a result, the project created the ATS2020 Competences and Skills framework (see Figure 5). The framework was developed after reviewing a series of competence frameworks, among others the DigComp (2013) framework developed by the Joint Research Centre (JRC).

The framework is composed of four core competences, together forming digital literacy: 1) Information literacy; 2) Collaboration and communication; 3) Creativity and innovation and 4) Autonomous learning. Each competence and skill includes a set of attainment goals.

The general approach taken to assess transversal skills is to provide students with a learning space, comprising a repository area, a working area where they create, collaborate, elaborate, reflect and assess, as well as a showcase area, where they share their learning achievements.

[10] Source: https://bit.ly/2k62ot7

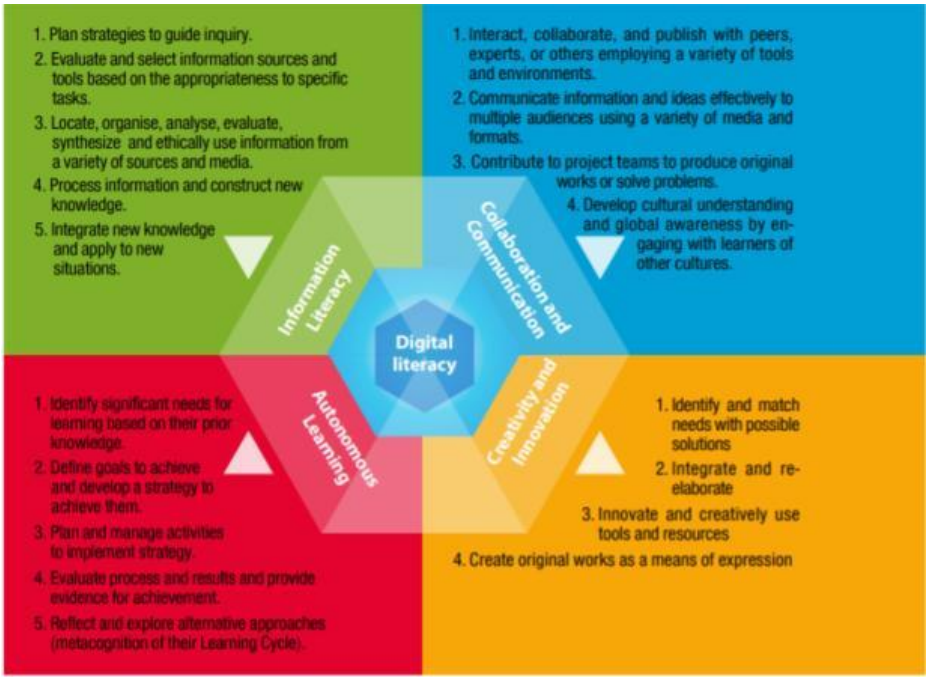

Figure 5. Areas of competences and skills addressed by ATS2020.[11]

Moreover, the "My Learning Journal" is designed for students' learning. This requires students to identify their existing knowledge on the subject, set their own goals, define their strategy on how to achieve these goals, describe the evidence they will gather to prove their learning and complete a self-evaluation to reflect on the learning cycle.

National trainings took the form of face-to-face and online workshops, as well as continuous online and school based support. The implementation in schools was considered as part of the training process. Participating teachers had a coach who guided them -online and through school visits- throughout the implementation of the learning model.

Implementation process

An in-depth, multi-phase implementation process

At project level, the implementation process started with training trainers from all 10 implementing countries (see Figure 6). For their training, trainers attended a two-day face-to-face workshop and participated in further online workshops and support. Trainers were also involved in the design of their national trainings, as well as the development of learning scenarios and educational resources. Teacher trainings were then continued at national level.

Once the teachers participating in the project were selected, information days were organised informing parents about the project. Teachers were required to obtain consensus forms of the parents agreeing to have their children participate in the project during one school year.

The implementation of the model in the classroom also promotes collaborative teaching, as two teachers with different profiles and expertise implement the model jointly in the same classroom. This usually incorporates teachers from different subject areas or disciplines, with one teacher having some expertise

[11] Source: https://bit.ly/2lGcsth

in technology. The collaborative teaching model was widely welcomed by participating teachers and students alike.

In each learning cycle, students conduct a self-assessment and are assessed by their peers and the teacher (summative assessment). Due to time constraints, three learning cycles were implemented. The e-portfolios contained the final product of each learning cycle and altogether serve as the final assessment product.

In a last step, a final conference (see Figure 7) allowed schools to share the portfolios created by each one of them as well as to exchange experiences, while a presentation of the experimentation results was prepared by each country.

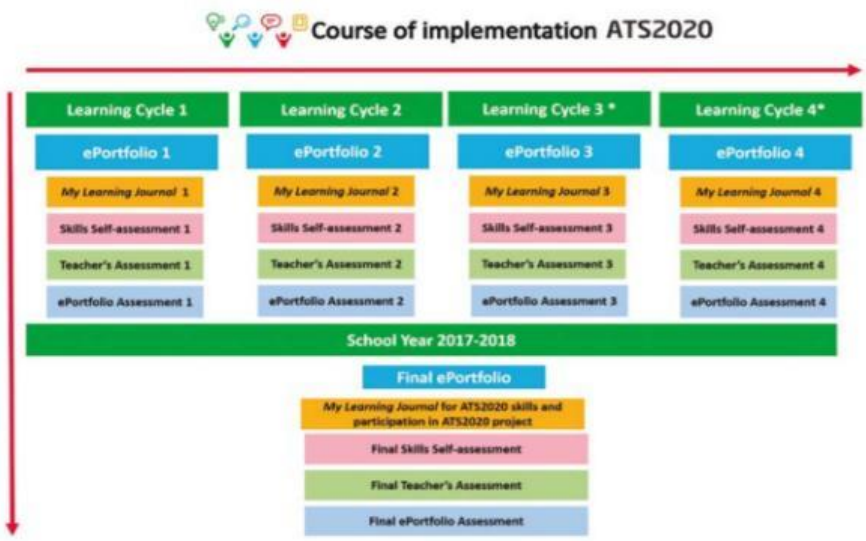

Figure 6. Course of implementation of ATS2020.[12]

Challenges faced

Time to embrace the learning model

The case studies provided country-focused insights into the challenges faced by the project's implementation. In Cyprus, issues existed in relation to the digital infrastructure, such as limited bandwidth; however, these issues are known to exist in most other European countries.

A further challenge -applicable to all countries- is related to the time factor. In theory, the time spent during the learning cycles was estimated to be equivalent to the time teachers would devote typically in their learning unit. In the curriculum of most countries this amounted to a weekly average of six hours. Although attempts were made to comply with these time frames, some activities, required more time. The time factor proved to be particularly challenging for teachers, as they needed to prepare the activities for the learning scenarios.

Moreover, teachers reported that the overall implementation schedule of the planned learning cycle was equally challenging since it started in November. Meanwhile, the students needed more time in order to understand the model and related concepts; they needed to acquaint themselves with the concept of

[12] Source: https://bit.ly/2lGcsth

transversal skills and competences, the methodology and the concept of assessment as a whole. Nevertheless, students adapted somewhat better to the new situation and managed to devote extra time to their learning journals.

Figure 7. The ATS2020 final conference.[13]

Furthermore, a barrier emerged due to differences in terms of curricula. While certain country systems were more suitable to adapt to the student-centred approach of the model, other education systems faced more difficulties. As a result, their transition to the ATS2020 model turned out to be more time consuming.

As the learning model involves a number of new and complex ideas and pedagogical practices, some teachers had difficulty to understand and adapt their teaching practice.

Success factors & lessons learned

Driven by a sound methodology and teacher collaboration

One of the main success factors was linked to teachers' enthusiasm for the project's goals and methodology. This enthusiasm was driven by a consensus that transversal skills should be fostered in schools and the positive outcomes of the project motivated them further. Through the project, teachers were provided with a comprehensive methodology including the tools and support to develop and assess such skills.

Moreover, the embedding of digital technologies represented an additional motivating factor, for students and teachers alike. Although the integration of technology in classrooms is mostly a reality, teachers greatly welcomed the support provided by tech-savvy coaches and peer-educators.

An additional success factor lied in the collaboration among teachers. Teachers from the same school gathered and engaged in co-teaching and the co-design of the learning scenarios. What is more, teachers across schools engaged in sharing their experiences at national and European level. Altogether, teacher collaboration resulted in a fruitful exchange and many lessons were learned.

Finally, giving students a voice in sharing their learning experiences turned out to be a success. They had the opportunity to present their e-portfolios and with their pride and enthusiasm, they became ambassadors of the project.

[13] Source: http://www.ats2020.eu/content/446-ats2020-final

Lessons learned

Although the scaffolding tools have been well-received by teachers, the coordinators of ATS2020 seek to enhance these assessment tools alongside additional standardised - summative assessment tools. Both can be used as complementary at various points of the learning cycle and in this way, the summative assessment tools of the transversal skills are incorporated into the continuous process of formative assessment. The implementation of both is expected to support the student and the teacher to redesign learning more efficiently.

Achieved results and impacts

Noticeable improvement of skills in Cyprus

The large-scale experimentation of the project was designed to be evaluated both quantitatively and qualitatively. Thus, a pre and post evaluation took place to assess the impact of the project on student performance. The assessment used experimental and control groups and conducted two case studies in each country in order to support and analyse the results further. The quantitative data analysis did not point out a significant development of transversal skills. Such an outcome may be due to the short period of implementation (three months) which is considered not enough for substantial changes to appear.

Yet, the qualitative data analysis revealed improvement of student's transversal skills through the implementation of the ATS2020 learning model. According to the data collected from the case studies, students on average showed a higher level in collaboration, communication and information literacy skills. Moreover, students' and teachers' artefacts showed considerable progress from the first learning cycles towards the latter ones.

The possibility that some countries already had a consolidated level of transversal skills equally needs to be taken into account. In the final stage of the project, these differences are targeted to be further explored.

Potential for scale-up and replication

Further replicability potential

Although the implementation of the learning model dates back to the school year 2016-2017, teachers in Cyprus and some other countries continue to apply the model. With a view to encouraging further replication in schools, a toolkit was developed -a step-by-step implementation guide- supporting schools and teachers. Moreover, some countries contacted the project leaders of ATS2020 requesting further information on the project and educational authorities consider its implementation at national level.

5.1.2 e-Assessment of Prior Learning in Swedish elderly care (e-APL)

As a joint initiative of academia and local government, a pilot project initiated to assess the knowledge, skills and abilities of staff with no-formal education in Swedish elderly care. Using an interactive assessment method, the practical and theoretical skills utilised in elderly care were evaluated. In line with the results, tailored e-training programmes were developed to address skill gaps.

Fact box	
Education level(s)	Employment/VET
Target group(s)	Elderly care workers
Specific subject(s) addressed	Healthcare; Elderly Care
Transversal skills addressed	Communication skills; Planning and organisation skills; Ergonomics
Assessment objectives	Certification of skills/recognition of prior learning
Assessment methods	e-assessment; Teachers' summative assessment
Assessment format	Quiz/multiple choice; Authentic learning/real life tasks

Rationale and assessment objectives

A joint initiative of local authorities and research

The increasingly ageing populations of Sweden and other EU countries, create a growing demand for elderly care staff, in terms of quantity but also in terms of qualification level. Being often self-taught or trained informally, elderly care workers tend to have relatively low levels of education or no-formal education at all. At the start of the project, the national share of all elderly care professionals without formal education amounted to 21%.

Against that background, the e-assessment of prior learning project was implemented as a pilot study in 2009-2011. The goal was to create a quality assured and time-efficient national model for validation and competence development of personnel working in elderly care without formal education.

Thanks to the assessments, knowledge and skills areas in need of further development and training were identified. In line with the results, tailored e-training plans were drawn up to address these areas. The model was recognised by formal vocational education and provided an opportunity for caregivers to obtain a formal certificate demonstrating their competence (see Figure 8).

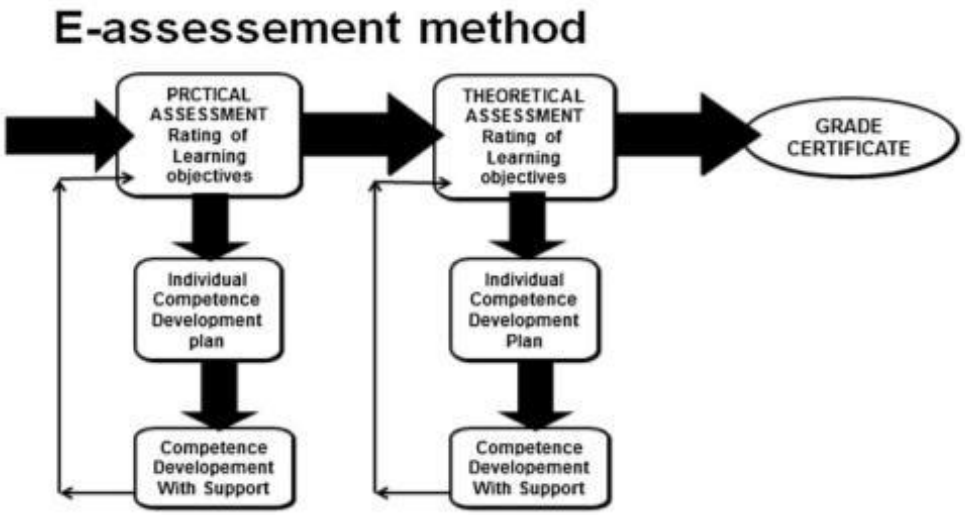

Figure 8. E-assessment method.[14]

Surprisingly, none of the 87 participating elderly caregivers passed all the assessment parts. According to the creators of the assessment model, the low results of the assessments may be due to the difference between vocational education standards and practical on-the-job training.

Following a change in government, the validation of prior learning of workers with low levels of qualification was temporarily dropped off the lists of education policy priorities. However, in the course of a recent reawakened interest in the validation of prior learning the creators of the model are working on an adult education project which uses elements of the e-assessment model. Beyond this project, there are intentions to re-establish the model adapting it to the changing polictical environment of validation of prior learning.

Key features

The pilot project applied an interactive assessment method in combination with e-assessment. The criteria applied in the assessment complied with the national occupational standards, as applied in vocational education.

The e-assessment was divided into a practical and theoretical assessment, each having specific learning goals. The two assessment parts were linked through a computer programme with a personal log-in for participating caregivers. The programme offered access and information about both assessments.

The practical part took place in a specifically designed apartment equipped with video cameras and ICT technology simulating everyday situations in elderly care, yet on the basis of an adult-sized doll and an actor posing as an elder care recipient. The practical assessment was divided into morning, lunch and evening parts, each lasting 40 minutes.

Theoretical assessments were performed through a tailored, computer-based test at the workplace. The test comprised two levels with each consisting of 8 learning objectives and was validated by formal education trainers and experienced staff. Before carrying out the assessments, participants received a 2-hour, face-to-face information session conducted in groups.

[14] Source: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3998952/

The session served to inform participants about the method and the learning objective criteria for the assessments. During the session, participants also received an introduction about the computer programme and scheduled their practical assessments. While completing the practical examination, participants were supported by the test manager in any parts of the assessment needed. On average, participants needed approximately 8 hours to complete the entire assessment process. Immediately after the theoretical assessment, participants could see their results on the computer screen.

Following the practical assessment, some general explanations and verbal feedback were given by the test leader so that staff members do not experience the assessment as something unpleasant. In line with the knowledge and skill gaps detected through the assessment, a tailored e-training programme consisting of various modules was developed for each caregiver.

The experience of the involved staff was gathered and evaluated. The underlying objective was to draw conclusions concerning the impact of the training on the degree of well-being of the caregivers through interviews.

1. Basic healthcare*	Check "Asta's" general condition
	Give her cream and sandwich.
	Put her in bed.
	High under head.
	Make sure she has something to drink that is easy to reach.
	Report to the nurse.
	Make sure she has the alarm on her left wrist.
1. Basic healthcare**	Meet her needs and satisfy her desire to eat supper in bed.
	On the report suggest supplementary nurse supervision at night.
1. Basic healthcare***	At all times secure her safety.
	Inform her that she will get extra supervision during the night.
	Make sure she knows how the alarm works and how to use it.
4. Communication*	Presentation. Coming from home care to help her through night-time preparations.
	Talk to "Asta" about what help she needs and how the help will be performed before it all starts.
	Through conversation encourage her to move with the walking table.
5. Rehabilitation*	Take advantage of her own resources during the movement.
	Encourage her to participate in the transfer movements based on her own ability.
	Ask her if she wants anything else before you leave.
6. Ergonomics, hygiene, esthetic environmental*	Preparation for the transfer movement. Current aids are available.
	Go through the transfer movement with "Asta".
	Perform the movement in an ergonomically correct manner.
8. Assistive technology*	Engage the break on walking table before transfer.
	Perform the transfer in an adequate way.
	Adjust the height of the bed.
	Check hearing aid status.
8. Assistive technology**	Implement the transfer in interaction with her.
	Meet her needs for safety at the moment of moving.

Numbering is related to the eight learning objectives.
Grades: fail; *= pass; **= pass with distinction; ***pass with great distinction.

Figure 9. Learning goals from the evening assessment.[15]

[15] Source: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3998952/

Transversal skills focus

Even though some of the skills and abilities utilised in the elderly care sector belong to the field of transversal skills, the project is not focused on the assessment of these skills specifically. The skills assessed in the project are (see Figure 9):

- Communication skills;
- Planning and organisation skills;
- Ergonomics;
- Ethics.

Implementation process

Piloting the interactive assessment model

The initial development process of the assessment model dates back to 2005 when official validation funding was introduced by the government in 2003. It was against that background that the project manager and her working group created the interactive assessment method together with Prof. Ingemar Wedman from the University of Gävle.

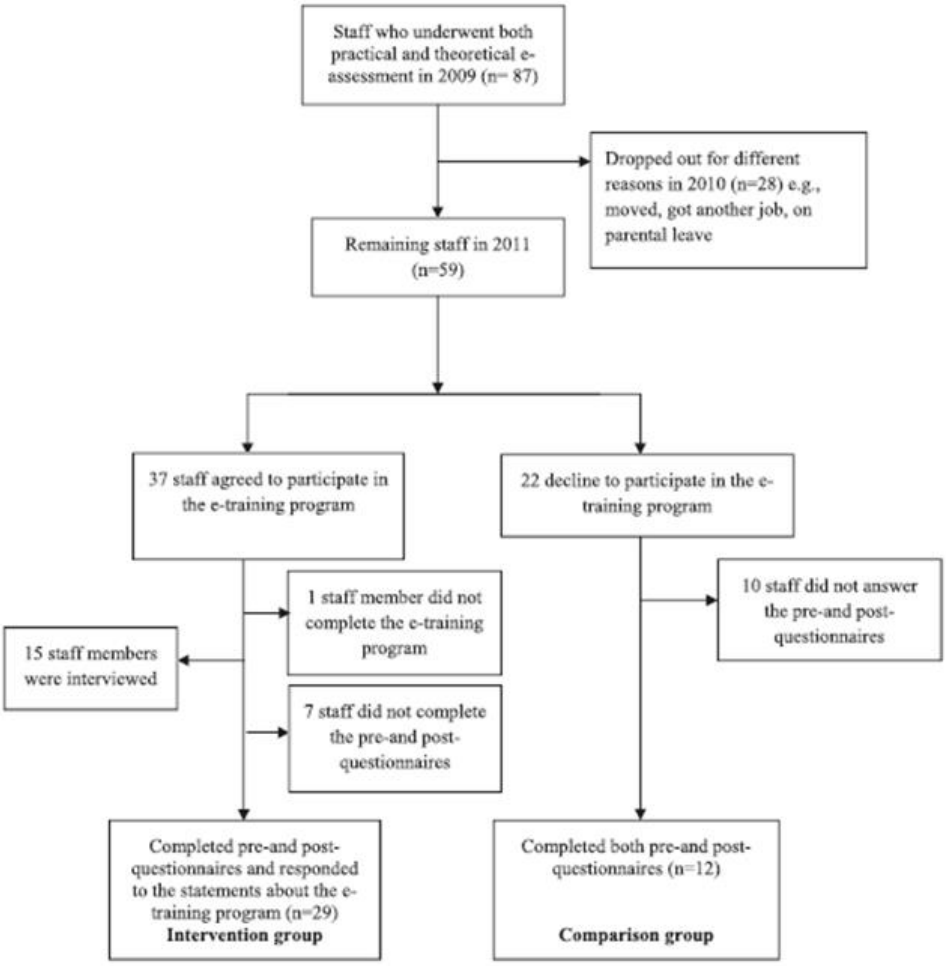

Figure 10. Data gathering process.[16]

[16] Source: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4438631/

The method was designed to reflect everyday practice in elderly care. It was developed on the basis of a user-centred approach through several workshops. The feedback of formally trained staff and qualified teachers was collected (e.g. drafting the questions of the theoretical assessment). The workshops helped ensure the authenticity of the practical assessments as well as the clarity of instructions received during the theoretical assessments. Once developed, the method was applied in two previous projects: the Off-e project (2005-2007) and Cluster-E senior living industry project (2007-2009). Throughout these projects the method has improved continuously (see Figure 10).

Both adult education centres and local government showed an interest in piloting the assessment of caregivers without formal education in Sandviken. The participants for the pilot were recruited from the pool of elderly caregivers employed in Sandviken.

The caregivers were contacted by their employers through personal meetings or e-mail on the basis of an elderly care inventory completed in 2009. While participation in the assessment was voluntary, caregivers were reassured that their results would not have any impact on their employment. On the contrary, the assessments allowed participants to obtain a formal recognition of their competences and participate in e-training to overcome potential difficulties.

Challenges faced

One step ahead of its time

A key challenge was due to the timing of the project's implementation. When the data of the project was collected in 2011, competence development and e-learning were not very widespread. As a result, the project faced some skepticism at individual and systemic level.

On the one hand, some participants, professional education and training institutions and elderly care providers were hesitant to see the benefits of the project. Additionally, the professional education and training system lacked the flexibility to adapt to the project's assessment-based approach.

On the other hand, insufficient linguistic and IT skills proved to be a minor challenge. While some caregivers had difficulties understanding the Swedish instructions, others felt insecure about their computer skills to take the e-assessments. Besides nervousness, some employees feared that the results of the assessments would affect their employment (e.g. a bad performance may have a negative effect) which was obviously not the case.

Success factors & lessons learned

Driven by beneficial prospects from all sides

Despite the project's challenges, a number of factors helped its implementation. At first, the impetus given by the Municipality of Sandviken was decisive for the implementation of the project. From the start, the Municipality showed great interest in the assessment of prior learning in local elderly care. Hence, the municipality jointly designed the pilot project and supported it financially.

Beyond the support of the local authorities, the assessments and tailored competence development courses provided various potential benefits for the elderly care sector. Next to the prospect of improved competence levels and more professional elderly care, it also included safer employment conditions and a reinforced role of vocational education to ensure better professional quality.

Table 2. Staff members' ratings of the assessments.[17]

Statements about the assessment	Totally disagree	Partly disagree	Neither agree nor disagree	Partly agree	Totally agree
1. It was good to do the practical assessment in the simulated apartment.	3	2	1	7	17
2. It was good to do the theoretical assessment using the IT *tool.				5	25
3. I would recommend this way of testing one's knowledge to my colleagues.				4	26
4. If possible, I would like to continue testing my knowledge, skills and abilities in this way.		2	3	5	20

*IT = Information technology; One person did not respond to statements 1, 2, 3, 4.

Finally, professionals were able to see benefits, as they were able to utilise, improve and test their real, practical skills acquired by different means. For some, this opportunity was a motivating factor along with the prospect of further training and higher wages. Furthermore, compared to conventional assessment methods participants acquired increased qualification levels in a short time frame.

Learning from applying a user-centred approach

Looking back at the process of its implementation, one of the lessons learned lied in the user-centered, participative approach. Elderly care professionals piloted the assessments and provided feedback. Thereby, they were given the opportunity to shape the design of the interactive assessments. As a result, their motivation has increased and they felt that their opinion was valued. Secondly, effective cooperation and high trust levels between vocational education institutions and elderly care providers proved to be vital to assessment implementation.

Furthermore, the designers of the interactive assessment point out the importance of creating a common target image, an open and transparent approach involving all concerned stakeholders in the shaping of the implementation process.

Applying a more holistic view of learning outcomes

While at the time of the model's development, learning outcomes were validated at more in-depth levels, in recent years there is a shift towards transversal skills with less emphasis on technical details. Therefore, the initiators of the model would have applied a more holistic view of learning outcomes at both theoretical and practical level.

Achieved results and impacts

High approval rates of involved staff members

The piloting of the interactive assessment model provided a series of quantitative and qualitative results. In quantitative terms, 87 staff members participated, with 63 completing the entire assessment. The practical assessment was passed by almost all participants with the exception of the morning hygiene part where several failed. Surprisingly, none of the staff members passed all components of the assessment. In addition, more than 50% of participants failed specific elements of the theoretical assessment.

[17] Source: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3998952/

The creators of the assessment model trace the rather low results of the validation back to the difference of vocational education standards and practical on-the-job training. The results show that when on the-job-training is not complemented by the theoretical basis of formal education training, the applied knowledge or skills regularly deviate from those described in the qualification framework.

© shutterstock / 577135267

In a similar way, the impact of the professional experience in terms of years working in elderly care and staff grades on the acquisition of the learning objectives proved to be insignificant.

Regardless of the low average performance, participating staff members evaluated the assessments positively. As shown in Table 2, 80% of the involved caregivers either fully or partially approved the practical assessments in the simulated apartment. Concerning the use of the IT tool for the theoretical assessments, all 30 surveyed staff members valued this approach (80% fully and 20% partially). In total, 84% of participants stated to highly recommend the assessments to their colleagues, while 83% were willing to continue the assessments, if possible.

Furthermore, the assessment model allowed for significant time savings. On average, staff members completed the assessments in 8 hours compared with 42-60 hours commonly used for traditional validation. While the duration of the validation of the theoretical part in school ranged between 10-20 hours, the practical assessment at the workplace typically took between 32 and 40 hours. In addition, the flexibility of the tool was valued by several participants as they had the possibility of taking the assessments during working hours and at their own workplace.

With regard to the qualitative results, the creators of the model admit that initially the model was not fully embraced by the vocational education system. Along with decreasing political interest in the validation of prior learning, the circumstances to further develop the model were less favourable at the time. As a result of its temporary standstill, empirical evidence is lacking in support of the e-training for staff members who completed it.

Potential for scale-up and replication

Further replicability potential

The interactive assessment model has a great potential to be replicated in other professional domains (e.g. the construction sector).

Given Sweden's increased reception of refugees with low qualifications, the discussion around the validation of prior learning has been reignited in recent years. In order to further develop the e-assessment model in line with current national requirements for validation and the Swedish Qualification Framework (SeQF), the County Council of Gävleborg is using practical validation for an adult education project on workers with disabilities.

Furthermore, the creators intend to utilise a revised version of the model in the near future. One of the main objectives is to find a more time-saving solution through on-the-job training. Given recent political developments at EU and national level for establishing a national strategy on prior learning, the initial model would undergo substantial changes. The content will be adjusted and updated in line with current national guidelines such as the SeQF. In addition, a closer cooperation with employers, and the education system on a whole needs to be ensured. Finally, the County Council of Gävleborg is using the gained experience from the project for a new adult education project, called ValidX, targeting personnel working with disable people.

5.1.3 Objective Structured Clinical Examination (OSCE)

A New Efficient Approach to Evidencing Clinical Skill Acquisition

Sheffield Hallam University found an effective way to transition away from traditional clinical assessment for its undergraduate paramedic trainees. The approach has surpassed all the expectations and provided a sustainable model for the future.

Fact box	
Education level(s)	Tertiary education
Target group(s)	Students participating in the OSCE examination of Paramedic training
Specific subject(s) addressed	Cross-subject; Behaviour; Attitude; Clinical skills; Soft or transversal skills
Transversal skills addressed	Critical thinking; Analysis; Reflection; Problem solving; Communication skills; Civic and social competences
Assessment objectives	Certification of skills; Recognition of prior learning; improve authenticity of learning
Assessment methods	Examiner's summative assessment; Students' peer- & self-assessment
Assessment format	Simulation; Authentic learning/real life tasks

Rationale and assessment objectives

Driven by the need to accommodate students' needs

Before introducing the innovative assessment in 2016, paramedic science students at Hallam Sheffield University had their practical skills assessed via a practical examination called Objective Structured Clinical Examination (OSCE). Students were demonstrating medical skills (e.g. managing a trauma patient) in front of the assessor, who would then mark them. At the time, 40 to 60 paramedics were in the course and went through 5 or 7 different assessments. The programme required roughly 9 members of staff (in 4 rounds) to complete the assessments within 4-5 days.

The majority of students found these practical exams very stressful: Performing a certain skill in front of somebody put additional pressure on them and made the examination more difficult. Even though the teachers knew that examinees were able to demonstrate the required skill after seeing it performed in class, trivial mistakes were often made once the live demonstration started.

120-130 students

... are currently using the new assessment approach in place

Source: Interview with Andrew Kirke, Sheffield Hallam University

There was a need for practitioners to come up with a solution to make OSCE more student-friendly. Additionally, in recent years class size has increased considerably. More staff, space, examination days and equipment were needed to allow OSCE examinations to continue. The cost of exams was also increasing due to the need to bring in external assessors to manage the workload.

The programme coordinators agreed that the assessment procedure had to be examined and revised. As a result, the programme leader came up with the idea to get students to record a video of them performing the skill, which would then be watched by the assessors.

Key features

Videos assessed not only by the examiners

The new assessment procedure foresees that students submit two videos of themselves performing a certain skill. The first video is self- and peer-assessed. Here, minor mistakes are allowed as long as the other video subject to examination is correct, which shows improvement, learning and development (see Figure 11). After some time, students are requested to provide another video performing a certain paramedic skill that is assessed by the examiners. The submission is followed by a 3-week marking period with the workload distributed among current staff members.

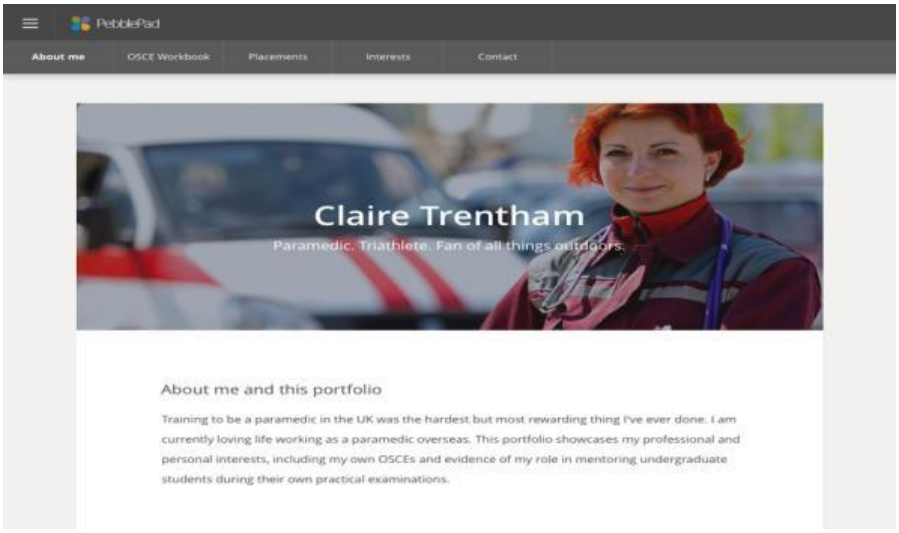

Figure 11. Assessment video example.[18]

After the initial attempt to use recorded videos instead of the formerly face-to-face OSCE examination, there was a need to set up some ground rules involving the following:

- All videos should start with a face shot in which the student states the name, the date and the OSCE examination that they were performing.
- Students performing the skill should be visible.
- Videos should not be edited.
- No prompts, writing instructions or scripts should be used in the submitted videos.
- Uniform must be worn if issued.

[18] Source: https://bit.ly/2kytpFR

- Videos should be performed in a professional manner.

Furthermore, the lecturers have provided some exemplary videos of the skills being performed that have been tagged with specific QR codes. Hence, students could easily access them and observe the correct execution of the skill, while one or two tutors are still available to assist students whilst they practice the skills.

Transversal skills focus

Addressing a diverse pool of transversal skills

The new assessment process focuses on a diverse set of transversal skills such as reflection, self-awareness and being honest about one's individual performance.

The assessment of other students' performance through peer support, peer coaching and feedback fosters social competences such as compassion and encouragement.

The research outcomes indicate that students were spending significantly more time practising for the required video with less tutor support. Most significantly, the skills were mastered and students could replicate them successfully in their respective careers.

Moreover, practitioners were able to see significant differences on time spent for learning for the practical skills examination in the two assessment types. This was demonstrated through two groups of students, who underwent a practical skills assessment in a traditional way (first year of study) and via video (second year of study).

Last but not least, the submission of the recorded videos helped students develop skills such as transferability and even entrepreneurship as medical blogging is expanding and has a positive impact on difficult to reach communities. Therefore, students can use their video as teaching resources for other citizens.

> *So I asked the fundamental question - why do we need to see the student perform the skill in front of us in that one off situation to be able to say yes, they can do it.*

Source: Interview with Andrew Kirke, Sheffield Hallam University

Implementation process

Initial technical issues have been overcome

At the onset, the first idea was to use blackboard, however the university had purchased an online platform called PebblePad that seemed to better accommodate the needs of the programme. Alongside students having an online workbook put in place, the platform allowed them to upload videos automatically after recording.

At the beginning, programme leaders requested students to use their own devices to record the videos; however, that created issues of file storage space, as students had to delete other files from their personal phones. Therefore, the faculty invested in recording devices and made them available for use at the University. However, students are still free to record with and upload videos from their own devices.

The faculty developed a number of student led teaching spaces where groups of about 6-8 students can book on an hourly basis.

Challenges faced

Minor concerns from the university and a new challenge to the faculty

Generally, the university was very supportive of the new approach. Initially, some concerns touched upon questions such as *"if a student has not performed in front of the lecturer, how do they know the student can actually do it?"*.

The programme coordinators argued that the final result was all that mattered: Even if the student recorded the video 10 times before getting it right, what mattered was that they performed it correctly and produced the evidence of the successful delivery. They argued that performing a skill in front of the examiner, does not give additional credibility to assessment.

Initially, one or two staff members were technologically less literate and were hesitant about using this approach, but with a little bit of training and support, they became very comfortable. Overall, the shift from the old approach was relatively smooth, with minor issues, which -according to the faculty- were overcome very quickly. Additionally, the Health Care Professional's Council was very impressed with the new assessment procedure when they validated the new 3-year Bachelor Degree. In fact, other health professionals and peers requested more information and consider its implementation within their own University.

85%

... is the first time pass target for the course set by the university which is now frequently reached, whereas before the numbers were often under 85%.

Source: Interview with Andrew Kirke, Sheffield Hallam University

Success factors & lessons learned

Exceeding all expectations

The key success factor of the new assessment of paramedic clinical skills lies in the number of benefits the approach has been able to provide. First, it allows genuine internal and external moderation. When lecturers finish the initial marking, a second person will look at all the videos of students who failed and at a percentage of those passed to reach consensus with the marker's comments. Before introducing the new method, it was not feasible to do so as only one examiner sat in the exam, while now an external marker has access to the videos and can refer to any of them as evidence to back up the decision made.

Moreover, a better clinical governance structure with a more robust evidence system has been achieved. If a graduate in, for example, 5 years' time is requested to prove that they have the ability to perform a certain skill (i.e. chest decompression), the University has the video to prove the competence.

Additionally, lecturers do not have to be at the University watching at students' performance. Assessors can now be anywhere they want and have the possibility to log in for marking as long as they have an internet connection. As a result, the marking is now easily done within the standard 3-week timeframe.

Moreover, students save commuting time and resources. Instead of traveling to University, they can log in and access training facilities without actually being physically present at the University.

Achieved results and impacts

Major impact on the wellbeing of the students

Following the OSCE submission, a survey was distributed to the students and 16 completed ones were received (80% return rate). Student feedback showed that most trainees found the process technically "easy" or "very easy". While significantly more time was spent on practising skills prior to submission, students found the experience significantly less stressful than the previous examination approach. In fact, 90% of the students preferred the new process over the previous one and furthermore, appeals against assessment decision were reduced. In case of an appeal both lecturers and a student can simply observe the video and see what went wrong.

Additionally, no external staff for the examination was needed and the faculty managed to save a significant amount of money that were invested elsewhere in the programme. Most significantly, students' success rate has improved. More people are passing first time whereas before the percentage was much lower.

In the end, the faculty highlights many benefits resulting from the new procedure: Next to the entire process being more relaxed, students have proven to be more satisfied. Furthermore, they enjoy the better working conditions at academic and faculty levels, achieve higher pass marks and have filed fewer appeals against University.

Potential for scale-up and replication

Possibilities for both scalability and transferability

Thanks to this assessment process, the faculty is now able to deal with greater amounts of student assessments without increasing staff numbers or additional external help or other negative impacts on the faculty.

With regards to transferability, the faculty has shared this new innovative assessment practice with other Universities and some are already implementing it (see Figure 12).

Figure 12. Learning technologies award ceremony.[19]

[19] Source: https://bit.ly/2kytpFR

However, the system is also used in completely different domains. For example, one of the Universities in Australia is using the process for their pilots training. The University has invested into go pro cameras in the flight simulation cockpits. This allows trainee pilots to upload their test videos onto their workbook for assessment after performing their circuits. The main benefit is that assessors can mark the students without actually sitting next to them in the cockpit. As Australia has a widespread territory, such a system has greatly improved students' experience.

5.1.4 Finnish Matriculation Examination (FME)

The Finnish Matriculation Examination, the only nation-wide high-stakes standardised exam taken by students at the end of upper secondary education, is going through a process of digitalisation. In 2019, students will be using their personal laptops to take their matriculation examination ending the gradual update process started in 2016.

Fact box	
Education level(s)	General upper secondary schools
Target group(s)	Teachers, examination board, and students
Specific subject(s) addressed	German language; Geography; Philosophy; French; Social studies; Psychology; Second national language (Swedish, Finnish); Religion; Ethics; Health Education; History
Transversal skills addressed	Thinking skills and learning to learn; Multi-literacies; ICT competence; Competence for the world of work, entrepreneurship
Assessment objectives	Modernise assessment
Assessment methods	Student's summative assessment
Assessment format	Digital assessment; Quiz/multiple choice; Portfolio assessment

Rationale and assessment objectives

Driven by the need to guarantee versatile skills for students to use ICT for learning and living

The Finnish Matriculation Examination (ME) is a standardised high-stakes exam where, over the course of 3 weeks, students take exams in their mother tongue and subjects of their choice. It is taken at the end of the upper secondary education for the purpose to discover whether students have assimilated the knowledge and skills, and reached an adequate level of maturity, in line with the goals of their education level. About half of the age group takes the examination, entitling the candidate to continue to tertiary studies.

4,000 school staff members

... were visited during the first year of the project in order to ensure its better uptake

Source: Interview with representatives from the Examination Board

In 2010, a ministerial working group proposed that ICT should be introduced gradually in the ME. Changing the way the ME was conducted it was seen as a catalyst for change in pedagogy and the general use of ICT for learning. In 2013, the timetable for the process of upgrading the ME was fixed with a step-by-step introduction of digital exams in various subject areas (see Figure 13).

The examination sessions are organised twice a year, in spring and in autumn, over a period of 3 weeks on 9 days. The Matriculation Examination consists of a minimum of four exams that can be split on three consecutive examination periods. Examination of the candidate's mother tongue is compulsory; in Finland, this means either Finnish, Swedish or one of three Sami languages. The other examination options are among the following: the second national language, foreign languages, mathematics, subjects in humanities and natural sciences. The maximum duration of each exam is 6 hours.

The learning objectives assessed by the Matriculation Examination are set by the National Core Curriculum for General Upper Secondary Schools. The importance ICT skills was already stressed in the core curriculum of 2003, whereas the recent one from 2015 defines more specific learning objectives in areas such as: 1) Thinking skills and learning to learn; 2) Multi-literacies; 3) ICT competence; 4) Competence for the world of work, entrepreneurship. Taking the ME in a digital form further supports the development and acquisition of these competences.

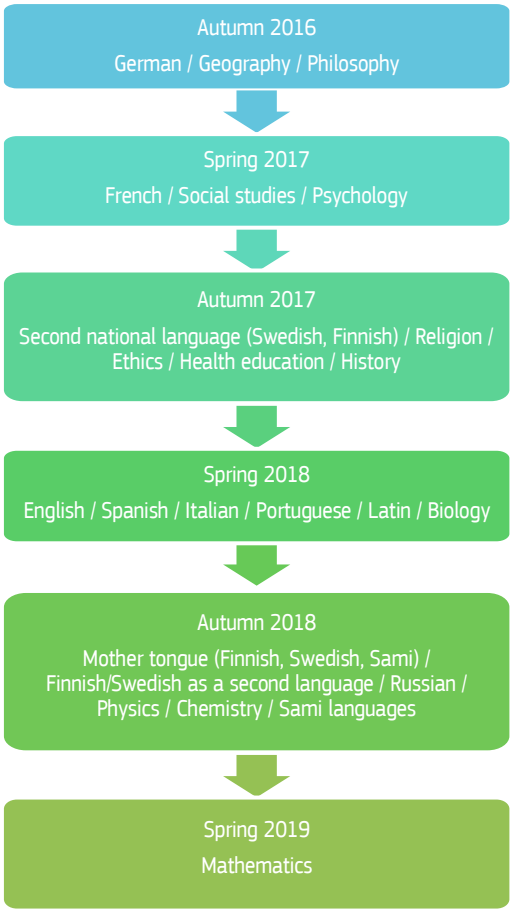

Figure 13. Digitalisation timeline.[20]

Key features

From a paper format to digital with more variety for assessment

The key features of the digital ME are based on the affordances that digital technologies offer; on the one hand, there is the development of the digital examination environment, and on the other hand, the fact that a vast amount of new assessment options are made possible by adding more variety to the examination questions (e.g. images and video, raw digital data, authentic footage in terms of news, voice recordings).

The digital examination environment is a tailor-made Linux operating system that can be started from a USB drive and includes a set of programs from generic utility suits for word processing and number crunching to image editing (LibreOffice, GIMP), but also specific ones for vector graphics (Dia, Pinta InkSpace) and symbolic computation (GeoGebra, wxMaxima). Some commercial programmes are also made available (Casio ClassPad Manager, MarvinSketch, LoggerPro, Texas Instruments TI-Nspire CAS, MAOL digital tables). Students can use these programmes to answer questions and complete tasks requiring data processing, and for instance, calculation, editing or drawing graphic presentations. The idea is to offer students an authentic variety of applications that they may encounter.

[20] Britschgi, V. (2015). *The Finnish Matriculation Examination*.

To run the ME on the given dates, the Matriculation Examination Board (MEB) prepares and delivers the USB drives with the operating system to schools where a local network has been set up in order for the ME to take place (see Figure 14).

At the beginning of the examination, students start their own laptops using the operating system from the USB drive. Candidates' computers are connected to the examination system via a local network and they get the examination material through a browser.

After the examination, answers are sent to the MEB's web service where they are marked and scored first by the teachers of students' own school and then by the assessors of the MEB. At the end, the candidates get their results through the system and receive the diplomas from their schools.

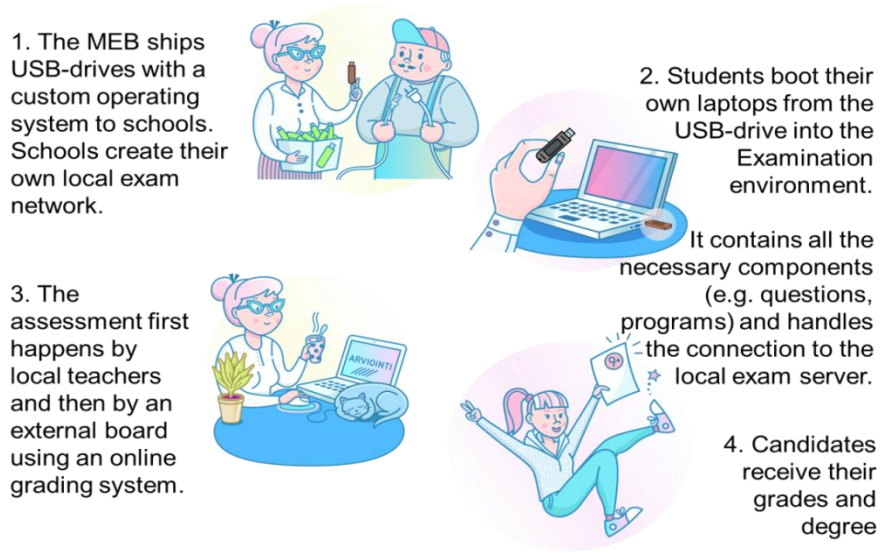

Figure 14. Process diagram of the MEB examination.[21]

Transversal skills focus

Assessing student´s readiness for the future of work, society and leisure

The digital examination environment allows candidates to show their expertise at the end of upper secondary education in a versatile way. As the examination questions can include video, audio, data tables, maps and other, this opens new possibilities for the assessment of transversal and subject specific skills that are not possible with the traditional paper format. Open ended questions are used for writing and science assignments.

Figure 15 presents an example of a type of an examination item that could appear in Geography. It focuses on the topic of air pollution. Question 9, with 5 sub-questions, includes various data sources that students have to use to answer. The screen capture shows some of the data sources, for example, number 9.1 provides a data file in ODS format, from Beijing in 2014 displaying the values for the average particulate matter (PM) in the air per date. Additionally, 9.2 shows another data

200,000 matriculation examinations

... are taken in Finland by 30,000 candidates on a yearly basis

Source: Britschgi, V. (2015). *The Finnish Matriculation Examination.*

[21] Sources: https://digabi.fi/kuvapankki/
Britschgi, V. (2015). *The Finnish Matriculation Examination.*

source, which is an image from a Real-time Air Quality Index in Beijing and some other data. Students can upload both text and diagrams or other evidence that they have produced using the sources to answer the question. More examples are available in Finnish and Swedish, through YLE, the National Broadcasting Company that publishes media versions of the tests after exam days.

Implementation process

Framing the user requirements and goals for the digitalisation process

The key requirements for the digital examination system were strict; it was important that the independent exams could be arranged simultaneously in approximately 400 locations throughout the country. It was also deemed important that students could take exams using their own computers and that they should have access to the same software applications and programmes to offer a levelled and equal 'playground' to all. As for privacy and further security reasons, no files should be stored on students' computers. Last, the system should be fail-safe (e.g. in case of computers break or electrical power failure data is automatically protected).

The ME reform process has been executed step by step (see Figure 13). The first digital exams were administered in German language, Geography and Philosophy in autumn 2016, while the other subjects were still in paper format. The digitalization process is expected to be finished by 2019 when Mathematics, the most challenging subject to digitalise, is completed.

An important and integral part of the implementation is the possibility to offer a training environment that contains the same tools and functions as the examination environment. The system, called Abitti, is offered to schools by the MEB so that they can use it to organise their regular exams during the school year (see section *Success factors & lessons learned*). This helps teachers and students to get used to the examination environment and its function, and importantly, also to new types of questions.

Challenges faced

Solvable challenges related to social, technical, human and digital aspects

The digitalisation of the only national-wide high-stakes standardised examination has pushed through transformation at different levels of the school system in Finland. Support for teachers and education providers at the municipal level has been provided jointly by education authorities, but a key role in the successful transformation has been teachers' peer support and local support solutions.

From the technical point of view, the requirements for implementation of the digital ME meant that a number of issues had to be tackled. For example, using an operating system from a USB drive has proven more unstable than first foreseen. For that reason, an additional 20% of USB drives are delivered to schools for preventive purposes.

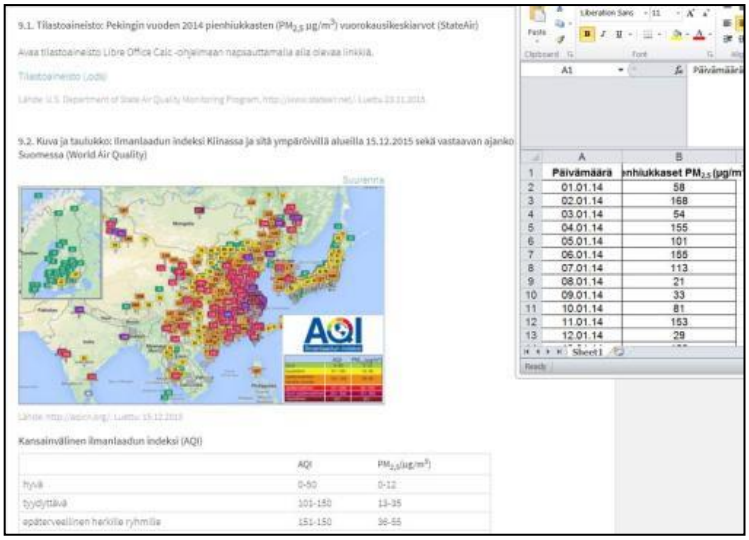

Figure 15. Example from Geography matriculation examination.[22]

Additionally, each school provides spare laptops for 10% of their students in case of technical difficulties. Last, a number of compatibility issues have emerged, however, at present, the digital ME has been successfully operated using more than 2,000 different computer models, suggesting that technical challenges are solvable. The MEB conducts user-surveys after each examination to gauge emerging issues.

The biggest challenges, however, for the digital Matriculation Examination are related to the change in educational culture that new digital tools and subsequent new working methods bring to schools.

Success factors & lessons learned

Practice makes perfect – ways to ease adoption

One of the key success factors of the digitalisation process and its adoption by learners and teachers is the environment for training and practicing purposes. This enables candidates, schools and the MEB to properly prepare for the digital ME. It also establishes a good feedback loop between the system and the MEB to ensure that necessary support measures are put in place.

The MEB released an environment called Abitti, which is similar to that used in the Matriculation Examination. The upper secondary schools can use it throughout the school year to organise regular exams, and thus practice and be fully prepared for the ME. It allows teachers authoring exam items, creating the USB drives for students, carrying out the exam in the local network, and evaluating students' answers (see Figure 16). As oppose to the digital Matriculation Examination system, Abitti can be used by schools and teachers unlimitedly. By early 2018, over 1,500,000 exams have been taken using the Abitti system.

[22] Sources: All official training material at https://yle.fi/aihe/abitreenit/ https://www.ylioppilastutkinto.fi/ext/harjoitus2016/fi_maantiede/attachments/#9.1

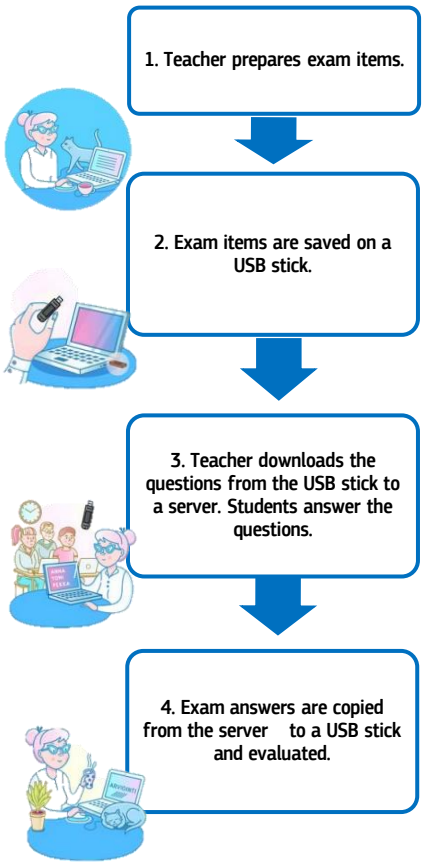

Figure 16. Practising digital tests.[23]

For example, for mathematics, all necessary tools needed for creating digital examination items are already in place in the Abitti environment and thousands of answers have been created using these tools. This allows the MEB to further develop the ME environment itself so that better items and smoother administration processes can be created. The MEB estimates that without the launch of Abitti, the digitalisation process, or parts of it, would not have succeeded.

On the other hand, teaching and assessment practices seem to have been also influenced by the digitalisation process and that teachers are changing the way they use digital tools to support teaching, learning and assessment. These new teaching practices do not only influence teachers in compulsory education but even higher education institutions show increased interest in organising their evaluations digitally.

Achieved results and impacts

Timely delivery and results at municipal level

The goals set up in 2011 for digitalisation of the Finnish ME have so far been met, the process is on schedule and budget. By spring 2018, 106,105 digital test submissions have been completed by candidates of the ME. All submissions have been aptly submitted, collected and graded.

The digitalisation of the ME has shown that the municipalities, who in Finland are responsible for upper secondary education, have taken steps to develop their infrastructure and created support processes for the digital ME, hence, the entire education system has been prepared for digital education. There is also some anecdotal evidence that textbook publishers have noted some increased use of digital resources in the upper secondary education and that, for example, schools no longer invest in physical calculators but rather move towards licenced software.

[23] Britschgi, V. (2015). *The Finnish Matriculation Examination.*

Potential for scale-up and replication

Facilitating the authoring of examination questions and their assessment through digitalisation

Regarding the Finnish ME, creating and authoring the examination questions and carrying out their evaluation is a large-scale operation. Moreover, proofreading and translations require rigid processes. The authoring process is carried out by some of the most prominent experts in their subject area in Finland, most of whom work in universities. For the assessment, around 400 assessors are needed, as most of the examination items are broad and open ended. The MEB is working to minimise the work load needed in technical assembly of the examination items and setting up background services to support these processes.

Building up the technical environment for examination setup in schools with servers and local networks is somewhat burdensome, and currently, no plug-and-play solutions exist. However, the examination environment has been successfully set up in all Finnish upper secondary high schools and even in atypical places such as prisons where inmates take the ME. As for the laptops used by students to take the ME, the requirements can be fulfilled even by relatively old computers and no purchase of licenses or software is needed by students or schools.

The examination environment seem to be a very scalable solution, proof is also the Abitti environment which is used successfully in all Finnish upper secondary schools by thousands of teachers who create their exams and grade them through the accompanied web services.

5.1.5 Multipoly Next

To recruit college students, the Hungarian division of PricewaterhouseCoopers (PwC) has created an online simulation called Multipoly. Over 5 years of its existence, this game-based assessment solution has not only increased the number of applications but also helped shape a more positive view of the firm.

Fact box	
Education level(s)	In-company training/Job-based learning
Target group(s)	University graduates
Specific subject(s) addressed	ICT; Technology; Computing
Transversal skills addressed	Business acumen; Digital competence; Relational skills
Assessment objectives	Recognition of prior learning; Certification of skills
Assessment methods	Recruitment game; Data-driven assessment; e-assessment
Assessment format	Simulation

Rationale and assessment objectives

Driven by the need to improve the recruitment process

According to the surveys of the Hungarian division of PricewaterhouseCoopers (PwC), potential job candidates were spending on average no more than 10 minutes on PwC's career website. In order to tackle the issue, PwC initiated the development of a simulation game based on the concept of a previous recruitment game for universities called "Unisafe". The game was developed for the Universities of Pecs and Szeget in order to attract students and counteract students' tendency to study in Budapest. Through Multipoly, PwC aimed to retain and more fully engage the national pool of interested talent. The game was funded solely by proprietary resources.

The main goal was to improve the branding of the company and make PwC stand out among its competitors. However, the end product went beyond a simple online simulation serving as a professional competition targeting business school graduates. It soon became clear that Multipoly provides PwC with an innovative means to assess various professional skills and knowledge as well as the 'players'' motivation to work at the firm.

During the game, the players learn what the company offers and how it differs from other professional service firms. Moreover, the participants can compete against other players and win prizes such as trips and e-gadgets. The ranking of participants determines how well they are doing. The points earned in the game allow the players to purchase PwC merchandise in the game online shop.

2.5 hours

... is the average time the Multipoly tool was used by the target audience

Source: Interview with Games for Business

Eventually, some of the best players have either been invited to job interviews or they were offered internship positions in the company. Although not intended, the game became a mean of assessing candidates' capabilities and their motivation to work in PwC.

Key features

From a short look at the website to an average 2.5h simulation

Multipoly began as a two-part game, with the first part being an online competition and a 'virtual traineeship' followed by the very best candidates being invited to a face-to-face finale. The online competition takes the format of a virtual interview, for which the potential candidate or 'player' answers a series of questions. After passing the interview, the candidate is invited to complete a virtual traineeship over a period of 12 game days. The traineeship simulation divides one year of professional experience into four quarters with different objectives and exercises.

Participants are requested to take part in a series of virtual job activities (see Figure 17), similar to those they would encounter in the job, and to join the virtual community of PwC employees. In the second stage of the game, the best players participate in a real life finale. The finale takes place in the Budapest office of PwC.

Over the years, the Multipoly game has been improved and adjusted in line with the needs and feedback obtained from the users. The second edition of the game -Multipoly Next- was released in 2015. The idea was to redesign the game as a more realistic recruitment experience. The second edition started with the virtual interview and then provided options to define the precise area of interest (risk assurance, audit, advisory services etc.) of the candidate to be explored further. While in the first edition of Multipoly candidates had to complete a full simulation cycle, the second edition allowed players to jump into different job profiles.

Figure 17. Screenshot of Multipoly Next.[24]

The interview was designed in the form of a series of questions that the candidates had to answer. The 'players' even had the opportunity to listen to the real life experience of staff of the Hungarian office. In

[24] Source: http://multipoly.hu/en/demo-video.html

the course of the game, players were provided with relevant information and their competences were assessed through automated testing solutions at multiple stages. In addition, these solutions also included the provision of feedback to potential candidates.

Once the interview stage is over, the players are given the opportunity to virtually meet other employees of the company including future mentors in a tour around the offices of the company's headquarters. Moreover, useful guidelines and advice are available at any stage of the game which participants are strongly encouraged to follow. Additionally, participants have the possibility to complete information on their professional profile including their CV and skills and competences.

Thanks to their participation in the game, players are able to gather useful information on the recruitment procedures, get to know the company and even improve their CV. The provision of different memory games and simulations also allowed participants to explore a multitude of available career paths. The duration of the game can last 3-4 hours. Users can choose to either finish the game in one go or quit and come back to it at a later stage. However, the average time that users spent on the game was 2.5 hours, which is substantial in comparison to other marketing tools available.

Transversal skills focus

Indirect result - assessment of competences

As an indirect outcome of the Multipoly game, potential candidates undergo a professional skills assessment including transversal and other professional competences. To begin with, English language skills are thoroughly assessed as the entire game is in English language, making proficiency in English a requirement for participation.

The game *"pre-educated [them] about PwC and its vision, services and skills needed for success."*

Source: Noémi Biró, PwC Hungary's regional recruitment manager

The game involves a simple competency test covering a total of 48 different questions on workplace competency, type of workforce, personal strengths, weaknesses and other qualities in relation to decision making. Furthermore, the game features an office management test. The test assesses candidates' ability to read, comprehend and organise information. It included imaginary situations like the following: the participant is a manager of a department store encountering a pile of documents on his/her desk. The task is to prioritise and organise the documents and then later on make the decision about who to approach according to the issues that arose.

Moreover, a video simulation test was also incorporated in the game. In this exercise, candidates watch a video segment such as a sales person presentation for a commercial sale. Participants need to listen to the client and make a recommendation in line with the information presented in the video. In the end, the participant is informed about the outcome of the attempted sale.

The professional assessment content of these automated solutions -integrated in the simulation- was provided by Structured Home Learning (SHL), while the game's content was developed by MarkCon Informatics.

Implementation process

Speedy and successful implementation

The game underwent a successful and speedy implementation. Although the preparations and planning

only started in 2011, Multipoly was successfully released in 2012. Indeed, the entire creation process took only 6 months in total. Under the supervision of PwC, the technical operation of the game was commissioned to MarkCon Informatics who developed the storyline and content of the game based on the concept of Games for Business and the game was accessible via the www.multipoly.hu webpage. Moreover, the game is linked to social media allowing players to share the content, like the posts and invite friends to participate (see Figure 18).

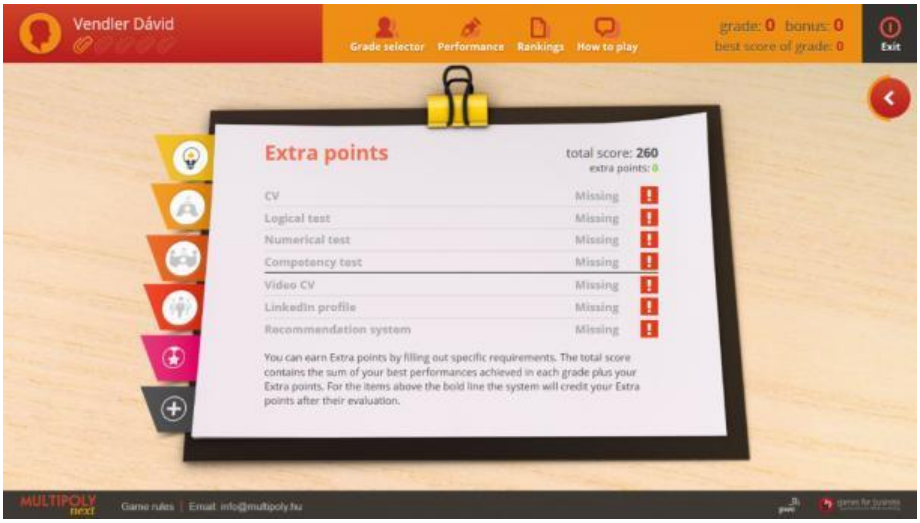

Figure 18. Screenshot of Multipoly Next.[25]

Challenges faced

The complexity of the game

Even though the implementation of Multipoly Next went very smoothly, receiving positive overall feedback, some 'players' faced difficulties. As a result, some participants have advocated that Multipoly should be simplified and made more approachable. Participants also claimed that mistakes had been made in the creation of the game. However, PwC confirmed that these issues were dealt with quickly and effectively after being reported. Overall, Multipoly can be considered a very successful employer-branding tool, which did not face any major barriers or obstacles in the course of its development and implementation.

Success factors & lessons learned

Unique proposal to communicate about the company

The key success factor of Multipoly lies in the uniqueness of the game. Before the introduction of Multipoly by PwC Hungary, no other comparable tool was used for promotion and recruitment.

Multipoly represents a fresh approach to improve a company's image by increasing transparency and openness to the public. Unlike commonly used business simulations in the market,

85%

... of Multipoly Next users confirmed that the game positively impacted their views of PwC.

Source: Interview with Games for Business

[25] Source: http://multipoly.hu/en/demo-video.html

Multipoly was engaging players in a way unlike any other business simulation.

The game itself has managed to improve the transparency of PwC Hungary's recruitment process, while making it more efficient. Instead of having companies spend a lot of money on human resources to make sure that suitable talent is hired and retained for the firm in the long run, the Multipoly game allowed candidates to make up their minds and set realistic expectations. While in some cases 'players' may change their mind about working in PwC, the tool allowed the company to filter the initial pool of candidates, thereby reducing time and costs.

Speeding up the team integration

Since the Multipoly game allowed participants to experience the actual working day in the company and experience what working for PwC is actually like, former Multipoly participants already knew about the company and its processes and therefore they were effectively integrated in the team. This was also confirmed by PwC Hungary: Employees that were hired after being successful in the game were on-boarded much faster during the first months of employment.

Achieved results and impacts

Strengthening the image of the company

As a recruitment tool, the game brought significant improvements to the company. The interest of Multipoly players in learning more about working at PwC increased by 78% and the time spent over the company´s website by around 18 times. Moreover, since the introduction of Multipoly Next, around 85% of job candidates said they had a more positive view of the company. As a result, PwC recorded a significant increase in the number of job applications. In particular, the job candidate pool grew by almost 190%. From a business perspective, by going virtual, the employer saves time and money, learning about the candidate and their working style.

Figure 19. Results overview of Multipoly Next.[26]

Easing the transition to becoming full employees

Bringing benefits not only to the company, the game makes it easier for job candidates to transition to full-time employees. By playing the game, they already had a taste of the company´s culture, tested out

[26] Source: https://bit.ly/2m1FLGQ

different roles, found out about various departments and faced “real life” business problems (see Figure 19).

Furthermore, Multipoly Next allowed players to get a virtual taste of different career paths at PwC without any attached risks. Providing this opportunity has helped reduce candidates' anxiety level. According to the company, applicants who played Multipoly were better prepared for live face-to-face interviews. Moreover, participation in the game also informed and prepared them for PwC by emphasising the skills needed for success.

Potential for scale-up and replication

Uncertainty over the future

Multipoly Next was used in the recruitment process of PwC Hungary during a period of 5 years. Each year the necessary improvements were made to reflect the reported issues and the growing number of participants, new functions and design aspects were added to improve the user's experience. In total, several thousands of 'players' have participated in the game Due to the successful implementation of the tool there are plans for replicating the game at regional level.

The reason Multipoly is currently not in use is -as for any marketing tool- related to the life cycle of corporate branding. After using the tool for a period of five years, there was a need to update the branding strategy of the company. However, there is a chance that an updated edition of the game will be released in the near future.

5.1.6 Neo Alta

Neo Alta is a practitioner-driven initiative, which has established an educational community focused on pedagogical learning and assessment. Assessment being a core component of this pilot project, the initiative applies the so-called "positive assessment" method, a combination of both, daily practice and end-of-period assessments.

Fact box	
Education level(s)	Lower secondary education (grades 3-5 – 10-14 years old)
Target group(s)	Learners (students)
Specific subject(s) addressed	Geography; Technology; English; Spanish; German; French; History; Sports; Visual Arts; Theatre
Transversal skills addressed	Critical thinking, analysis, reflection; Team work/Collaboration; Learning to learn
Assessment objectives	Improve student engagement; Improve feedback to students; Improve feedback to parents
Assessment methods	Student self-assessment; Teachers' formative & summative assessment
Assessment format	Portfolio assessment; Student response system

Rationale and assessment objectives

Towards a positive educational assessment

In France -and also in many other countries- educational assessment tends to follow a rather negative tradition.

Neo Alta started in 2012/2013 as a counter -practitioner-driven- experiment in the French middle school "Anatole France" (10-14 year old students) by introducing a "positive assessment" approach. The idea behind the concept is to keep a positive spirit by regarding failures as inevitable. Neo Alta emphasises the importance of preventing students from entering a negative spiral, which could lead them to learning disengagement.

As a complementary -yet voluntary- initiative, Neo Alta sets out to experiment with different assessment forms in two -later three- lower secondary education classes. The main mission of the experiment was to demonstrate that innovating within a normal school with normal teachers is possible (see Figure 20).

At the centre of the initiative is the student-teacher-parent relationship. They all actively participate in setting learning

Practitioner-driven innovation with relatively simple means

Neo Alta started as a bottom-up experiment of a group of inspired educators seeking to change assessment practices in the classroom

Source: Interview with Neo Alta initiator

targets and accompanying the learning path of their children. By means of different classroom activities the experiment seeks to improve student engagement and provide feedback on their learning progress.

Key features

A series of practices and principles

Neo Alta is composed of several components, combining pedagogical with institutional aspects. One of the first points of the initiative is Neo Alta's school certificate which is complementary to the conventional school certificate. Unlike the latter, the Neo Alta certificate is based on four cross-curricular competences.

The initiative follows several pedagogical principles such as the absence of rewards and sanctions in the classroom. Besides the abolishment of warnings and extra work, the experiment equally bans congratulations, compliments and encouragements from its classroom practices.

© Vincent Jarousseau

Figure 20. Teacher demonstration during Neo Alta.

The “second-chance principle” is an additional integral component of the French initiative. This principle allows students -after a moment of reflection on their achieved work/assignment- to re-do/modify their work as many times as they want to, if they are not yet satisfied with the final grade. If the new score is better after the second submission, it replaces the first; if it is worse, the first remains, while in some cases students have asked for a third chance.

Institutional changes for student, teacher and parent involvement

Neo Alta puts in place several institutional/structural changes strengthening the teacher-student relationship as well as enhancing the involvement of parents in the school education of the child.

One institutional arrangement is the Neo Alta's class council. What distinguishes this council from traditional ones is its composition, comprises all students and parents as well as a maximum of teachers. The class council allowed every member to voice concerns and discuss potential problems.

Further structural arrangements concern the monitoring of students' learning goals. Next to an individual meeting of students with the referent teacher, parents equally participate in the (re-)negotiation of students' progress goals. These goals are individualised and appear at the bottom of the school certificate. Bi-weekly all students do a self-assessment, which is discussed collectively or individually. This way students take stock of the achievement of his/her objectives.

Once per week, Neo Alta organises the so-called LABOs as interdisciplinary classes to emphasise the link between several disciplines in co-animation with two or three teachers. Examples from the past have included linking Geography with Technology -in terms of sustainability- and communication classes bringing together aspects from History and Technology (using digital tools). In 2017-18, the focus was on group work satisfaction and on co-evaluation of oral productions using descriptive scale compositions.

Transversal skills focus

The key competences targeted in Neo Alta are: 1) French language proficiency (oral, written, reading); 2) “Putting oneself to work” organising oneself to complete a task; 3) "Autonomy and creativity" and 4) “College life”. These four competences appear on the Neo Alta school certificate with comments for all subjects (see Table 3).

The focus on these competences stems from the school's experiences. Educators have pointed out general difficulties of students to express themselves in French, to not work enough, to lack a sense of initiative and to stay inactive during the lessons and extra-curricular activities. As a result, the key competences were designed with the purpose to help students overcome difficulties in these fields.

Table 3. Examples from Neo Alta school certificate.

Autonomy and creativity	English: Excellent written production on the end of sequence chapter "Into the wild"! Student very autonomous in the work. Club FSE Theater: A quality performance on stage! Spanish: Seeking to reuse his knowledge and deepen it on his own! French: A presentation of personal reading - quite creative and original, bravo! Able to take initiative, follow instructions and look after his work. Neo Alta: Prepared and presented upon his own initiative a theater dialogue at the presentation evening.

Implementation process

Innovating bottom-up by inspired educators

Neo Alta started as a practitioner-driven experiment developed during the traineeship of an educator at the Collège Anatole. Three teachers approached the school Principal in 2013 to propose an educational experiment. The Principal approved the experiment and a team of volunteers was formed. In the course

of an initial training, Neo Alta was born. Importantly, the education authorities validated the experiment granting the school four hours per week of extra tuition to pay participating teachers -yet, teachers dedicated several hours of their free time to ensure Neo Alta's operations. Neo Alta at first started to become a reality in two classes in the 5th (13-14 years old) and 3rd (10-11 years old) grade before being extended to the 4th grade (12-13 years old; 77students in total).

Challenges faced

Dependent on the drive of engaged individuals

Being a practitioner-driven initiative, the main challenge faced by Neo Alta is its dependency on engaged educators, students and parents. This became clear when the leadership of the institution changed. The flexible approach of Neo Alta gradually declined towards a more rigid implementation of the intervention activities.

Figure 21. Neo Alta activity.[27]

In relation to educators, challenges concern adapting to new teaching practices and staff fluctuations. Involved educators report to be in need of further training. This is even the case for volunteer teachers who meet every two weeks and exchange pedagogical practices or engage in co-training.

As part of Neo Alta's principles -"to innovate in a normal school with normal teachers"- it is not possible to recruit educators externally on the basis of specific skills for Neo Alta. In fact, Neo Alta relies on reforming educational assessment from within focusing on a number of feasible changes.

Rather than job recruitment, Neo Alta is operated on the basis of volunteer engagement. Yet, this results in many leaving the programme after one or two years.

Although some extra hours may be paid, educators report that the programme requires significant additional workload (e.g. for the morning reception, tutoring, evening study, co-animation). Currently, the Neo Alta coordinators lack human resources. That is why the programme is now limited to two classes.

[27] Source: Neo Alta.

Despite of some financial support, more teacher training and development is needed for educators to become motivated and engaged to the programme.

Success factors & lessons learned

Trust and teamwork between adults and students

The key success factor of the programme lies in the relations between children and adults. Students' and parents' feedback show the great level of trust to the aim of the programme. Students feel that educators are investing in them and parents are disappointed that the intervention will come to an end after 2018-19 (see Figure 21).

The greatest strength of Neo Alta is its team of educators. Unlike the occasional collaboration among teachers, the exchange of teaching practices and experiences occurs on a daily basis. Besides the more informal ad-hoc collaboration, there is project-specific teacher training. Altogether, four training courses have been provided over the duration of the programme, each one focusing on a different topic (e.g. the last two years the focus was on cooperative practices).

Notwithstanding the challenges tied to volunteering (see section *Challenges faced*), it can equally unfold a strong dynamic, if applied by engaged individuals. All involved actors -teachers, parents and students- are free to stay or choose to end his/her participation. At the time of implementing Neo Alta, an educational reform of middle schools ("collèges") was introduced. Since Neo Alta was already experimenting with some aspects of the reform (e.g. personalised support and interdisciplinary practical lessons) it could assist Neo Alta and vice-versa; unfortunately, this did not occur. On the contrary, implementing the reform without any alignment, Neo Alta was time and energy consuming and contributed to the fatigue of the educators.

Achieved results and impacts

Since its first launch in 2013, Neo Alta is by now in its 5th year of implementation. Although the initiative has not been subject to an in-depth evaluation, substantial feedback from parents and students allow for some conclusions to be driven.

Benefiting from monitored learning progress and setting learning goals

Feedback received from parents and students alike indicates that the programme allows for more precise follow-up of students' learning status and progress. As one parent commented, "Neo Alta's school certificate and parental involvement helps clarify the particular strengths and weaknesses of students".

Focusing on short-term learning goals can serve as a useful orientation point for students. Moreover, throughout the process of working towards their set targets they received help and ideas from their peers or teachers.

Positive assessment impacts

At the beginning of the programme, the Neo Alta team has repeatedly been confronted with parents' concerns that the removal of rewards and warnings will alter students' behavior and result in reduced

work levels. However, there is no such effect. Good students keep on performing well and their results were improved when work was resubmitted.

Students with difficulties had also improved although to a smaller scale. Yet, Neo Alta values more, students' learning process than their final grade.

In general, the very low drop out quota -only one student left the programme at the very beginning- proves that students see advantages in it. However, it should be noted that Neo Alta is not imposing what students should learn - this is decided by the student; neither does the programme force students to re-submit their work after a low personal result; yet a quick interview will take place as a way to understand their reasons and motivation.

Potential for scale-up and replication

At the outset of the initiative, Neo Alta was designed to be applied in a normal formal education environment without significant costs. Therefore, the initiative has potential for replication in other schools in France or even in other countries. Even though the initiatives' low financial costs may support its potential replication, the voluntary engagement of educators and parents as well as the motivation of students are essential for its success.

Despite being a local school initiative, Neo Alta has received nation-wide attention. The programme was presented during the Innovation Day in March 2017. A further indicator of the programme's success is that parents from other grades have requested the introduction of Neo Alta to other grades. Scaling up the programme in the current format would need to address the shortcomings of volunteering and the lack of human resources due to staff fluctuation. Moreover, a greater level of teacher training and financing is also required for its successful implementation. However, there is a possibility for the project to continue under a different format.

5.1.7 Teach for Slovakia (Teach4SK)

As an initiative financed by the government and private donors, Teach for Slovakia aims to improve access to quality education in Slovakia. During a 2-year programme, university graduates teach primary students making use of specific assessment strategies. Moreover, selected candidates receive targeted training and coaching to improve their skills.

Fact box	
Education level(s)	Primary education (grade 1-9, age 6-15)
Target group(s)	Youth learners (high share from Roma communities)
Specific subject(s) addressed	Cross-subject (support for the development of basic literacy and reading comprehension)
Transversal skills addressed	Civic and social competences; Critical thinking, analysis, reflection; Problem-solving; Communication skills; Planning and organisation skills; Entrepreneurship
Assessment objectives	Improve student engagement; Improve students' learning; Personalise learning
Assessment methods	Teachers' formative assessment; Student self- & peer-assessment
Assessment format	Project assessment; Collaboration/competition; Presentation

Rationale and assessment objectives

Countering inequality in access to education

According to the PISA 2012 study, Slovakia has one of the highest inequality rates in access to quality education among OECD countries. The socio-economic background of children has a high impact on pupils' school results: Only 4% of higher education students in Slovakia are the children of parents without a higher education degree.

Teach for Slovakia aims to build a community of leaders working both inside and outside schools to improve the education system in Slovakia. Targeting university graduates across disciplines, participants complete an intensive 2-year programme.

Next to being full-time primary school teachers in schools -where they teach 6-15 year-old pupils- participants also complete other customised assignments that equally address inequality in education. In return for their engagement, the prospective teachers obtain in-depth training and coaching/mentoring as well as access to a highly motivated community seeking to improve the education system in Slovakia.

The programme focuses on underserved communities giving it a social inclusion character. Although many of the classes show a higher percentage of Roma children, one cannot speak of a specific Roma focus.

Key features

The Teach for Slovakia programme has a 2-year duration and combines intensive teaching of Slovak pupils with targeted competence prospective teachers are assigned one primary school class with full teaching timetable where they will apply a set of specific teaching practices.

At the start of the programme, the selected participants are trained in a basecamp, a 6-week intensive workshop (see Figure 22). The basecamp aims to prepare the recruits, equipping them with the baseline skills and a mindset to succeed. The prospective teachers become part of the programme's community which interacts and supports each other and works towards the common goal of improving Slovakia's education system.

Figure 22. Pilot teachers from Teach for Slovakia in 2015.[28]

During basecamp the recruits are trained on teaching and motivation techniques including self management. Given the diversity of participants' education backgrounds and lack of formal teacher training for the majority of candidates, recruits receive in-depth training on class teaching skills and assessment techniques during one week of the summer programme.

The prospective teachers implement teaching practices in the classroom. For three weeks, they are supported by an assigned mentor who provides them with daily feedback on their progress. In addition, each participant sets ambitious personal goals to be achieved in the course of the programme.

Throughout the programme, a variety of assessment formats are employed such as self- and peer-assessment. The formative assessments focus on building students engagement with a view to improve their learning with techniques such as the use of mini whiteboards and the coloured cups method.

The use of mini whiteboards in combination with open and closed questions serves as an instant feedback tool, which increases student engagement. Once the question is raised, students note down their solutions and raise the whiteboards; in closed ended questions, a voting system can be deployed. In this way, teachers know who has understood the concept and who needs more instruction and may equally serve to show how confident students are.

[28] Source: http://skolskyservis.teraz.sk/skolstvo/teach-for-slovakia/17058-clanok.html

The colours cups method is a non-verbal feedback tool for students which can be used in multiple instructional situations (e.g. for whole group and small group instruction or as group work/task and independent work/task). The green cup signals comfort with the pace of the lesson/task, the yellow cup indicates a need for review and the red cup indicates lack of comprehension or a question to be posed.

According to the skills students need to develop, unit and lesson objectives are set up. Students are provided with a list of success criteria per activity and self-assessments are completed in the middle or towards the end of an activity. Periodically self-assessments are changed for peer-assessments, which enable students to give feedback to their classmates.

Beyond the teaching activities, the recruits also complete an internship in line with their professional development plans, the so-called leadership challenge project. After graduating from the programme, participants can join the ambassadorial's programme which offers them additional training and networking opportunities as well as individual career support (see Figure 23).

Figure 23. Ceremony of Teach for Slovakia programme graduates.[29]

Transversal skills focus

Focus on working with disengaged learners

The transversal skill focus is connected to the theoretical approach of the programme. Tough (2016) describes successful interventions with more active and engaging educational processes which improve the educational outcomes of students with disadvantaged social and economic backgrounds.

In contrast, teachers who work with such students believe that dominant control is needed and opt to a more teacher-centered approach. As a result, students feel alienated, they become disengaged from learning and further disempowered.

However, when children feel a sense of belonging, experience autonomy and are given the opportunity to engage deeply in their own learning they become motivated to work harder (see Figure 24).

Based on such a perspective, the work with disengaged learners focuses on fostering the following transversal skills:

[29] Source: https://bit.ly/2k9qkMk

- Civic and social competences.
- Critical thinking, analysis, reflection.
- Problem solving.
- Communication skills.
- Planning and organisation skills.
- Entrepreneurship.

Implementation process

Inspired by practices in the United Kingdom

At the beginning, two of the initial team members worked in the UK and observed a diversity of assessment strategies that are not commonly used in Slovakia. They decided to design a teacher training and leadership development programme, which integrated some of these strategies.

What the initial team observed during the creation process of the programme was that once teachers gain confidence in the application of assessment strategies, they tend to use them incrementally. Therefore, the initial team wanted to ensure that teachers are equipped with effective tools and enough support in using them. The initial training and support was particularly important in the Slovakian context since the introduction of new assessment strategies could cause some friction.

Figure 24. Teach for Slovakia classroom.[30]

Challenges faced

Transition time to master the new approaches

One of the main challenges of the programme is related to students' unfamiliarity with the assessment methods although training was provided.

The new assessment techniques were not used for the school certificates' final grade, i.e. as summative assessment, due to the fear of uncertainty. For this reason, the majority of approaches used in Teach for Slovakia are formative. In the future, the programme intends to increase the share of summative

[30] Source: https://bit.ly/2k6eYbX

assessments. Yet, the question remains whether alternative summative assessment forms would be considered rigorous enough to be accepted by education authorities.

Success factors & lessons learned

Tailored teaching practices and a supportive environment

One of the key success factors of the programme is the flexible adaptation of teaching practices to the particular environment of each school. Rather than deploying a rigorous method, teachers are supported through tailored training and coaching to apply the most effective teaching practices for each course and subject.

Moreover, the participants have access to a supportive environment, which consists of training seminars and workshops, personal mentors and a cohort of fellow participants. As a result, teachers are familiarised with assessment strategies, they consolidate their skills and the impact of their teaching is increased.

Secondly, the work on growth mind-sets is seen to make a difference compared to other programmes. According to participating teachers, this strand of work has a decisive impact on students and teachers. Therefore, it is considered as one of the most important pedagogical components for the success of the programme.

Achieved results and impacts

Positive feedback from students and teachers alike

The programme measures impacts in terms of student performance and assess the programme internally for each school year. Yet, there is not any comprehensive monitoring of a specific assessment technique.

According to monitoring data of student performance from 2016/2017, classes in grade 4-8 on average improved their basic literacy or numeracy levels by 11% over the school year. In some classes of grade 5, the increase in student performance over the school year even reached 24%.

Overall, students' feedback has been very positive with teachers confirming that students find the practices interesting and engaging. In order to avoid student discomfort in relation to the new assessment practices, teachers often introduce them gradually. They start with low-stake activities and once students are more comfortable the techniques are slowly built up.

Teachers also share the view that the programme has been successful. Many of them continue using the assessment techniques in their professional career in order to engage students and increase the ownership of their learning process.

Potential for scale-up and replication

High replicability potential

Teach for Slovakia is part of a global network spanning 48 countries on six continents. In principle, there are attempts to gradually spread the assessment practices to other schools. Through the programme alumni, school principals are encouraged to advocate their use. In fact, our hitherto strategy is to keep on gradually expanding the number of classes using innovative approaches of assessment.

5.1.8 Assessment based on an Online Collaborative Project (AOCP)

Online teachers and course designers (Montse Guitert, Teresa Romeu and Marc Romero) from the Open University of Catalonia (UOC) have installed a continuous assessment model. Combining multiple assessment methods, their 360° model assesses individual and collaborative performance and is highly valued by students and teachers alike.

Fact box	
Education level(s)	Tertiary education
Target group(s)	Online University students
Specific subject(s) addressed	Transversal/soft skills
Transversal skills addressed	Digital competence; Collaboration; Dynamics; Interaction
Assessment objectives	Assess transversal or soft skills; Improve student engagement & learning; Improve feedback to students
Assessment methods	Teachers' formative & summative assessment; Students' peer- & self-assessment
Assessment format	Collaboration; Project assessment; e-assessment

Rationale and assessment objectives

A continuous e-assessment model

The ICT Competences Course is a mandatory transversal subject which students take in the first year of their online studies. The aim of the course is to gradually and integratedly equip students with transversal competences considered essential for studying at a 100% online university. The key activity of the course is that students develop a collaborative digital project in wiki, web or video format. Developed in work groups of four, the projects are assessed in terms of students' individual and collective performance.

Emerging from academic research, the assessment approach -the 360° model- is based on continuous e-assessment. Through self- and peer assessments the model puts students in the centre of the assessment providing them with a stronger voice in the evaluation process.

The model uses several assessment methods, ranging from online self- and peer-assessment, to reflection on teamwork. Thanks to the platform, teachers can monitor data per student including their perception of the course as well as their performance.

Key features

The 360° assessment model

The assessment of the course is based on the 360° assessment model aiming to reinforce and boost students' learning process and help them acquire digital competences. The e-assessment model developed for the course relies on the concept of a 360° communication promoting the participation of online teachers and students, assessing individual as well as group processes and results (see Figure 25).

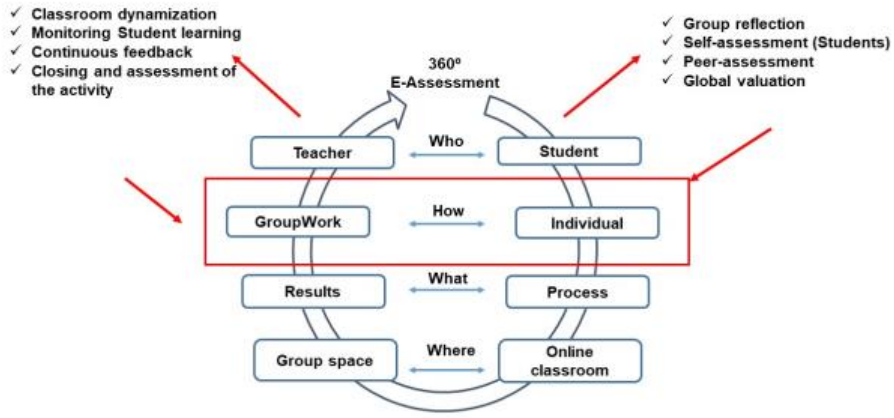

Figure 25. 360° assessment model.[31]

Diverse set of assessment methods

The assessment design of the ICT Competences Course is grounded on continuous e-assessment featuring several assessment methods. In summary, students are assessed in relation to two dimensions: 1) the process assessment executed in all phases of the project and 2) the assessment of the final outcome. While the teacher's role resembles that of a facilitator guiding and advising the work groups, students' active role is central for the assessment of both dimensions.

The process assessment includes a self-assessment and peer-assessment in which both individual's contribution and participation within group are assessed. Apart from individual performance, the group's overall dynamics and processes are also assessed.

Developed in the context of online teamwork, students' complete self-assessments in the last two phases. On the basis of these assessments, teachers can monitor the group work dynamics and act if problems are detected. In addition, teacher has the possibility to tone their grade up or down for the final assessment and students can equally reflect on the quality of teamwork after every project phase.

In the final assessment, students from other work groups assess the project of their peers. Enabled by UOC's virtual communication space, the presented projects are commented in an interactive setting. Thereafter, every student participates in the joint defence of the group work as a whole.

The criteria and indicators for the course assessment are set up jointly by course coordinators and teachers. These criteria are used to assess both individual and group activity and students receive separate feedback on both aspects.

[31] Source: UOC
https://bit.ly/2k9di1p

The activity: Developing a digital project

In the course students form groups of four to develop a digital project, the content of which can vary. In terms of the product, most of the projects are delivered as wikis, google sites or videos. Each group has access to an own virtual group space providing a series of tools (communication tools and tools to develop their projects).

The development of the digital project follows a structured process of four main phases, each subject to the multiple assessment methods:

> ***Giving a stronger voice to students***
>
> *The 360° assessment model promotes a continuous engagement of students in the assessment process.*
>
> Source: https://bit.ly/2kyjWOR

1) In the starting phase, the working teams are created and students conduct an initial search to set out the general theme of the project.
2) In the second phase, students perform research and define the precise structure of their project.
3) Once the structure is set up, the actual development phase starts. Here, students carry out the data and develop a first version of the project.
4) Finally, the project is closed and disseminated with students sharing and discussing the final outcome of their work (see Figure 26).

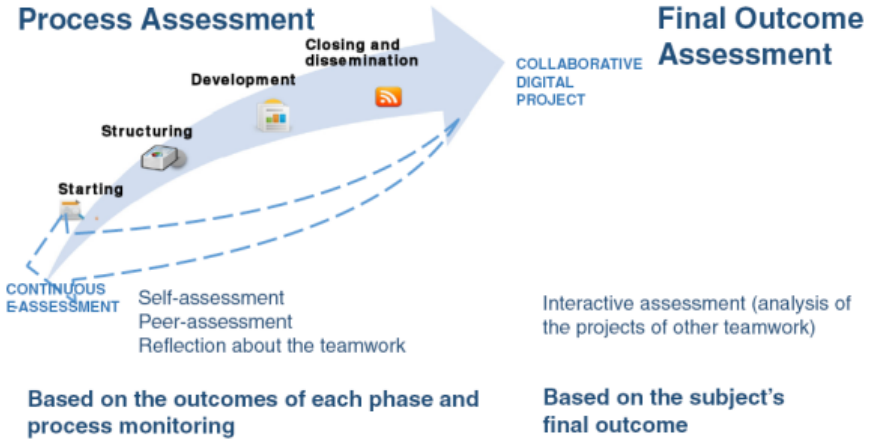

Figure 26. Dimensions of continuous assessment.[32]

Transversal skills focus

Digital competence focus

The skills are related to the definition of digital competence of the European Commission in relation to the "Use and application of ICT in an academic and professional environment" and "Online team work".

These competences are outlined as:

- Search and selection of information online.
- Processing and development of digital information.

[32] Source: https://bit.ly/2kyjWOR

- Presentation and dissemination of digital information.
- Notions of digital technology.
- Work and study planning in a virtual environment (see Figure 28).
- Communication strategies in the Net.
- Teamwork in an online environment.
- Digital attitude.

The competences/skills obtained through the course are also officially recognised by the Catalan Government as Digital Competence Level 2, which requires a streamlined assessment process and meeting further more specific conditions (see Figure 27).

Figure 27. UOC Virtual Space for peer-assessment of digital projects.[33]

Implementation process

Constant monitoring and adaptations

Since the set-up of the online course, its assessment has continuously evolved extensively over time improving and re-adapting contents. This evolution was enabled through constant monitoring.

When the course was introduced in 1995, its main purpose was to equip students with the skills to study online. In the context of the Bologna reform and the increased focus on digital competence, the course was re-designed to be competence-focused (see section *Transversal skills focus*).

Meanwhile, the introduction of the 360° assessment model in 2011 was planned with the purpose to enhance student learning and did not consider economical or efficiency gains. Therefore, there is no data on such an impact.

Today, 3,000 students and 75 teachers are involved in the course and different stages of the model are taken into account for its assessment.

[33] Source: UOC

Challenges faced

The limitations of the current tool

The main challenges are related to technological limitations, as the current online environment does not yet provide standardised information. Even though some progress has been made, an effective learning analytics system that collects systematic evidence of students' activity has not been developed to date.

In order to overcome this problem, the University is currently promoting the development of learning analytics tools and the improvement of strategies for more systematic data analysis. These solutions are expected to further improve the skills as well as to identify areas for improving student's learning.

The need for teacher training

The increased complexity of the online course and the assessment process has augmented the need for teacher training and corresponding documentation. In addition, the teachers' coordinators are working on the development of an assessment guide to facilitate its implementation. Further, experience has shown that the online teacher training allows educators to make proposals on how the assessment tools and processes can be improved.

Success factors & lessons learned

Research-based course design

The design of the course evolved on the basis of in-depth academic research in the field of collaborative learning in virtual environments undertaken by a group of UOC researchers. Factors such as technological progress and teachers' students' online experiences also shaped the design of the course assessment.

Motivating students by engaging them

An additional driver for the project can be attributed to the active role of the participants and the development of students' awareness. The fact that students are at the centre of the assessment practice, participating at several stages, increased their motivation. More significantly, the transparency of the assessment (e.g. in terms of criteria) allows them to see on what basis they are being assessed and how their own learning evolves over the course.

85%

of students evaluate the assessment method of the course positively

Source: https://bit.ly/2kyjWOR

A lesson learned in relation to the student-centred assessment design concerns the learners' increasing appreciation of the assessment methods over time. Although students initially show reluctance to assess themselves, they often report that the assessment experience has become valuable and assisted them in other areas. The process of self-assessment, peer-assessment and final assessments empower students as they become more aware of their learning process and achieved goals.

Some learners have reported to undergo a change of mind: The realisation of being in the centre of the assessment lead to increased reflection of their own learning process.

Furthermore, the positive feedback received from students may have an impact on how course participants approach the course. Overall, more than 3,000 students have completed the course and the

grand majority considers it very useful.

Engaged online teachers

The online course relies mainly on the collaboration of a team of engaged online teachers and much less on financial resources. Indeed, the fruitful exchange among teachers' of knowledge and good practices among teachers regarding course contents and as well as the assessment processes can be seen as a driver of the project. Without the support of a team of professionals, the online course would have not been able to evolve as extensively.

Achieved results and impacts

Comprehensive monitoring in support of the assessment model

The online monitoring tools of the course combined with satisfaction surveys provide a comprehensive dataset on student performance and perception/satisfaction of the continuous assessment model. The dataset has been used in academic research conducted by the course designers to the validity of the assessment model.

The data from the online surveys shows that a very high number of students (85%) evaluate the assessment model positively despite being more time consuming.

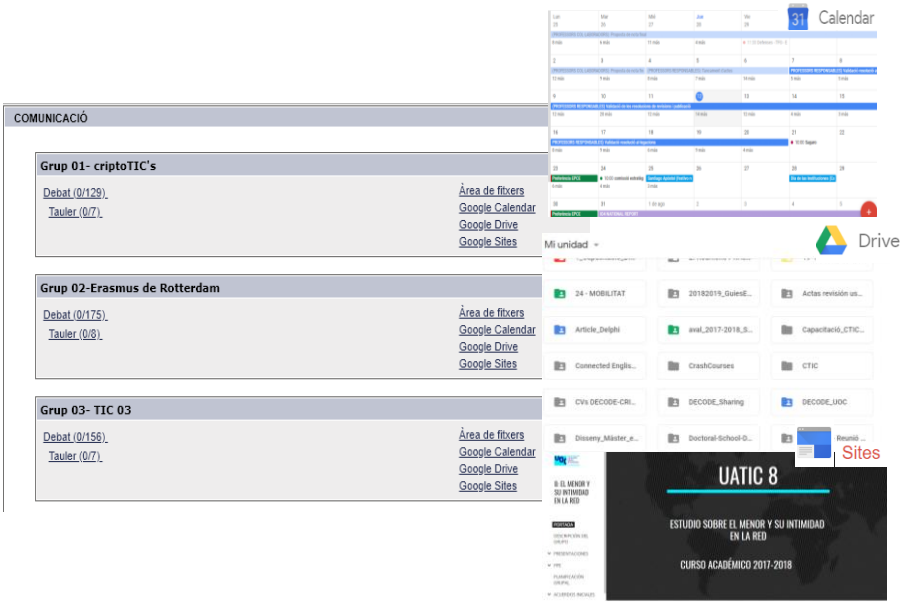

Figure 28. UOC's virtual group workspace.

With regard to the process assessment, the results show that students highly value self-assessment and peer-assessment. While self-assessment allows them to reflect on their individual participation and learning process throughout the project activities, peer -assessment encourages them to reflect on their role as a member of the group. As students have repeatedly pointed out, comparing their actions with those of their peers enhances their learning process.

Furthermore, the findings reveal that students' reflection on their groupwork allowed a better understaning of the dynamics/interaction taking place in the process of collaboration. They also valued the monitoring, guidance and feedback provided by teachers.

Moreover, students who express satisfaction with process assessment were more likely to claim the acquisition of the competences targeted by the course (see Figure 28).

Items of the questionnaire	Mean	Std. Deviation
Self-assessment lets me be aware of my competency acquisition.	4.30	0.858
Peer-assessment allows me to objectively specify my valuation of my peers	4.20	0.889
Reflection about the teamwork during the different phases helped us to improve our group's dynamics	4.08	1.039
Teacher's individual feedback helped me to improve my learning process	4.32	0.958
Teacher's group feedback helped us to improve group dynamics	4.22	0.976

Figure 29. Results of the online questionnaire (scale 1-5 =Do not agree - Fully agree).[34]

Final outcome assessment

Concerning the final assessment, students were highly satisfied with the assessment methodology on a whole. They also advocated that the assessment of other projects was an important experience for them.

In addition, students consider that the provision of continuous assessment facilitates the performance of self-and peer-assessment. In particular, interactive assessment was found to have two major advantages: The promotion of critical analysis of students' own projects and the detection of project's weaknesses through peer evaluation (see Figure 29).

Potential for scale-up and replication

High replicability

The partricular assessment model is potentially suitable to assess other competences. The assessment methodology can be replicated in online universities or in face to face academic courses. In fact, some online teachers of the institution have trasferred the assessment methodology to their face to face courses.

Most significantly, replicability depends more on the motivation of availabe human resources and less on financing.

[34] Source: https://bit.ly/2kyjWOR

5.2 Comparison of the case studies

5.2.1 Demographics of case studies

Table 4 illustrates the main characteristics of the eight examined cases on innovative assessment. The included studies originate from various European countries and cover all educational levels from Primary to Tertiary, with the exception of two that are related to professional contexts (company, VET). Most studies are practitioner driven apart from Multipoly which is technology driven and FME which is policy driven. The implementing organisation are mostly educational institutions (universities, colleges, pedagogic institutes) beside one (Multipoly). Their target groups are mainly students, teachers, university graduates and elderly care workers. Funding is mainly public except Teach4SK, which is mixed, and Multipoly, which is private.

Table 4. Overview of case studies' characteristics.

Assessment practice[35]	Country	Education level	Implementing organisation	Target group	Type of innovation	Funding source
ATS2020	10 EU countries	Upper Primary/ Lower Secondary	Cyprus' Pedagogical Institute (coordinator)	Teachers/ students	Practitioner driven	Public
e-APL	Sweden	Employment/ VET	University of Gävle; Sandviken municipality	Elderly care workers	Practitioner driven	Public
OSCE	United Kingdom	Tertiary	Sheffield Hallam University	Tertiary students	Practitioner driven	Public
FME	Finland	General Upper Secondary	The National Matriculation Examination Board	Teachers/ students	Policy driven	Public
Multipoly	Hungary	In-company training/ job-based learning	Pricewaterhouse Coopers	University graduates	Technology driven	Private
Neo Alta	France	Lower Secondary	Collège Anatole France	Students	Practitioner driven	Public
Teach4SK	Slovakia	Primary & Secondary	Teach4SK	Youth learners (Roma communities)	Practitioner driven	Mixed
AOCP	Spain	Tertiary	Universitat Oberta de Cataluña	Online University students	Practitioner driven	Public

[35] Case studies examined:
- ATS2020 : Assessment of Transversal Skills project;
- e-APL : E-Assessment of Prior Learning in Swedish elderly care;
- OSCE : Objective Structured Clinical Examination;
- FME : Finnish Matriculation Examination;
- Multipoly : Multipoly Next;
- Neo Alta;
- Teach4SK: Teach for Slovakia;
- AOCP : Assessment based on an Online Collaborative Project.

5.2.2. Rationale and assessment objectives

Depending on the nature of the initiative, the rationale and objectives behind the launch of the assessment practices illustrate a range of different motives and courses of action. In most initiatives, the beginning is marked with a perceived need/problem statement, which triggers the development of an improved practice/new product (Table 5).

Table 5. Motive and assessment objectives of examined initiatives.

Assessment practice	Motive behind launch	Assessment objectives
ATS2020	Inefficient tools to assess transversal skills; Teachers often not provided with effective tools for the development and assessment of transversal skills	Assess transversal or soft skills; Improve feedback to students, their engagement & learning
e-APL	High share of elderly care workers without formal education; Lack of quality assurance of skills and possibilities to participate in further training/certification	Certification of skills/recognition of prior learning
OSCE	For the university: Time-consuming, costly, personnel-intensive to conduct examination; For students: Decrease exam stress and panic	Certification of skills; Recognition of prior learning; Improve authenticity of learning
FME	Paper-based system lacked flexibility and efficiency	Modernise assessment
Multipoly	Low retention of local talent; Insufficient attractiveness and interest in the company	Recognition of prior learning; Certification of skills
Neo Alta	Negative assessment culture; Innovating with simple means and low resources is possible	Improve students' engagement, collaboration & learning; Improve feedback to students & parents
Teach4SK	Inequality in access to education	Improve students' engagement & learning; Personalise learning
AOCP	Lack of diversity and transparency of assessment methods (in particular, digital competence)	Assess transversal or soft skills; Improve feedback to students their engagement & learning

Several of the examined practices were launched with the aim to increase the feedback and engagement of involved learners (e.g. Neo Alta, AOCP, ATS2020) and encourage the development of transversal skills. The certification of skills and recognition of prior knowledge were among the objectives of e-APL and OSCE. In the case of Sheffield Hallam's examination for paramedics (OSCE), traditional assessment schemes were time-consuming, costly and personnel-intensive and induced increased stress levels and panic when learners were tested on their motoric skills. Teach4SK is a special case since Slovakia's high inequality in access to quality education was the main motivation to start the programme. On the other hand, Multipoly was launched due to a recruitment process, which was inefficient and costly for PwC. Finally, the FME seek to modernise assessment through digitalisation since traditional assessment lacked efficiency and flexibility.

5.2.3 Assessment design

The examined assessment practices comprise a variety of different methods and formats, as shown in Table 6. With regard to applied assessment methods, the majority (five out of eight cases) of them applied 3 or 4 methods on average. In terms of assessment, formats it becomes clear that innovative practices applied today are no longer limited to a quiz format/multiple choice structure, but incorporate formats such as portfolio, project assessment and collaborative assessment tasks. Interestingly, real life tasks and simulations were applied as alternative assessment formats in three of the studies (Multipoly, e-APL, OSCE).

Table 6. Overview of assessment methods and formats per case study.

Assessment practice	Assessment methods	Assessment formats
ATS2020	Teachers' formative & summative assessment; Students' peer- & self-assessment	Collaboration; Portfolio assessment
e-APL	e-assessment; Teachers' summative assessment	Quiz/multiple choice; Authentic learning/real life tasks
OSCE	Examiner's summative assessment; Students' peer- & self-assessment	Simulation; Real life tasks
FME	Digital summative assessment	Quiz/multiple choice; Portfolio assessment
Multipoly	Digital summative & formative assessment	Simulation
Neo Alta	Students' self-assessment; Teachers' formative & summative assessment	Portfolio assessment; Student response system
Teach4SK	Teachers' formative assessment; Students' self- & peer-assessment	Project assessment; Collaboration/competition
AOCP	Teachers' formative & summative assessment; Students' peer- & self-assessment	Collaboration; Project assessment

5.2.4 Subject and transversal skills focus

In terms of subject focus, the formal education cases cover a variety of subjects (e.g. FME, Neo Alta, Teach4SK). All case studies address transversal and soft skills and their range is wide and diverse. Four practices target on digital competence (i.e. ATS2020, FME, Multipoly, AOCP) while certain skills such as critical thinking, reflection, analysis and communication skills appear to be addressed more often (Table 7).

5.2.5 Implementation process

The implementation process followed similar patterns across the examined initiatives. In general, practitioner-driven initiatives show a more gradual implementation process often supported by pilots, information days or workshops.

Table 7. Subject and transversal skills focus of examined cases.

Assessment practice	Subject	Transversal skills
ATS2020	Transversal/soft skills	Digital competence
e-APL	Healthcare; Elderly care	Communication, planning & organisation skills; Ergonomics
OSCE	Cross-subject; Behaviour; Attitude; Clinical skills; Soft/transversal skills	Critical thinking; Analysis; Reflection; Problem solving; Communication skills; Civic and social competences
FME	German; Geography; Philosophy; French; Social studies; Psychology; Second national language (Swedish, Finnish); Religion; Ethics; Health education; History	Critical thinking, analysis, reflection/problem-solving/digital competence
Multipoly	ICT; Technology; Computing	Business acumen; Digital competence; Relational skills
Neo Alta	Geography; Technology; English; Spanish; German; French; History; Sports; Visual Arts; Theatre	Critical thinking, analysis, reflection; Team work/collaboration; Learning to learn
Teach4SK	Basic literacy and reading comprehension	Civic and social competences; Critical thinking, analysis, reflection; Problem-solving; Communication, planning & organisation skills; Entrepreneurship
AOCP	Transversal/soft skills	Digital competence; Collaboration; Dynamics; Interaction

The training of trainers and teachers' trainings (face to face and online) were central to the implementation of ATS2020. In addition, information days were organised in order to obtain consensus forms from parents agreeing to their child's participation. In each learning cycle self-, peer- and teacher assessment were conducted and students' products were stored in e-portfolios. Overall, three learning cycles were implemented. Final conferences were organised for the exchange of experiences among schools.

The model created for the assessment of prior learning in Swedish elderly care was piloted with a group of elderly caregivers and workshops were organised with a view to refine the authenticity of the practical assessments. The practical part took place in a specifically designed apartment with cameras and caregivers had to simulate their skills on an adult sized doll or an actor. Apart from that, theoretical assessments were also applied through a computer-based test at their workplace. For the pilot's implementation, the good cooperation between the University and Adult Education Centres was considered crucial.

The founders of Teach for Slovakia designed a teacher training and leadership development programme, which integrated a diversity of assessment strategies. The selected participants were trained in a basecamp for six weeks on teaching and motivation techniques. For three weeks, teachers were supported by an assigned mentor who provided them with daily feedback on their teaching progress. Moreover, formative assessment focused on increasing students' engagement with the use of various techniques.

Neo Alta started as a teachers' initiative. Individual meetings were organised among students, the referent teacher and their parents during which individual progress goal were negotiated. Students self-assess their progress bi-weekly and then further discussed. Students could modify their work as many times as they wanted until they were satisfied with the result. Furthermore, interdisciplinary classes were organised with the collaboration of two or three teachers once a week. Initially, the project was applied on two classes and then extended to a third one with volunteered teachers.

Other initiatives had to deal with issues of technical infrastructure and its scaling up during the implementation phase. In the case of Sheffield Hallam University, senior staff initiated the use of video to demonstrate the acquisition of clinical skills for paramedics. The implementation process was dominated by the search for the right technical solution and the provision of recording devices and practice space. In addition, lecturers provided some exemplary video demonstrating the correct execution of skills and tutors were available to assist students to practice their skills. Students could also use their own devices to record the videos and then upload them on an online platform.

In the Finnish matriculation examination candidates upload an operating system from a USB drive in their own computers. They are connected to the examination system via a local network and get the examination material through a browser. After the examination, answers are sent to the MEB's service where they are marked and scored by the teachers' own school and then by the assessors of the MEB. Student's get the results through the system and receive the diplomas from their schools. Most importantly, schools are provided with a system called Abitti, which has the same tools and functions as the examination environment. Teachers can use it for the regular exams during the school year in order students to get used to the examination environment and the new types of questions. The digital exams are being implemented gradually. They started in 2016 with certain subjects and will be completed by 2019 when the last subject Mathematics is to be digitalised.

Amidst the practitioner-driven approaches, Multipoly, the implementation approach of the gamified assessment solution of PwC Hungary stands out Multipoly was developed as a recruitment tool and has been implemented rapidly. The participants have to take part on an online competition, which includes a virtual interview and a virtual traineeship, which last for a period of 12 game days. During that period, they take part to a series of virtual job activities where they get to know the company and explore a multitude of available career paths. The company gets the best candidates and invites them for a face to face interview. After 18 months, a new edition of the game was developed on the basis of previous experiences and feedback received from the Multipoly players and PwC.

The Open University of Catalonia designed an ICT Competences Course based on continuous e-assessment (the 360° assessment model). Students have an active role assessing all phases of its execution during the implementation of a digital project, in groups of four. It included self- and peer-assessment, in addition to the assessment of the group's overall dynamics and processes. Teachers have a facilitative role guiding and advising, providing separate feedback on individual and group work and furthermore, they can grade up or down students' final assessment. Since the setup of the online course, the course design has been considerably evolved and its content is more competence focused.

5.2.6 Challenges faced

Getting familiarised with new practices/time and energy to adapt

Across half of the examined assessment practices, the time and efforts to implement the new approaches and adapt previous practices proved to be challenging. In the case of ATS2020, most of the implementation partners saw the introduction of innovative practices in the educational setting as

relatively time consuming. Although attempts were made to comply with set time frames, some activities (e.g. those with active participation of students) at times required more time for teaching and learning. Furthermore, certain education systems faced more difficulties to adapt to a student-centered approach due to differences in term of curricula.

Adapting to new teaching practices was also an obstacle for the implementation of Neo Alta. Teachers pointed out to be in need of further training to change their teaching habits and make up for additional work due to frequent staff fluctuations. Similarly, some teachers in AT2020 had difficulties to understand a number of new and complex ideas. During the implementation of OSCE, training and support was given to few members of staff who were technologically less literate. Some participants faced difficulties with Multipoly and advocated that it should be simplified.

In a similar fashion, the Matriculation Board of the FME indicated that the greatest challenges were related to the change in educational culture that new digital tools and subsequent new working methods bring to schools. Moreover, the prospect of applying the new approaches in classes with vulnerable learners (Roma) created uncertainty for some of the participating teachers from the Teach4SK programme.

Technical/technological issues

A second challenge concerns technical/technological issues. For example, challenges occurred with the digital infrastructure of the AT2020 programme due to the limited bandwidth of certain educational organisations. The FME equally had to face various challenges of technical nature. Using an operating system from a USB drive has proven more unstable than first foreseen. As a result, spare laptops and additional USB drives were given to each school.

The coordinators of AOCP also emphasised current technological limitations of the assessment tool. The online environment did not provide standardised information; therefore, there was a need for an effective learning analytics system, which collects systematic evidence of student activity. However, solutions are in development and it is likely that in the near future there will be a tool fulfilling the aforementioned criteria.

Lack of digital skills and language barriers

In the Swedish e-APL some caregivers had difficulties understanding the Swedish instructions, others felt insecure about their computer skills to take the e-assessments. Besides nervousness, some employees also feared that low results of the assessments would affect their employment, which was not the case.

Staff commitment

Another interesting challenge resulted from Neo Alta's volunteering culture. Since the initiative relied on committed teachers who participated on a volunteering basis, it was inevitable that success was depended on the engagement of the particular teachers. When school leadership and staff changed, there was a decline of Neo Alta's learning culture.

5.2.7 Success factors and lessons learned

Teacher cooperation/exchange of practices

Clearly, the most prominent success factor detected across the examined cases is related to the effective cooperation and exchange of practices of those involved. In the case of ATS2020, the cooperation of teachers happened at multiple levels. Teachers co-designed and taught in the classroom and exchanged experiences virtually as well as face-to-face at school, national and EU-level. Unlike the occasional collaboration among teachers happening at most schools, the exchange of teaching practices and experiences is also part of Neo Alta's daily practice.

The coordinators of AOCP equally emphasised the key role of educators. In their view, rather than financial resources, the online course relies on the collaboration and fruitful exchange of a team of engaged online teachers. Without the support of a team of professionals, the online course would have not been able to evolve as extensively.

Teachers' professional development

Besides the cooperation of programme participants through regular workshops and training, the detailed planning and methodical teacher training was seen as the main success factor of Teach4SK. The trainings assisted teachers to become familiarised with assessment strategies, to practice related procedures and eventually master those strategies. The structured and regular trainings, the systematic feedback from mentors and peers facilitated the consolidation of their skills.

In-service training of educators and further support from educational authorities (face to face/online) seems an indispensable component of many programmes (e.g. ATS2020, FME, Neo Alta).

User-engagement and ownership

In the case of ATS2020 teachers' enthusiasm on the project's goal and methodology was a major success factor as there was a consensus that transversal skills should be fostered in schools. Furthermore, students were proud to present their e-portfolios and share their learning experiences.

Across two assessment practices (i.e. e-APL, AOCP) the active involvement of learners in the assessments is highlighted as a core driver. Concerning e-APL, elderly care professionals were involved in piloting the assessments, giving them the opportunity to provide feedback and shape the design of the assessments. The fact that users were able to participate in this process increased their motivation making them feel that their opinion is valuable. Meanwhile, students' active role in the continuous assessments assisted to the evolvement of such processes in AOCP. The fact that students are in the centre of the assessment practice participating at several stages was very motivating for them.

Technical solutions/digital technologies

The examples of the FME and ATS2020 show how the technical solutions and technology integration in classrooms can be a success factor.

In Finland, an environment called Abitti was released for training and practicing purposes. The Upper Secondary Schools can use it throughout the school year to organise regular exams and prepare students for the Matriculation Examination. Its adoption by learners and teachers was crucial in enabling

candidates, schools and the MEB to properly prepare for the digital examination. It also managed to establish a good feedback loop between the system and the MEB to ensure that necessary support measures were put in place.

Meanwhile, the embedding of technologies in the ATS2020 project was an additional motivating factor, for students and teachers alike. Although the integration of technology in classrooms is already a reality in most classrooms, teachers greatly welcomed the support provided by tech-savvy coaches and peer-educators provided through the project.

The key success factor of Multipoly lied in the uniqueness of the game (virtual working experience of a company). In fact, before the introduction of the game by PwC Hungary, no other comparable tool was used for recruitment purposes. Moreover, the implementation of Multipoly increased transparency of the recruitment process and facilitated candidates' integration into the team.

In the case of AOCP, the course design in virtual environments was based on in-depth academic research, which apart from technological progress took into consideration users' online experiences contributing to the success of the project.

Further success factors

In the case of e-APL the interest and **financial support** from the local authorities were decisive for the implementation of the project. Furthermore, the project offered them increased qualifications in a short time frame and the prospect of higher wages became a strong motivating factor for the caregivers.

With regard to OSCE, the new assessment scheme include certain benefits which assisted to a successful outcome: (1) A more **efficient** and genuine internal and external moderation; (2) The opportunity to store videos as evidence for skill acquisition for hospitals and (3) Remote marking and reduced travel time for commuting students, which served as a main driver for both them and teachers.

Another key success factor is the **quality of relationship** between stakeholders. The success of Neo Alta lied in the relations of trust established between children and adults. Similarly, the high level of trust between vocational educational institutions and elderly care providers in e-APL contributed to its successful implementation. Transparency and collaboration between stakeholders in creating a common target image are additional success factors.

5.2.8 Achieved results and impacts

The results of the assessment practices are important indicators of their success. However, the detail to which the quantitative and qualitative results are analysed varies greatly across initiatives making comparisons significantly difficult. Large-scale policies such as the FME and European projects such as ATS2020 are obliged to set up more comprehensive monitoring and evaluation systems, while practitioner-driven initiatives hardly collect any quantitative data. Moreover, the extent to which digital solutions are incorporated in the assessments equally impacts the availability of results data (e.g. AOCP, Multipoly). The following section presents the results in more details.

Comprehensive data collection

As a major education policy reform in Finland, the FME carried out a substantial monitoring and evaluation of results. To this point, the goals set up in 2011 for the digitisation of the FME have been

met, while the process is on schedule and budget. By spring 2018, 106,105 digital test submissions have been completed by candidates of the FME. All submissions have been aptly submitted, collected and graded.

In a similarly comprehensive fashion, the large-scale experimentation of ATS2020 was designed to be evaluated both quantitatively and qualitatively. Thus, a pre and post evaluation took place while experimental and control groups were used for comparison. At the same time, 2 case studies in each country were also conducted to further support and analyse the results. Due to the short period of implementation, the quantitative data analysis did not point out significant outcomes. Yet, the data collected from the case studies showed that students had acquired a higher level in collaboration, communication and information literacy skills while throughout the project students' and teachers' artefacts showed considerable progress.

Digital tools facilitating the monitoring of results

The creators of the e-assessment of prior knowledge in elderly care mention low results as none of the participants passed all components of assessment and more than 50% failed to pass the theoretical part. The poor outcomes are probably due to the difference of vocational education standards and practical on the job training. However, 80% of caregivers either fully or partially approved the practical assessment in the simulated apartment and all staff members valued the use of the IT tool for the theoretical assessment. Interestingly, 83% of the participants were willing to continue the assessment. Furthermore, the assessment model was flexible and allowed for significant time savings compared to traditional validation.

The OSCE examination based on video recording has improved students' success rate while appeals against assessment decision were reduced. Although, students spend more time on practising skills prior to submission, 90% of them state that they prefer the new process over the previous one.

Meanwhile, the online monitoring tools of the AOCP combined with satisfaction surveys provide a comprehensive dataset on students' performance and perception/satisfaction concerning the continuous assessment model. The data from the online surveys shows that the majority of students (85%) evaluate the assessment model positively. Despite being more time consuming, surveyed students consider that self- and peer-assessment had considerable impact on the promotion of their critical ability as well the monitoring, guidance and feedback provided by teachers.

Since the creation of the game, the impact on PwC's talent recruitment was in-depth assessed providing quantitative and qualitative results. In terms of quantitative results, the job candidate pool grew by 190% while 85% indicated that the game affected their view of the company positively. Moreover, the experience of participating to the game facilitated their transition to full-time employees, since they were more familiar to the company culture compared to other candidates.

Practitioner-driven cases focusing on qualitative results

Neo Alta and Teach4SK belong to the practices, which provided only qualitative results. The outcomes rely on students' and teachers' feedback gathered over the last years. In the case of Neo Alta, parents and students confirmed that the programme had facilitated students' learning status and progress. The low drop out proves that students are satisfied and see advantages in it. With regard to Teach4SK, teachers shared the view that the programme has been successful since students have improved their literacy and numeracy level by 11% on average over a school year. Even after completing the

programme, they continue using the assessment techniques in order to engage students and increase the ownership of their learning process.

5.2.9 Potential for scale up and replication

The majority of case studies have either already been replicated or have received expressions of interest from other organisations. The transferability potential of the initiatives can serve as an indication of how easily organisations in other countries/regions can replicate them The main criteria used for the replicability assessments are: (1) Whether the assessment has already been subject to replication including potential ones; (2) The financial cost of setting up the schemes; (3) The complexity of the assessment and (4) The required non-financial resources (e.g. needed efforts).

As indicated in Table 8, the replicability potential is of medium level on average. The main obstacles hindering high replicability lie in relatively high financial investments or human resource efforts required to set up the assessment schemes.

ATS2020 project was designed with a view that schools beyond the piloting ones would implement the particular model. For this reason, a toolkit was developed, a step-by-step implementation guide, supporting both them and the teachers. The toolkit explains in detail how to introduce the learning scenarios along with other tools needed for the assessment. It has been mentioned that teachers continue to apply the model in Cyprus and other countries. Furthermore, some countries contacted the project leaders to express an interest on implementing the model at national level.

E-APL has been temporary stand still, yet a revised model in line with the current national requirements for validation and the Swedish SeQF is being prepared for personnel working with disable people. The creators maintain that the interactive assessment model has great potential for the validation of prior learning and can been replicated in various professional domains.

The video assessment method of OSCE was shared by the faculty with other universities. As a result, some of them are now using the same system for the assessment of paramedic skills. Moreover, it has been used in completely different domains. For example, a university in Australia is using it for their pilots training, yet in combination with action cameras. The cost of resources (e.g. faculty recording devices, practice facilities) is not high, which facilitates its replication.

In the case of FME, the technical environment for examination setup in schools requires significant financial investments. On the other hand, resources are saved since students can use their laptops to take the ME. Indeed, the technical requirements can be fulfilled even by relatively old computers and no purchase of licenses or software is needed by students or schools.

After five years of implementation, Multipoly is currently not in use, as the company needs to update its branding strategy. However, there are plans to replicate Mutipoly at regional level.

The case of Neo Alta shows how required human resources and voluntary engagement can play a role in reducing the scheme's replicability potential. Although financial resources are not needed to operate the programme, the experiment relies significantly on the voluntary engagement of teachers and parents. Without their engagement, replication seems difficult.

Meanwhile, Teach4SK is committed to gradually spread the assessment practices at classroom level to other schools. This process is promoted by the programme's alumni during which principals are encouraged to use more varied assessment practices. Instead of a school level approach, the hitherto strategy is to keep on gradually expanding the number of classrooms using alternative assessment approaches. However, teacher training is extensive throughout the programme in terms of workshops as

well as personal mentoring and coaching. Therefore, the teacher training of Teach4SK is equally regarded as the main obstacle to a higher replicability potential. The programme could be run with relatively low financial resources including savings on the whiteboards, which could be replaced, by laminated carton as a low cost solution. Apart from that, the programme has high potential to be replicated in other countries, in particular to countries with unequal access to education.

Table 8. Assessment of initiatives' replicability potential.

Assessment practice	Replication so far	Financial cost	Complexity	Human resources	Replicability potential
ATS2020	No	Medium	Medium	Yes	Medium
e-APL	Possible	Medium	Medium	Yes	High
OSCE	Yes	Medium	Low	Yes	High
FME	No	Medium/ High	Medium	Yes	Medium
Multipoly	No	Medium	Medium	No	Medium
Neo Alta	No	Low/ Medium	Medium	Yes	Medium
Teach4SK	Yes	Medium	Medium	Yes	Medium
AOCP	Yes	Medium	Medium	Yes	Medium

The continuous assessment model of AOCP is potentially suitable for online and face-to-face courses and for the assessment of other competences. In fact, some online university teachers have used the assessment methodology for their face-to-face courses in other institutions. Rather than the financing, the initiative's founders see challenges in terms of needed human resources and teacher collaboration to set up the 360° assessment model in other institutions.

5.3 Insights from the case studies

This section provides insights into the eight cases of innovative assessment practices, which have been obtained by the involved actors. In particular, the interviews illuminate the main drivers and challenges and contribute to a better knowledge of their potential for replication.

The main findings show that despite different backgrounds (e.g. job-based assessment vs. formal education) and different education levels, the examined cases display various similarities. On a broader scale, the practices were primarily initiated in the context of assessment quality concerns as well as inefficiencies and cost savings. The case of Teach4SK, however, is an exception being launched for political reasons, i.e. inequality in access to education.

Furthermore, the examined assessment practices implement diverse assessment methods with an average of three to four methods applied per initiative. A variety of assessment formats such as simulation, portfolio and project assessment co-exist alongside the traditional quiz/multiple choice format.

Rather than in particular assessment methods or formats, innovation in educational assessment lies in

the particular context and complex processes tied into pedagogical components. In particular, students' participation through self- and peer assessment stands out as a mean to improve the awareness of their learning progress.

The subject and transversal skills focus illustrate the diversity of the examined practices. While subjects vary considerably, transversal skills take on a prominent role in most cases. Transversal skills such as digital skills, critical thinking, communication and reflection are recurring across most initiatives.

Practitioner-driven schemes follow a more gradual implementation approach being often supported by pilots, information days and/or workshops. In general terms, the development process of the schemes are dominated by teacher training and optimising the provided technical solutions.

Despite the apparent differences of the cases, the challenges and success factors faced during the set-up of the schemes show considerable overlaps. The identified shared challenges primarily concern the adjustments of teachers to new teaching and assessment practices as well as new digital tools. A second challenge was related to technical and technological issues; more precisely, in terms of lacking learning analytics tools which would ease the implementation and the higher instability of computer operating systems.

Meanwhile, the shared success factors supporting the implementation of several initiatives can be structured into three core groups. The first group concerns the co-operation of teachers and corresponding exchange of good practices Secondly, the participative, user-centred approach through self-, peer-assessment and simulation acted as a main driver by giving students/users a stronger role and increased their motivation. Thirdly, the technical solution/product as well as technology integration in the classroom equally succeeded to increase students' motivation.

A look at the achieved results of the examined initiatives shows different levels of available results data. While policies and EU projects are often obliged to set up rigorous monitoring and evaluation systems, those initiatives with a strong practitioner-driven character have not collected quantitative data systematically. Contrary, the data collection process is significantly facilitated for those practices with a stronger digital component (e.g. AOCP, Multipoly). The provision of quantitative results helps consolidate the credibility of outcomes. Finally, the examined initiatives display on average a medium level replicability potential.

6. Conclusion and recommendations

The following sections display the key conclusions of the study and makes recommendations based on the empirical data and the case studies being reviewed.

6.1. Conclusion

One of the main objectives of this study was to define innovative assessment, elucidate how it can successfully be implemented with the purpose to enhance learning and support the development of key competencies. Apart from a synthesis of research evidence on the effectiveness of a variety of approaches, the study presents eight case studies that integrated innovative assessment, highlighting the challenges and success factors of such schemes.

Innovative assessment refers to a form of assessment that is novel, flexible and adaptable to different

contexts and approaches, which integrate a variety of methods, and techniques (e.g. new, contemporary, traditional) while giving enough space to learners to engage in reflective practices and actively accommodate their learning needs. Digital technology can increase the validity and reliability of assessment practices, yet decisions should be based on pedagogy. If digital tools are to used, they should be tailored to the needs of educators and learners. Moreover, innovative assessment needs to be sustainable and prepare learners for their future needs. It should encourage deep learning through the development of certain disposition for practice such as reflection and evaluation, developing assessment criteria, engage in peer-assessment, working with peers and use of transversal skills.

Self- and peer-assessment are essential components of innovative assessment and valuable means of empowerment. Learners gain ownership of their learning, their motivation is increased and they become self-sustained learners. Such processes cause a redistribution of power, as learners assume the responsibility of their own learning by assessing themselves and others. Research outcomes highlight the positive effect of self- and-peer-assessment on learning and achievement, on learners' motivation and engagement, on their self-efficacy and use of self-regulated strategies and on the quality of student-teacher relationship, which have also been observed in most of the case studies.

Digital tools however contain certain qualities, which facilitate both processes of formative and summative assessment and the development of key competencies. Simulations are ideal instruments for situated learning, training in decision making within complex and dynamic situations and transferring of knowledge in the workplace as shown in the Multipoly case study. Digital badges can recognize diverse learning trajectories and competences from formal, non-formal and informal education and become credentials of learning. Learning analytics can optimize learning as Learning Management Systems are dynamic and can support reflection, adaptation, personalization and recommendation according to current learning state. However, there is still lack of knowledge on how learning analytic tools can improve learning and teaching although there is agreement for their great potential. Moreover, there is considerable lack of evidence on the effectiveness of learning analytic tools and open badges in accommodating users' needs.

A number of success factors were reported by the innovative assessment case studies. The cooperation of teachers and exchange of good practices, the active role of the learners during self/peer-assessment and simulation and technology integration in the classroom equally succeeded to increase learners' motivation and engagement. Additional enablers are the commitment of the involved participants, a supportive school leadership and high quality professional training for teachers. On the other hand, participants need time and energy to adapt and be familiarized with the new practices while the lack of confidence in the digital skills of some participants is an additional barrier. Finally, some of the examined innovative assessment schemes have been subject to successful replication. The obstacles that reduce the potential replicability are mostly related to teacher training. Allocation of more time for teachers training and development activities besides their teaching can boost the replicability of assessment schemes.

6.2. Recommendations

The overview of innovative assessment practices presented in this report illustrate their potential as well as the challenges to be encountered. These challenges indicate that many more steps have to be followed for the effective implementation of innovative assessment at all education levels, as illustrated below.

Educational policy and governance practice

- Assessment practices determine the priorities of an education system. The rationale of innovative pedagogy for the 21st century should be disseminated to all stakeholders and provide a framework for the introduction of a comprehensive assessment approach with an emphasis on formative assessment. Stakeholders should realise the necessity of such a reform for innovation to occur.
- The micro-, meso-, macro- level of an educational system should be aligned to develop clear goals and reference points to guide innovative assessment. Empirical evidence should inform the effectiveness of various assessment methods/formats within specific contexts, illuminate their strengths and limitations and guide accordingly educational policy and practice.
- Constant evaluation of assessment systems (meta-assessment) is also needed and amendments for improvement should ensure the effectiveness of innovative assessment practices.
- Policies should aim to increase awareness of stakeholders of the potential of innovative tools to motivate, direct and recognize authentic learning.

Institutional leadership and governance practice

- Educational organisations should be provided with comprehensive guidance and support when implementing an innovative approach in order to blend/integrate methods and benefit from various approaches. A structured implementation of increased difficulty can maximize the benefits of an innovative assessment framework.
- Leaders should encourage experimentation as well as bottom-up approaches during the implementation of innovative assessment procedures.
- Performance criteria in relation to national curriculum goals should be better clarified and illustrated when using innovative assessment methods.

Collaboration and networking

- The cost of designing digital tools (e.g. simulations) can be high. Governments, research centres, companies from the private sector can collaborate in designing such tools. Learners can have an active role and propose innovative ideas that meet their own learning needs and therefore, improve the effectiveness of such tools.
- Networks of instructors and learners for the exchange of experiences and good practices of innovative assessment should be encouraged.

Teaching and learning practices

- Assessment should become an active part of the learning process with emphasis on self- and peer-assessment.
- Self-reflection is a central competence and should be an integral part of self-assessment. Learners should be given time and space to practice self-reflection, as it is an essential life-long learning skill. Models of effective reflective practices and debriefing should be implemented.

- Digital assessment practices should provide real time feedback and measurement of learning and skills over time, adapted to learners' progress and their individual needs.

Capacity building

- Teachers' initial training should be modified in such a way to facilitate innovative assessment practises.
- Professional development of educators/instructors should become a top priority with the aim to develop high value belief on innovative assessment. It can also address a wide range of objectives such as becoming competent users of digital tools, knowing the benefits and barriers of new learning environments, how and when to use different assessment formats, how to provide constructive feedback on task performance and how to scaffold the development of learners' self- and peer-assessment.

Infrastructure

- Priority should be given to the development of authentic learning environments and open source digital tools adaptable to educators' and learners' needs.
- E-assessment toolkits in Europe are rather fragmented. Systematic recording and coordination of the available e-assessment tools would be beneficial. An open access platform of such tools with evidence about their effectiveness can facilitate their dissemination to learners and instructors.

Research

- Research is needed into the principles that ensure quality and the effectiveness of innovative assessment practices.
- Collaboration among universities, professional organisations and institutions should be encouraged for the development of innovative assessment tools (e.g. digital badges, simulations) providing evidence of informal learning while increasing the weight and importance of such credentials.

6.3. Concluding remarks

This study examined a variety of innovative assessment practices however, this report should not be considered as a 'guide' to innovative assessment implementation. Rather it should be perceived as a starting point of what assessment possibilities exist and how their use can trigger change in various educational settings. The information draws a picture of the field and can be useful to various stakeholders as it suggests areas where interventions can prompt the implementation of innovative assessment. Finally, the recommendations concern all educational levels and call for action of the following dimensions:

- Policy makers should promote investment on initial training and continuous professional development of teachers with reference to innovative assessment.
- A minimum of consent among parties should be achieved by policy makers for a strategic approach to be followed when implementing innovation in assessment.

- Policy should increase efforts and invest in infrastructure, encourage networks of practitioners and bottom-up implementation efforts.
- Research should provide evidence of effective methods and techniques applicable to various contexts and support both policy and practice.

The recommendations presented in this report might provide an incentive for the modernisation of assessment systems and stimulate further research, policy and practice in the field.

References

Abramovich, S. (2016). Understanding digital badges in Higher Education through assessment. *On the Horizon, 24*(1), 126-131. https://doi.org/10.1108/OTH-08-2015-0044

Abramovich, S., Schunn, C., & Higashi, R.M. (2013). Are badges useful in education?: It depends upon the type of badge and expertise of learner. *Educational Technology Research and Development, 61*(2), 217-232. https://doi.org/10.1007/s11423-013-9289-2

Alonso-Tapia, J., & Pardo, A. (2006). Assessment of learning environment motivational quality from the point of view of Secondary and High School learners. *Learning and Instruction, 16*(4), 295-309. https://doi.org/10.1016/j.learninstruc.2006.07.002

Anderson, A., Huttenlocher, D., Kleinberg, J., & Leskovec, J. (2013, May 13-17). Steering user behaviour with badges. *Proceedings of World Wide Web Conference, Rio de Janeiro, Brazil.* Available at http://cs.stanford.edu/people/jure/pubs/badges-www13.pdf

Anderson, D.M., & Staub, S. (2015). Postgraduate digital badges in Higher Education: Transforming advanced programs using authentic online instruction and assessment to meet the demands of a global marketplace. *Procedia - Social and Behavioral Sciences, 195*(3), 18-23. https://doi.org/10.1016/j.sbspro.2015.06.165

Andrade, H.L., Du, Y., & Mycek, K. (2010). Rubric-referenced self-assessment and Middle School students' writing. *Assessment in Education: Principles, Policy & Practice, 17*(2), 199-214. https://doi.org/10.1080/09695941003696172

Andrade, H.L., Wang, X., Du, Y., & Akawi, R.L. (2009). Rubric-referenced self-assessment and self-efficacy for writing. *The Journal of Educational Research, 102*(4), 287-301. https://doi.org/10.3200/JOER.102.4.287-302

Arias Aranda, D., Haro Domiguez, C., & Romerosa Martinez, Mª.M. (2010). An innovative approach to the learning process in management: The use of simulators in Higher Education. *Revista de Educación, 353,* 707-721. http://www.revistaeducacion.mec.es/re353/re353_27.pdf

Arnold, K. (2010, March 3). *Signals: Applying academic analytics.* Available at https://er.educause.edu/articles/2010/3/signals-applying-academic-analytics

Arnold, K.E., & Pistilly, M.D. (2012). Course signals at Purdue: Using learning analytics to increase student success. *Proceedings of the International Conference on Learning Analytics and Knowledge* (pp. 267-270). New York, NY: ACM.

Babik, D., Gehringer, E.F., Kidd, J., Pramudianto, F., & Tinapple, D. (2016). Probing the landscape: Toward a systematic taxonomy of online peer assessment systems in education. *Teaching & Learning Faculty Publications, 22*. https://digitalcommons.odu.edu/teachinglearning_fac_pubs/22/

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 & Education, 55*(2), 846-857. https://doi.org/10.1016/j.compedu.2010.03.016

Béguin, A. (2016). Education can benefit from flexible and non-flexible assessment. In J., Vlasblom (Ed.), *Thematic issue innovations in digital assessment. Assessment as an integral part of the learning process* (pp. 6-7). SURFnet. Available at www.surf.nl/surfnet

Bevitt, S. (2015). Assessment innovation and student experience: A new assessment challenge and call for a multi-perspective approach to assessment research. *Assessment & Evaluation in Higher Education, 40*(1), 103-119. https://doi.org/10.1080/02602938.2014.890170

Biles, M.L., Plass, J.L., & Homer, B.D. (2014). *Good badges, evil badges? An empirical inquiry into the impact of digital badge design on goal orientation and learning. Report on 2013-2014 HASTAC Digital Media and Learning Research Grant Competition.* Available at https://bit.ly/2SPxp1l

Black, P., & Wiliam, D. (1998). Assessment and classroom learning. *Assessment in education: Principles, Policy & Practice, 5*(1), 7–74. https://doi.org/10.1080/0969595980050102

Black, P., Harrison, C., Lee, C., Marshall, B., & Wiliam, D. (2003). *Assessment for learning: Putting it into practice.* Maidenhead, Berkshire: Open University Press.

Boud, D. (1995). Assessment and learning: Contradictory or complementary? In P. Knight (Ed.), *Assessment for learning in Higher Education* (pp. 35-48). London: Kogan Page. Available at https://bit.ly/2Tw13c9

Boud, D., & Associates (2010). *Assessment 2020: Seven propositions for assessment reform in Higher Education.* Sydney: Australian Learning and Teaching Council. Available at http://bit.ly/2ZO6dmC

Boud, D., & Falchikov, N. (1989). Quantitative studies of student self-assessment in Higher-Education: A critical analysis of findings. *Higher Education, 18*(5), 529-549. https://doi.org/10.1007/BF00138746

Boud, D., & Soler, R. (2016). Sustainable assessment revisited. *Assessment and Evaluation in Higher Education, 41*(3), 400-413. https://doi.org/10.1080/02602938.2015.1018133

Boud, D., Lawson, R., & Thompson, D.G. (2013). Does student engagement in self-assessment calibrate their judgement over time? *Assessment & Evaluation in Higher Education, 38*(8), 941-956. https://doi.org/10.1080/02602938.2013.769198

Bourke, R. (2014). Self-assessment in professional programmes within tertiary institutions. *Teaching in Higher Education, 19*(8), 908-918. https://doi.org/10.1080/13562517.2014.934353

Brauer, S., & Siklander, P. (2017, December 4-6). Competence-based assessment and digital badging as guidance in vocational teacher education. Paper presented at *5th International Conference on Educational Technologies 2017 (ICEduTech 2017).* Available at https://bit.ly/2FCfFnb

Brečko, B.N., Kampylis, P., & Punie, Y. (2014). *Mainstreaming ICT - enabled Innovation in Education and Training in Europe: Policy actions for sustainability, scalability and impact at system level.* Luxembourg: Publications Office of the European Union. http://bit.ly/2vs7ylg

Broadfoot, P., Timmis, S., Payton, S., Oldfield, A., & Sutherland, R. (2012/2013). *Integrating the formative and summative through technology enhanced assessment* (Discussion paper 2). University of Bristol, Graduate School of Education. Available at https://bit.ly/2OZdSrY

Brown, G.T.L., & Harris, L.R. (2014). The future of self-assessment in classroom practice: Reframing self-assessment as a core competency. *Frontline Learning Research, 2*(1), 22-30. https://doi.org/10.14786/flr.v2i1.24

Brydges, R., Hatala, R., Zendejas, B., Erwin, P.J., & Cook, D.A. (2015). Linking simulation-based educational assessments and patient-related outcomes: A systematic review and meta-analysis. *Academic Medicine, 90*(2), 246-256. doi: 10.1097/ACM.0000000000000549

Buchanan, T. (2001). The efficacy of a world-wide web mediated formative assessment. *Journal of Computer Assisted Learning, 16*(3), 193–200. https://doi.org/10.1046/j.1365-2729.2000.00132.x

Carey, K.L., & Stefaniak, J.E. (2018). An exploration of the utility of digital badging in Higher Education settings. *Educational Technology Research and Development, 66*(5), 1211-1229. https://doi.org/10.1007/s11423-018-9602-1

Casilli, C., & Hickey, D. (2016). Transcending conventional credentialing and assessment paradigms with information-rich digital badges. *The Information Society, 32*(2), 117-129. http://dx.doi.org/10.1080/01972243.2016.1130500

Castillo-Merino, D., & Serradell-Lopez, E. (2014). An analysis of the determinants of students' performance in e-learning. *Computers in Human Behaviour, 30*, 476-484. https://doi.org/10.1016/j.chb.2013.06.020

Chang, Y.C., Peng, H.Y., & Chao, H.C. (2010). Examining the effects of learning motivation and of course design in an instructional simulation game. *Interactive Learning Environments, 18*(4), 319-339. https://doi.org/10.1080/10494820802574270

Cheng, A., Eppich, W., Grant, V., Sherbino, J., Zendejas, B., & Cook, D.A. (2014). Debriefing for technology-enhanced simulation: A systematic review and meta-analysis. *Medical Education in Review, 48,* 657-666. https://doi.org/10.1111/medu.12432

Cheng, A., Grant, V., Dieckmann, P., Arora, S., Robinson, T., & Eppich, W. (2015). Faculty development for simulation programs: Five issues for the future of debriefing training. *Simulation in Healthcare, 10*(4), 217-222. 10.1097/SIH.0000000000000090

Cheng, Z., Watson, S.L., & Newby, T.J. (2018). Goal setting and open digital badges in Higher Education. *TechTrends, 62*(2), 190-196. https://doi.org/10.1007/s11528-018-0249-x

Chiappe, A., Pinto, R., & Arias, V. (2016). Open assessment of learning: A meta-synthesis. *International Review of Research in Open and Distributed Learning, 17*(6), 44-61. http://www.irrodl.org/index.php/irrodl/article/view/2846/3965

Cho, K., & MacArthur, C. (2011). Learning by reviewing. *Journal of Educational Psychology, 103*(1), 73-84. http://dx.doi.org/10.1037/a0021950

Cho, Y.H., & Cho, K. (2011). Peer reviewers learn from giving comments. *Instructional Science, 39*(5), 629-643. https://doi.org/10.1007/s11251-010-9146-1

Chou, C.C., & He, S.-J. (2017). The effectiveness of digital badges on student online contributions. *Journal of Educational Computing Research, 54*(8) 1092-1116. https://doi.org/10.1177/0735633116649374

Chua, Y.P. (2012). Effects of computer-based testing on test performance and testing motivation. *Computers in Human Behaviour, 28*(5), 1580-1586. https://doi.org/10.1016/j.chb.2012.03.020

Chua, Y.P., & Don, Z.M. (2013). Effects of computer-based educational achievement test on test performance and test takers' motivation. *Computers in Human Behaviour, 29*(5), 1889-1895. https://doi.org/10.1016/j.chb.2013.03.008

Collimore, L.-M., Paré, D.E., & Joordens, S. (2015). SWDYT: So What Do You Think? Canadian students' attitudes about peerScholar, an online peer-assessment tool. *Learning Environments Research, 18*(1), 33-45. https://doi.org/10.1007/s10984-014-9170-1

Colthart, I., Bagnall, G., Evans, A., Allbutt, H., Haig, A., Illing, J., & McKinstry, B. (2008). The effectiveness of self-assessment on the identification of learning needs, learner activity, and impact on clinical practice: BEME Guide No. 10. *Medical Teacher, 30*(2), 124-145. https://doi.org/10.1080/01421590701881699

Cook, D.A., Brydges, R., Hamstra, S.J., Zendejas, B., Szostek, J.H., Wang, A.T., Erwin, P.J., & Hatala, R. (2012).

Comparative effectiveness of technology-enhanced simulation versus other instructional methods: A systematic review and meta-analysis simulation in healthcare. *Simulation in Healthcare: The Journal of the Society for Simulation in Healthcare, 7*(5), 308-320. doi: 10.1097/SIH.0b013e3182614f95

Cook, D.A., Hamstra, S.J., Brydges, R., Zendejas, B., Szostek, J.H., Wang, A.T., Erwin, P.J., & Hatala, R. (2013). Comparative effectiveness of instructional design features in simulation-based education: Systematic review and meta-analysis. *Medical Teacher, 35*(1), e867-e898. https://doi.org/10.3109/0142159X.2012.714886

Cook, D.A., Hatala, R., Brydges, R., Zendejas, B., Szostek, J.H., Wang, A.T., Erwin, P.J., & Hamstra, S.J. (2011). Technology enhanced simulation for health professions education: A systematic review and meta-analysis. *JAMA, 306*(9), 978-988. doi:10.1001/jama.2011.1234

Crocco, F., Offenholley, K., & Hernandez, C. (2016). A proof-of-concept study of game-based learning in Higher Education. *Simulation & Gaming, 47*(4), 403-422. https://doi.org/10.1177/1046878116632484

Crossouard, B. (2011). Using formative assessment to support complex learning in conditions of social adversity. *Assessment in Education: Principles Policy & Practice, 18*(1), 59-72. https://doi.org/10.1080/0969594X.2011.536034

Ćukušić, M., Garaća, Z., & Jadrić, M. (2014). Online self-assessment and students' success in Higher Education institutions. *Computers & Education, 72*, 100-109. http://dx.doi.org/10.1016/j.compedu.2013.10.018

da Rocha Seixas, L., Gomes, A.S., & de Melo Filho, I.J. (2016). Effectiveness of gamification in the engagement of students. *Computers in Human Behavior, 58,* 48-63. https://doi.org/10.1016/j.chb.2015.11.021

Davies, P. (2004). Don't write, just mark: The validity of assessing student ability via their computerized peer-marking of an essay rather than their creation of an essay. *Research in Learning Technology, 12*(3), 261-277. https://doi.org/10.1080/0968776042000259573

Dawson, S., Gašević, D., Siemens, G., & Joksimovic, S. (2014). Current state and future trends: A citation network analysis of the learning analytics field. *LAK '14: Proceedings of the 4th international conference on Learning Analytics and Knowledge.* Available at https://dl.acm.org/citation.cfm?id=2567585

Decker, S., Fey, M., Sideras, S., Caballero, S., Rockstraw, L.(R.), Boese, T., ... Borum, J.C. (2013). Standards of best practice: Simulation standard VI: The debriefing process. *Clinical Simulation in Nursing, 9*(6), S26-S29. https://doi.org/10.1016/j.ecns.2013.04.008

Dega, B.G. (2019). Cognitive diagnostic assessment of students' responses: An example from energy and momentum concepts. *European Journal of Physics Education, 10*(1), 13-23. https://doi.org/10.20308/ejpe.v10i1.219

Denny, P. (2013, April 27 - May 2). The effect of virtual achievements on student engagement. *Proceedings of the SIGCHI conference on human factors in computing systems* (pp.763-772). Paris, New York: ACM Press. Available at https://bit.ly/2YwfulC

Desjarlais, M., & Smith, P. (2011). A comparative analysis of reflection and self-assessment. *International Journal of Process Education*, *3*(1), 3-18. http://www.ijpe.online/archive.html#simple7

Devedžić, V., & Jovanović, J. (2015). Developing Open Badges: A comprehensive approach. *Educational Technology Research and Development, 63*(4), 603-620. https://doi.org/10.1007/s11423-015-9388-3

Domínguez, A., Saenz-de-Navarrete, J., de-Marcos, L., Fernández-Sanz, L., Pagés, C., & Martínez-Herráiz, J.J. (2013). Gamifying learning experiences: Practical implications and outcomes. *Computers and*

Education, 63, 380-392. https://doi.org/10.1016/j.compedu.2012.12.020

Dunning, D., Heath, C., & Suls, J.M. (2004). Flawed self-assessment: Implications for health, education, and the workplace. *Psychological Science in the Public Interest, 5*(3), 69-106. https://doi.org/10.1111/j.1529-1006.2004.00018.x

Eaglen Bertrando, S.L. (2017). *Rethinking workplace learning in the digital world: A case study of open badges.* D.Ed. Thesis, Pepperdine University, Graduate School of Education and Psychology. Available at https://eric.ed.gov/?id=ED578496

Eggen, T. (2016). Adaptive assessment proven to measure better. In J., Vlasblom (Ed.), *Thematic issue innovations in digital assessment. Assessment as an integral part of the learning process* (pp. 10-11). SURFnet. Available at www.surf.nl/surfnet

Ellis, L.E., Nunn, S.G., & Avella, J.T. (2016). Digital badges and micro-credentials: Historical overview, motivational aspects, issues, and challenges. In D. Ifenthaler, N. Bellin-Mularski, & D.-K. Mah (Eds.), *Foundation of digital badges and micro-credentials. Demonstrating and recognizing knowledge and competencies* (3-21). Switzerland: Springer International Publishing. https://doi.org/10.1007/978-3-319-15425-1_1

Eppich, W., & Cheng, A. (2015). Promoting Excellence and Reflective Learning in Simulation (PEARLS). Development and rationale for a blended approach to health care simulation debriefing. *Simulation in Healthcare, 10*(2), 106-115. http://bit.ly/2J2U9rT

Eppich, W.J., Hunt, E.A., Duval-Arnould, J.M., Siddall, V.J., & Cheng, A. (2015). Structuring feedback and debriefing to achieve mastery learning goals. *Academic Medicine, 90*(11), 1501-1508. doi: 10.1097/ACM.0000000000000934

Ertmer, P.A., Ottenbreit-Leftwich, A.T., Sadik, O., Sendurur, E., & Sendurur, P. (2012). Teacher beliefs and technology integration practices: A critical relationship. *Computers & Education, 59*(2), 423-435. https://doi.org/10.1016/j.compedu.2012.02.001

European Commission (2017). *Reflection paper on the social dimension of Europe.* http://bit.ly/2VyLBQk

European Commission (2018). Commission staff working document accompanying the document proposal for a Council Recommendation on Key Competences for LifeLong Learning {COM(2018) 24 final}. https://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=SWD:2018:0014:FIN:EN:PDF

European Parliament and the Council of the European Union (2006). *Recommendation of the European Parliament and of the Council of 18 December 2006 on key competences for lifelong learning (2006/962/EC). Official Journal of the European Union, L 394/10.* http://bit.ly/2Wa37au

Falchikov, N., & Goldfinch, J. (2000). Student peer assessment in Higher Education: A meta-analysis comparing peer and teacher marks. *Review of Educational Research, 70(3),* 287-322. https://doi.org/10.3102/00346543070003287

Farmer, T., & West, R.E. (2016). Opportunities and challenges with digital open badges. *Educational Technology, 56*(5), 45-48. https://bit.ly/2QNsKOS

Ferguson, R., Brasher, A., Clow, D., Cooper, A., Hillaire, G., Mittelmeier, J., Rienties, B., Ullmann, T., Vuorikari, R. (2016). *Research evidence on the use of learning analytics - implications for education policy*. R. Vuorikari, J. Castaño Muñoz (Eds.), Joint Research Centre Science for Policy Report; EUR 28294 EN. Luxembourg: Publications Office of the European Union. http://bit.ly/2V5z3Af

Frederiksen, L. (2013). Digital badges. *Public Services Quarterly, 9*(4), 321-325. http://dx.doi.org/10.1080/15228959.2013.842414

Garnett, T., & Button, D. (2018). The use of digital badges by undergraduate nursing students: A three-year study. *Nurse Education in Practice, 32*, 1-8. https://doi.org/10.1016/j.nepr.2018.06.013

Gegenfurtner, A., Quesada-Pallarès, C., & Knogler, M. (2014). Digital simulation-based training: A meta-analysis. *British Journal of Educational Technology, 45*(6), 1097–1114. https://doi.org/10.1111/bjet.12188

Gibson, D., Ostashewski, N., Flintoff, K., Grant, S., & Knight, E. (2013). Digital badges in education. *Education and Information Technologies, 20*(2), 403-410. https://doi.org/10.1007/s10639-013-9291-7

Gielen, M., & De Wever, B. (2012). Peer assessment in a Wiki: Product improvement, students' learning and perception regarding peer feedback. *Procedia - Social and Behavioral Sciences, 69*, 585-594. https://doi.org/10.1016/j.sbspro.2012.11.450

Gielen, M., & De Wever, B. (2015). Structuring peer assessment: Comparing the impact of the degree of structure on peer feedback content. *Computers in Human Behavior, 52*, 315-325. https://doi.org/10.1016/j.chb.2015.06.019

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. https://doi.org/10.1016/j.learninstruc.2009.08.007

Gikandi, J.W., & Morrow, D. (2016). Designing and implementing peer formative feedback within online learning environments. *Technology, Pedagogy and Education, 25*(2), 153-170. https://doi.org/10.1080/1475939X.2015.1058853

Gozuyesil, E., & Tanriseven, I. (2017). A meta-analysis of the effectiveness of alternative assessment techniques. *Eurasian Journal of Educational Research, 70*, 37-56. http://dx.doi.org/10.14689/ejer.2017.70.3

Griffiths, M., & Davies, C. (1993). Learning to learn: Action research from an equal opportunities perspective in a junior school. *British Educational Research Journal, 19*(1), 43-58. https://doi.org/10.1080/0141192930190104

Güler, Ç. (2016). Use of WhatsApp in Higher Education. What's up with assessing peers anonymously? *Journal of Education Computing Research, 55*(2), 272-289. https://doi.org/10.1177/0735633116667359

Hakulinen, L., Auvinen, T., & Korhonen, A. (2013, March 21-24). Empirical study on the effect of achievement badges in TRAKLA2 online learning environment. *Proceeding of Learning and teaching in computing and engineering (LaTiCE)* (pp. 47–54). doi: 10.1109/LaTiCE.2013.34

Halpern, D.F. (2008). Is intelligence critical thinking? Why we need a new definition of intelligence. In P.C. Kyllonen, R.D. Roberts, & L. Stankov (Eds.), *Extending intelligence. Enhancement and new constructs* (pp. 293-310). New York, NY: Routledge.

Hämäläinen, R., & Oksanen, K. (2014). Collaborative 3D learning games for future learning: Teachers' instructional practices to enhance shared knowledge construction among students. *Technology, Pedagogy and Education, 23*(1), 81-101. https://doi.org/10.1080/1475939X.2013.838451

Hamari, J. (2017). Do badges increase user activity? A field experiment on the effects of gamification. *Computers in Human Behavior, 71*, 469-478. https://doi.org/10.1016/j.chb.2015.03.036

Hamson-Utley, J., & Heyman, E. (2016). Implementing a badging system faculty development. In D. Ifenthaler, N. Bellin-Mularski, & D.-K. Mah (Eds.), *Foundation of digital badges and micro-credentials. Demonstrating and recognizing knowledge and competencies* (236-258). Switzerland: Springer International Publishing. https://doi.org/10.1007/978-3-319-15425-1_13

Harrison, K., O'Hara, J., & McNamara, G. (2015). Re-thinking assessment: Self- and peer assessment as drivers of self-direction in learning. *Eurasian Journal of Educational Research, 15*(60), 75-88. https://doi.org/10.14689/ejer.2015.60.5

Hattie, J. (2012). *Visible learning for teachers maximizing impact on learning.* New York, NY: Routledge.

Hattie, J., & Timperley, H. (2007). The power of feedback. *Review of Educational Research, 77*(1), 81-112. https://doi.org/10.3102/003465430298487

Hatzipanagos, S., & Code, J. (2016). Open badges in online learning environments: Peer feedback and formative assessment as an engagement intervention for promoting agency. *Journal of Educational Multimedia and Hypermedia, 25*(2), 127-142. Available at http://www.learntechlib.org/noaccess/173261/

Hofer, M., Duggan, J., & Moellendick, T. (2018, March 26-30). Promise and parameters: Digital badges for ongoing professional learning. Paper presented at the *29th International Conference of the Society for Information Technology & Teacher Education.* Available at https://www.learntechlib.org/primary/p/182521/

Hou, H.T. (2015). Integrating cluster and sequential analysis to explore learners' flow and behavioral patterns in a simulation game with situated-learning context for science courses: A video-based process exploration. *Computers in Human Behavior, 48*, 424-435. https://doi.org/10.1016/j.chb.2015.02.010

Hounsell, D., Falchikov, N., Hounsell, J., Klampfleitner, M., Huxham, M., Thomson, K., & Blair, S. (2007). Innova*tive assessment across the disciplines. An analytical review of the literature. Final Report.* The Higher Education Academy. Available at http://bit.ly/2QlmcHi

Hounsell, D., Xu, R., & Tai, C.M. (2007). *Enhancing practice. Integrative assessment. Blending assignments and assessments for high-quality learning. Guide no 3.* Gloucester: The Quality Assurance Agency for Higher Education. Available at http://bit.ly/2UR4d9N

Howard, C.D., Barrett, A.F., & Frick, T.W. (2010). Anonymity to promote peer feedback: Pre-service teachers' comments in asynchronous computer-mediated communication. Jou*rnal of Educational Computing Research, 43*(1), 89-112. https://doi.org/10.2190/EC.43.1.f

Hsia, L.-H., Huang, I., & Hwang, G.-J. (2016). A web-based peer-assessment approach to improving junior high school students' performance, self-efficacy and motivation in performing arts courses. *British Journal of Educational Technology, 47*(4), 618-632. https://doi.org/10.1111/bjet.12248

Husamah, Fatmawati, D., & Setyawan, D. (2018). OIDDE learning model: Improving higher order thinking skills of Biology teacher candidates. *International Journal of Instruction, 11*(2), 249-264. https://doi.org/10.12973/iji.2018.11217a

Hwang, G.-J., & Chang, H.-F. (2011). A formative assessment-based mobile learning approach to improving the learning attitudes and achievements of students. *Computer and Education, 56*(4), 1023-1031. https://doi.org/10.1016/j.compedu.2010.12.002

Hwang, G.-J., Hung, C.-M., & Chen, N.-S. (2014). Improving learning achievements, motivations and problem-solving skills through peer assessment-based game development approach. *Educational Technology Research and Development, 62*(2), 129-145. https://doi.org/10.1007/s11423-013-9320-7

Ilgen, J.S., Sherbino, J., & Cook, D.A. (2013). Technology-enhanced simulation in emergency medicine: A systematic review and meta-analysis. *Academic Emergency Medicine, 20,* 117-127. https://doi.org/10.1111/acem.12076

Iwata, J., Clayton, J., & Saravan, S.-J. (2013, November 29-December 1). Using self-reflection and badges in moodle-based medical English review courses for enhancing learners' autonomy. Paper presented at

International conference on Educational Technologies 2013 (ICEduTech 2013). Available at https://files.eric.ed.gov/fulltext/ED557168.pdf

Iwata, J., Telloyan, J., Murphy, L., Wang, S., & Clayton, J. (2013, December 11-13). The use of a digital badge as an indicator and a motivator. Paper presented at the *5th International conference on educational technologies 2017 (ICEduTech 2017).* Available at https://files.eric.ed.gov/fulltext/ED579282.pdf

Jones, W.M., Hope, S., & Adams, B. (2018). Teachers' perceptions of digital badges as recognition of professional development. *British Journal of Educational Technology, 49,* 427-438. https://doi.org/10.1111/bjet.12557

Jovanovic, J., & Devedzic, V. (2014). Open Badges: Novel means to motivate, scaffold and recognize learning. *Technology, Knowledge and Learning, 20*(1), 115-122. https://doi.org/10.1007/s10758-014-9232-6

Kablan, Z. (2014). Comparison of individual answer and group answer with and without structured peer assessment. *Research in Science & Technological Education, 32*(3), 251-262. https://doi.org/10.1080/02635143.2014.931840

Kim, J., Park, J.-H., & Shin, S. (2016). Effectiveness of simulation-based nursing education depending on fidelity: A meta-analysis. *BMC Medical Education, 16.* https://doi.org/10.1186/s12909-016-0672-7

Knight, P. (2002). Summative assessment in Higher Education: Practices in disarray. *Studies in Higher Education, 27*(3), 275-286. https://doi.org/10.1080/03075070220000662

Kostons, D., van Gog, T., & Paas, F. (2009). How do I do? Investigating effects of expertise and performance-process records on self-assessment. *Applied Cognitive Psychology, 23*(9), 1256-1265. https://doi.org/10.1002/acp.1528

Kovalik, C.L., & Kuo, C.L. (2012). Innovation diffusion: Learner benefits and instructor insights with the Diffusion Simulation Game. *Simulation & Gaming, 43*(6), 803-824. https://doi.org/10.1177/1046878112444577

Lai, C.L., & Hwang, G.J. (2015). An interactive peer-feedback criteria development approach to improving students' art design performance using handheld devices. *Computers & Education, 85*, 149-159. https://doi.org/10.1016/j.compedu.2015.02.011

Lau, A.M.S. (2016). 'Formative good, summative bad?' - A review of the dichotomy in assessment literature. *Journal of Further and Higher Education*, *40*(4), 509-525. https://doi.org/10.1080/0309877X.2014.984600

Law, P., & Law, A. (2014, October 23-24). Digital badging at The Open University: Recognition for informal learning. *The Open and Flexible Higher Education Conference 2014: 'New Technologies and the Future of Teaching and Learning'.* Krakow, Poland. http://oro.open.ac.uk/41354/

Li, L., & Chen, N. (2013). Study on agent based intelligent feedback system in online teaching and interactive learning. *Information Technology Journal, 12*(4), 763-769. 10.3923/itj.2013.763.769

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. https://doi.org/10.1111/j.1467-8535.2009.00968.x

Lin, Y.L., & Tu, Y.Z. (2012). The values of college students in business simulation game: A means-end chain approach. *Computers & Education, 58*(4), 1160-1170. https://doi.org/10.1016/j.compedu.2011.12.005

Lius, E. (2016, July 1-4). Gamify and recognize prior learning: How to succeed in educators' further professional training with open badges. Paper presented at *International conference e-learning 2016. Part of the MultI conference on Computer Science and Information Systems 2016.* Available at https://files.eric.ed.gov/fulltext/ED571430.pdf

Liyanagunawardena, T.R., Scalzavara, S., & Williams, S.A. (2017). Open badges: A systematic review of peer-reviewed published literature (2011-2015). *European Journal of Open, Distance and e-Learning, 20*(2), 1-16. http://www.eurodl.org/materials/contrib/2017/Liyanagunawardena_et_al.pdf

Lock, J., Kim, B., Koh, K., & Wilcox, G. (2018). Navigating the tensions of innovative assessment and pedagogy in Higher Education. *The Canadian Journal for the Scholarship of Teaching and Learning, 9*(1). https://doi.org/10.5206/cjsotl-rcacea.2018.1.8

Lucas, M. (2018). External barriers affecting the successful implementation of mobile educational interventions. *Computers in Human Behavior,* 1-7. https://doi.org/10.1016/j.chb.2018.05.001

Lukosch, H., Kurapati, S., Groen, D., & Verbraeck, A. (2016). Microgames for situated learning: A Case study in interdependent planning. *Simulation & Gaming, 47*(3), 346-367. https://doi.org/10.1177/1046878116635468

Marion, S.F., Lyons, S., & Pace, L. (2017). Evaluating and continuously improving an innovative assessment and accountability system. Available at https://bit.ly/2FAT2zy

McGaghie, W.C., Issenberg, S.B., Cohen, E.R., Barsuk, J.H., & Wayne, D.B. (2011). Does simulation-based medical education with deliberate practice yield better results than traditional clinical education? A meta-analytic comparative review of the evidence. *Academic Medicine, 86*(6), 706-711. doi: 10.1097/ACM.0b013e318217e119

Melrose, S. (2017). Balancing reflection and validity in health profession students' self-assessment. *International Journal of Learning, Teaching and Educational Research, 16*(8), 65-76. https://www.ijlter.org/index.php/ijlter/article/view/979

Merchant, Z., Goetz, E.T., Cifuentes, L., Keeney-Kennicutt, W., & Davis, T.J. (2014). Effectiveness of virtual reality-based instruction on students' learning outcomes in K-12 and Higher Education: A meta-analysis. *Computers & Education, 70,* 29-40. https://doi.org/10.1016/j.compedu.2013.07.033

Moore, C., & Teather, S. (2013). Engaging students in peer review: Feedback as learning. *Issues in Educational Research, 23*(2), 196-211. http://www.iier.org.au/iier23/moore.html

Motheeram, P., Herselman, M., & Botha, A. (2018). A scoping review of digital open badge ecosystems in relation to resource-constrained environments. *The Journal for Transdisciplinary Research in Southern Africa, 14*(1), 1-10. https://doi.org/10.4102/td.v14i1.463

Mowl, G. (2006). *Red Guides Paper 17. Innovative student assessment: What's the point?* Available at http://bit.ly/2L8osA2

Mulder, R., Baik, C., Naylor, R., & Pearce, J. (2014). How does student peer review influence perceptions, engagement and academic outcomes? A case study. *Assessment & Evaluation in Higher Education*, *39*(6), 657-677. https://doi.org/10.1080/02602938.2013.860421

Munns, G., & Woodward, H. (2006). Student engagement and student self-assessment: The REAL framework. *Assessment in Education: Principles, Policy and Practice, 13*(2), 193-213. https://doi.org/10.1080/09695940600703969

Nicol, D., Thomson, A., & Breslin, C. (2014). Rethinking feedback practices in Higher Education: A peer review perspective. *Assessment & Evaluation in Higher Education, 39*(1), 102-122. https://doi.org/10.1080/02602938.2013.795518

Nikou, S.A., & Economides, A.A. (2016). The impact of paper-based, computer-based and mobile-based self-assessment on student' science motivation and achievement. *Computers in Human Behaviour, 55*(Part B), 1241-1248. https://doi.org/10.1016/j.chb.2015.09.025

Nkhoma, M., Calbeto, J., Sriratanaviriyakul, N., Muang, T., Ha Tran, Q., & Kim Cao, T. (2014). Towards an understanding of real-time continuous feedback from simulation games. *Interactive Technology and Smart Education, 11*(1), 45-62. https://doi.org/10.1108/ITSE-03-2013-0005

O'Byrne, W.I., Schenke, K., Willis, J.E. III, & Hickey, D.T. (2015). Digital badges. Recognizing, assessing, and motivating learners in and out of school contexts. *Journal of Adolescent & Adult Literacy, 58*(6), 451-454. https://doi.org/10.1002/jaal.381

OECD (2016). *Innovating education and educating for innovation: The power of digital technologies and skills.* Paris: OECD Publishing. http://dx.doi.org/10.1787/9789264265097-en

Oldfield, A., Broadfoot, P., Sutherland, R., & Timmis, S. (n.d.). *Assessment in a digital age: A research review.* University of Bristol, Graduate School of Education. Available at http://bit.ly/2DId0oF

Olina, Z., & Sullivan, H.J. (2002). Effects of classroom evaluation strategies on student achievement and attitudes. *Educational Technology, Research and Development, 50*(3), 61-75. https://doi.org/10.1007/BF02505025

Panadero, E. (2016). Is it safe? Social, interpersonal, and human effects of peer assessment: A review and future directions. In G.T.L. Brown, & L.R. Harris (Eds.), *Handbook of social and human conditions in assessment* (pp. 247-266). New York: Routledge. https://bit.ly/2Ucb8tA

Panadero, E., & Romero, M. (2014). To rubric or not to rubric? The effects of self-assessment on self-regulation, performance and self-efficacy. *Assessment in Education: Principles, Policy & Practice, 21*(2), 133-148. https://doi.org/10.1080/0969594X.2013.877872

Panadero, E., Brown, G., & Courtney, M. (2014). Teachers' reasons for using self-assessment: A survey self-report of Spanish teachers. *Assessment in Education: Principles, Policy & Practice, 21*(4), 365-383. https://doi.org/10.1080/0969594X.2014.919247

Panadero, E., Brown, G.T.L., & Strijbos, J.-W. (2016). The future of student self-assessment: A review of known unknowns and potential directions. *Educational Psychology Review, 28*(4), 803-830. https://doi.org/10.1007/s10648-015-9350-2

Panadero, E., Jonsson, A., & Alqassab, M. (2018). Providing formative peer feedback: What do we know? In A.A. Lipnevich, & J.K. Smith (Eds.), *The Cambridge Handbook of Instructional Feedback* (pp. 409-431). Cambridge: Cambridge University Press. https://doi.org/10.1017/9781316832134.020

Panadero, E., Jonsson, A., & Botella, J. (2017). Effects of self-assessment on self-regulated learning and self-efficacy: Four meta-analyses. *Educational Research Review, 22,* 74-98. https://doi.org/10.1016/j.edurev.2017.08.004

Papamitsiou, Z., & Economides, A. (2014). Learning Analytics and educational data mining in practice: A systematic literature review of empirical evidence. *Educational Technology & Society*, *17*(4), 49-64. https://www.jstor.org/stable/jeductechsoci.17.4.49?seq=1#page_scan_tab_contents

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. https://doi.org/10.1111/j.1365-2729.2008.00290.x

Paré, D.E., & Joordens, S. (2009). peerScholar: Tired of marking? Using peerScholar to explore the change in peer grading reliability as a function of increased number of peer evaluations. Presentation at the Society for Computers in Psychology Conference (SCiP), Long Beach, CA.

Pasin, F., & Giroux, H. (2011). The impact of a simulation game on operations management education. *Computers & Education, 57*(1), 1240-1254. https://doi.org/10.1016/j.compedu.2010.12.006

Perrotta, C. (2014). Innovation in technology-enhanced assessment in the UK and the USA: Future scenarios and critical considerations. *Technology, Pedagogy and Education, 23*(1), 103-119. https://doi.org/10.1080/1475939X.2013.838453

Price, M., & O'Donovan, B. (2006). Improving student performance through enhanced student understanding of criteria and feedback. In C. Bryan, & K. Clegg (Eds.), *Innovative assessment in Higher Education* (pp. 100-109). London: Routledge.

Raemer, D., Anderson, M., Cheng, A., Fanning, R., Nadkarni, V., & Savoldelli, G. (2011). Research regarding debriefing as part of the learning process. *Simulation in Healthcare, 6*(7), s52-s57. doi: 10.1097/SIH.0b013e31822724d0

Regehr, G., & Eva, K. (2006). Self-assessment, self-direction, and the self-regulating professional. *Clinical Orthopaedics and Related Research, 449*, 34-38. http://bit.ly/2PFeRPR

Reinholz, D. (2016). The assessment cycle: A model for learning through peer assessment. *Assessment & Evaluation in Higher Education*, *41*(2), 301-315. https://doi.org/10.1080/02602938.2015.1008982

Rotsaert, T., Panadero, E., Schellens, T., & Raes, A. (2018). "Now you know what you're doing right and wrong!" Peer feedback quality in synchoronous peer assessment in Secondary Education. *European Journal of Psychology of Education, 33*(2) 255-275. https://doi.org/10.1007/s10212-017-0329-x

Rust, C. (2007). Towards a scholarship of assessment. *Assessment & Evaluation in Higher Education, 32*(2), 229-237. https://doi.org/10.1080/02602930600805192

Rutten, N., van Joolingen W.R., & van der Veen, J.T. (2012). The learning effects of computer simulations in Science Education. *Computer & Education 58*(1), 136-153. https://doi.org/10.1016/j.compedu.2011.07.017

Sadler, D.R. (1989). Formative assessment and the design of instructional systems. *Instructional Science, 18*(2), 119-144. https://doi.org/10.1007/BF00117714

Sadler, P.M., & Good, E. (2006). The impact of self- and peer-grading on student learning. *Educational Assessment, 11*(1), 1-31. https://www.tandfonline.com/doi/abs/10.1207/s15326977ea1101_1

Sawyer, T., Eppich, W., Brett-Fleegler, M., Grant, V., & Cheng, A. (2016). More than one way to debrief. A critical review of healthcare simulation debriefing methods. *Simulation in Healthcare: The Journal of the Society for Simulation in Healthcare, 11*(3), 209-217. doi: 10.1097/SIH.0000000000000148

Schuetze, H.G. (2007). Individual learning accounts and other models of financing lifelong learning. *International Journal of Lifelong Education, 26*(1), 5-23. https://doi.org/10.1080/02601370601151349

Schunk, D.H. (1996). Goal and self-evaluative influences during children's cognitive skill learning. *American Educational Research Journal, 33*(2), 359-382. https://doi.org/10.3102/00028312033002359

Siarova, H., Sternadel, D., & Mašidlauskaitė, R. (2017). *Assessment practices for 21st century learning: Review of evidence (NESET II Report)*. Luxembourg: Publications Office of the European Union. http://bit.ly/2VypWrn

Siewiorek, A., Saarinen, E., Lainema, T., & Lehtinen, E. (2012). Learning leadership skills in a simulated business environment. *Computers & Education, 58*(1), 121-135. https://doi.org/10.1016/j.compedu.2011.08.016

Silvia, C. (2012). The impact of simulations on higher-level learning. *Journal of Public Affairs Education, 18*(2), 397-422. https://doi.org/10.1080/15236803.2012.12001690

Sin, K., & Muthu, L. (2015). Application of big data in education data mining and learning analytics - A literature review. *Journal on Soft Computing, 5*(4), 1035-1049. http://ictactjournals.in/paper/IJSC_V5_I4_paper6_1035_1049.pdf

Sitzmann, T. (2011). A meta-analytic examination of the instructional effectiveness of computer-based simulation games. *Personnel Psychology, 64,* 489-528. https://doi.org/10.1111/j.1744-6570.2011.01190.x

Sitzmann, T., & Ely, K. (2011). A meta-analysis of self-regulated learning in work-related training and educational attainment: What we know and where we need to go. *Psychological Bulletin, 137*(3), 421-442. http://psycnet.apa.org/doiLanding?doi=10.1037%2Fa0022777

Sitzmann, T., Ely, K., Brown, K.G., & Bauer, K.N. (2017). Self-assessment of knowledge: A cognitive learning or affective measure? *Academy of Management Learning & Education, 9*(2), 169-191. https://doi.org/10.5465/amle.9.2.zqr169

Sluijsmans, D.M.A., Brand-Gruwel, S., van Merrienboer, J.J.G., & Martens, R.L. (2004). Training teachers in peer-assessment skills: Effects on performance and perceptions. *Innovations in Education and Teaching International, 41*(1), 59–78. https://doi.org/10.1080/1470329032000172720

Söderström, T., Häll, L., Nilsson T., & Ahlqvist J. (2014/2015). Computer simulation training in Health Care Education: Fuelling reflection-in-action? Simulation & Gaming, 45(6), 805-828. https://doi.org/10.1177/1046878115574027

Steiner, C., Kickmeier-Rust, M., & Türker, M.A. (2014). *Deliverable D3.1 Review article about LA and EDM approaches.* Available at http://css-kmi.tugraz.at/mkrwww/leas-box/downloads/D3.1.pdf

Tannenbaum, S.I., & Cerasoli, C.P. (2012/2013). Do team and individual debriefs enhance performance? A meta-analysis. *Human Factors: The Journal of the Human Factors and Ergonomics Society, 55*(1), 231-245. https://doi.org/10.1177/0018720812448394

Taras, M. (2010). Student self-assessment: Processes and consequences. *Teaching in Higher Education, 15*(2), 199-209. https://doi.org/10.1080/13562511003620027

The Assessment Reform Group (1999). *Assessment for learning: Beyond the Black Box.* University of Cambridge School of Education. Available at http://bit.ly/2DUEcSG

Thummaphan, P. (2017). *Innovative assessments that support students' STEM learning.* Ph.D. Thesis, University of Washington. Available at http://bit.ly/2Btt6Sc

Tierney, W.G. (2014). *Creating a culture of innovation. The challenge in becoming and staying a world-class university.* Pullias Center for Higher Education, Rossier School of Education, University of Southern California. Available at https://bit.ly/2r9e5iI

Tolgfors, B., & Öhman, M. (2015). The implications of assessment for learning in Physical Education and Health. *European Physical Education Review, 22*(2), 150-166. https://doi.org/10.1177/1356336X15595006

Topping, K.J. (1998). Peer assessment between students in colleges and universities. *Review of Educational Research, 68*(3), 249-276. https://doi.org/10.3102/00346543068003249

Topping, K.J. (2003). Self and peer assessment in school and university: Reliability, validity and utility. In M. Segers, F. Dochy, & E. Cascallar (Eds.), *Optimising new modes of assessment: In search of qualities and standards* (pp. 55-87). New York, NY: Kluwer Academic Publishers.

Topping, K.J. (2017). Peer assessment: Learning by judging and discussing the work of other learners. *Interdisciplinary Education and Psychology, 1*(1), 1-17. https://doi.org/10.31532/InterdiscipEducPsychol.1.1.007

Torrance, H. (2007). Assessment as learning? How the use of explicit learning objectives, assessment criteria and feedback in post-secondary education and training can come to dominate learning. *Assessment in Education, 14*(3), 281-294. https://doi.org/10.1080/09695940701591867

Torrance, H. (2012). Formative assessment at the cross roads: Conformative, deformative and transformative assessment. *Oxford Review of Education, 38*(3), 323-342. https://doi.org/10.1080/03054985.2012.689693

van Gennip, N.A.E., Segers, M.S.R., & Tillema, H.H. (2009). Peer assessment for learning from a social perspective. The influence of interpersonal variables and structural features. *Educational Research Review, 4*(1), 41-54. https://doi.org/10.1016/j.edurev.2008.11.002

Van Heukelom, J.N., Begaz, T., & Treat, R. (2010). Comparison of postsimulation debriefing versus in-simulation debriefing in medical simulation. *Simulation in Healthcare, 5*(2), 91-97. doi: 10.1097/SIH.0b013e3181be0d17

Vanderhoven, E., Raes, A., Motrieux, H., Rotsaert, T., & Schellens, T. (2015). What if pupils can assess their peers anonymously? A quasi-experimental study. *Computer & Education*, *81*, 123-132. https://doi.org/10.1016/j.compedu.2014.10.001

Veiga Ávila, L., Leal Filho, W., Brandli, L., Macgregor, C.J., Molthan-Hill, P., Gökçin Özuyar, P., & Martins Moreira, R. (2017). Barriers to innovation and sustainability at universities around the world. *Journal of Cleaner Production, 164,* 1268-1278. http://dx.doi.org/10.1016/j.jclepro.2017.07.025

Vincent-Lancrin, S., Kärkkäinen, K., Pfotenhauer, S., Atkinson, A., Jacotin, G., & Rimini, M. (2014). *Measuring innovation in education: A new perspective.* OECD: OECD Publishing. http://dx.doi.org/10.1787/9789264215696-en

Vlachopoulos, D., & Makri, A. (2017). The effect of games and simulations on Higher Education: A systematic literature review. *International Journal of Educational Technology in Higher Education, 14*(22). https://doi.org/10.1186/s41239-017-0062-1

Vogel, J.J., Vogel, D.S., Cannon-Bowers, J., Bowers, C.A., Muse, K., & Wright, M. (2006). Computer gaming and interactive simulations for learning: A meta-analysis. Journal of Educational Computing Research, 34(3), 229-243. https://doi.org/10.2190/FLHV-K4WA-WPVQ-H0YM

Vongkulluksn, V.W., Xie, K., & Bowman, M.A. (2018). The role of value on teachers' internalization of external barriers and externalization of personal beliefs for classroom technology integration. *Computers & Education, 118*, 70-81. https://doi.org/10.1016/j.compedu.2017.11.009

West, R.E., & Randall, D.L. (2016). The case for rigor in open badges. In L., Muilenburg, & Z., Berge (Eds.), *Digital badges in education, trends, issues and cases* (pp. 21-29). New York, NY: Routledge.

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 & Education, 57*(2), 1493-1500. http://dx.doi.org/10.1016/j.compedu.2011.02.011

Wouters, P., & van Oostendorp, H. (2013). A meta-analytic review of the role of instructional support in game-based learning. *Computers & Education, 60*(1), 412-425. https://doi.org/10.1016/j.compedu.2012.07.018

Yang, M., Badger, R., & Zhen, Y. (2006). A comparative study of peer and teacher feedback in a Chinese EFL writing class. *Journal of Second Language Writing, 15*(3), 179-200. https://doi.org/10.1016/j.jslw.2006.09.004

Zacharis, N.Z. (2010). Innovative assessment for learning enhancement: Issues and practices. *Contemporary Issues in Education Research, 3*(1), 61-70. https://doi.org/10.19030/cier.v3i1.162

Zimmerman, B.J. (1989). A social cognitive view of self-regulated academic learning. *Journal of Educational Psychology*, *81*(3), 329-339. http://dx.doi.org/10.1037/0022-0663.81.3.329

References from the case studies

ATS2020

ATS2020. *Assessment of transversal skills.* www.ats2020.eu

ATS2020. *Assessment of transversal skills. Learning designs.* https://bit.ly/2lGbzRt

ATS2020. *Toolkit.* https://resources.ats2020.eu/resource-details/ADM/toolkit

Ferrari, A. (2013). *DIGCOMP: A framework for developing and understanding digital competence in Europe.* Y., Punie, & B.N., Brečko (Eds.), EUR 26035 EN. Luxembourg: Publications Office of the European Union. http://dx.doi.org/10.2788/52966

e-Assessment of Prior Learning in Swedish elderly care (e-APL)

Nilsson, A., Ardén, M., & Engström, M. (2014). E-assessment of prior learning: A pilot study of interactive assessment of staff with no formal education who are working in Swedish elderly care. *BMC Geriatrics, 14*(52), 1-10. https://doi.org/10.1186/1471-2318-14-52

Nilsson, A., & Engström, M. (2015). E-assessment and an e-training program among elderly care staff lacking formal competence: Results of a mixed-methods intervention study. *BMC Health Services Research, 15*(189), 1-11. https://doi.org/10.1186/s12913-015-0843-y

Objective Structured Clinical Examination (OSCE)

Holmes, D. (2017, December 13). *Sheffield Hallam University and PebblePad lift learning technologies award.* https://bit.ly/2k1Bgvg

Kirke, A. (2017, July). *A new efficient approach to evidencing clinical skill acquisition.* [Video file].

https://www.youtube.com/watch?v=QhkowLvK2ds

Kirke, A., Ball, A., & Edwards, S. (2016). *A new approach to confirmation of skill acquisition.* https://bit.ly/2m7fzuD

Finnish Matriculation Examination (FME)

Britschgi, V. (2015). *The Finnish Matriculation Examination.* https://bit.ly/2kqbUYr

European Schoolnet (2017). *Finland: Country report on ICT in Education.* https://bit.ly/2koKphX

Lakkala, M., & Ilomäki, L. (2013). *Lukioiden valmiudet siirtyä sähköiseen ylioppilastutkintoon: Kahden lukion tapaustutkimus*. Published by city of Vantaa, Finland. https://bit.ly/2lCLKSu

Ministry of Foreign Affairs (2017). *Education in Finland: Key to the nation´s success.* https://bit.ly/2koKRwE

Training material for all subjects available by YLE https://yle.fi/aihe/abitreenit

Multipoly Next

Coppens, A. (2014, June 14). *Gamification stuff we love: Multipoly – recruitment game* [Blog post]. https://bit.ly/2koJQVk

Dungan, R. (2015). *PwC Hungary offers Multipoly.* https://bit.ly/2lORYid

Heong W.M. (2015). *PwC's Multipoly boosts employee recruitment and retention* [Blog post]. https://bit.ly/2k3XNrm

Meister, J. (2015). *Future of work: Using gamification for human resources.* https://bit.ly/2lChMxZ

http://multipoly.hu/

www.gamesforbusiness.com

Neo Alta

Di Martino, A. (n.d.). # Neo Alta, enseigner et évaluer autrement pour mieux accompagner les élèves 2017A. *Expérithèque*. https://bit.ly/2k626T3

Piquemal, M. (2015, March 10). Au collège des Clayes, les profs marient les matières. *Liberation.* https://bit.ly/2k622mh

https://bit.ly/2kz7ePS

Teach For Slovakia (Teach4SK)

Miyanaga, G. (2012, October 12*). Mini whiteboards for formative assessment.* https://www.sd43.bc.ca/Resources/StaffResources/BIG/default.aspx#/view/32

Tough, P. (June 2016). How kids learn resilience. *The Atlantic.* https://bit.ly/2m6e8wt

https://bit.ly/2k6dtdP

http://www.sstr2.org/Downloads/Cups as student feedback.pdf

http://www.teachforslovakia.sk/

Assessment based on an Online Collaborative Project (AOCP)

Pérez-Mateo, M., Romero, M. & Romeu, T. (2014). Collaborative construction of a project as a methodology for acquiring digital competences. *Comunicar, 42*(XXI), 15-23. https://doi.org/10.3916/C42-2014-01

Romeu, T., Romero, M., & Guitert, M. (2016). E-assessment process: Giving a voice to online learners. *International Journal of Educational Technology in Higher Education, 13*(20), 1-14. https://bit.ly/2k9di1p

List of abbreviations

AOCP	Assessment based on an Online Collaborative Project
ATS2020	Assessment of Transversal Skills
CPD	Continuous Professional Development
e-APL	E-Assessment of Prior Learning in Swedish elderly care
FME	Finnish Matriculation Examination
ICT	Information Communication Technology
INSET	In Service Training
IT	Information Technology
ITE	Initial teacher Education
MEB	Matriculation Examination Board
n.d.	No date
OSCE	Objective Structured Clinical Examination
PwC	PricewaterhouseCoopers
SeQF	Swedish Qualification Framework
Teach4SK	Teach for Slovakia
VET	Vocational Education and Training

Figures

Tables

Annex 1. Examples of digital assessment tools

Software/ Application	Description	Website
Examples of game-based learning platforms		
Use Your Brainz EDU	An educational version of a video game called Plants vs. Zombies 2	https://www.glasslabgames.org/games/PVZ
The Detective: Verona	An educational detective game where students differentiate credible information from unreliable information	https://k20center.ou.edu/games/detective-verona/
The Detective: Bavaria	An educational detective game for learning to identify and interpret different kinds of data	https://k20center.ou.edu/games/detective-bavaria/
Advance U: The Talent Machine	A role-playing video game for teaching growth mindset	https://k20center.ou.edu/games/advance-u/
Kahoot!	A free educational game-based learning platform	https://kahoot.com/
Brainology®	A blended-learning curriculum, learning about study techniques, self-regulation strategies, and other essential non-cognitive skills that help students to become effective learners	https://www.mindsetworks.com/
Video games used in educational settings		
Pokémon GO	A game for catching virtual creatures (Pokémons) in real-world locations via mobile device's GPS ability	https://www.pokemongo.com/en-us/
Assasin's Creed	A video game based on historical eras	https://assassinscreed.ubisoft.com/game/en-us/home
Minecraft	A sandbox video game where players can create constructions with blocks in 3D world	https://education.minecraft.net/
Impulse	A game for learning about Newton's first and second laws of motion, designed to foster and measure implicit learning	https://www.brainpop.com/games/impulse/
Quantum Spectre	A puzzle-style game, scientifically accurate simulations	https://play.google.com/store/apps/details?id=air.com.edge.quantumspectre&hl
Examples of computer software linked to learning analytics		
Dota2	A team-based multiplayer online game	http://www.dota2.com/international/overview/
Graphistry	A platform of graph reasoning for visual investigation	https://www.graphistry.com/
Lea's Box	A learning analytics toolbox	http://css-kmi.tugraz.at/mkrwww/leas-box/

eAdventure	Platform is a research project	http://e-adventure.e-ucm.es/
WEKA	A collection of machine learning algorithms for data mining tasks	https://www.cs.waikato.ac.nz/ml/weka/
Caffe	An open deep learning framework	http://caffe.berkeleyvision.org/
Blackboard	Learning management systems	http://www.blackboard.com/about-us/index.html
Desire2Learn	For administering the delivery of online learning and training	https://www.d2l.com/
Examples of educational video games		
aleks	A web-based, artificially intelligent assessment and learning system	https://www.aleks.com/
Education Galaxy	Online assessment, practice, and instruction for K-5 elementary students	http://educationgalaxy.com/
Plickers	Teachers collect real-time formative assessment data without the need for student devices	https://www.plickers.com/
Crystals of Kaydor	Aimed at teaching children prosocial behaviors, including recognizing others' emotions	http://www.gameslearningsociety.org/kaydor_microsite/
seppo	Authoring tool for creating educational games	http://www.seppo.io/en/
Cognify	Measure problem solving, numerical reasoning, processing speed and verbal knowledge	http://help.revelian.com/kb/cognify-game-based-assessment/
MindX	Mind mapping tool	https://www.xmind.net/
lumosity	Brain-training with cognitive games designed by scientists	https://www.lumosity.com/
Games4Sustainability	Teaching learning and practicing sustainability through serious games	https://games4sustainability.org
SpeedGrader	Provides feedback to students	https://facdev.e-education.psu.edu/node/389
Chinese Character a Day	Apps for learning Mandarin Chinese	http://chinesecharacteraday.com/
FunBrain	Online educational games for kids of all ages	https://www.funbrain.com/
edutopia	Formative assessment tools	https://www.edutopia.org/
Fun Atomic	Developers of educational game analytics services	https://funatomic.com/
Common Sense Education	Lots of different educational apps	https://www.commonsense.org/education/reviews/all
Technology-supported peer- and self-assessment tools		
Moodle	A learning management system	https://moodle.com/
DiscussionBoard	An online forum (provided by Blackboard)	https://help.blackboard.com/Learn/Student/Interact/Discussions

Qridi	A comprehensive digital assessment tool for teacher evaluation and self-, peer- and group-assessment	http://www.qridi.com/
SkillzzUp	A digital assessment tool for teacher, student, parent and school-level use	http://skillzzup.com/en/
peerScholar	A digital assessment tool for peer-assessment and teacher evaluation	http://peerscholar.com/
Peergrade	An interactive online platform for peer-feedback sessions	https://www.peergrade.io/
eRubric	A digital rubric tool for Higher Education teachers and students	https://gteavirtual.org/rubric/
WhatsApp	A social media and communication platform that can be used in peer-assessment	https://www.whatsapp.com/?l=en
Commercial software providers with electronic home-school collaboration tools		
Educamos	A school and classroom management system	www.qualitaseducativa.com/
Wilma	A school and classroom management system	www.visma.fi/inschool/wilma/
Eschool	A school and classroom management system	https://sites.google.com/hawaiidoe.k12.hi.us/e-school/home
ParentLocker	A school and classroom management system	www.parentlocker.com/
RenWeb	A school and classroom management system	www.renweb.com/
PowerSchool	A school and classroom management system	www.powerschool.com/
ClassDojo	A classroom communication app	https://www.classdojo.com/en-gb/?redirect=true

GETTING IN TOUCH WITH THE EU

In person

All over the European Union there are hundreds of Europe Direct information centres. You can find the address of the centre nearest you at: https://europa.eu/european-union/contact_en

On the phone or by email

Europe Direct is a service that answers your questions about the European Union. You can contact this service:
- by freephone: 00 800 6 7 8 9 10 11 (certain operators may charge for these calls),
- at the following standard number: +32 22999696, or
- by electronic mail via: https://europa.eu/european-union/contact_en

FINDING INFORMATION ABOUT THE EU

Online

Information about the European Union in all the official languages of the EU is available on the Europa website at: https://europa.eu/european-union/index_en

EU publications

You can download or order free and priced EU publications from EU Bookshop at: https://publications.europa.eu/en/publications. Multiple copies of free publications may be obtained by contacting Europe Direct or your local information centre (see https://europa.eu/european-union/contact_en).

KJ-NA-29882-EN-N

The European Commission's science and knowledge service

Joint Research Centre

JRC Mission

As the science and knowledge service of the European Commission, the Joint Research Centre's mission is to support EU policies with independent evidence throughout the whole policy cycle.

EU Science Hub
ec.europa.eu/jrc

@EU_ScienceHub

EU Science Hub - Joint Research Centre

EU Science, Research and Innovation

EU Science Hub

Publications Office
of the European Union

doi:10.2760/552774
ISBN 978-92-76-12116-9