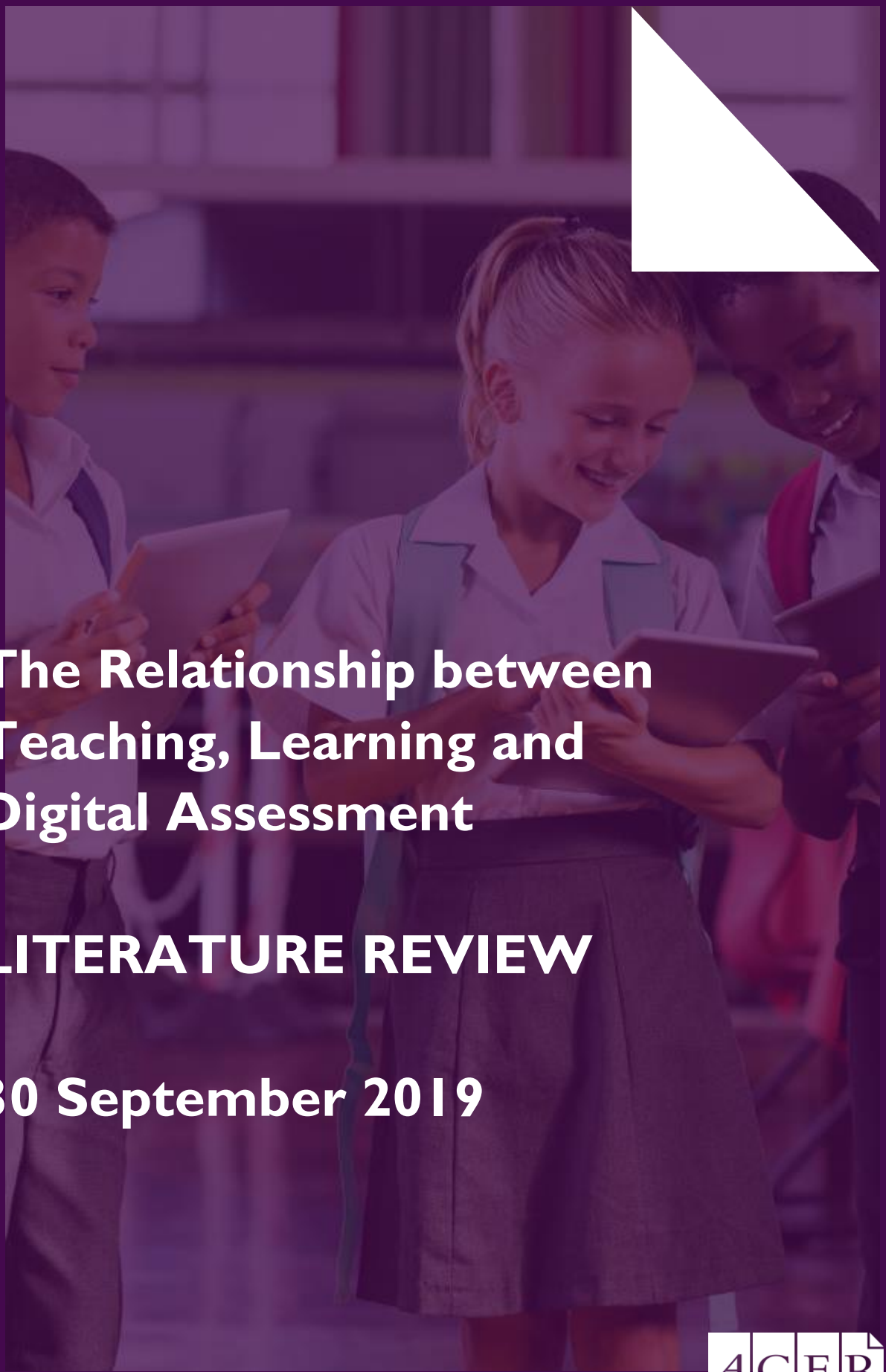

The Relationship between Teaching, Learning and Digital Assessment

LITERATURE REVIEW

30 September 2019

ACER

Report submitted to the International Baccalaureate Organisation
By the Australian Council for Educational Research

EXECUTIVE SUMMARY

Introduction

The International Baccalaureate (IB) is at an interesting time in the use of digital assessment, with the use of e-Assessment in the Middle Years Programme (MYP) but retaining traditional paper-and-pen summative assessment in the Diploma Programme.

The IB also has a unique approach to education, and key principles running through all IB programmes include inquiry, contextualisation of learning, learning how to learn, community service and global understanding. A recent IB publication states that '*technology should support assessment*' (IBO, 2018a, 6) but further states that this should not be done at the expense of these principles but rather to better channel them into practices.

In this context, the literature review has focused on addressing the question '*What is the current state of the art understanding of digital assessment use within the future educational landscape?*' It looks at the contemporary context within the evolution of key movements in education and identifies the role that digital assessment can and should play, and some key considerations for the IB to address moving forward.

Throughout the review we have retained a focus on the fundamental significance of purpose and on finding ways to ensure that the three components of learning, teaching and assessment work in harmony. This means the critical importance of retaining a clear view of objectives in determining the form and function of digital assessment, as well as ensuring that there is alignment with teaching and learning.

Methodology

The literature review has focused on three key sources of information. First, academic journal articles and books written by educational researchers that reflect on the key philosophies and movements in education and that also report on initiatives to enhance educational practices.

One of the limitation of academic literature is the time lag. It can be several years between research taking place and the publication of an article reporting on it. In a rapidly moving sphere such as digitalisation, this means that the newest practices may not yet be reflected.

To address this, the literature review has also incorporated so-called '*grey literature*' from organisations and governments who are engaged in relevant activities in the areas of teaching, learning and digital assessment. These include reports, blogs, websites, interviews and presentations which engage with the most contemporary movements in education.

Finally, the literature review has also engaged with publicly-available sample assessment materials from organisations engaged in some of the most innovative and forward-looking initiatives. These include large international bodies such as the Organisation for Economic Cooperation and Development (OECD) as well as some of the well-known assessment bodies.

The researchers involved in this literature review undertook an approach which was informed by a combination of their own knowledge in this field, key word searches in academic databases, a web search of key organisations involved in assessment and utilising key themes that arose in the initial sources perused to lead them to other important areas of consideration.

With attention to digital assessment, the review considered a range of ways in which this has evolved and is practiced in contemporary education, including the opportunities that it offers and the key design considerations. Reviewers looked at a range of examples of digital assessment usage in

international assessments, language assessments, in national assessment programmes and in initiatives to measure twenty-first century skills. Both academic and stakeholder insights were drawn on to contextualise each of these examples.

Finally, the research focused on some of the key components of the relationship between teaching, learning and digital assessment in the context of best practice approaches to education. This raised the issue of the importance of the skills of learners and teachers, and the reviewers drew on research to identify best practice approaches in supporting these key stakeholders to navigate new dynamics in educational practice.

Contemporary contexts

This balance needs to be found in an educational landscape in which some growing understandings are reshaping nature of teaching and learning. The first of these is a shift towards knowledge creation rather than acquisition, and this is combined with the emergent understanding that all learners are on an individual learning pathway.

This pathway is demarcated by a series of learning progressions, along which all learners can make progress. Key to learning progressions is the regular collection of empirical evidence to help learners and teachers identify where they are and where they need to go next.

Learning progressions raise the need for ongoing assessment, moving away from the mutual exclusivity of formative and summative approaches. Instead, assessment becomes a '*present to future*' model in which a regular flow of insights to teachers and learners is enabled and assessment fundamentally embedded in the everyday nature of teaching and learning rather than separate from these.

Part of the potential of digital assessment is its ability to generate a whole series of data-informed insights into the way that learners are navigating the curriculum. This is often thought of in terms of the responses that learners have given to particular test questions.

But there is much greater scope within digital platforms, including to measure time-on-task, misconceptions that lead to errors, and decisions that learners make about order of responses. When data is also able to be collected from digital platforms that learners use in learning activities, a wealth of insights are generated. And in a relatively closed, and big in size, ecosystem such as the IB's suite of programmes - the potential of using this data to enrich education is enormous.

This brings data analytics and artificial intelligence (AI) to the fore. AI in this context refers to the ability of digital devices and systems to undertake tasks normally performed by humans (such as recognising speech), while data analytics is the capacity to draw together a number of sources of data in order to draw conclusions.

These offer vast potential for not only identifying learner proficiency within traditional subject boundaries but also across them and in the sphere of 21st century skills. This has the potential to significantly reconfigure how education is enacted but all three of these movements (learning progressions, ongoing assessment and the use of data analytics and AI) are next-to-impossible to achieve in a non-digital environment.

Thus the contemporary context of education calls for digitalisation to reflect the broader digitalisation of our lives. On the one hand, digital tools are increasingly mainstream in teaching, learning and assessment. On the other hand, the purpose of the use of digital tools in education remains somewhat unresolved.

Moreover, any mention of digital tools raises

important ethical questions about the unequal access of learners and teachers to digital tools – both in school and at home – and concerns about inequalities in access to quality education as a result. The sorts of data collection activities that result from movements such as learning progressions, data analytics and AI also raise issues of learner's privacy, especially with regards to how they will be used to judge performance.

A further issue in the context of the IB relates to the agency of schools. Schools offering IB programmes differ significantly around the world and in this context it is important to consider what the ideal balance is for the IB in setting requirements without being prescriptive.

This returns us to a need to focus on the purpose that needs to be achieved. If there is a desire for enhanced learning then to what extent can digitalisation support that move? How can digital assessment and digital pedagogy be best configured in order to allow for deeper, more effective or more efficient learning?

Digitalisation

The addition of the prefix *'digital'* to elements related to education is the contemporary expression of a long history of evolution in educational technology. In recent times it has been characterised by exponential growth. There is enormous scope to enhance educational opportunities but these are often forgotten in the excitement of innovation.

We are now at a point in which digital tools used for teaching, learning and assessment embody a range of exciting possibilities. Computer based testing has evolved to enable adaptive testing, the integration of data analytics the evaluation of 21st century skills, virtual reality, gamification and the use of avatars. All of these have begun to be used in educational contexts with initial research indicating positive impact on learner motivation and learning.

Nevertheless, despite all of these stimulating advances, most digital assessment remains in a phase where it is basically '*paper-on-screen*'. It is largely unchanged from traditional assessment other than learners typing rather than handwriting their responses. This is largely a consequence of the cost and skill required to create more forward looking tools, but is also due to a degree of anxiety around any change potentially jeopardising high-stakes assessments.

This is unfortunate since digital platforms can enhance assessment in a number of ways, from providing learners with new forms of representing knowledge and skills, to using peer- and self-assessment, the opportunity to move away from time specific testing, to supporting collaboration, assessing complex problem solving and enhancing feedback to learners.

Interestingly, these opportunities are much more likely to be utilised in formative types of assessment rather than in summative examinations due to the greater flexibility and lower stakes that characterise formative models.

Yet even in formative assessment they are often unable to be fully realised due to practical reasons such as the digital proficiency of teachers and learners, the availability of digital devices, the ability of teachers and learners to use data to inform practice and ethical issues around the collection, usage and storage of data.

Another challenge lies in digital assessment design. There are multiple formats that could possibly be used for assessing learners but these require good assessment design. Digital assessment is not necessarily good assessment and getting it right means starting from its purpose and recognising that the best methods are those that are able to collect useful information about the position of learners in their learning journey. Moreover, assessment also needs to be seen

for what it is - merely a vehicle to gather useful information.

These understandings focus the attention on how to manage innovation. This means avoiding the trap of using digital assessment for its own sake and instead only deciding to use it because doing so offers clear value that alternatives are unable to provide.

In addition, the flaws of traditional assessment practices do not disappear when a digital device is used. Assessment is still assessment and much of it involves placing learners in stressful situations in which the entirety of their experience in a particular domain is expected to be summarised in a few hours of testing, with results exerting a profound influence on their future lives. The traditional criticisms of this model in terms of its very limited ability to capture the gamut of learners' capacities remain intact regardless of the tools used.

While digital assessment offers a range of possibilities for rethinking assessment, these can only be optimised if there is an appetite for transformation. Many education professionals are very uncomfortable with changing the way that '*things have always been done*' and this means that much digital assessment looks very much like traditional assessment, with the opportunity of evolving approaches to the evaluation of learners entirely missed.

Best practice in digital assessment

Innovations in digital assessment are flourishing but robust academic investigation of their impact on learning is lagging far behind. In terms of assessment as a whole, best practice means that the learning of individuals is made visible through the use of efficient and well-targeted tools, in a way that can identify where learners currently are and the future steps for them to take to continue their learning journey.

With this understanding, '*best practice*' in digital assessment means that digital technologies are used in a way to enhance the achievement of this goal. This could mean that digital assessment allows for more efficient or better targeted tools that are able to identify learner's progress. Equally, best practice in digital assessment could mean that assessment is able to measure types of learning that have previously been impossible to make visible (such as 21st century skills).

The lack of a strong core of empirical evidence on the value of digital assessment, or indeed any aspects of a digital educational environment, places organisations such as the IB in a difficult situation. Making the best possible judgement about the format of assessment that will best meet the needs of its programmes and stakeholders inevitably needs to take place in something of a vacuum in which many of the necessary answers remain opaque.

To address this, however, the IB can look to what initiatives are occurring in digital assessment, the extent to which assessment bodies have been able to use them to achieve valuable purposes and the ways in which they have been implemented. To inform this, the reviewers considered the most innovative approaches to assessment that are taking place around the world. This involved extensive research in academic and '*grey*' literature.

One of the challenges to this research – as well as in evolving the field of digital assessment practice – is that many innovations are hidden behind walls constructed to protect intellectual property. It is almost impossible to find any details about innovative approaches used by many commercial organisations, for example.

Hence, the review has focused on examples of digital assessment where there is sufficient information available in the public domain to enable a clear picture of their characteristics and the way in which they are used.

Freely-available examples of state of the art use of digital assessment tend to be found in large-scale and often national or international assessment programmes. In addition to the open nature of these programmes, this is often due to their ability to spread the cost of innovation and their low stakes nature for individual learners, allowing greater room for experimentation.

Examples from international assessments include the use of avatar-like agents to support the assessment of collaborative problem solving in PISA 2015; a range of interactive stimuli including posts from a web forum in PISA 2018; the use of interactive scenarios in eTIMSS 2019; the requirement for learners to navigate a series of curated websites in ePIRLS 2016; and the use of drag-and-drop tasks in ICILS.

From national assessments examples, learners in Denmark have free access to the internet during online assessments while in Finland learners and teachers have prior access to the testing system to enable familiarisation. In Norway, all marking is fully digitised, and in Scotland, a form of adaptive assessment is used, with instant results provided.

In English language proficiency tests, IELTS has used a computer delivered version of its test for some time and is currently looking into the possibility of using video-conferencing to assess speaking skills; PTE Academic is investigating the use of AI to assess speech and/or writing; and speaking and writing tasks for TOEFL are marked through online networks.

At the same time, there is great interest in using digital assessment to measure 21st century skills but this has been stymied by a lack of definition of the constructs themselves. Nevertheless, ongoing efforts are being made such as the inclusion of modules on critical thinking, problem solving and global competency in PISA.

All of these examples have elements that the IB can learn from but, overall, the e-assessment used in the IB's MYP programme is close to global best practice, and indeed eclipses some of these much larger-scale examples.

Backwash and forewash

For digital assessment to work in harmony with teaching and learning it is important that there is a focus on the dynamics involved between them. Key to these are the participants, processes and products that comprise them. Both backwash (the influence of assessment on teaching and learning practices) and forewash (the influence of teaching and learning practices on assessment) are mediated through a range of variables, including teacher and learner attitudes and experiences, access to resources and self-efficacy. Importantly, both backwash and forewash can be both positive and negative influence.

In pedagogy, as in assessment, there are certainly many benefits in using digital tools. Similarly to assessment, however, the benefits are often different to the ambitions. Moreover, there is little evidence for the assumption that digital natives (whether teachers or learners) can more easily adapt to digitalisation than others.

The backwash of digital assessment into teaching can put immense pressure on teachers and they are increasingly required to have Technological Pedagogical Content Knowledge (TPACK) and to be digital learning designers.

Achieving this requires a multiplicity of types of support, resourcing, changes in institutional cultures, scaffolding and the understanding that there is no one-size fits all approach. This poses interesting challenges for the IB as it considers how best to support teachers in its programmes, particularly in light of its 7 year curriculum review cycle.

Seven years is a very long time in the world of digitalisation and many educational practices can be transformed within this period, potentially leading to guides becoming out-of-date increasingly quickly. Equally, constant innovation and change can be profoundly disruptive to learning and teaching and hence charting a sensible course between the two dynamics is profoundly important.

Resourcing digitalisation

A further consideration for the IB lies in the resourcing that schools require for digital teaching, learning and assessment to work in harmony with each other. All IB schools must have digital devices but the number, type and capabilities of these varies significantly. This means that teachers and learners across IB programmes have differing access and exposure to digital tools and modes of learning. Tailoring support in this context requires a deft approach.

There is extensive academic and '*grey*' literature about the digital divide between learners with greater access to digital tools and those with less, and the consequences that this can have in terms of disadvantaging some learners in relation to others. This is a major concern for educational bodies around the world and is an aspect that will not be easy for the IB to resolve. Even when learners have equal access to devices, irregular internet connectivity can undermine their opportunities to use them.

Moreover, the resources (particularly in terms of skills) required to design and develop digital assessment that makes the most of its promise, cannot be understated. Moving to interactive, augmented reality, scenario-type assessment is very attractive in theory but the practicalities of bringing this to reality are very different.

The skill levels required, not only in terms of programming expertise design but also in terms of assessment design, greatly transcend those available within most educational assessment bodies. Rapid developments in educational/instructional technology mean that skills also need to evolve continuously as educational bodies such as the IB continuously evolve their approaches to digital assessment, as well as the digitalisation of teaching and learning.

Even at the level of simple '*paper on screen*' type assessments, a high degree of expertise is required to develop the types of items that can evaluate higher-order skills in a way that both enables automatic marking and the identification of misconceptions underlying errors. Taken together, these understandings point to a need for continued upskilling of IB staff and stakeholders and ongoing debate about how best to meet the needs of learners and other stakeholders. They also raise questions for the IB about who is best to fill the roles of assessment developers going forward.

In addition, psychometricians are required if robust learning progressions are to be created. While well-defined learning progressions for domain areas – or indeed across them – offer great potential to support the tracking of learner's growth, they are not easy to develop. The research- and practice- informed initial stages can be done by those with a range of skills but collecting and analysing the statistical data to validate learning progressions requires sophisticated psychometric approaches.

Both of these have important implications for the IB's overall deliberations around digitalisation and the way that the IB addresses this within its curriculum documentation, curriculum review processes and systems to support schools.

Conclusion

This review has emphasised the need for approaches to assessment to work in harmony with teaching and learning and to be driven by a clear purpose around the

enhancement of learning.

The world of digital assessment is rapidly changing and there is enormous potential for the IB to innovate its approach to assessment. This includes moving to a model in which assessment is embedded within teaching and learning on an ongoing basis.

There are certainly a number of exciting ways in which digital assessment can provide positive support for teaching and learning but moving beyond a context in which '*new technologies are used to do old things*' (Dolan, 2013, npn) requires a vision and learning from the approaches already used in formative assessment.

[This report is a shorter version of a longer body of work that was delivered to the IB at an earlier stage in this project. As such some of the references included in the reference list are not referred to in the text. The decision has been made to keep them in this report, however, in order to provide a resource of relevant literature for the IB to draw on].

CONTENTS

ACKOWLEDGEMENTS

This study, *The Relationship between Teaching, Learning, and Digital Assessment* has been commissioned by the International Baccalaureate Organization (IBO). The first draft of this literature review was developed by Professor Kathryn Moyle, Dr Antonio Mercurio, Pru Mitchell and Kevin McRae.

Revisions were made by Anit Cherian, Dr Sarah Richardson, Dr Sladana Krstic and Dr Shani Sniedze-Gregory. Further input has been provided by Daniele Cavassa, Bethany Davies, Stewart Monkton, Abha Bhagat, Bikramjit Sen, Dr Hilary Hollingsworth, Dr Mariya Khatoon, Dr Vijaya Raghavan and Jenny Trevitt, while Gail Erwood has provided administrative support.

The authors would like to thank the IBO for the opportunity and funding to undertake such timely research. In particular, we would like to thank Dr Sarah Manlove (Research Manager, IBO), who has provided thoughtful direction, and access to key people and resources, to enable this literature review to be prepared. Thanks is also given to the Advisory Group for their insightful review and input.

ABBREVIATIONS & ACRONYMS

ACACA	Australian Curriculum Assessment and Certification Authorities
ACARA	Australian Curriculum, Assessment and Reporting Authority
ACCT	Analysing Children's Creative Thinking
ACER	Australian Council *for* Educational Research
AEA	Association for Educational Assessment – Europe
AES	Automated Essay Scoring
AGAT	ACER General Ability Tests
ALC	Assessment of Languages Competence
AP	Advanced Placement – USA
ART	Abstract Reasoning Test
ATC21S	Assessment and Teaching of 21st Century Skills
ATL	Approaches to Learning
CAA	Computer-Assisted Assessment
CARI	Centre for Assessment Reform and Innovation
CAT	Computer Adaptive Testing
CCSSO	Council of Chief State School Officers
CIL	Computer and Information Literacy
ECD	Evidence-Centred Design
ETS	Education Testing Service
EVS	Electronic Voting Systems
FLORA	Fluent Oral Reading Assessment
HL	High Level
IAEA	International Association for Educational Assessment
IAEEA	International Association for the Evaluation of Educational Achievement
IB	International Baccalaureate
IBO	International Baccalaureate Organization
ICCS	International Civic and Citizenship Study
ICILS	International Computer and Information Literary Study
ICT	Information and Communication Technologies
IEA	International Association for Evaluation of Educational Achievement
IELTS	International English Language Testing System
ILSA	International Large-Scale Assessments
MCQ	Multiple Choice Questions
MOE	Ministry of Education
MOOC	Massive Open Online Course
MyST	My Science Tutor
NAPLAN	National Assessment Program Literacy and Numeracy
NCEA	National Certificate of Educational Achievement
NCLB	No Child Left Behind

ND	No Date
NPN	No Page Number
NZQA	New Zealand Qualifications Authority
OECD	Organisation for Economic Cooperation and Development
OEI	On-screen Examination Interface
P21	Partnership for 21st Century Learning
PA	Progressive Achievement
PAT	Progressive Achievement Tests
PIRLS	Progress in International Reading Literacy Study
PSPE	Personal, Social and Physical Education
PTE	Pearson Test of English
QCAA	Queensland Curriculum and Assessment Authority
QSA	Queensland Studies Authority
SACE	South Australian Certificate of Education
SAMR	Simulation Augmentation Modification Redefinition
SAT	Suite of Assessments – USA
SEAB	Singapore Examinations and Assessment Board
SEW	Social-Emotional Wellbeing Survey
SL	Standard Level
STEAM	Science, Technologies, Engineering, Arts, Maths
TEA	Technology Enhanced Assessment
TIMSS	Trends in International Mathematics and Science Study
TIMSS A	TIMSS Advanced Mathematics and Physics
TOEFL	Test of English as a Foreign Language
TOK	Theory of Knowledge
TPACK	Technological Pedagogical Content Knowledge
UNESCO	United Nations Educational, Scientific and Cultural Organization
VCE	Victorian Certificate of Education

INTRODUCTION

Background

The relationship between teaching, learning and assessment has always been a topic of curiosity to those with a passion for education. It is a three-way dynamic that works best when in equilibrium, with each component feeding into the others and with the needs of learners central to all.

The International Baccalaureate (IB) is at an interesting time in the use of digital assessment, with the use of e-Assessment in the Middle Years Programme (MYP) but retaining traditional paper-and-pen summative assessment in the Diploma Programme.

The IB also has a unique approach to education and key principles running through all IB programmes include inquiry, contextualisation of learning, learning how to learn, community service and global understanding. A recent IB publication states that '*technology should support assessment*' (IBO, 2018, 6) but further states that this should not be done at the expense of these principles but rather to better channel them into practices.

As '*digital*' is becoming regarded as an increasingly fundamental prefix to all three components, the dynamic is inevitably shifting. But in order to leverage the benefits that digital teaching, digital learning and digital assessment can offer to enhance the quality of learning, it is essential that one understanding remains constant – that the needs of learners remains paramount.

In this context, the literature review has focused on addressing the question '*What is the current state of the art understanding of digital assessment use within the future educational landscape?*' It looks at the contemporary context within the evolution of key movements in education and identifies the role that digital assessment can and should play, and some key considerations for the IB to address moving forward.

Approach

In this literature review we examine the key trends shaping education in the second decade of the 21st century and identify how these can inform and strengthen practices within IB programmes. We explore the way in which notions of learning, teaching and assessment have evolved over time and look at how notions of assessment have both reflected and informed ideas about learning, particularly with relation to learning progressions.

We consider ways in which the introduction of digital technologies into the learning and teaching space has transformed the understandings of what knowledge, application, understanding, achievement and success mean. Moreover, we look at the impact that this has had on teaching practice.

We look at the practical and educational benefits of the inclusion of digital assessment (both formative and summative) into educational activities and investigate the field of 21st century skills, looking at evolving efforts to measure critical elements such as problem-solving, creativity and critical thinking.

Throughout the review we have retained a focus on the fundamental significance of purpose and on finding ways to ensure that the three components of learning, teaching and assessment work in harmony. This means the critical importance of retaining a clear view of objectives in determining the form and function of digital assessment, as well as ensuring that there is alignment with teaching and learning.

Considering all of these key trends, we look at the unique perspective of the IB suite of programmes and considers how best the IB can navigate the arising opportunities and threats posed by emerging dynamics.

We consider how the IB can approach the synthesis of new understandings of education with its focus on inquiry- and concept- based, contextualised, collaborative, differentiated and assessment-informed teaching and learning. Within the IB context, we focus on the evolution of digital technologies – particularly in the realm of assessment – and discuss how the IB can move forward in a way that best balances technological advances with an approach that stays true to the philosophy of the IB and its programmes.

While recognising the meta trends that are shaping contemporary education, we consider approaches that the IB can use to support the day-to-day efforts of teachers and schools to provide learners with a supportive and enabling environment in which to grow and thrive, and prepare themselves for their adult responsibilities.

We do not aim to be prescriptive around how the IB can select particular tools or resolve how best to assess learners in specific subjects. This is a very dynamic space and new approaches and tools are emerging on almost a daily basis. Nor do we set out to suggest how the IB should approach the assessment of particular domains.

We are also not working on the assumption that digital approach necessarily adds value. Poor teaching with a digital devices is still poor teaching. And poor assessment on a computer is still poor assessment. The fundamentals of good education remain constant but in the 21st century the digitalisation of the world around us cannot be ignored, and this clearly has important implications for IB programmes.

Overall, we aim to provide the IB with a review of the key considerations that need to be brought into decisions around how best to serve their schools and enhance learning opportunities for learners in IB programmes. In this literature review we set a foundation for the other activities in this project – using interviews, surveys and audits to explore the current and potential role that digital technologies can play in shaping the relationship between teaching, learning and assessment within IB programmes.

Methodology

The literature review has focused on three key sources of information. First, academic journal articles and books written by educational researchers that reflect on the key philosophies and movements in education and that also report on initiatives to enhance educational practices.

However, one of the limitation of academic literature is the time lag. It can be several years between research taking place and the publication of an article reporting on it. In a rapidly moving sphere such as digitalisation, this means that the newest practices may not yet be reflected.

To address this, the literature review has also incorporated so-called '*grey literature*' from organisations and governments who are engaged in relevant activities in the areas of teaching, learning and digital assessment. These include reports, blogs, websites, interviews and presentations which engage with the most contemporary movements in education.

Finally, the literature review has also engaged with publicly-available sample assessment materials from organisations engaged in some of the most innovative and forward-looking initiatives. These include large international bodies such as the Organisation for Economic Cooperation and Development (OECD) as well as some of the well-known assessment bodies.

The researchers involved in this literature

review undertook an approach which was informed by a combination of their own knowledge in this field, key word searches in academic databases, a web search of key organisations involved in assessment and utilising key themes that arose in the initial sources perused to lead them to other important areas of consideration.

With attention to digital assessment, the review considered a range of ways in which this has evolved and is practiced in contemporary education, including the opportunities that it offers and the key design considerations. Reviewers looked at a range of examples of digital assessment usage in international assessments, language assessments, in national assessment programmes and in initiatives to measure twenty-first century skills. Both academic and stakeholder insights were drawn on to contextualise each of these examples.

Finally, the research focused on some of the key components of the relationship between teaching, learning and digital assessment in the context of best practice approaches to education. This raised the issue of the importance of the skills of learners and teachers and the reviewers drew on research to identify best practice approaches in supporting these key stakeholders to navigate new dynamics in educational practice.

THE IB CONTEXT

It is important that this review is read in the context of the IB's four programmes and their current approach to teaching, learning and assessment, implemented in almost 5,000 schools around the world (IB, 2005-2018a). This includes the programme and subject guides that the IB creates and the range of professional development options the IB offers to teachers and school leaders in each region.

The educational philosophy of the IB has been heavily influenced by the theories of John Dewey, A.S. Neil, Jean Piaget, Jerome Bruner, Maria Montessori, Rudolf Steiner and Reggio Emilia. The four programmes focus are structured quite differently but there are common threads running through each of them.

The **Primary Years Programme** (PYP) is designed for learners aged 3 to 12 years. It is an inquiry based, transdisciplinary curriculum framework, organised through a school's Programme of Inquiry, which includes knowledge, concepts, skills, attitudes and actions around six transdisciplinary themes (IB, 2005-2018d).

The PYP provides a curriculum framework and emphasises pedagogy, concepts and skills rather than content (IBO, 2008, 1). The PYP aims to equip learners to become creative, critical and reflective thinkers who inquire into a wide range of issues and ideas of significance locally, nationally and globally. Assessment in the PYP is predominantly formative, with learners also completing an exhibition at the end of the programme.

The **Middle Years Programme** (MYP) is intended for learners aged 11 to 16 years. It is an inquiry- and concept-based curriculum framework consisting of eight subject groups or curriculum areas. Each year, learners in the MYP must participate in at least one collaboratively-planned interdisciplinary unit that involves at least two subject groups (IB, 2005-2018e).

There is an emphasis within the MYP to contextualise learning and this is done through six '*global contexts*' (IBO, 2014). Another unifying thread within the MYP are the skills known as Approaches to Learning (ATL). These skills are embedded across all subject groups of the MYP and aim to help learners to '*learn how to learn*' (IB, 2005-2018g).

Learners who complete the MYP in years three or four undertake a community project, while learners who complete the MYP in the fifth year undertake a personal project which is externally moderated. In the final year of the programme, learners can opt to complete MYP eAssessment, which enables IB-validated grades based on both the eAssessments and course work (IB, 2005-2018e). Learners who do not opt to complete eAssessments are still eligible for school-based results. The eAssessment is currently the IB's only digital assessment activity.

The **Diploma Programme** (DP) is a two-year senior secondary programme designed for learners planning on a university pathway beyond secondary school. It is for learners aged 16 to 19. It comprises a core of Theory of Knowledge (TOK); an independently-written extended essay; and the completion of a community project that directly addresses creativity, activity and service (IB, 2005-2018h). Learners take six subjects at either Standard or Higher levels.

In 2010 the IB signed an agreement with Pamoja Education for the development and delivery of a range of DP courses online (IB, 2005-2018i). At this stage there is no digital assessment within the DP but there are plans

underway to trial digital assessment in a small number of subjects in the future.

The **Career-related Programme** is the newest IB programme and it designed for senior secondary learners who wish to study a career focused course. Learners take at least two DP subjects from any area, plus a career related study and the CP core. The core consists of personal and professional skills, service activities, a reflective research project related to their career study, and an additional language development course.

The career related study is not assessed by the IB but by an approved outside provider (IB, 2005- 2018j). There is currently no digital assessment in the CP.

While **digital assessment** remains relatively limited in scope across IB programmes at present, digital technologies could easily become a more significant part of the pedagogical frameworks and assessment processes. The key here, however, is to identify the extent this would better support learners and teachers to achieve programme outcomes.

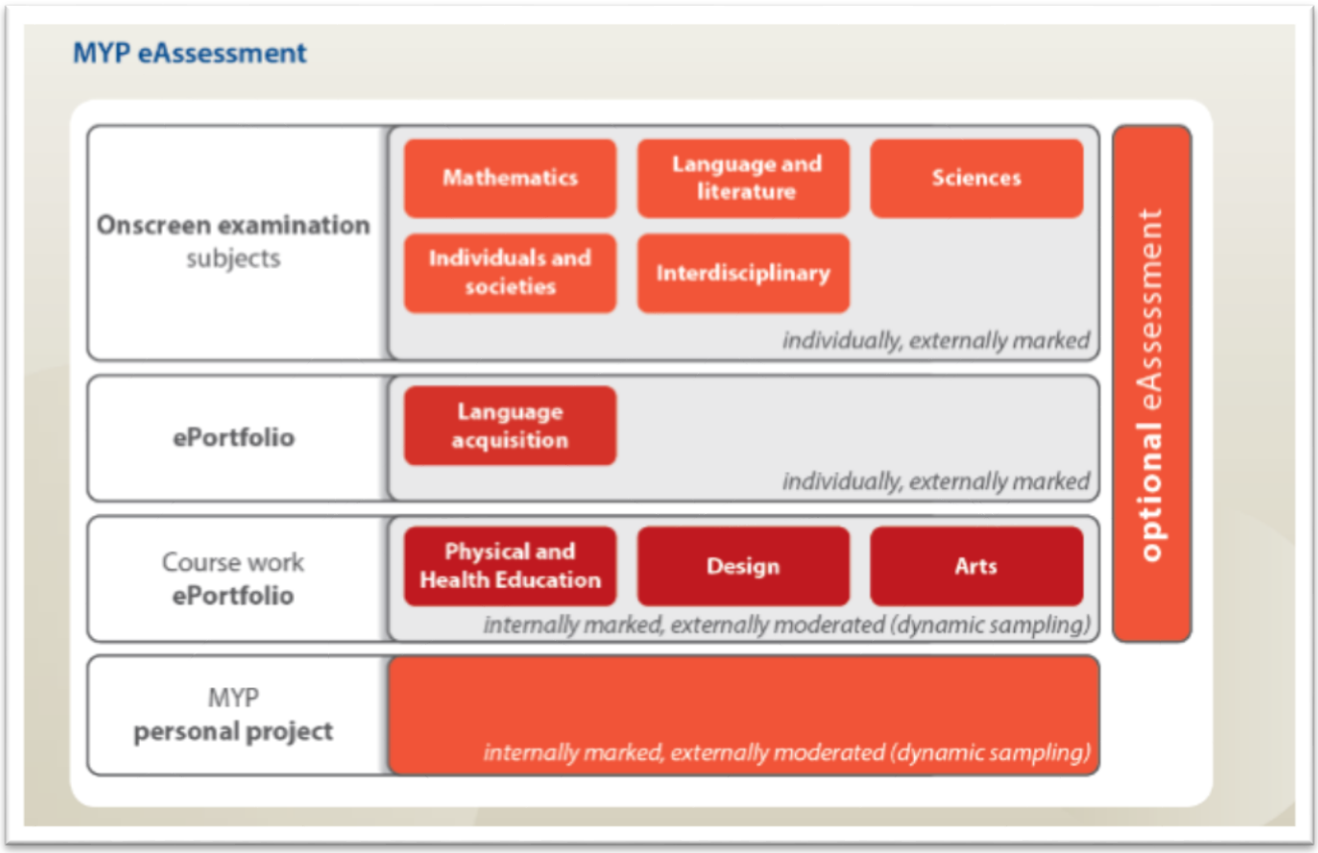

Figure 1 Availability of IBMYP eAssessments (IB, 2014, 1)

A present, the MYP has two forms of e-assessments for learners who wish to receive IB-validated grades: on-screen examinations in *Mathematics, Science, Individuals and Societies, Language and literature* and *Interdisciplinary*; and assessment via e-portfolios in *Language acquisition, Design, Arts* and *Physical and Health Education*. An e-portfolio is the process that schools use to upload internally assessed coursework for external moderation (IB 2018a, 257).

In addition, year 5 learners must complete a compulsory e-portfolio for their personal project, which is externally moderated. **Figure 1** illustrates the on-screen

assessments currently in place.

A recent publication provides details about the planned directions for IB examinations (IB, 2018a, 6). It states: "…*we strongly believe that technology should support assessment and the move towards computerized on-screen examinations will not change our principles; but it may open up new possibilities in turning these principles into practices*".

In terms of digital assessment specifically, the IB is currently in the process of increasing the opportunities for learners to undertake their examinations on-screen instead of using pen and paper.

The IB currently conducts two examination sessions per year. The DP and MYP have an examination session in May for schools in the northern hemisphere and there is also a November examination session for southern hemisphere schools. In 2017, 157,488 candidates sat for their DP examinations in May (IB, 2017b) and 16,535 candidates did so in the November 2017 examination session (IB, 2017c).

One of the options that digital assessment provides, as discussed here, is an alternative approach in which assessment is ongoing, and is built in to teaching and learning activities throughout an academic year. This could mean moving away from a notion of summative vs formative assessment, to one in which assessment throughout the year contributes to a final grade.

There are a number of considerations for the IB around digital assessment and the extent to which this does, or should, align with current practices in the digitalisation of teaching and learning. This review aims to inform these considerations through a reviews of some of the broad contexts that are of greatest pertinence in contemporary education.

One of the key considerations for the IB is that although its programmes operate in a similar way in schools around the world, schools equally need to take account of the educational contexts in which they sit. This includes a range of requirements and standards which inevitably have an impact on how IB programmes play out.

For example, the length of the MYP varies with some schools offering years 1-3, other just the first two years and others only years 4 and 5. This variability inevitably adds complexity to the context in which the IB operates and in which it develops summative assessment.

In addition, it is essential that assessment in IB programmes, as in any other educational programmes both reflects and informs learning and teaching. In the IB there is a participatory model of both curricula review and exam setting and the IB also provides support to schools through training and teacher support materials.

All of these processes and the stakeholders involved in them will inevitably be impacted as approaches to education across the IB evolve, including any increase in the use of digital assessment. This may require a reconsideration of the way in which processes and practices are managed and the skills that those involved in them require.

Finally, the traditional model of assessment in many IB subject areas involves having learners write extended, open-ended responses on the basis of authentic scenarios. This is a reflection of the focus on inquiry in IB programmes and is also a reflection of traditional methods of assessment in many education programmes around the world.

This approach has many advantages, particularly in enabling the evaluation of the ability of learners to build sustained arguments. One significant downside, however, is the ability of these forms of assessment to generate data that is suitable for data analytics.

As approaches to assessment evolve, there is

growing understanding of different ways to assess skills such as, for example, inquiry and the pros and cons of different approaches to doing so. This is taking place in parallel with, and partly driven by, the development of new tools which allow for novel approaches. As with evolutions in teaching and learning, this requires the IB to consider how best to approach assessment.

This is particularly the case if the use of the most sophisticated approaches to the analysis of assessment data (e.g. item response theory) are desired. This requires assessment to be able to generate particular types of data and inevitably has consequences for the forms of assessment that can be used.

Taken together, the unique characteristics of the IB, together with the ongoing evolution of approaches to education and the rapid development of technology, pose a number of questions for the IB. How best to navigate these will need to be determined over time as the IB continues its pursuit of the best possible ways to support learner achievement and learning. The contribution of this literature review is to raise some of the considerations that need to be addressed.

THE EVOLUTION OF TEACHING, LEARNING AND ASSESSMENT

Paying due attention to the wider context in which the curriculum, learning, and assessment systems are situated, what their origins are and how they function, is not only informative, but necessary as these wider contexts afford us a context in which to interpret contemporary changes.

Theories of learning and assessment have influenced and reinforced each other over time, with a transformation in understanding of what happens when learning takes place and what it means to know something (Baird, et al., 2017).

As education has evolved there has been a general move from '*knowledge acquisition, through participation, to knowledge creation*' (Klenowski & Wyatt-Smith, 2014, 3). Each of these stages brings with it a different concept of assessment, as Table 1 illustrates.

View of Learning	Dominant mode	Assessment form
Acquisition and adoption of what has been taught – subject matter, knowledge and representations	Monologic Teacher-dominated Didactic	Written exams Correct answers How much has been learnt Answers reflect instruction
Process of participation in learning communities with co-construction of knowledge and individual sense-making	Dialogic Participation in learning communities	Performance & demonstration Dialogue Evidence of sense-making and meaning
Process of creativity and developing new materials and conceptual artefacts, building knowledges as individuals and in groups using mediating artefacts	Trialogic ICT-mediated Digital learning environment	Collaborative, generative Creating and developing new material and conceptual solutions Critical and reflective forms of learning Authentic tasks using ICTs

Table 1: Towards Future-Oriented Assessment, Klenowski & Wyatt-Smith 2014, 3

Inevitably this shift has not been linear in nature, and much education continues to conform to a didactic model in which the purpose of learning is acquisition and the purpose of assessment is to identify how much '*has been learnt*'. This model is gradually being transcended, however, and we are now seeing greater focus on participation and creativity, with an associated focus on meaning, authenticity and application.

In all of these models, key underlying questions are '*why learn*' and also '*why assess*' (Rowntree, 1987). The focus on purpose persists in contemporary education and should be fundamental to any consideration of digitalisation of teaching, learning and assessment.

Three of the key themes in contemporary learning that are important for the IB in the context of teaching, learning and digital assessment are: learning progressions, ongoing assessment and the role of data analytics and Artificial Intelligence (AI). The IB is well placed to incorporate all of these in its programmes.

Learning progressions

In considerations of teaching, learning and assessment a major disrupter in educational practice is the notion of '*learning progressions*'. In contrast to previously widely-accepted notions about fixed or innate learner ability, educational professionals are increasingly understanding that all learners make progress

in their learning at different paces. An understanding of learning progressions gives rise to a focus on personalising learning journeys (Dweck, 2015). It also highlights new responsibilities for assessment.

Underpinning theories about learning progressions, is the view that individual learners can improve their learning from whatever position they may find themselves on the roadmap of learning. The proposition is that all learners, with the right mindset, effort, support, and tailored teaching, can make a year's progress every year and that '*every learner is capable of making further progress*' (Masters, 2016).

A central principle is that learners acquire a '*growth mindset'* in order to foster gains in achievement (Dweck, 2015). From this perspective, '*ability*' or '*intelligence*' or '*capability*' is not seen as a fixed trait in any child, but rather all learners are able to continue learning. There should be no plateauing to learners' abilities and capabilities.

The metaphor of the roadmap is useful here for conceptualising learning progressions (Heritage, 2009). Learners are on a journey. All learners can take this journey, but they require a roadmap. At various points along the journey, they may have to evaluate where they have been and where they might want to go next. They also require descriptors to map their progress along the journey.

The role of assessment within this context becomes that of monitoring progress and advising both learners and teachers on next steps. Assessment can also generate the data required to build learning progressions to lay out the important ideas, principles, and skills of the domain in a sequence that represents how competence in a domain develops (Hayward et al., 2018; Jackson & Turner, 2018; Mosher & Heritage, 2017).

Learning progressions are of interest in the context of the IB, in that the structure of its programmes are ones that would be relatively easily able to incorporate learning progressions. Depending on how the IB chooses to frame, or even adopt, learning progressions this could allow for a very different model in IB programmes.

For example this could lead to a flexible and evolutionary nature of learning in IB programmes that would allow learners to move through the curriculum at different speeds, increasing competences in different areas of learning as they progress.

It is often thought that the adoption of learning progressions could allow education to break away from age-bound grades. This is indeed possible in theory but in reality it is also essential to consider the importance of the social dimensions of learning and of the value of peer relationships. The reality is also such that a break from age-bound grades must be considered within current qualification models for assessment where the goal is to work with systems of accreditation and recognition.

One of the key assumptions of learning progressions is that regular data is available to shed light on learner's progress. This points to a need for ongoing assessment – not in terms of regular formal examinations (although this is certainly one approach) but more in terms of assessment being an everyday part of teaching and learning.

Ongoing assessment

Assessment is often defined in mutually exclusive ways – for example formative and summative; or assessment for learning, of learning or as learning, are common distinctions. There is a growing realisation, however, that assessment has multiple purposes – for multiple stakeholders - which often have to be satisfied simultaneously (Klenowski & Wyatt-Smith, 2014; Newton, 2007).

Alternatively, assessment purposes are

sometimes collated in the literature to suggest there may be one overriding purpose that contains multiple sub-purposes, for example, Masters (2013, 6-7) describes: '*… a simple unifying principle; namely, that the fundamental purpose of assessment is to establish where learners are in their learning at the time of assessment'.*

A focus on learning progressions changes considerations around assessment and highlights the need for it to be embedded in teaching and learning as an ongoing activity. This enables a regular flow of insights to teachers and learners to guide next steps.

As Heritage (2009, 5) suggests, the fundamental role of assessment is increasingly being understood as moving on from the '*past to present models*' of the past that concentrated on accountability and certification, to the provision of empirical insights to enhance teaching and learning, or *'present to future models'.*

In the context of IB programmes, other than in the PYP, the current focus is very much on formative and summative assessment. Formative assessments is seen as something that happens in schools and between teachers and learners while summative assessment is regarded as something externally mediated by the IB to make a judgement about the candidate (IB, 2018a).

The distinction is highlighted by the shift in nomenclature here - from learner to candidate. Moreover, the IB identifies how formative and summative assessments are used differently (IB, 2018a). The differing dynamics of the two assessment processes are conveyed as follows: formative assessment is portrayed as inclusive, consultative, iterative and continuous; summative assessment is independent, evaluative, definitive and once off.

As the IB emphasises, however, it is vital that all assessment has integral links to teaching and learning and that summative assessment should not just be '*an activity conducted after learning has taken place'* (IB, 2018a, 45). This indicates that a model of ongoing assessment should be easily adopted within the IB, bringing together the purposes currently understood as formative or summative to an approach which focuses on giving learners the information they require to advance in their learning.

Within this, Rowntree's 5 dimensions of assessment (why assess, what to assess, how to assess, how to interpret, how to respond) (1977) remain paramount, as does a requirement to ensure validity, reliability, fairness and feasibility.

Data analytics

One of the interesting areas of innovation in education are the possibilities wrought by the use of data analytics, which can be considered as a renaissance (Hill & Barber, 2014) or part of an '*an emergent reality*' (Cope & Kalantzis 2016, 13), a situation that is already with us, but one where the journey has barely begun.

Data analytics and the use of artificial intelligence (AI) opens up the increased ability to collect evidence on learning. This includes: identifying misconceptions; the ongoing availability of empirical evidence; the ability to record interactions, hence opening up the possibility of measuring 21st century skills such as collaboration; and new opportunities in data analytics to inform learning resource design.

Such capabilities immediately highlight the importance of the relationship between learning, teaching and assessment, and ways in which it can be reconfigured. The greatest potential is for personalised learning - directly targeting resources to meet the individual needs of learners.

Significant potential lies in the ability of technology to take on some of the roles traditionally played by teachers. While automated marking of multiple-choice type

assessment items is already very common, automated marking of essays and of speech are already beginning to be used at what is currently the cutting edge of contemporary assessment.

Moreover, the recommendation of resources based on assessment outcomes, computer adaptive testing to pinpoint a learner's strengths and weaknesses and give them progressively easier or harder assessment items, and virtual worlds that challenge learners to solve problems, are already in existence.

In this, as in the other key themes, the IB is uniquely well placed to leverage the value offered by data analytics. Arguably, the network of schools around the world that offer IB programmes form an enormous ecosystem. With the right approach this ecosystem could be drawn on to generate dynamic insights into teaching, learning and assessment that could support improvements in learning and teaching around the world.

Leveraging data analytics would allow educational activities within IB programmes to focus on setting high expectations for learners and fostering and nurturing a growth-mindset in learners while supporting forward-looking curricula that enable personalisation and the nurturing of 21st century skills (Masters, 2013).

Learning analytics could also provide individual learners with the insights they need to enable targeted support and progression at different rates, while teachers can push aside some of the burden of administration in order to become dynamic activators and coaches.

This potential does, however, have to be couched with a degree of caution about ethical issues. There are concerns around the extent to which it is socially desirable to be measuring and collecting data on everything that learners do and the way in which this could drive undesirable impacts on learners. Ultimately the benefits of data analytics need to be weighed against these important considerations.

Digital underpinnings

The three key themes identified here – learning progressions, ongoing assessment and data analytics– hold out great potential to enrich the educational offerings within IB programmes, as they do in other educational systems. But doing so in a context which continues to rely on hand-written assessments, will prove nigh-on-impossible.

Reflecting the digitalisation of human interactions, professional practices and global markets, education is in an interesting position in relation to digitalisation. On the one hand, digital tools are increasingly mainstream in teaching, learning and assessment. On the other hand, the purpose of the use of digital tools in education remains somewhat unresolved.

As EdTech grows in its reach and significance, more and more questions are being asked about its impact on education and the extent to which it adds value to teaching and learning – or does not. Indeed, when technological tools are integrated into educational practices (regardless of their level of sophistication) the question '*to what end*' needs to be asked.

As notions of learning, teaching and assessment have progressed over time, so have the ways in which they are transacted. It is easy to forget how recently, and how quickly, digitisation in education has grown. For example, the first use of the word MOOC was only a decade ago (Stommel, 2015). Inevitably, digitisation of education has occurred in a step-wise manner over a lengthy trajectory (Cox 2018; 2013).

As technologies have evolved over time they have gradually progressed from ones that enable static or instructive affordances (in which information is transmitted one way) to

those that enable collaborative and productive affordances, allowing flexible representations to be adjusted and shared (Bower, 2008).

Digital technologies allow for the blending of learning and assessing so that learners can be assessed as they interact with learning materials, and provided with step-by-step feedback to assist with their learning. Digital technologies enable communication and collaboration across electronic networks, locally and globally.

In this context, '*digital*' refers to electronic technologies that generate, store and process data. This includes video, film, multimodal texts, podcasts, voice recognition, sketching programs, animation, scientific calculators, graphic calculators, and music-writing software programs. Devices include mobile phones with their wealth of functionality, 3D printers and scanners, as well as computers, tablets and other digital equipment.

As with prior technology, there is a tendency to be distracted from the purpose of the use of devices in teaching, learning and assessment and instead to focus on their functionality. But the only value of digital technologies lies in how they are used and in their ability to '*make educational opportunities possible*' (Blundell, Lee & Nykvist, 2016).

In line with this, previous research into the use of technology in IB programmes suggested that as digitalisation grows, attention needs to continue to be paid to the purpose and to the goal of enhancing learning. For example, Cooker, Crook and Ainsworth (2010) identified the danger of focusing on technology use per se, rather than pedagogy through technology.

Similarly, Bergeron (2014, 54) identified that in PYP documents, technology is a '*vehicle for supporting learning and not a singular tool to be mastered*'. Cooker, Crook and Ainsworth (2010, 91) further noted that the big challenge in technological integration is to give teachers time and space for '*invention, reflection and sharing*'

Driving positive change

As this section of the literature review has highlighted, a number of the key trends in contemporary education – learning progressions, ongoing assessment and data analytics – are greatly facilitated by a digital setting. And yet '*digital for its own sake*' is not a positive direction for education to take.

Hence, a key consideration for the IB is how to best chart a forward direction for its programmes that optimise the potential that digitalisation have to offer, but does so in a way that enables the IB to fulfil its philosophy and objectives, including that of inclusivity.

The IB now incorporates digital assessment in its MYP summative assessment, although schools can choose whether or not to sign up to this mode of assessment. In order to support discussions within the IB about the future for digital assessment within its programmes, the following section considers good practice in digital assessment and the implications that this has for the IB's ability to support positive backwash into teaching and learning.

DIGITAL ASSESSMENT

In the context of constant evolutions in the education space, digital assessment can be seen as an exemplar of all that the digitalisation of education has to offer, and where it can go wrong. The issues involved present important lessons for the IB in how it innovates assessment practices while maintaining a firm hold on the purpose of assessment and keeping the needs of its constituents in mind.

This section reviews digital assessment from a number of perspectives – its evolution over time, the opportunities it can offer, digital assessment design and approaches to educational innovation. It identifies key considerations for the IB, and recognises the need for a focus on the purpose of assessment.

It is important to note the use of the term '*digital assessment*'. This term encompasses assessment that is delivered through a range of digital devices but that does not necessarily require an internet connection.

Crucially, the notion of '*best practice*' in digital assessment cannot be separated from the notion of '*best practice'* in assessment. Digital assessment is simply assessment using digital tools. As this section identifies, the rapidly evolving nature of digital technologies means that scholarly work to identify the impact on learner's achievement and learning has not been able to keep up.

Hence, the arena of digital assessment is very much an evolving one in which identification of those technologies that can support enhanced forms of assessment is ongoing. In this context, this section looks at how digital assessment has evolved over time, how it is used around the world, digital assessment design and the implications for the IB

The evolution of digital assessment

Classical methods of assessment involved the use of a paper question form and handwritten responses. This has been the dominant format of assessment throughout much of the history of education and – globally - remains so today. Gradually, however, digital assessment has begun to gain a foothold, not only due to the widening availability of technology, but also due to the benefits that it can offer.

Digital assessment can be defined as '*the presentation of evidence, for judging learner achievement, managed through the medium of computer technology*' (NZQA, 2015, npn). Importantly, this should not be taken to be synonymous with online assessment, as internet access is not a requirement for digital assessment, with the focus instead on hardware, software and networks.

Other common terms for digital assessment are *'on-screen assessment', 'e-assessment'* and *'technology enhanced assessment'*.

An attempt to characterise the likely evolution of digital assessment was made as far back as 1989 when Bunderson, Inouye and Olsen (1989) predicted four generations of computerised educational measurement, namely:

- **Generation 1:** Computerised testing (administering conventional tests by computer);
- **Generation 2**: Computerised adaptive testing (tailoring the difficulty or contents or an aspect of the timing, on the basis of examinees' responses);
- **Generation 3**: Continuous measurement (using calibrated measures to continuously and unobtrusively estimate dynamic changes in the learner's achievement trajectory);
- **Generation 4**: Intelligent measurement

(producing intelligent scoring, interpretation of individual profiles, and advice to learners and teachers by means of knowledge bases and inferencing procedures).

As Redecker and Johannessen (2013) suggest, the first two generations are now relatively common - the challenge faced currently is to transition to Generations 3 and 4. This points to a future in which data analytics and personalisation are likely to become increasingly mainstream.

Digital assessment has evolved since the 1970s when Computer-Based Testing (CBT) began to be introduced (Burkhardt & Pead, 2003). As the cost of digital technologies has fallen, and the benefits of including multimedia tools and of automatic marking has begun to be recognised, digital assessment has become increasingly common, particular in large scale assessments where its logistical convenience is of a significant benefit (Klerk, 2012).

For much of its evolution, however, digital assessment has looked, and performed, much like traditional pen and paper models. The **format** of assessment items (as assessment experts refer to questions) has remained largely the same, with learners either responding to multiple choice items by clicking on one of a selection of options or by typing a response to an open-ended item. Beyond the skills to type rather than handwrite, and to be able to use the computer, there has been very little impact on the experience of learners. This overwhelmingly remains the case today.

Where the impact has been significant however, is for the entities responsible for assessments. Depending on how it is implemented, digital assessment has numerous advantages over pen-and paper assessment in terms of costs, security, ease of marking responses and flexibility. Although any form of assessment has its detractors, these benefits have made digital assessment highly popular with various education systems, and in particular it is widely used in large-scale assessment.

In the 2000s, developments in digital assessment began to incorporate **Computer Adaptive Testing** (CAT). This is a new paradigm in digital assessment in which it is possible to alter the level of difficulty of an upcoming item depending on the response of a candidate to a previous item (Papanastasiou, 2003).

Due to the level of sophistication required, however, CAT remains relatively limited in scope and – where it is used (for example in the Scottish National Standardised Assessments) – it is often done using a modular approach. Moreover, CAT is not suitable for open-ended responses such as essays and other descriptive type responses (Clarke & Dede, 2010).

Artificial Intelligence (AI) is able to overcome the latter challenge (Chung and Baker, 2003) through the use of Natural Language Processing (NLP) to analyse and compare open-ended responses, hence enabling a combination of essays and short response items in CAT (Looney, 2009).
Beyond summative assessment, AI is also able to interrogate data collected in teaching and learning systems in order to provide learners with real-time feedback. This helps educators and learners to identify areas for improvement and pathways towards improvement (Knight, 2009).

A further development in digital assessment is in the evaluation of **21st century skills** such as critical thinking, problem-solving and teamwork/collaboration (Ananiadou & Claro, 2009). In theory, digital assessment may provide opportunities for both developing and assessing these skills (Bennett, 2001; Mislevy *et al*, 2001).

In the last decade or so, digital technologies have opened up some possibilities for increased complexity, sophistication and

intellectually-challenging assessments. There are pockets of research where investigations are being undertaken into how digital technologies may assist with the assessment of multimodal observations (e.g. classroom interactions, collaboration, interviewing skills), and content-based constructed-responses such as mathematical equations (O'Leary et al, 2018).

The assessment of 21st century skills is, however, problematic as the constructs themselves remain relatively vague, inhibiting the development of assessment instruments that can truly be proven to measure the skills that they claim to measure. One approach that may help overcome this is the advances in **Virtual Reality** (VR).

When applied in a learning environment, VR can provide the opportunity to immerse learners into a real-world environment in which they are, for example, required to problem solve in collaboration with other learners. This can be used to measure skills in communication and collaboration or provide the opportunity for the simulation of a real-world application of an experimental procedure (Psotka, 2013; Reiners, Gregory & Dreher, 2011).

For example VR has been used in the assessment of emergency medical students (McGrath et al., 2017) and to measure cognitive load among Japanese learners (Hashimura et al., 2018).

Gamification is another trend in digital assessment, as well as in teaching and learning. Gamification allows a range of complex situations to be simulated for both individuals and groups to engage with (Ifenthaler at al., 2012).

Again, this provides the potential for assessing 21st century skills and for presenting learners with real world situations in which they are required to apply their skills and knowledge. In both VR and in games, every move, interaction and decision can be tracked, producing a rich seam of data that can diagnose strengths and areas for improvement.

Gamification of peer-assessment has been shown to increase the quality and quantity of evaluations (Tenório et al., 2017) while gamification can also enable data to be collected on learner anxiety, boredom and flow in relation to performance, as Ketamo et al. (2018) have shown by using SmartKid Maths.

One of the more recent advances in digital education is the use of **Avatars** – 3D computer animated characters who can interact with learners. Avatars can be used for supporting teaching and learning as well as assessment (Adamo-Villani & Dib, 2016) and Avatar-like approaches have been incorporated in PISA to assess problem-solving and collaborative capabilities (Conley, 2015). Avatars have also been used in other education contexts, such as to evaluate communication skills among medical students (Kava et al., 2017).

While it is easy to get excited by the potential that CAT, AI, VR, gamification and avatars have to offer, the reality is that the majority of digital assessment currently in existence does not make use of any of these. Instead, as identified above, much digital assessment looks very much like paper-and-pen based assessment.

Opportunities of digital assessment

The common approach to digital assessment, also referred to as '*paper on screen*' simply substitutes a computer screen for a test form and learners respond by clicking on, or typing in, their answer. Substitution of this kind does not involve any functional change.

Paper on screen is often referred to as the substitution component of the SAMR framework devised by Puentedura (2006) (the other elements are Augmentation, Modification and Redefinition). It is important

to note, however, that the SAMR model has only limited application in education as it does not consider purpose but only form, and in education, ***purpose*** is of critical importance.

The reason for the dominance of '*paper on screen*' in digital assessment around the world lies in its simplicity – it is relatively straightforward and cost effective to simply replace a paper-based test with the same format on screen. '*Paper on screen*' tends to be the first step in digital assessment, for example it is regarded as the '*transition'* phase by the New Zealand Qualifications Authority and the IB refers to this as a '***starting point***' for on-screen assessment (IBO, 2018a)

While 'paper on screen' is probably the most common form of digital assessment, it misses many of the benefits that digital assessment can deliver. This brings us back to the purpose of assessment. If the use of digitalisation is to enhance education within the IB – add value for learners and teachers, deepen inquiry, enhance globalisation and better prepare learners for the future – there are big questions around whether a '*paper on screen*' approach is adequate.

When digital assessment is used as a way to add value to the learning-teaching-assessment cycle, it is able to offer a range of benefits. For example, Timmis et al (2016) present eight distinct areas of opportunity that may enhance assessment:

- **New forms of representing knowledge and skills** - text, image, video, audio, data visualisations and haptics (touch). Such new tools allow assessments to be designed in different forms and for achievements and progress to be documented differently (e.g. e-portfolios, virtual worlds and immersive environments, breaking simulations using haptics).
- **Crowd sourcing and decision-making opportunities in assessment** - there is the potential to provide learners with new or alternative decision-making opportunities in assessment (e.g. peer- and self-assessment, use of electronic voting, ranking methods and crowdsourcing of grades in schools).
- **Increasing flexibility** - the potential for assessment to be less time critical and location specific. (E.g. assessments can take place in multiple locations and over different timescales).
- **Supporting and enhancing collaboration** - digital technologies can support collaborative learning and assessment practices (e.g. co-evaluation and peer-to-peer assessment).
- **Assessing complex problem solving skills** - new opportunities to assess more complex skills, such as hypothesis testing, role-playing and problem-solving through the use of simulations and immersive environments (e.g. the SimScientists project, EcoMuve, Quest Atlantis).
- **Enhancing feedback to learners** - improving the quality of feedback and the way in which it is delivered, is critical to improving the use of assessment for learning. (E.g. using of online discussion tools, immersive environments and social media such as blogs and wikis; novel forms of feedback for individuals and for groups of learners, through interactive or rich media elements, stealth assessment where feedback is integrated into learning activities).
- **Innovation in recording achievement** - the use of online badges accreditation system (e.g. Mozilla's Open Badges Infrastructure).
- **Exploiting learning analytics** - combining different datasets to provide aggregated assessment data. The interrogation of increasingly large and complex data sets about an individual learner can be used to understand their needs better.

In many education programmes, these benefits are more often seen in formative

rather than in summative models of assessment where there is more flexibility for experimentation and for models of assessment (such as peer and self-evaluation as well as group work) that do not need to generate a single grade or mark for individual learners.

Truly leveraging the opportunities that digital assessment can offer to enhance learning and teaching requires not only technological advances – many of which are already in place – but also paradigm shifts among educators as their roles evolve. Importantly, it also requires ensuring that digital assessment does not deepen the divides that already exist between different groups of learners. In addition, digital assessment raises a number of practical consideration (Oldfield et al., 2012):

- Digital proficiency of teachers and learners:
- Availability of digital devices;
- Collection and use of feedback to support learners to improve (Draper, 2009);
- Ability of teachers and learners to use data to inform practice (Goldstein, 2012) ; and
- Ethical issues around the collection, usage and storage of data (Facer, 2011, 2012).

Digital assessment design

The potential value that lies in digital assessment is only able to be optimised if assessment tasks themselves are well-designed. This means that the starting point should be not the functionality of a digital system but the purpose of the assessment.

Two points about the design of assessment tasks are important. First, Masters (2013, 6-7) urges educators, assessment designers, and policymakers not to make too much of the possible differences in assessment tasks and focuses instead on the purpose of assessment itself.

Rather than beginning with a belief that some methods of assessment are intrinsically superior to others, this principle recognises that the most appropriate method(s) in any assessment context are those that provide the most practicable, valid and reliable information about where learners are in their learning.

Second, whilst there is no doubt that the design of assessment tasks is of critical importance to the quality of information that is elicited from them, it is useful to remember that the assessment procedure itself has no intrinsic value of its own. Instead, what is vital is the inferences that are made - with assessment tasks merely vehicles to gather useful information (Masters, 2014a).

As such, digital assessment should be regarded as a **means to an end,** not an end in itself, with a focus on the value of the insights that can be drawn from it (Cronbach & Meehl 1955; Messick, 1989). Moreover, it is important to emphasise that *'assessment'* is not synonymous with 'examination'. Instead, there are many potential ways of assessing learners' skills and knowledge beyond examinations.

In terms of assessment items, in 2009, Scalise developed a taxonomy of 28 different types of digital assessment items which incorporate those with differing degrees of openness or constraints (from multiple-choice options to fully open essays) and differing degrees of complexity. From most constrained to least constrained, item types are: multiple-choice questions (MSQ); selection or identification; rearrangement or reordering; substitution or correction; completion; construction and, finally, presentation.

In terms of complexity, the simplest MCQ would be one with true or false response options, while the most complex would be one using new media distractors. For presentations, the least complex would be one presenting a project and the most

complex would be one presenting a diagnosis or teaching something. For something that is now a decade old, there are clearly limitations to the coverage of the taxonomy, but it remains one of the most comprehensive.

More recently, efforts to list item types include the following identified by Digital Chalkboard (2018): *Technology-enhanced items*: where learners use technology to respond to an item, for example dragging and dropping text; reordering text or highlighting text; and *Performance tasks*: where learners have to apply their skills and knowledge to respond to complex real-world problems.

Bearing the enormous array of opportunities in mind, the IB's approach to assessment design needs to return to the key points around purpose and insights. This implies a need to carefully consider the role that assessment (whether digital or not) plays throughout IB programmes. While much of the focus (other than in the PYP) is on summative assessment in order to demonstrate achievement, this does not mean that the structure of examinations that is currently in use is a requirement.

Instead, deep thought needs to be given to the role that assessment can play in enhancing IB programmes beyond an end-of-programme grade and beyond leaving formative approaches up to schools and teachers. Instead, a consideration of the needs of learners, and the best way to prepare them for their future lives, can usefully drive a debate around the form and function of assessment.

There are many ways in which to achieve the purpose of assessment, that go beyond the common duality of closed and open response tasks. Assessment tasks can include for example, multiple-choice questions, short-constructed responses, extended responses, essays, reports, journals, case studies, examinations, vivas, portfolios, research assignments, performances (e.g. dance, musical performance, presentation, exhibition) and artefacts (e.g. paintings, products, designs).

Crucially, the choice of task needs to be driven by consideration of what insights into learner proficiency it is able to elicit, and the value of these insights to inform teaching and learning practices. While digital assessment is a focus here, another element to be carefully considered is whether – in light of the purpose of a particular assessment – paper-and-pen, digital tools or other formats are most suitable to achieve the ultimate goal.

Managing innovation

One of the key issues in digital assessment, and in the extent to which it is able to positively influence teaching and learning, lies in how innovation is managed. The education sector has been burnt by enthusiastic uptake of failed innovations too often, and the challenge of digitalisation is in being innovative in a cautious way.

Innovation is often considered as the use of something new but in education it is more usefully conceptualised as '*the* successful *introduction of a new thing or method*' (Brewer &Tierney, 2012, 15). Innovations in education often drive efficiency, and this is equally the case in digital assessment. But driving enhancements in the quality or value of learning is a more fundamental – indeed essential – goal.

One of the challenges in educational innovation is that evidence of successful application elsewhere can be difficult to find, with empirical evidence of success often unavailable and best practices not widely disseminated (Polikov, 2017). Two particular challenges are that firstly trials are often too small to be generalizable, and secondly, much of the technology behind digital assessment is subject to proprietary protection.

These factors mean that the education sector tends to lack understanding of how learners

learn in a digital environment. For example, measures such as learning outcomes and self-assessment can indicate whether digitalisation has led to positive outcomes (Serdyukov, 2017) but its impact on learning processes is often poorly understood.

There are, however, useful sources of insights into innovation. For example European Schoolnet (2019) use rigorous research methods to develop innovative educational tools and resources which are then tested in order to arrive at results that can be effectively mainstreamed. Their 'Living Lab' seeks evidence for the effective use of technology in schools and highlights the implications for using technology in teaching and learning (Marcus 2012; Morais et al., 2004). A key focus of research should also be the way in which learning activities using technology can best be implemented (UNESCO, 2013).

STATE OF THE ART DIGITAL ASSESSMENT EXEMPLARS

As the previous section has noted, despite all of the talk about AI, VR and avatars, much digital assessment is '*paper on screen*' in nature. This is partly because the cost of using anything more innovative can be prohibitive but also because the appetite for taking risks in the world of assessment is often constrained by the more pressing demands of reliability, validity and the high stakes of summative examinations.

This needs to be understood further in a context in which there is very little empirical evidence of the impact of digital assessment on learners. While the notion of '*best practice*' in assessment has had decades of research to identify beneficial approaches, the relatively limited amount of digital assessment that is used, the conservative nature of the formats that are employed and the reluctance of educational bodies to take risks combines to a context in which '*best practice in digital assessment*' has yet to be determined.

In this context, the notion of what constitutes '*state of the art*' digital assessment remains extremely opaque. Literature focuses on initiatives in digital assessment but not on subjecting these to robust critique. The field is too immature for any conclusions to be drawn about what does and does not work. Moreover, the small scale of most studies – sometimes with only a handful of learners – has not enabled the development of a robust empirical analysis of existing evidence.

In the absence of definitive insights into the benefits or disadvantages that digital assessment implies for learners, this review focuses on the use of digital assessment in large scale assessment activities.

The most fertile ground for innovation tend to lie in these large-scale assessments, often at the international level, and in those conducted by prominent examination boards for practical reasons - at a global scale, the costs of using innovative approaches to digital assessment can be spread across countries.

Moreover, the fact that much large-scale assessment is low stakes for learners (albeit high stakes for countries) means that there tends to be more appetite for risk-taking. In addition the relatively lengthy assessment cycles provide time and space for innovation that can be squeezed out in other assessment contexts.

The advantages of reviewing these digital assessment activities here are twofold. First, the assessment instruments tend to be subject to lengthy periods of gestation in which large numbers of education and assessment experts are able to have input. This means that the approaches used have not just been designed in isolation but have been debated and critiqued from many different angles. This offers the resulting assessment instruments the best chance of being *'state of the art'* in the current context.

Second, the assessment tools used are developed by the world's leading assessment experts. These are organisations who have been developing assessment over many decades and who have some of the foremost global experts in assessment design to hand. Their access to the world's best know-how and experience means that their digital assessment designs are likely to be as good as it is currently possible to achieve.

It should be noted, that even in the case of prominent digital assessment there remains a paucity of research to definitively identify what works well and what is able to lead to enhanced experiences and outcomes for learners.

Nevertheless, the examples provided below are of direct relevance to the IB since the

enormous scale of its assessment activities (particularly in the Diploma Programme) mean that there is scope for incorporating many of the innovative practices. Indeed, the approach used in the Middle Years Programme on-screen assessment already mirrors many of the innovative approaches described below.

International assessments

The 1990s and 2000s saw the arrival of three major international testing programs:

- The International Association for Evaluation of Educational Achievement (IEA)'s Trends in International Mathematics and Science Study (TIMSS) (introduced in 1995);
- The Organisation for Economic Cooperation and Development (OECD)'s Programme for International Student Assessment (PISA) (introduced in 1997); and
- The IEA's Progress in International Reading Literacy Study (PIRLS) (introduced in 2001).

These have since been supplemented by two more IEA studies, the International Civic and Citizenship Education Study (ICCS) and the International Computer and Information Literacy Study (ICILS). Meanwhile the OECD also carried out the Assessment of Higher Education Learning Outcomes (AHELO) Feasibility Study between 2010 and 2012.

National testing programs, such as the *No Child Left Behind Act (2001)* (introduced in 2002) and the *Race to the Top* reforms introduced in 2010 in the United States of America (USA), and the National Assessment Plan – Literacy and Numeracy (NAPLAN) introduced in 2008 in Australia, are smaller-scale, national iterations of these international testing programs.

The main rationale for large scale assessments is education system accountability, and they tend not to provide results at the level of individual learners, or schools (Tobin et al., 2015). Important work has been done to develop new technology-based forms of digital assessment, including for national assessments of ICT literacy (Ainley et al, 2012) and international assessments of digital reading (Thomson & De Bortoli, 2012).

PISA

Digital assessment has been used in PISA twice – in 2015 and in 2018. In 2015, a module on collaborative problem solving was included. The Assessment Framework for this module includes tasks such as 'The Aquarium' and 'Class Logo' which introduce the use of avatar-like characters that enable learners to engage in collaborative problem solving in order to design and make judgements (OECD, 2018).

Learners are required to engage in a typed discussion with two other characters (or agents): one who represents behaviour that is supportive of collaboration and another who represents behaviour more associated with someone who is unsupportive of collaboration.

These tasks are designed to measure collaborative skills such as discussing perspectives, collaborative interaction to solve problems, communicating with team members and negotiating meaning. This is an example of one of the most advanced forms of digital assessment that has been implemented on a large scale.

In 2018, PISA incorporated digital assessment for the second time, following its use in 2015 when it was used to assess collaborative problem solving. This involved the inclusion of stimuli about a real situation and a series of multiple-choice and short constructed response tasks linked to each one.

In contrast, in tasks for the science assessment such as the 'Bee Colony Collapse Disorder' unit presented learners with a text and graph and the 'Fossil Fuels' unit

presented learners with a diagram of carbon cycles, a short text and a graph (OECD, 2015) which are more traditional forms of digital assessment.

In 2018, PISA focused on reading and released items (OECD 2018) include the 'Chicken Forum' unit which presented learners with posts from a web forum and again gave learners a series of multiple choice and short constructed response tasks.

Another unit, 'The Galapagos Islands', presents a context in which learners are reviewing a website, with similar question types. While attractive and stimulating, these items are similar to the type of items that the IB uses in its MYP on-screen assessment.

TIMSS

In 2019, TIMSS is commencing a transition to a digital format, with around half of participating countries choosing e-TIMSS and the other half using the traditional paper-based test. The digital version includes problem solving and inquiry tasks which aim to have learners apply their knowledge of mathematics and science to real world problems. As the IEA (2018a, npn) states, these tasks:

'involve visually attractive, interactive scenarios that present learners with adaptive and responsive ways to follow a series of steps toward a solution … which learners find … engaging and motivating … [and which] may provide information to help improve instruction'.

While sample items from e-TIMSS are not available at the time of writing, the description indicates that the approach is similar to the one used in MYP e-assessment.

PIRLS

The introduction of eTIMSS follows the introduction of ePIRLS in 2016. Similarly to eTIMSS, ePIRLS focuses on using engaging authentic tasks, this time in the domain of reading (IEA, 2018b). Interestingly, e-PIRLS involved learners being required to navigate to a set of curated websites in order to assess how well they are able to integrate information. This approach is not currently used in the IB's e-assessment but could be considered as a valuable tool in subjects such as history, literature and social studies.

The IEA is also planning to offer the 2023 sitting of the ICCS in a digital version (IEA, 2018c) and the ICILS focuses on digital literacy and the ability of learners to use computers to investigate, create, and communicate in order to participate effectively at home, at school, in the workplace, and in the community (IEA, 2018d).

ICILS

The ICILS uses a range of item formats across a spectrum of complexity, including the more traditional multiple choice and constructed response tasks, as well as drag and drop tasks and tasks that require learners to generate a digital product using multiple software applications (IEA, 2015).

Taking all of these examples of large-scale assessment together, it is interesting to note the range of approaches used towards digital assessment, from the use of avatar-like agents to much more traditional '*paper on screen*' type assessment tasks. While the scale of these assessments means that they are able to utilise innovations that may be out of the reach of many education programmes, they inevitably exert an influence on the way in which the digitisation of assessment evolves.

Moreover, there is a move to align national assessments internationally. Adams, et al. (2018) make the case for the development of '*global learning progressions*' in various fields as a reference points for individual nations, and others (presumably organisations), to use as a basis to interpret their own devised tests (rather than necessarily using a single, cross-national test such as TIMSS).

National Assessment Programs

A number of countries use digital assessment in their national assessment programmes. It is not possible to include examples of all of these here so those with particular relevance to the IB context are included here.

Denmark uses digital assessment for all of its ten national tests, with 800,000 tests completed each year (Thonbo, 2017). These are marked automatically and are adaptive in nature in that difficulty of items adjust to learner performance. Upper-secondary assessment is also digital and this facilitates digital distribution of tests, digitally supported marking and digital delivery of marks. The digital assessment tools allow learners to have free access to the Internet, something that the MYP e-assessment does not currently do.

In **Finland**, the Digabi project was started by the Matriculation Examination Board in 2013 and involves the gradual adoption of digital assessment starting with German, geography and philosophy in 2016 and leading to all examinations being digital in 2019 (Kimmo, 2017).

To prepare learners and teachers for the format of the matriculation examinations, they are able to access the test system that is used. This is an important model for the IB to use going forward, expanding on the prior access of learners to the MYP e-assessment and also in considering a step-wise approach to any expansion of digital assessment.

In **Norway**, the Norwegian Directorate for Education and Training (UDIR) conducts national standardised assessments in the 5th, 8th, and 9th grades (Aranbarri, 2019). The assessments were first implemented in digital format in 2004, at which point education stakeholders raised a number of concerns about digital assessment – particularly in relation to its use for schools' accountability. As a consequence, the emphasis was placed on informing teacher practices in a formative way and this has met with broad support (International Education News, 2016).

At present all elements of assessment – authoring, piloting, delivery and scoring – are digitally delivered and emphasis is placed on accessibility requirements for all learners. This example is interesting for the IB in that early concerns from stakeholders first had to be overcome before a unified digital system could be developed.

In **New Zealand** there is a push to make assessment more relevant to the way in which learners have learned and to encourage the development of personalised learning pathways. To this end, the New Zealand Qualifications Agency (NZQA) uses the slogan '***digital first***'. This refers to the desire for assessment to be designed for digital devices rather than adapted to them, with an emphasis on usability and accessibility, online marking and the generation of data to inform teaching and learning.

The aim is to have the National Certificates of Educational Achievement (NCEA) (national qualifications for senior secondary school learners) online by 2020, with trialling ongoing and an intent of going beyond simple '*paper on screen*'.

Interestingly, however, the NZQA has emphasised that if the current technology is unable to provide a good experience for learners, they will keep trialling rather than rushing to introduce it, reporting that *'if the teaching and learning is not typically digitally enabled, then we will consider the wisdom of a digital examination in that subject*' (Poutasi, 2018, 7).

This is an interesting insight for the IB as it indicates that there needs to be alignment between the digitalisation of teaching and learning and assessment in particular subjects before digital assessment should be introduced.

In **Scotland**, the Scottish National Standardised Assessments (SNSA) are adaptive, online standardised assessments for all learners in Scottish state schools in Primary years 1, 4 and 7 and Secondary year 3. The SNSA is part of the National Improvement Framework and the assessments are designed to support teacher's professional judgement and monitor learner progress.

Teachers can administer the assessments at any time during the school year. Schools can generate instant diagnostic reports for individuals and groups of learners. This model is an interesting one for the IB to consider as teachers have control over when to administer the assessment and the tools are specifically designed to provide teachers with an instant diagnostic report to inform their teaching practices.

Another interesting characteristic of the SNSA for the IB to consider is the fact that the adaptive nature of the assessments means that each learner follows their own pathway through the assessments and avoids facing items that are either too difficult or too easy resulting in a more satisfying experience for the learner.

English Language Proficiency Tests

Another international sphere of assessment in which digital assessment is widely used is in English language testing. Again, this is partly due to the enormous scale and global reach of English language testing. Interestingly, the approach to digital assessment in English language testing is much more conservative than the approach used in PISA, for example, likely reflecting the extremely high stakes nature of these assessments. Three main tests are:

- International English Language Testing System (IELTS)
- Pearson Test of English Academic (PTE).
- Test of English as a Foreign language (TOEFL).

IELTS was established in 1989, with a computer delivered version introduced in 2016 in 20 countries for test takers requiring a UK visa, and is now increasingly widespread. IELTS states that there are '*no differences in content, test timings, structure, marking, question types, security, speaking test or test report form*' (IELTS, nd, p3), indicating that this is basically '*paper on screen*'.

There are tools available in the digital version that would also be accessible to those taking a paper test, including highlighting facilities, the ability to make notes, turn pages and return to previous questions. At present, IELTS speaking tests are always face-to-face with an IELTS Examiner although researchers are exploring whether video-conferencing could be used (Nakatsuhara et al., 2017).

PTE Academic was launched in October 2009 and is fully computer based (Pearson, 2017b; Pearson 2011). Unlike the IELTS and TOEFL, in PTE the writing and speaking sections are combined. Pearson is looking ahead to ways to leverage digital tools, including the incorporation of artificial intelligence, '*using machine learning to build systems and algorithms which accurately and speedily assess a person's speech or samples of their writing*' (Booth, 2019, npn).

Since its inception in 1964, the **TOEFL** test has evolved from a paper-based test to a computer-based test and, in 2005, to an internet-based test, requiring the integration of receptive and productive skills (Norris, 2018). A review of publicly released test questions (ETS, 2015) indicates that while the TOEFL may be delivered digitally the format of the items are largely '*paper on screen*', with the advantages lying in the efficiency of scoring writing and speaking through online networks and an automated scoring system (ETS, 2018a).

The examples of IELTS, PTE Academic and TOEFL are instructive for the IB in that they indicate the degree of caution that is being used in high-stakes digital assessment, even

where the scale is enormous, with benefits more linked to efficiencies in implementation and marking and not in innovation in the format of the assessment tasks themselves.

Assessing twenty-first century skills

As previously mentioned, one of the potential benefits of digital assessment is the ability to assess twenty-first century skills such as creativity, critical thinking and problem solving, which are notoriously difficult to assess (Beller, 2013; Riggio, 2014; Shute, Leighton, Jang & Chu, 2016).

21st century skills remain difficult to both teach and assess (Scoular, 2018; Lamb et al., 2017), largely because the constructs themselves remain vague and poorly defined. Nevertheless, many education systems incorporate 21st century skills in mission statements and curriculum documents (Care et al., 2017) and this supports Tucker (2018) argument that it is important to measure what we deem important, regardless of the difficulty of doing so.

In a summary of their work on assessment of 21st century skills, the Partnership for 21st Century Learning (P21) suggest that a combination of technology-enhanced, formative and summative assessments are required to measure the development of 21st century skills.

There are a number of commercially and freely available products that schools can use to measure certain 21st century skills. One approach is to combine a 21st century skill with subject knowledge. For example, the online ACER General Ability Tests (AGAT) are designed for learners in year 2 to year 10 and assess verbal, numerical and abstract **reasoning** (ACER, 2019).

In terms of assessing **creativity**, a range of approaches are taken. Some examples include observing the exploration, involvement, enjoyment and persistence of early-grade learners during class (Robson, 2014); a sixth grade instrument of nine tasks to measure divergent thinking (Pásztor, Molnár, & Csapó, 2015); and a gamified assessment using artificial assessment to identify creativity and critical thinking (Fittes, 2018).

In terms of assessing **critical thinking** and **problem solving**, PISA incorporated a problem-solving unit in 2012 (OECD, 2014). This used simulated real-life problem situations, such as an unfamiliar vending machine, or a malfunctioning electronic device, and measured learners' reasoning skills, ability to regulate problem-solving processes, and their willingness to do so. It involved exploring interactive items in an online environment and the OECD suggested that '*a paper-and-pencil assessment of problem solving could not have measured the same construct*' (2014, 33).

In terms of assessing **collaboration**, ongoing work is continuing to identify what this construct actually involves, and how it can be assessed. Wright et al. (2013, 12) define it as '*a social process of knowledge building that requires learners to work as an interdependent team towards a clear objective resulting in a well-defined final product, consensus, or decision*'.

In the Assessment and Teaching of 21st Century Skills (ATC21STM) project, Griffin & Care (2014) undertook a research and development plan that included conceptualization of collaborative problem solving. This involved definition and identification of sub-skills, as well as development of a method for its assessment.

The assessment task they designed incorporates both the cognitive and social aspects of collaborative problem solving, and ability to generate formative feedback for teachers to identify levels of ability and support tailoring of instruction.

Another 21st century skill that could be considered is **global competence** or **global citizenship** (UNESCO, 2015) (also

termed as a **cosmopolitan disposition** (Richardson, 2017)). A review of assessment instruments to measure this construct (Singh & Qi, 2015) identified that this is an under-researched area with only a very limited number of assessments.

In 2018, the OECD's Programme for International Student Assessment (PISA) introduced a global competency assessment (OECD, 2018). This assessment looks at how 15-year-olds can critically process information on key sustainability issues. They are required to use and connect multiple sources of evidence, identify biases and gaps in information, and manage conflicting arguments. The test also assesses learners' capacity to recognise other's perspectives and the factors that might influence them.

Ensuring learners are prepared for the world of work in the **digital economy** has connections through curriculum areas such as Health and Physical Education. Hendry et al., (2018) released a roadmap entitled 'YeS Project educator guide' which provides an overview of research in young people's digital cultures. Digital culture terms of interest to this curriculum project include '*social norms, values, ideas and practices that relate to the internet, social media and technology*.' This has implications for all programmes and all subjects.

Taken together, all of these initiatives have implications for the IB. First, to recognise that the assessment of 21st century skills is a work in progress, something that the IB can make an important contribution to due to the focus of its programmes. Second, that while digital tools can offer ways to measure 21st century skills, the focus should be on the purpose of the assessment and consideration given to whether ongoing project- or task- based approaches may be a more suitable means of assessing 21st century skills than a model in which digital assessment is used in a summative way.

CONNECTING DIGITAL ASSESSMENT TO LEARNING AND TEACHING

In the first section of this review the need for teaching, learning and assessment to work in harmony was identified. Considerations of digital assessment hence have important implications for digitalisation of teaching and learning.

While it is important for the IB to understand the dynamics between these three components, it is equally important to identify the ways in which assessment practices need to draw on approaches to teaching and learning used in IB programmes. This section looks at the relationship between these three elements.

Backwash and forewash

Returning to the critical interaction of teaching, learning and assessment, it is well known that the form of assessment has a profound impact on teaching and learning. Alderson and Wall (1993) suggest that the format of assessment is likely to influence teaching and learning in terms of approaches, content, prioritisation, sequencing and depth of focus.

Hughes (1994, 2) places the spotlight of the backwash effect on three constituents: the '*participants', 'processes',* and *'products*' of an educational system:

- **Participants** include classroom teachers, learners, educational administrators, textbook developers, and publishers.
- **Processes** refer to '*any actions taken by the participants which may contribute to the process of learning*', such as materials development, syllabus design, changes in teaching methodology, the use of Test-taking strategies, etc.
- **Products** refer to '*what is learned and the quality of the learning*'.

As the works of Alderson and Wall, and of Hughes suggest, backwash is multifaceted and the agency of a range of variables beyond the assessment are likely to shape its influence (Tsagari, 2007). Importantly, these include teacher and learner attitudes and experiences and the social context.

The notion of backwash has important implications for the digitalisation of teaching, learning and assessment, but it would be wrong to assume that it is a one-way process. Instead, assessment designers are inevitably influenced by their experience in teaching and learning, and hence the integration of digital tools into classrooms practices is likely to influence the form of assessment (or, at the very least, expectations around assessment). Therefore '*forewash*' is a phenomenon that should not be neglected.

Beyond linear backwash and forewash, an important challenge for the IB is to consider the possibility of unintended or unpredicted consequences of assessment. It is vital for the IB to navigate the ways in which assessment can influence whole-school teaching and learning systems and structures to see if the assessment is truly beneficial, and is in harmony with teaching and learning.

Digital assessment and digital pedagogy

If digital assessment is to influence, or be influenced by, teaching and learning, then it is important to unpack some of the misconceptions that surround the relationship, and indeed the digitalisation of education, more broadly.

There is much focus on the transformational ways in which digital technologies can be used in teaching and learning. As with digital assessment, however, the reality tends to be quite different to the ambitions.

There is certainly much discussion in the literature of the advantages that digital pedagogies can lead to. These include:

- The scope for more **learner centred** approaches and greater personalisation of learning (FitzGerald et al., 2017; Yarbo et al., 2016);
- The enhancement of learners' **critical thinking** skills (Yarbo et al., 2016; Baylor & Ritchie, 2002; Hopson, Simms, & Knezek, 2001);
- The ability for learners to develop **different ways of thinking** (Scanlon, Anastopoulou, & Kerawalla, 2012; McCabe & Meuter, 2011);
- An increase in the opportunity for **learner collaboration** and group work (Gokhale & Machina, 2018; Inan, Lowther, Ross, & Strahl, 2010; Waxman & Huang, 1996);
- The ability of teachers and learners to receive **instant feedback** (Tancock, et al., 2018; Caldwell, 2007).
- The capacity of learners to engage in **informal learning** through games (Iacovides, Aczel, Scanlon, Taylor, & Woods, 2011).
- The value of **learning analytics** for informing approaches to learning and teaching (Nussbaumer et al., 2015; Hernandez-Lara, Perera-Lluna & Serradell-Lopez, 2018; Vieira, Parsons & Byrd, 2018).

All of these are certainly incredibly valuable in their potential to enhance the value of teaching and learning activities and to engage learners. What is important to question, however, is the extent to which the potential in each of these is drawn out in practice.

Unfortunately there is considerable evidence that the aspirations for the digitalisation of teaching, learning and assessment outlined in the literature too often fall short of their goals. An important reason for this is the key role that teachers and learners play in mediating digital approaches.

Digitalisation and teachers

Interestingly, the literature of backwash of digital assessment on teachers often focuses on practical benefits, such as freeing teachers up from some of the administrative burdens of implementing and marking assessment to spend more time on teaching and learning (McKnight et al., 2016; Smith & Gray, 2016).

When positive backwash is discussed in terms of impact on teaching and learning, this tends to emphasise the impact on classrooms in which learners have their own digital devices and hence have access to technology in class. In this context, methods used in digital assessment can be echoed in the classroom, facilitating some of the benefits highlighted above.

In this context, for example, teachers are able to use the example set by digital assessment to model and support digital learning, as well as actively manage learners' use of technology (Murray, Luo, & Franklin, 2019). Moreover, they can encourage cognitive processes that stimulate inquiry and deep learning (Glassett & Schrum, 2009).

An important implication of these examples for the IB is the reliance on learners having access to digital devices in the classroom, raising questions about how this can be made possible in under-resourced schools.

The potential negative impact of digital assessment onto classrooms without digital devices readily available is currently underexplored.

Beyond access to resources, learner and teacher attitudes, skills and comfort with technology is often discussed as critical in ensuring that there is positive backwash from digital assessment to digital teaching and learning. It is common for assumptions to be made about so-called '*digital natives*' and their

easy accommodation of the digitalisation of education.

Assuming that certain cohorts of teachers, particularly those new to the profession who would be regarded as '*digital natives*', are comfortable to integrate digitisation into their teaching (or, indeed, that non digital natives are less comfortable) is erroneous (Gill et al, 2015).

For example, studies show that regardless of their comfort with digital tools such as social media and positive beliefs in technology, teachers lack experience and confidence in using digital resources in the classroom (Evers et al., 2018; Buzzard et al., 2011; Lei, 2009).

Moreover, even when teachers are impressed by the positive impact of digital pedagogy on learners' skills and engagement, they may not go on to implement this in their teaching (Dogan & Robin, 2008). This may be because they perceive it as potentially too difficult to use (Rientjes et al., 2014).

For digitalisation of education to be successful, and to truly add value to teaching and learning, it needs to be carefully selected and mediated. Thus the role of teachers is critical. Inevitably, therefore, digitalisation of education has a profound impact on the skills and knowledge required by teachers, and the support and professional learning opportunities that they need.

The research of Mishra and Koehler (2006) on *Technological Pedagogical Content Knowledge* (TPACK), a framework for teacher knowledge for technology integration has been widely cited as a useful basis for focusing professional learning in this area. The focus on TPACK builds on the construct of pedagogical content knowledge (PCK) developed by Shulman (1986) and emphasises the need for interactions between the three bodies of knowledge of pedagogy, content and technology.

Achieving this requires both experience and professional learning, as well as the agency to take risks and (sometimes) to fail. It can be difficult for teachers to '*stay ahead of the curve*' (King et al, 2016, 8), particularly since the pace of technological change is so rapid.

Increasingly, teachers are called on to be designers, not only of analogue but also of digital learning materials (Kali et al, 2015). This is often done individually through trial and error and critical reflection followed by customisation. Support such as teamwork, process guidance (including source materials and task templates) and a shared vision '*stimulate teachers to apply their (technological pedagogical content) knowledge, especially when tackling new topics*' (Kali et al., 2015, 174).

McKenney et al. (2015) identify that teacher design knowledge can be thought of as a combination of their knowledge base, productive beliefs, repertoire for action, ability to consult those more experienced and understanding of design work more broadly. In addition, it requires the ability to judge '*which ideas and processes make the most sense under certain circumstances, at certain points in time, with certain people*' (McKenney et al., 2015, 190). Supporting these needs requires a multiplicity of responses and there is no one-size fits all approach.

While teacher skills in integrating digital technologies are critical, research suggests that '*wide-scale transformation of teacher practice and digital learning remain unrealized*' (Blundell et al., 2016, 535). Challenges include those that are external, including access to digital resources, access to technical support, access to professional learning opportunities and institutional cultures.

Challenges can also be intrinsic, including attitudes to innovation, pedagogical beliefs and self-efficacy, with suggestions that these are '*the true gatekeepers*' (Gill et al., 2015, 539). Gunn and Hollingsworth (2013) found that sustained input in the form of extensive professional learning opportunities, the easy

availability of digital tools and a culture that encourages teachers to take risks, are required for teachers to gain expertise and confidence in the integration of digital technologies.

In addition, teacher background plays a significant part in their approach to the integration of technology. Those teachers who are most comfortable with technology are able to integrate it seamlessly but others require a greater deal of support and the ability to receive advice from more knowledgeable colleagues (Ruggiero & Mong, 2015). Moreover, teachers need the space to constantly adjust their use of technology based on their own reflections and how learners respond.

Where integration of digital tools has been found to have a positive impact on teaching and learning, there tends to be active leadership and management support, pedagogical and technical support for staff and solid staff development (Whitelock & Brasher, 2006) Individuals who champion tools and practices have also been seen as important to success (JISC, 2010) as had a clearly established pedagogic need (Beevers et al., 2011).

Scaffolding is essential in any learning, and this includes in introducing learners and teachers to digital pedagogies. The use of AI can help through the analysis of education data which supports personalised scaffolding, such as through the EDUCATE Educational Technology (EdTech) programme (Luckin & Cukurova, 2019).

Together, these findings have important implications for the IB. If the IB considers the use of digital assessment, teaching and learning to be important in the achievement of the goals set for IB programmes, then what kind of support should the IB provide to schools? And what should the IB promulgate as the purpose of digitalisation?
Is it a question of leaving approaches to digital teaching, learning and formative assessment up to schools while the IB's main responsibility is to focus on the possible expansion of digital approaches to summative assessment? Or does the IB need to provide teachers and programme coordinators with targeted support to ensure that they are able to mediate the digitalisation of education in their schools?

As Kirschner (2015, 321) suggests:

'teachers as professionals […] need to be professional learners and to function as researchers into their own practices. Their initial and continuing training must equip them to learn and work as co-researchers and co-designers of the technology and tools they want and need to use ... The key to this is the dynamic interaction between the respective networks, communities and actors involved in the knowledge chains in the practice of education'.

One of the challenges for the IB lies in the 7-year curriculum review cycle, during which a whole number of new opportunities and challenges is likely to arise regarding the use of technology. This indicates the need for the IB to provide teachers with TPACK resources that are parallel to the programme and subject guides that teachers use.

Nevertheless, the more recent IB guides do indicate a shift in focus in approaches to learning. For example the latest PYP: From Principles into Practice (2018) set of documents has the most detailed exemplars of ATL skills, including emphasis on: digital and non-digital critical and ethical information and media literacy skills; communication using a range of digital technologies and media in different contexts; and using technology effectively to manage time and tasks.

Since these approaches to learning are both cross-curricular and cross-programme skills, there is clear scope in this area for the IB to showcase teaching and learning of digital technology knowledge and skills that are already embedded in the written, taught and

assessed curricula.

Resourcing digitalization

Debates around the pros and cons of digital assessment, teaching and learning tend to take place in something of an echo chamber in which one of the key determinants can easily be overlooked, and this relates to the resources needed for, and financial restrictions around, facilitating digitalisation.

At present, schools delivering IB programmes must have in place '*teaching and learning [that] incorporates a range of resources, including information technologies*.' (IB, 2010/2016a, 16) and schools are required to confirm that '*there are appropriate information technology facilities to support the implementation of the programme*' (2010/2016a, 22).

This means that there should be no schools taking IB programmes in which no digital devices are available, but there are nevertheless likely to be many schools in which the notion that all learners have digital devices for personal use in the classroom is simply not the reality. Nor can it be assumed that all teachers have personal digital devices to use in planning their teaching or that all classrooms are equipped with projectors and digital devices.

In this context the IB needs to tread very carefully in its support and advice to schools. If it is determined that digitalisation does confer benefits to learners and teachers that cannot be achieved without digital devices, then how prescriptive should the IB be in setting requirements for schools? There is no *'right'* answer to this and every education system and programme around the world has to identify what is most appropriate within a given context.

In terms of digital assessment, there are certainly cost savings to be made in putting all assessments online, taking away the need to ship examination materials around the world and facilitating the automatic marking of any closed response items.

Going beyond '*paper on screen*' approaches, however, does require significant investment. The fundamental reason why the majority of digital assessment remains in the realm of '*paper on screen*' is that innovation costs money.

The costs of developing simple elements such as static graphics alone can be prohibitive if there is not sufficient scale to make this worthwhile. Moving to interactive, augmented reality, scenario-type assessment is very attractive in theory but the practicalities of bringing this to reality – in terms of cost and skills required – are simply beyond the budgets of most assessment bodies.

Hence much of contemporary research around digital assessment is into the areas that could generate further cost savings – such as in automated assessment of speech and extended written responses – rather than in how a digital environment can be better used in more fundamental ways, such as to assess learners' higher order or 21st century skills, or to provide teachers with valuable insights.

In addition, while much is made of the ability of digital assessment systems to provide automatic results to learners and diagnostic data to teachers, this depends entirely on the way that digital assessments are built.

A high degree of expertise is required to develop the types of items that can evaluate higher-order skills in a way that both enables automatic marking and the identification of misconceptions underlying errors. This therefore demands both up-front investment in capacity development and ongoing training. It is essential that these kind of considerations are part of the IB's overall deliberations around digitalisation.

CONCLUSION

This review has provided a foundation for the IB in its considerations about the role of digital assessment in its programmes, and its relationship with teaching and learning. It has emphasised the need for approaches to assessment to work in harmony with teaching and learning and to be driven by a clear purpose around the enhancement of learning.

The world of digital assessment is rapidly changing and there is enormous potential for the IB to innovate its approach to assessment. The way in which contemporary lives and professions have been transformed by digitalisation has to be taken into account by any educational body interested in helping to prepare learners for their future lives.

Equally important is that the key ways in which education is being reshaped are taken into account. These include the understanding that all learners are on their own learning pathways and that mapping out their next steps requires empirical evidence about where they currently are and where they need to go. This means that assessment is increasingly being understood as something that is embedded within teaching and learning and is very much an ongoing activity.

As the review has affirmed, any desire to rush headlong into the digital transformation of education needs to be tempered by a range of considerations: cost and resources; teacher skills and attitudes; the availability of assessment expertise within the IB; current practices in teaching and learning, and a careful consideration of the way in which the use of digital technologies can enhance educational practices. This is particularly important in a world where there is a paucity of evidence of impact of digitalisation on the quality of learning.

There are certainly a number of stimulating ways in which digital assessment can provide positive support for teaching and learning and there are exciting developments in digital assessment on a regular basis, with opportunities of virtual reality, gamification and the use of avatars yet to be optimised.

Equally, the fact is that most digital assessments that are currently used have not yet strayed far from '*paper on screen*' in which '*new technologies are used to do old things*' (Dolan, 2013). This is largely a consequence of the high cost of creating more advanced forms of assessment and it is noteworthy that where innovations have been implemented this has largely taken place in contexts in which the scale of assessment can justify the expense involved.

Moreover, when the format of items used in digital assessment is considered, there is little evidence of a great leap into new functionality, despite the wide array of possibilities available.

Formative assessment tends to be regarded as possessing greater scope for innovation since the stakes are lower and there is the possibility of including collaborative work and a range of ways of capturing insights into learner performance such as e-portfolios (Yang et al., 2017), peer review (Colbeck et al., 2014) and social media (Cochrane et al., 2015). In high-stakes examinations, however, there is much less appetite for moving away from what is tried and tested.

The implication for the IB, then, is how best to ensure that its assessment practices, just as much as teaching and learning practices, reflect the underlying and unique purpose of its suite of programmes with their focus on inquiry, trans-disciplinarity, 21st century skills and global-mindedness. In this context it would seem that the IB has a responsibility to consider how to enhance its assessment practices, going beyond what other

educational programmes currently accept.

As the first Director General of the IB, Alec Peterson, made clear in 1971, a fundamental responsibility rests on the IB. That is to ensure that assessment '*assesses the whole endowment and personality of the pupil in relation to the next stage of his life*' (quoted in IB, 2018c, 7). Arguably, in the 21st century it is no longer possible to do this with pen and paper tests. Equally, the '*whole*' learner is not likely to be described in a single set of final examinations. Hence the consideration of ongoing assessment that draws on some of the advantages that digital technology offers seems like a sensible approach.

The other key consideration for the IB, however, is the extent to which innovation in assessment meets the needs, and appropriately guides, its large number of stakeholders around the world. As this review has emphasised, there needs to be synergy between assessment, learning and assessment. While all IB schools are required to have digital devices this does not mean that all learners in IB programmes have equal opportunities to learn digitally, and nor that all teachers are up to speed with learning design.

The notion of backwash and forewash indicate that digital tools used in assessment need to both encourage greater digitalisation of teaching and learning, while also reflecting the practices currently used in classrooms. This is complex in a situation in which practices in schools are thought to vary considerably. Returning to Alec Peterson, it is important that backwash from assessment does not '*distort good teaching*' (quoted in IB, 2018a, 7) or distract from the important purpose of learning.

This brings us to the other key component of any shift to greater digitisation of assessment within the IB – the need to provide support to teachers as demands on them increase, particularly in terms of digital learning design. This means a consideration of the role that IB does, and should, play in supporting teachers in IB programmes to enhance their skills.

Practice tells us that teacher learning is not a linear process and nor is it influenced by single actors. Instead, it is a multi-faceted and multi-agent process that is not yet fully understood and the IB's approach needs to be designed with that in mind. Moreover, the ability of schools to resource and embrace digitalisation of teaching and learning will continue to shape what is possible.

In many ways the answers lie in the formative assessment practices of teachers in IB programmes around the world. These often remain unseen and unremarked but they are a powerful resource in thinking about ways to leverage digital tools in order to bring teaching, learning and assessment into greater harmony, and to enhance the value of an IB education for learners. As Dolan (2013) suggests, the critical challenge ahead of educators is to find the best ways to '*harness the potential offered by digitisation for the benefit of learners*'.

For the IB, useful future research will include:

- an in-depth look at which approaches to teachers support and training across IB programmes yield the most sustained results in enhanced practices;
- an evaluation of current practice in data analytics across IB schools and how this could be enhanced;
- a review of approaches to supporting a growing understanding of learning design across IB stakeholders;
- research into optimal approaches to multimedia design; and
- research into the cognitive impact of using digital devices for teaching, learning and assessment on learners/

Moreover, if good practices are to be shared across the IB ecosystem, then optimal ways of sharing examples of what works in different programmes and curricula areas will be important, with the IB Educator Network a good basis for this to take place.

REFERENCES

[This report is a shorter version of a longer body of work that was delivered to the IB at an earlier stage in this project. As such, some of the references included in the reference list are not referred to in the current text. The decision has been made to keep them in this report, however, in order to provide a resource of relevant literature for the IB to draw on as required].

Adamo-Villani, N., & Dib, H. N. (2016). A study of the effects of teaching avatars on learners' learning of surveying mathematics. International Journal of Information and Communication Technology Education (IJICTE), 12(2), 1-13.

Adams, R., Jackson, J., & Turner, R. (2018). Learning progressions as an inclusive solution to global education monitoring, Australian Council for Educational Research (ACER) Global Education Monitoring (GEM). Retrieved from https://research.acer.edu.au/monitoring_learning/32/

Adamson, F. & Darling-Hammond, L. (2015). Policy pathways for twenty-first century skills. In P. Griffin & E. Care (Eds.), Assessment and teaching of 21st century skills: Methods and approach. pp. 293–310. Dordrecht, the Netherlands: Springer Science and Business Media.

Ainley, J., Fraillon, J., Gebhardt, E., & Schulz, W. (2012). National Assessment Program – ICT Literacy Years 6 and 10 Report. Sydney: Australian Curriculum, Assessment and Reporting Authority (ACARA).

Alderson, J. C. & Wall, D. (1993). Does Washback Exist? Applied Linguistics, 14: 115-129.

Ananiadou, K., & Claro, M. (2009). 21st-century skills and competencies for new millennium learners in OECD countries. OECD Education Working Papers, No. 41.

Anderson, C. W. (2008). Conceptual and empirical Validation of Learning Progressions: Response to Learning Progressions: Supporting Instruction and Formative Assessment, retrieved from http://edr1.educ.msu.edu/EnvironmentalLit/index.htm

Anderson, L., Krathwohl, D., Airasian, P., Cruikshank, K., Mayer, R., Pintrich, P., Raths, J. & Wittrock, M. (2001). A taxonomy of learning, teaching and assessing: A revision of Bloom's taxonomy of educational objectives, New York. Pearson, Allyn & Bacon

Aranbarri, L. (2019). Digitise the Norwegian National Tests at Scale. Retrieved 1 June 2019 from http://www.inspera.com/blog/digitise-the-norwegian-national-tests-at-scale

Assessment Reform Group (2002). Assessment for Learning: 10 Principles. Retrieved from https://www.aaia.org.uk/content/uploads/2010/06/Assessment-for-Learning-10-principles.pdf.

Australian Council for Educational Research (2019). ACER General Ability Tests. Retrieved 21 January 2019 from https://www.acer.org/au/agat

Australian Council for Educational Research (ACER) (2018a). New technology. Centre for Assessment Reform and Innovation. Retrieved from https://www.acer.org/cari/about-cari/new-technology

Australian Council for Educational Research (ACER) (2018b). Online Assessment and Reporting. Retrieved from https://www.acer.org/oars

Baird, J. A., Andrich, D., Hopfenbeck, T. N., & Stobart, G. (2017). Assessment and Learning: Fields Apart?, Assessment in Education: Principles, Policy & Practice, 24:3, 317-350.

Baird, J. A., Hopfenbeck, T. N., Newton, P., Stobart, G., Steen-Utheim, A. T. (2014). Assessment and Learning: State of the Field Review. Oslo: Knowledge Center for Education.

Bargh, S. (2014). A Model for the implementation of digital examinations: New Zealand's vision for digital assessment. Paper presented to the 40th Annual Conference of the International Association for Educational Assessment (IAEA), Singapore. Retrieved from https://iaea.info/documents/a-model-for-the-implementation-of-digital-examinations-new-zealands-vision-for-digital-assessment/

Baylor, A. L., & Ritchie, D. (2002). What factors facilitate teacher skill, teacher morale, and perceived learner learning in technology-using classrooms. Computers & Education, 39(4), 395–414.

Beevers, C. et al (2011). 'What can e-assessment do for learning and teaching? Part 1 of a draft of current and emerging practice review by the e-Assessment Association expert panel'. International Journal of e-Assessment, 1 (2).

Beller, M. (2013). Technologies in large-scale assessments: New directions, challenges, and opportunities. In M. von Davier, E. Gonzalez, I. Kirsch, & K. Yamamoto (Eds.), The role of international large-scale assessments: Perspectives from technology, economy, and educational research. pp. 254–46. Dordrecht, the Netherlands: Springer.

Bergeron, E. (2014). The role of technology in fostering agency and leadership in the PYP: Literature Review and Document Analysis. The Hague: International Baccalaureate Organisation.

Bigum, C., Durrant, C., Green, B., Honan, E, Lankshear, C., Morgan, W., Murray, J.,Snyder, I., Wild, M., McKenna, A., Burbules, N., Kapitzke, C., Doneman, M., Bleicher, B., Knobel, M., Muspratt, S., Callister, T. & Lemke, J. (1997). Digital Rhetorics: Literacies and Technologies in Education - Current Practices and Future Directions (Volumes 1-3). Canberra ACT Australia: Department of Employment, Education, Training and Youth Affairs.

Billingsley, G., Smith, S., Smith, S. & Meritt, J. (2019). A Systematic Literature Review of Using Immersive Virtual Reality Technology in Teacher Education. Journal of Interactive Learning Research, 30(1), 65-90. Waynesville, NC: Association for the Advancement of Computing in Education (AACE). Retrieved June 18, 2019 from https://www.learntechlib.org/primary/p/176261/ .

Bingimlas, K. A. (2009). Barriers to successful integration of ICT in teaching and learning environments: A review of the literature. Eurasia Journal of Mathematics, Science, & Technology Education, 5(3), 235–245.

Black, P. & Wiliam, D. (2018). Classroom assessment and pedagogy, Assessment in Education: Principles, Policy & Practice, DOI: 10.1080/0969594X.2018.1441807.

Blamires, M. & Peterson, A. (2014). Can creativity be assessed? Towards an evidence-informed framework for assessing and planning progress in creativity. Cambridge Journal of Education, 44(2), pp. 147-162. doi: 10.1080/0305764X.2013.860081

Blikstad-Balas, M. (2016), You get what you need: A study of learners' attitudes towards using Wikipedia when doing school assignments, Scandinavian Journal of Educational Research, Vol. 60/6, http://dx.doi.org/10.1080/00313831.2015.1066428, pp. 594-608.

Bloom, B., Engelhard, M., Furst, E., Hill, W., & Krathwohl, D (1956). Taxonomy of educational objectives, Handbook 1, The cognitive domain, New York, David McKay Co Inc.

Blundell, C., Lee, K.T. & Nykvist, S. (2016). Digital Learning in Schools: Conceptualizing the Challenges and Influences on Teacher Practice. Journal of Information Technology Education: Research, 15, 535-560. Retrieved from http://www.informingscience.org/Publications/3578

Booth, D. (2018). Computer Based Testing: The Future is Now. Retrieved 20 January 2019 from https://pearsonpte.com/articles/computer-based-testing-future-now/

Boulder Language Technologies, (2018). FLORA Retrieved from http://boulderlearning.com/products/flora/ .

Bower (2008). Affordance analysis—matching learning tasks with learning technologies. Educational Media International, 45 (1), 3–15.

Brewer, D. and Tierney, W. (2012). 'Barriers to innovation in the US education', in Wildavsky, B., Kelly, A. and Carey, K. (Eds), Reinventing Higher Education: The Promise of Innovation, Harvard Education Press, Cambridge, MA, pp. 11-40.

Bunderson, C. V., Inouye, D. K., & Olsen, J. B. (1989). The four generations of computerized educational measurement. In R. L. Linn (Ed.), The American Council on Education/Macmillan series on higher education. Educational measurement (pp. 367-407). New York, NY, England: Macmillan Publishing Co, Inc; American Council on Education.

Burkhardt, H., & Pead, D. (2003). Computer-based assessment: a platform for better tests. Whither assessment, 133-148.

Buzzard, C., Crittenden, V. L., Crittenden, W. F., & McCarty, P. (2011). The Use of Digital Technologies in the Classroom: A Teaching and Learning Perspective. Journal of Marketing Education, 33(2), 131–139.

Caldwell, J. E. (2007). Clickers in the large classroom: Current research and best-practice tips. CBE Life Science Education, 6(1), 9–20.

Care, E., Kim, H., Anderson, K., & Gustafsson-Wright, E. (April 2017). Skills for a Changing World: National Perspectives and the Global Movement. Center for Universal Education at Brookings. Retrieved from https://www.brookings.edu/wp-content/uploads/2017/03/global-20170324-skills-for-a-changing-world.pdf

Carranza, R. (2016). Blooms digital taxonomy, Arizona State University, https://teachonline.asu.edu/2016/05/integrating-technology-blooms-taxonomy/

Chung, G.K.W.K. and E.L. Baker (2003). 'Issues in the Reliability and Validity of Automated Scoring of Constructed Responses' in M.D. Shermis and J. Burstein (eds.), Automated Essay Scoring: A Cross-Disciplinary Perspective, pp. 23-40, Lawrence Erlbaum, Mahwah, NJ.

Cizek, G. (2010). An Introduction to Formative Assessment: History, Characteristics, and Challenges. In: HL Andrade & G J Cizek (eds.) Handbook of Formative Assessment. New York, Routledge.

Clarke-Midura, J., & Dede, C. (2010). Assessment, technology, and change. Journal of Research on Technology in Education, 42(3), 309-328.

ClassTech. (2006). ClassTech 2005–2006 assessment conclusions report. Retrieved from http://www.ncsu.edu/classtech/survey_results/2005-6/ClassTech_Assessment_Conclusions.pdf

Cochrane, T., Narayan, V., & Oldfield, J. (2015). Emerging technologies in New Zealand: A pedagogical framework for mobile social media. In V. Bozalek, D. Ng'ambi, D. Wood, J. Herrington, J. Hardman & A. Amory, Activity theory, authentic learning and emerging technologies(Eds.), (pp.126-143). New York, U.S.A.: Routledge.

Colbeck, Douglas & Sakulwichitsintu, Songlak & Turner, Paul & Ellis, Leonie. (2014). Online Peer Learning: Understanding Factors Influencing Learners' Learning Experience.

Commonwealth of Australia Department of Education and Training, (2018). Through Growth to Achievement: The Report of the Review to Achieve Excellence in Australian Schools. Retrieved from https://docs.education.gov.au/documents/through-growth-achievement-report-review-achieve-educational-excellence-australian-0

Conley, D. (2015). A new era for educational assessment. Education policy analysis archives, 23, 8.

Cooker, L., Crook, C. and Ainsworth, S. (2010). The Integration of Technology in the International Baccalaureate Diploma Programme: Final Report. Nottingham: The University of Nottingham.

Cookson, J. C. (2018). Assessment Terms Half a Century in the Making and Unmaking: from Conceptual Ingenuity to Definitional Anarchy, in Assessment & Evaluation in Higher Education, 43:6, pp. 930-940.

Cope, B., & Kalantzis, M. (2016). Big Data Comes to School: Implications for Learning, Assessment, and Research. AERA Open, April-June 2016, Vol. 2, No. 2, pp. 1-19.

Cox, M. (2013). What the research has to say. The impact of technology on teaching and learning. In N. Pachler & M. Leask (eds). Learning to teach using ICT in the secondary school: A companion to school experience, pp18-39, London, Routledge

Cox, M. (2018). Researching information technology in education: meeting the challenges of an ever-changing environment, in in J. Voogt, G. Knezek, R. Christensen & Kwok-Wing, L. (eds). Second Handbook of Information Technology in Primary and Secondary Education, Volume 1, Springer International Publishing.

Crockett, L. & Churches, A. (2017). Growing Global Digital Citizens: Better Practices that Build Better Learners. Bloomington, IN.: Solution Tree Press.

Cronbach, L.J. & Meehl, P.E., (1955). Construct validity in psychological tests. Psychological Bulletin, 52, pp. 281-302.

Cronbach, L.J. (1971). Validity. In R.L. Thorndike (ed.), Educational measurement. Second edition. pp. 443-597. Washington, DC, American Council on Education.

Crook, Charles. (2012). The 'digital native' in context: tensions associated with importing Web 2.0 practices into the school setting, Oxford Review of Education, 38:1, 63-80.

de Klerk, S. (2012). An overview of innovative computer-based testing. Psychometrics in Practice at RCEC, 137.

Department of Education and Skills, Ireland. (2015). Digital Strategy for School 2015-2020: Enhancing teaching, learning and assessments. Retrieved from https://www.education.ie/en/Schools-Colleges/Information/Information-Communications-Technology-ICT-in-Schools/Digital-Strategy-for-Schools

DiCerbo, K. E. and J. T. Behrens (2014). Impacts of the Digital Ocean on Education, London:

Digital Chalkboard (2018) Assessment Methods and item Types. Accessed 14 January 2019 from https://www.mydigitalchalkboard.org/portal/default/Content/Viewer/Content?action=2&scId=505706&sciId=17336

Dogan, Bulent & Robin, Bernard. (2008). Implementation of Digital Storytelling in the Classroom by Teachers Trained in a Digital Storytelling Workshop. 902-907.

Dolan, R. P. (2013). A Universal Design for Learning-based Framework for Designing Accessible Technology-Enhanced Assessments, Pearson: UK.

Donaldson, G. (2015). Successful Futures: Independent Review of Curriculum and Assessment Arrangements in Wales. Retrieved from https://gov.wales/docs/dcells/publications/150225-successful-futures-en.pdf

Dweck, C. (2015). Discussant: Growth, British Journal of Educational Psychology, Vol. 85 (2), pp. 242-245.

Eckstein, M. A. & Noah, H. J. (1993). Secondary School Examinations: International Perspectives on Policies and Practice, Yale University Press, New Haven.

Educational Testing Service (2015.) TOEFL iBT® Test Questions. Retrieved 20 January 2019 from https://www.ets.org/Media/Tests/TOEFL/pdf/SampleQuestions.pdf

Educational Testing Service (2018). TOEFL iBT Test Framework and Test Development. TOEFL® Research Insight Series, Volume 1. Retrieved 20 January 2019 from https://www.ets.org/s/toefl/pdf/toefl_ibt_research_insight.pdf

Educational Testing Services (ETS) (2018). TOEFL iBT® Test Content, Retrieved from https://www.ets.org/toefl/ibt/about/content/

Ercikan, K. & Oliveri, M. (2016). In Search of Validity Evidence in Support of the Interpretation and Use of Assessments of Complex Constructs: Discussion of Research on Assessing 21st Century Skills. Applied Measurement in Education, 29(4), pp. 310-318.

European Schoolnet (2019). Evidence for Innovation. http://www.eun.org/focus-areas/innovation

Evers, K; Golding, J; Grima, G; (2018). (Missed) opportunities for teaching and learning with digital resources: what and why. In: Golding, J and Bretscher, N and Crisan, C and Geraniou, E and Hodgen, J and Morgan, C, (eds.) Proceedings of the 9th British Congress of Mathematics Education: BCME9. (pp. pp. 48-55). BSRLM: Univeristy of Warwick, UK.

Eyre, J., Berg, M., Mazengarb, J., & Lawes, E. (2017). Mode equivalency in PAT: Reading Comprehension. NZCER. Retrieved from https://www.nzcer.org.nz/system/files/Eyre%20et%20al%202017,%20Mode%20equivalency%20in%20PAT.pdf

Fisher, M. (2019). How To Design A 21st Century Assessment, in https://www.teachthought.com/the-future-of-learning/how-to-design-a-21st-century-assessment/

Fittes, K. (2018, Sep. 4). What you should know about Project Lead The Way's 'gamified' new test. Indianapolis Star, Sept. 4, 2018. Retrieved from https://www.indystar.com/story/news/education/2018/09/04/standardized-test-indy-stem-group-creativity/972109002

FitzGerald, Elizabeth; Kucirkova, Natalia; Jones, Ann; Cross, Simon; Ferguson, Rebecca; Herodotou, Christothea; Hillaire, Garron and Scanlon, Eileen (2017). Dimensions of personalisation in technology-enhanced learning: a framework and implications for design. British Journal of Educational Technology, 49(1) pp. 165–181.

Fraillon, J., Ainley, J., Schulz, W., Friedman, T. & Gebhardt, E. (2014). Preparing for Life in a Digital Age: The IEA International Computer and Information Literacy Study International Report, International Association for the Evaluation of Educational Achievement (IEA), http://www.iea.nl/fileadmin/user_upload/Publications/Electronic_versions/ICILS_2013_International_Report.pdf

Freedman, M. & Houtz, J. (2004). A Glossary of Terms Used in Educational Assessment, Davidson Institute. Retrieved from http://www.davidsongifted.org/Search-Database/entry/A10461

Gibson, J. (1977). The theory of affordances in: R. Shaw & J. Bransford (eds) Perceiving, acting, and knowing, Hillsdale, NJ, Erlbaum.

Gill, L., Dalgarno, B., & Carlson, L. (2015). How Does Pre-Service Teacher Preparedness to Use ICTs for Learning and Teaching Develop Through Their Degree Program?. Australian Journal of Teacher Education, 40(1). http://dx.doi.org/10.14221/ajte.2015v40n1.3

Gokhale, A. & Machina, K. (2018). Guided Online Group Discussion Enhances Learner Critical Thinking Skills. International Journal on E-Learning, 17(2), 157-173. Waynesville, NC USA: Association for the Advancement of Computing in Education (AACE). Retrieved June 18, 2019 from https://www.learntechlib.org/primary/p/173291/ .

Greaves, T., Hayes, J., Wilson, L., Gielniak, M., & Peterson, R. (2010). The technology factor: Nine keys to learner achievement and cost-effectiveness. Project RED. Retrieved from http://pearsonfoundation.org/downloads/ProjectRED_TheTechnolgyFactor.pdf

Green, B. (1988). Subject-specific literacy and school learning: a focus on writing, Australian Journal of Education 32 (2), pp. 156-179.

Griffin, P. (2001). Using Indicators of Quality to Infer Competence, paper presented t Bringing Assessment and Curriculum Issues Together, National ACACA Conference 2001, Sydney, July 25-27.

Griffin, P., & Care, E. (Eds). (2014). Assessment and teaching of 21st century skills. Methods and approach. Dordrecht, Netherlands: Springer.

Griffin, P., McGaw, B. & Care, E. (Eds). (2012). Assessment and teaching of 21st century skills, Dordrecht, Netherlands: Springer.

Gunn, T. M., & Hollingsworth, M. (2013). The Implementation of a Shared 21st Century Learning Vision: A District-Based Approach. Journal of Research on Technology in Education, Vol. 45, No. 3, pp. 201-228.

Haertel, E. H. (2006). Reliability. In R. L. Brennan (ed.), Educational measurement (4th ed), pp. 65-110. Westport, CT: American Council on Education/Praeger.

Hamer, R., Manlove, S., & Adams, C. (2016). Innovative international eAssessements at the completion of the IB's Middle Years Programme (11-16 years). International Association for Educational Assessment (IAEA). Retrieved from https://iaea.info/documents/innovative-international-eassessments-at-the-completion-of-the-ibs-middle-years-programme-11-16-years

Harlen, W. & Johnson, S. (2014). A review of current thinking and practices in assessment in relation to the Primary Years Programme. Retrieved from https://www.ibo.org/globalassets/publications/ib-research/assessmentinthepypfinalreport.pdf

Hatlevik, O., Guðmundsdóttir, G.B. & Loi, M. (2015). Digital diversity among upper secondary learners: A multilevel analysis of the relationship between cultural capital, self-efficacy, strategic use of information and digital competence. Computers & Education, 81, pp. 345-353.

Hattie, J. (2009). Visible learning, Routledge: Oxford.

Hayward, L., Jones, D. E., Waters, J., Makara, K., Morrison-Love, D., Spencer, E., Barnes, J., Davies, H., Hughes, S., Jones, C., Nelson, S., Ryder, N., Stacey, D., Wallis, R., Baxter, J., MacBride, G., Bendall, R., Brooks, S., Cooze, A., Davies, L., Denny, H., Donaldson, P., Hughes, S., Lewis, I., Lloyd, P., Maitra, S., Morgan, C, Pellew James, S., Samuel-Thomas, S., Sharpling, E., Southern, A., Stewart, S., Valdera-Gil, F., Wardle, G. (2018). CAMAU Project: Research Report, Glasgow: University of Glasgow; Swansea: University of Wales Trinity Saint David.

Hendry, N., Walsh, J., & Patten, S. (2018). The Yes Project educator guide. eSafety Commissioner. Retrieved from http://apo.org.au/system/files/192936/apo-nid192936-1007461.pdf

Heritage, M. (2008). Learning progressions: Supporting instruction and formative assessment. Council of Chief State School Officers (CCSSO).

Heritage, M. (2009). Assessment for Teaching and Learning, Exploratory Seminar: Measurement Challenges With the Race to the Top Agenda, Center for K-12 Assessment & Performance Management.

Hernandez-Lara, A., A. Perera-Lluna, and E. Serradell-Lopez. (2018). 'Applying Learning Analytics to Learners' Interactions in Business Simulation Games. The Usefulness of Learning Analytics to Know What Learners Really Learn.' Computers in Human Behavior, 1–13. doi:10.1016/j.chb.2018.03.001.

Hill, P. & Barber, M. (2014). Preparing for a Renaissance in Assessment, London: Pearson.

Hirsh-Pasek, K., Zosh, J.M. Golinkoff, R., Gray, J., Robb, M., & Kaufman, J. (2015). Putting education in 'educational' apps: Lessons from the Science of Learning. Psychological Science in the Public Interest, 16, 3-34. doi:10.1177/1529100615569721.

Hollingsworth, H. & Heard, J. (2018). Communicating learner learning progress: What does that mean and can it make a difference? Research Conference, Retrieved from: https://research.acer.edu.au/cgi/viewcontent.cgi?article=1337&context=research_conference.

Hooft Graafland, J. (2018). New technologies and 21st century children: Recent trends and outcomes, OECD Education Working Papers, No. 179, Paris, France: OECD Publishing.

Hopson, M. H., Simms, R. L., & Knezek, G. A. (2001). Using a technology-enriched environment to improve higherorder thinking skills. Journal of Research on Technology in Education, 34(2), 109–119.

Howard, S. (2018). The practice of digital writing: Benefits, challenges and choice. NSW Department of Education. Retrieved from https://schoolsequella.det.nsw.edu.au/file/32218e44-8f9f-4eb5-b2f8-0e814a2d7bf2/1/T4Lengage.zip/images/DigitalWritingLitReview2018.pdf

Hoyles, Celia (2018). Transforming the mathematical practices of learners and teachers through digital technology, Research in Mathematics Education, 20:3, 209-228.

Hughes, A. (1994). Backwash and TOEFL 2000. Unpublished manuscript, commissioned by Educational Testing Service (ETS). University of Reading.

Iacovides, I., Aczel, J., Scanlon, E., Taylor, J., & Woods, W. (2011). Motivation, engagement and learning through digital games. International Journal of Virtual and Personal Learning Environments, 2, 1–16.

IBO (2018c). What is eAssessment? Retrieved from https://player.vimeo.com/video/90527836. Assessment Principles and Practices – Quality Assessments in a Digital Age, pp. 1 – 267.

IBO. (2014). MYP: From principles into practice. Cardiff, UK: International Baccalaureate Organization.

IBO. (2018b). PYP Learning and teaching. International Baccalaureate Organization: Cardiff, UK.

Ifenthaler, D., Eseryel, D., & Ge, X. (2012). Assessment for game-based learning. In Assessment in game-based learning(pp. 1-8). Springer, New York, NY.

Inan, F. A., Lowther, D. L., Ross, S. M., & Strahl, D. (2010). Pattern of classroom activities during learners' use of computers: Relations between instructional strategies and computer applications. Teaching and Teacher Education, 26(3).

International Association for Educational Assessment (2015). ICILS 2013 Technical Report. Amsterdam: IEA.

International Association for Educational Assessment (IAEA) (2018). A model for the implementation of digital examinations: New Zealand's vision for digital assessment, IAEA. Retrieved from https://iaea.info/documents/a-model-for-the-implementation-of-digital-examinations-new-zealands-vision-for-digital-assessment/

International Association for Evaluation of Educational Achievement (IEA) (nd). TIMSS 2019. Retrieved from https://www.iea.nl/sites/default/files/studies/T2019_TIMSS_BROCHURE.pdf

International Association for Evaluation of Educational Achievement (IEA). (2018a). TIMSS. Trends in International Mathematics and Science Study Retrieved from https://www.iea.nl/timss

International Association for Evaluation of Educational Achievement (IEA) (2018b). Progress in International Reading Literacy Study. Retrieved from https://www.iea.nl/pirls

International Association for Evaluation of Educational Achievement (IEA) (2018c) International Civic and Citizenship Education Study (ICCS). Retrieved from https://www.iea.nl/iccs

International Association for Evaluation of Educational Achievement (IEA) (2018d) ICILS. International Computer and Information Literacy Study, Overview. Retrieved from https://www.iea.nl/icils

International Association for Evaluation of Educational Achievement (IEA) (2018e) IEA International Computer and Information literacy study 2018. How well are learners prepared for study, work, and life in the digital age? Retrieved from https://www.iea.nl/sites/default/files/studies/IEA%20ICILS%202018%20leaflet.pdf

International Baccalaureate (2018a). Assessment principles and practices: Quality assessments in a digital age. The Hague: International Baccalaureate.

International Baccalaureate (IB) (2005-2018a). Facts and Figures, Retrieved from https://www.ibo.org/about-the-ib/facts-and-figures/
International Baccalaureate (IB) (2005-2018b). Office Locations, Retrieved from https://www.ibo.org/contact-the-ib/office-locations/
International Baccalaureate (IB) (2005-2018c). Primary Years Programme, Retrieved from
International Baccalaureate (IB) (2005-2018d). How the PYP works, Retrieved from https://www.ibo.org/programmes/primary-years-programme/what-is-the-pyp/how-the-pyp-works/
International Baccalaureate (IB) (2005-2018e). Middle Years Programme, Retrieved from https://www.ibo.org/programmes/middle-years-programme/
International Baccalaureate (IB) (2005-2018f). What is the MYP? Retrieved from https://www.ibo.org/programmes/middle-years-programme/what-is-the-myp/
International Baccalaureate (IB) (2005-2018g). Curriculum, Retrieved from https://www.ibo.org/programmes/middle-years-programme/curriculum/
International Baccalaureate (IB) (2005-2018i) DP Online, Retrieved from https://www.ibo.org/programmes/diploma-programme/dp-online/
International Baccalaureate (IB) (2005-2018j). Career-related Programme, Retrieved from https://www.ibo.org/programmes/career-related-programme/
International Baccalaureate (IB) (2008). Science scope and sequence, Retrieved from http://dypisworli.in/pdfs/Science%20scope%20and%20sequence.pdf
International Baccalaureate (IB) (2010/2016a). Guide to School Authorization: Primary Years Programme, Retrieved from
International Baccalaureate (IB) (2010/2016b). Guide to School Authorization: Diploma Programme, Retrieved from https://www.ibo.org/globalassets/publications/become-an-ib-school/dp-guide-school-authorization-en.pdf
International Baccalaureate (IB) (2013). IB Learner profile. Retrieved from https://www.ibo.org/contentassets/fd82f70643ef4086b7d3f292cc214962/learner-profile-en.pdf
International Baccalaureate (IB) (2014). Approaches to teaching and learning in the International Baccalaureate (IB) Diploma Programme. Retrieved from https://www.ibo.org/globalassets/digital-tookit/flyers-and-artworks/approaches-to-teaching-learning-dp-en.pdf
International Baccalaureate (IB) (2014). Middle Years Programme: Guide to Assessment, Retrieved from https://www.ibo.org/globalassets/digital-tookit/brochures/1503-myp-eassessment-factsheet.pdf
International Baccalaureate (IB) (2015/2016). Guide to School Authorization: Middle Years Programme, Retrieved from https://www.ibo.org/globalassets/publications/become-an-ib-school/myp-guide-to-authorization-en.pdf
International Baccalaureate (IB) (2017a). The history of the IB, Retrieved from https://www.ibo.org/globalassets/digital-tookit/presentations/1711-presentation-history-of-the-ib-en.pdf
International Baccalaureate (IB) (2017b). The IB Diploma Programme Statistical Bulletin, May 2017 Examination Session, Retrieved from https://www.ibo.org/contentassets/bc850970f4e54b87828f83c7976a4db6/dp-statistical-bulletin-may-2017-en.pdf
International Baccalaureate (IB) (2017c). The IB Diploma Programme Statistical Bulletin, November 2017 Examination Retrieved from https://www.ibo.org/contentassets/bc850970f4e54b87828f83c7976a4db6/dp-statistical-bulletin-november-2017-en.pdf

International Baccalaureate (IB) (2018a). Assessment Principles and Practices – Quality Assessments in A digital Age, pp. 1 – 267.

International Baccalaureate (IB) (IB) (2005-2018h). Curriculum, Retrieved from https://www.ibo.org/programmes/diploma-programme/curriculum/

International English Language Testing System (IELTS). (nd). Introducing computer-delivered IELTS Retrieved from https://ielts.org/-/media/publications/computer-delivered-ielts/computer-delivered-tt-brochure.ashx?la=en .

International Education News (2016). Testing and Assessment in Norway. Retrieved 1 July 2019 from https://internationalednews.com/2016/02/11/testing-and-assessment-in-norway/

JISC (2010) Effective Assessment in a Digital Age: A guide to technology-enhanced assessment and feedback. Retrieved 14 January 2018 from https://facultyinnovate.utexas.edu/sites/default/files/digiassass_eada.pdf

Kali, Y., McKenney, S. & Sagy, O. (2015). Teachers as designers of technology enhanced learning. Instructional Science, 43(2): 173-179.

Kava, B., Andrade, A. Marcovich, R. Idress, T. & Ruiz, J. (2017). Communication Skills Assessment Using Human Avatars: Piloting a Virtual World Objective Structured Clinical Examination. Urology Practice, 4 (1): 76-84.

Ketamo, H., Devlin, K. & Kiili, K. (2018). Gamifying Assessment: Extending Performance Measures with Gaming Data. Retrieved 1 June 2019 from https://web.stanford.edu/~kdevlin/Papers/Ketamo_Devlin_Kiili_2018.pdf

Kimmo, K. (2017). Finland Country Report on ICT in Education. Retrieved 1 June 2019 from http://www.eun.org/documents/411753/839549/Country+Report_Finland_2017.pdf/f106f29c-7092-44e3-9ecf-5ae24b521cab

Kincheloe, J. (1997) Foreword in I.F. Goodson, (ed). The changing curriculum: Studies in social construction, Studies in the Postmodern Theory of Education, Vol. 18, Peter Lang Publishing, London

King, J., South, J., & Stevens, K. (2016). Advancing Educational Technology in Teacher Preparation: Policy Brief. U.S. Department of Education.

Kingsbury, G. G., Freeman, E. H., & Nesterak, M. (2014). The Potential of Adaptive Assessment. Educational Leadership, 71. Retrieved from http://www.ascd.org/publications/educational-leadership/mar14/vol71/num06/The-Potential-of-Adaptive-Assessment.aspx

Kirschner, P. (2015). Do we need teachers as designers of technology enhanced learning? Instructional Science, 43 (2): 309-322.

Klenowski, V. & Wyatt-Smith, C. (2014) Assessment for Education, Standards, Judgement and Moderation, Sage Publications.

Knezek, G. & Christensen, R. (2018). The evolving role of attitudes and competencies in information and communication technology in education, in J. Voogt, G. Knezek, R. Christensen & Kwok-Wing, L. (eds). Second Handbook of Information Technology in Primary and Secondary Education, Volume 1, Springer International Publishing

Knight, S. (2009). Effective Practice in a Digital Age. A guide to technology-enhanced learning and teaching. Higher Education Funding Council for England (HEFCE).

LaBonte, R. (2016). Blending Assessment: Shifting Pen & Paper to the Digital World. Retrieved 1 June 2019 from https://CANeLearn.net/research

Lamb, S., Maire, Q. & Doecke, E. (2017). Key Skills for the 21st Century: An evidence-based review. Victoria University. Retrieved from https://education.nsw.gov.au/our-priorities/innovate-for-the-future/education-for-a-changing-world/research-findings/future-frontiers-analytical-report-key-skills-for-the-21st-century/Key-Skills-for-the-21st-Century-Analytical-Report.pdf

Lawlor, J., Marshall, K., & Tangney, B. (2016). BRIDGE21 - Exploring the potential to foster intrinsic learner motivation through a team-based, technology-mediated learning model. Technology, Pedagogy and Education, 25(2), pp. 187-206.

LEARN Research Briefing. Canadian eLearning Network. Retrieved from http://canelearn.net/wp-content/uploads/2016/05/CANeLearn-Digital-Assessment-Overview.pdf

Lei, J. (2009) Digital Natives As Preservice Teachers, Journal of Computing in Teacher Education, 25:3, 87-97.

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

Liu, C., Solis, S. L., Jensen, H., Hopkins, E. J., Neale, D., Zosh, J. M., Hirsh-Pasek, K., & Whitebread, D. (2017). Neuroscience and learning through play: a review of the evidence. Billund, Denmark: The LEGO Foundation. Retrieved from https://www.legofoundation.com/media/1064/neuroscience-review_web.pdf

Looney, J. (2009). Assessment and innovation in education. OECD Education Working Papers, No. 24.

Luckin, R & Cukurova, M. (2019). Designing educational technologies in the age of AI: A learning sciences-driven approach. British Journal of Educational Technology, First published: 21 July 2019 https://doi.org/10.1111/bjet.12861

Madaus, G. (2001). Educational testing as a technology. National Board on Educational Testing and Public Policy Statements, 2(1). Retrieved from http://www.bc.edu/research/nbetpp/publications/v2n1.html

Marcus, J. (2012), 'Old school: four-hundred years of resistance to change', in Wildavsky, B., Kelly, A. and Carey, K. (Eds), Reinventing Higher Education: The Promise of Innovation, Harvard Education Press, Cambridge, MA, pp. 41-72.

Masters, G. (2013). Reforming educational assessment: Imperatives, principles and challenges. Melbourne, Australia: Australian Council for Educational Research. Retrieved from https://research.acer.edu.au/cgi/viewcontent.cgi?article=1021&context=aer

Masters, G. N. (2013). Reforming educational assessment: imperatives, principles and challenges. Australian Council for Educational Research, Camberwell, Victoria.

Masters, G. N. (2014a) Towards a Growth Mindset, Address to the Queensland Curriculum and Assessment Authority, Mini-Conference, 2014. Reproduced and adapted from: The State of Queensland (Queensland Curriculum and Assessment Authority). Retrieved from https://www.acer.org/cari/videos/geoff-masters-on-assessment-qcaa-mini-conference-2014

Masters, G. N. (2014b) Towards a growth mindset in assessment, in Practically Primary, June. Retrieved from https://www.alea.edu.au/documents/item/896

Masters, G. N. (2016) Monitoring learner growth, in Australian Council for Educational Research ACER, Teacher Evidence + Insight + Action, Retrieved from https://www.teachermagazine.com.au/columnists/geoff-masters/monitoring-learner-growth.

Masters, G.N. (2018) Challenging our most able learners, in Teacher Evidence + Insight + Action, Retrieved from https://www.teachermagazine.com.au/columnists/geoff-masters/challenging-our-most-able-learners

McCabe, D. B., & Meuter, M. L. (2011). A Learner View of Technology in the Classroom: Does It Enhance the Seven Principles of Good Practice in Undergraduate Education? Journal of Marketing Education, 33(2), 149–159. https://doi.org/10.1177/0273475311410847

McCurry, D. (2013). Teaching critical thinking. The Research Digest, 9. Queensland College of Teachers. Retrieved from https://www.qct.edu.au/pdf/Research%20Periodicals/QCTResearchDigest2013-9.pdf

McGaw, B. (1996) Their Future: Options for Reform of the Higher School Certificate, Department for Training and Education Co-ordination, New South Wales, Sydney.

McGrath, J., Taekman, J. Dev. P., Danforth, D. Mohan, D. Kman, N. Crichlow, A. & Bond, W. (2017) Using Virtual Reality Simulation Environments to Assess Competence for Emergency Medicine Learners. Academic Emergency Medicine, 25(2): 186-195.

McKenney, S., Kali, Y., Markauskaite, L. & Voogt, J. (2015). Teacher design knowledge for technology enhanced learning: an ecological framework for investigating assets and needs. Instructional Science, 43(2): 181–202.

McKew, M. in conversation with Masters, G. (2018) Melbourne Graduate School of Education (MGSE), Talking Teaching Episode 3. https://education.unimelb.edu.au/talkingteaching/talking-teaching-episode-three/episode-3-transcript .

McKnight, K., O'Malley, K., Ruzic, R., Horsley, M., Franey, J., & Bassett, K. (2016). Teaching in a digital age: How educators use technology to improve learner learning. Journal of Research on Technology in Education, 48(3), 194–211.

McNamara, T. (2010) The use of language test in the service of policy: Issues of validity https://www.cairn.info/revue-francaise-de-linguistique-appliquee-2010-1-page-7.htm , Revue Francaise de Linguistique Applique, 2010/1/ (Vol. XV), Section 5.

Messick, S (1989) Validity, in R. L. Linn (ed.) Educational measurement (3rd ed), New York, American Council on Education & Macmillian, pp. 13-103.

Miller, G., Boehner, J., Kennedy, T., & Gregg, J. No Child Left Behind Act of 2001 (NCLB), Pub. L. No. 107-110, § 115, Stat. 1425.

Mishra, P. & Koehler, M.J., (2006). Technological Pedagogical Content Knowledge: A framework for teacher knowledge. Teachers College Record 108, 6, pp 1017-1054

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

Morais, A., Neves, I. and Pires, D. (2004), 'The what and the how of teaching and learning: going deeper into sociological analysis and intervention', in Muller, J., Davies, B. and Morais, A. (Eds), Thinking with Bernstein, Working with Bernstein, Routledge, London.

Mosher, F. & Heritage, M. (2017) A Hitchhiker's Guide to Thinking about Literacy, Learning Progressions, and Instruction. CPRE Research Report #RR 2017/2. Philadelphia: Consortium for Policy Research in Education. Retrieved from http://repository.upenn.edu/cpre_researchreports/97.

Mullis, I. & Martin, M. (eds). (2015). PIRLS 2016, Assessment Framework, 2nd Edition. TIMSS & PIRLS, International Study Center, Lynch School of Education, Boston College. Retrieved from https://timssandpirls.bc.edu/pirls2016/downloads/P16_Framework_2ndEd.pdf

Mullis, I. V. S. (2017). Introduction. In I. V. S. Mullis & M. O. Martin (Eds.), TIMSS 2019 Assessment Frameworks. Retrieved from Boston College, TIMSS & PIRLS International Study Center, pp. 1-10. Retrieved from http://timssandpirls.bc.edu/timss2019/frameworks/

Mullis, I., Martin, M., Ruddock, G., O'Sullivan, C., & Preuschoff, C. (2009). TIMMS 2011, Assessment Frameworks, TIMSS and PIRLS International Study Center, Lynch School of Education, Boston College. Retrieved from https://timss.bc.edu/timss2011/downloads/TIMSS2011_Frameworks.pdf

Murray, A., Luo, T. & Franklin, T. (2019). Embracing a Technologically Enhanced Environment: Teachers' Experience Educating Learners in an Always-on and Connected BYOD Classroom. International Journal on E-Learning, 18(1), 53-78. Waynesville, NC USA: Association for the Advancement of Computing in Education (AACE). Retrieved June 18, 2019 from https://www.learntechlib.org/p/180503.

Nakatsuhara, F., Inoue, C., Berry, V. and Galaczi, E. (2017). Exploring performance across two delivery modes for the IELTS Speaking Test: face-to-face and video-conferencing delivery (Phase 2). IELTS Partnership Research Papers, 3. IELTS Partners: British Council, Cambridge English Language Assessment and IDP: IELTS Australia. Available at https://www.ielts.org/teaching-and-research/research-reports

National Foundation for Education Research (nd) An introduction to formative and summative assessment. Retrieved 15 January 2019 from https://www.nfer.ac.uk/for-schools/free-resources-advice/assessment-hub/introduction-to-assessment/an-introduction-to-formative-and-summative-assessment/

National Research Council. 2010. State Assessment Systems: Exploring Best Practices and Innovations: Summary of Two Workshops. Washington, DC: The National Academies Press.

New Pedagogies for Deep Learning Global Partnership. (2014). Critical thinking deep learning progression. Retrieved from http://npdl.global/making-it-happen/deep-learning-progression

New Zealand Qualifications Authority (2015) Digital Assessment. Retrieved 15 January 2019 from https://www.nzqa.govt.nz/assets/About-us/Our-role/innovation/DAT-factsheet-May15.pdf

New Zealand Qualifications Authority (NZQA) (2018). About the 2018 Digital Trials and Pilots, Retrieved from https://www.nzqa.govt.nz/about-us/future-state/digital-assessment-trials-pilots/about-2018/

New Zealand Qualifications Authority (NZQA) (nd). Digital Assessment Vision: Design Principles Retrieved from https://www.nzqa.govt.nz/about-us/future-state/digital-assessment-vision/

Newton, P. E. (2007) Clarifying the purposes of educational assessment, in Assessment in Education: Principles, Policy & Practice, 14:2,

Newton, P. E. (2017) There is More to Educational Measurement Than Measuring: The Importance of Embracing Purpose Pluralism, in Educational Measurement: Issues and Practice, Summer 2017, Vol. 36, No. 2. pp. 5-15.

Norman, D. (1988) The psychology of everyday things, London, Basic Books

Norris, J. (2018) Test of English as a Foreign Language, (TOEFL) Research Insight Series, Volume 6. INSIGHTS, Education Testing Service (ETS). Retrieved from https://www.ets.org/s/toefl/pdf/toefl_ibt_insight_s1v6.pdf

Nussbaumer, A., Hillemann, E. C., G€uTl, C., & Albert, D. (2015). A competence-based service for supporting self-regulated learning in virtual environments. Journal of Learning Analytics, 2, 101–133.

O'Leary, M., Scully, D., Karakolidis A. & Pitsia, V. (2018). The state-of-the-art in digital technology-based assessment, European Journal of Education, March,

O'Reilly, E. N. (2016). Developing Technology Needs Assessments for Education Programs: An Analysis of Eight Key Indicators. International Journal of Education and Development using Information and Communication Technology, Vol. 12, Issue 1, pp. 129-143.

OCED (2018) Preparing our youth for an inclusive and sustainable world: The OECD PISA Global Competence Framework. Paris: OECD.

OECD Programme for International Learner Assessment (2015) PISA 2015 released field trial cognitive items. Retrieved 15 January 2019 from http://www.oecd.org/pisa/pisaproducts/PISA2015-Released-FT-Cognitive-Items.pdf

OECD Programme for International Learner Assessment (2018) PISA 2018 released field trial new reading items. Retrieved 15 January 2019 from http://www.oecd.org/pisa/test/PISA_2018_FT_Released_New_Reading_Items.pdf

Oldfield, A., Broadfoot,P., Sutherland, R. and Timmis, S. (2012) Assessment in a digital age: a research review . Available from: http://www.bristol.ac.uk/education/research/sites/tea/publications/index.html

Organisation for Economic Co-operation and Development, (OECD). (2014). PISA 2012 Results: Creative Problem Solving: Learners' skills in tackling real-life problems (Volume V). PISA, OECD Publishing. doi: 10.1787/9789264208070-en

Organisation for Economic Co-operation and Development, (OECD). (2018). Preparing our youth for an inclusive and sustainable world: The OECD PISA global competence framework. Retrieved from https://www.oecd.org/pisa/Handbook-PISA-2018-Global-Competence.pdf

Organisation for Economic Cooperation and Development, OECD (2016), PISA 2015 Results (Volume I): Excellence and Equity in Education, PISA, OECD Publishing, Paris. http://dx.doi.org/10.1787/9789264266490-en

Paniagua, A. & Istance, D. (2018), Teachers as Designers of Learning Environments: The Importance of Innovative Pedagogies, Educational Research and Innovation. Paris, France: OECD Publishing. doi: 10.1787/9789264085374-en.

Papanastasiou, E. (2003, July). Computer-adaptive testing in science education. In Proceedings of the 6th International Conference on Computer Based Learning in Science, Nicosia, Cyprus (pp. 965-971).

Partnership for 21st Century Learning. (2002). Framework for 21st century learning. Retrieved from http://www.p21.org/about-us/p21-framework

Pashler, N. (2001). Connecting schools and pupils: To what end?: Issues related to the use of ICT in school-based learning in M. Leask (ed) Issues in teaching using ICT, London: Routledge/Falmer.

Pásztor, A., Molnár, G., & Csapó, B. (2015). Technology-based assessment of creativity in educational context: the case of divergent thinking and its relation to mathematical achievement. Thinking Skills and Creativity, 18, pp. 32–42.

Patrick, S., Worthen, M. & Frost, D. (2017). Redesigning Systems of Assessments for Learner-Centered Learning. iNACOL. Retrieved from https://www.inacol.org/wp-content/uploads/2017/11/iNACOL-RedesigningSystemsOfAssessments.pdf

Pearson Test of English (PTE) (2011). Pearson Test of English Academic: Automated Scoring Pearson. Retrieved from https://pearsonpte.com/wp-content/uploads/2015/05/7.-PTEA_Automated_Scoring.pdf

Pearson Test of English (PTE) (2017a). Pearson Test of English, Academic, Pearson. Retrieved from https://pearsonpte.com/

Pearson Test of English (PTE) (2017b). Test Format, Pearson. Retrieved from https://pearsonpte.com/the-test/format/

Polikov, V. (2017) Innovation in Education Is More than a New Approach, in Getting Smart: Ed Tech and Data. https://www.gettingsmart.com/2017/04/innovation-in-education-is-more-than-a-new-approach/

Poutasi, K. (2018). SPANZ 2018: address by Dr Karen Poutasi, Chief Executive of NZQA Retrieved from https://www.nzqa.govt.nz/assets/Uploads/SPANZ-2018.pdf

Power, C. (1986) Criterion-based Assessment, Grading and Reporting at Year 12 level, in Australian Journal of Education, Vol. 30, No. 3, 1986, pp. 266-84.

Psotka, J. (2013). Educational games and virtual reality as disruptive technologies.

Puentedura, R. (2013, 2/2/2013). Technology in Education: A Brief Introduction. [Video]. Retrieved from http://www.hippasus.com/rrpweblog/archives/000080.html

Raaheim, A., Mathiassen, K., Moen, V., Lona, I., Gynnild, V., Ringlund Bunæs, B. & Trygve Hasle, E. (2019) Digital assessment – how does it challenge local practices and national law? A Norwegian case study, European Journal of Higher Education, 9:2, 219-231.

Redecker, C. and Johannessen, Ø. (2013), Changing Assessment — Towards a New Assessment Paradigm Using ICT. European Journal of Education, 48: 79-96

Reiners, T., Gregory, S., & Dreher, H. (2011). Educational assessment in virtual world environments. In Proceedings of the Australian Technology Network Assessment Conference 2011 (pp. 132-140). Curtin University.

Richardson, S. (2017) Cosmopolitan Education for a Global Era: Higher Education in an Interconnected World. Abingdon: Routledge.

Rienties, B., Giesbers, B., Lygo-Baker, S., Ma, H-W & Rees, R. (2014). Why some teachers easily learn to use a new virtual learning environment: a technology acceptance perspective. Interactive Learning Environments, 24 (3): 539-552.

Riggio, R. (2014). The 'hard' science of studying and developing leader 'soft' skills. Leader interpersonal and influence skills. New York, NY: Routledge.

Roach, J. (1971). Public examinations in England 1850-1900. Cambridge, United Kingdom: Cambridge University Press.

Robson, S. (2014). The Analysing Children's Creative Thinking framework: development of an observation-led approach to identifying and analyzing young children's creative thinking. British Educational Research Journal, 40(1), pp. 121-134.

Ruggiero, D., & Mong, C. J. (2015). The teacher technology integration experience: Practice and reflection in the classroom. Journal of Information Technology Education: Research, 14, 161-178.

Sadler, D. R. (1989). Formative Assessment and the Design of Instructional Systems, in Instructional Science 18 (2): pp. 118-144.

Scalise, K. (2009) Computer-Based Assessment: Intermediate Constraint Questions and Tasks for Technology Platforms. Accessed 13 January 2019 from https://pages.uoregon.edu/kscalise/taxonomy/taxonomy.html

Scanlon, E., Anastopoulou, S., & Kerawalla, L. (2012). Inquiry learning reconsidered: contexts, representations and challenges. In K. Littleton, E. Scanlon & M. Sharples (Eds.), Orchestrating inquiry learning. Abingdon, England: Routledge.

Schuwirth L, van der Vleuten C, Durning SJ. What programmatic assessment in medical education can learn from healthcare. Perspect Med Educ. 2017;6 (4):211-215.

Scoular, C. (2018). Equipping teachers with tools to assess and teach general capabilities. Research Conference, 2018. Retrieved from https://research.acer.edu.au/cgi/viewcontent.cgi?article=1343&context=research_conference

Scriven, M. (1967). The Methodology of Evaluation, in Perspectives of Curriculum Evaluation, edited by R. W. Tyler, R. M. Gagné, and M. Scriven, 39-83. IL: Rand McNally.

Serdyukov, P. (2017) 'Innovation in education: what works, what doesn't, and what to do about it?', Journal of Research in Innovative Teaching & Learning, Vol. 10 Issue: 1, pp.4-33.

Sergis, S., Sampson, D. G., Pelliccione, L. I. (2018) nvestigating the impact of Flipped Classroom on learners' learning experiences: A Self-Determination Theory approach, Computers in Human Behavior, Volume 78, 2018, Pages 368-378,.

Sfard, A. (1998) On Two Metaphors for Learning and the Dangers of Choosing Just One, in Educational Researcher, Vol. 27, No. 2, pp. 4-13.

Shota Hashimura, S., Shimakawa, H. & Kajiwara, Y. (2018). Automatic Assessment of Student Understanding Level Using Virtual Reality. 2018 Federated Conference on Computer Science and Information Systems, 9-12 Sept. 2018, Poznan, Poland.

Shulman, L. S. (1987). Knowledge and Teaching: Foundations of the New Reform. Harvard Educational Review (1), pp. 1-22.

Shute, V. J., Leighton, J. P., Jang, E. E., & Chu, M.-W. (2016). Advances in the science of assessment. Educational Assessment, 21, pp 34–59.

Singh, M. & Qi, J. (2013). 21st century international mindedness: An exploratory study of its conceptualisation and assessment. Centre for Educational Research, University of Western Sydney. Retrieved from https://www.ibo.org/globalassets/publications/ib-research/singhqiibreport27julyfinalversion.pdf

Smith, R. & Gray, L. (2016) Enhancing assessment and feedback with technology: a guide for FE and skills. Guidance for colleges and training providers on how technology can add value to assessment and feedback. JISC.

Sniedze-Gregory, S. (2018). Enhancing Interdisciplinary Teaching and Learning through Guided Assessment Design. (PhD thesis). Flinders University, Australia. Retrieved from https://theses.flinders.edu.au/view/072add24-fee6-4f07-bac1-e9b129de111c/1

Stanley, G. (2013) Foreword, in Masters, G. N. & Australian Council for Educational Research (2013). Reforming educational assessment: imperatives, principles and challenges. Australian Council for Educational Research, Camberwell, Victoria.

Stiggins, R. (2005). Assessment for Learning Defined. ETS/ Assessment Training Institute's International Conference. Promoting Sound Assessment in Every Classroom. Portland, OR. Retrieved 26 August 2018. Retrieved from http://downloads.pearsonassessments.com/ati/downloads/afldefined.pdf .

Stiggins, R. J. (1995) Assessment Literacy for the 21st Century in Phi Delta Kappan, Vol. 77l, Issue 3, pp. 238-245.

Stiggins, R.J. (2002). Assessment crisis: The absence of Assessment FOR Learning. Phi Delta Kappan, pp. 758-765.

Stobart, G. (2008) Testing times: The uses and abuses of assessment, Routledge, London.

Stommel. J. (2015) Digital Pedagogy: A Genealogy. Retrieved 13 January 2019 from https://www.tiki-toki.com/timeline/entry/392826/Digital-Pedagogy-a-Genealogy/

Sweeney, T., West, D., Groessler, A., Haynie, A., Higgs, B. M., Macaulay, J., Mercer-Mapstone, L. and Yeo, M. (2017) Where's the transformation? Unlocking the potential of technology-enhanced assessment', Teaching and Learning Inquiry, 5(1), pp. 1-16.

Tancock, S, Dahnoun, Y & Dahnoun, N (2018) Real-Time and Non-digital Feedback E-learning Tool. in 2018 International Symposium on Educational Technology (ISET 2018)., 8456190, Institute of Electrical and Electronics Engineers (IEEE), pp. 57-59.

Tenório, T., Bittencourt, I. Isotani, S., Pedro, A. Ospina, P. & Tenório, D. (2017). Dataset of two experiments of the application of gamified peer assessment model into online learning environment MeuTutor. Data in Brief, 12:433-437.

Test of English as a Foreign Language (TOEFL) iBT Test Content, Retrieved from https://www.ets.org/toefl/ibt/about/content/

Thai, T., De Wever, B., Valcke, M. (2017) The impact of a flipped classroom design on learning performance in higher education: Looking for the best 'blend' of lectures and guiding questions with feedback, Computers & Education, Volume 107, 2017, Pages 113-126,

Thomson, S., & De Bortoli, L. (2012). Preparing Australian learners for the digital world. Melbourne: ACER.

Thonbo, R. (2017). Denmark Country Report on ICT in Education. Retrieved 1 Jun 2019 from http://www.eun.org/documents/411753/839549/Country+Report+Denmark+2017.pdf/7a0b9045-cd44-4831-875a-e42306beeefe

Timmis, S, Broadfoot, P, Sutherland, R & Oldfield, A (2016) 'Rethinking assessment in a digital age: opportunities, challenges and risks' British Educational Research Journal, vol. 42, no. 3, pp. 454 -476.

Timmis, S., Broadfoot, P., Sutherland, R. & Oldfield, A. (2016). Rethinking assessment in a digital age: opportunities, challenges and risks. British Educational Research Journal, 42, (3), pp. 454–476.

Tobin, M., Lietz, P., Nugroho, D., Vivekanandan, R., & Nyamkhuu, T. (2015). Using large-scale assessments of learners' learning to inform education policy: Insights from the Asia-Pacific region. Melbourne: ACER and Bangkok: United Nations Educational, Scientific and Cultural Organization (UNESCO)

Toe, D, Lang, J, Paatsch, L, Yim, B, Jobling, W, Doig, B & Aranda, G. 2016. Assessment of learner development and learning in International Baccalaureate Primary Years Programme schools. The Hague, Netherlands: International Baccalaureate Organization. Retrieved from

Tsagari, D. (2007). Review of Washback in language Testing: What Has Been Done? What More Needs Doing? Lancaster University, Retrieved from https://files.eric.ed.gov/fulltext/ED497709.pdf

Tucker, M. (2018, Sep 21). On Nobel prizes that do not yet exist and other exciting matters. Tucker's Blog, 21 September, 2018. Retrieved from http://ncee.org/2018/09/on-nobel-prizes-that-do-not-yet-exist-and-other-exciting-matters

Turner, R., Adams, R., Schwantner, U., Cloney, D., Scoular, C., Anderson, P., Daraganov, A., Jackson, J., Knowles, S., O'Connor, G., Munro-Smith, P., Zoumboulis, S. & Rogers, P. (2018) Development of Reporting Scales for Reading and Mathematics: A report describing the process for building the UIS Reporting Scales HTTPs://research.acer.edu.au/monitoring_learning/33

Tyack, D. & Cuban, L. (1995) Tinkering Toward Utopia: A Century of Public School Reform, Harvard University Press, Cambridge, Mass.

UNESCO (2013), 'ITL – Innovative teaching and learning research: a global look at pedagogies for 21st century skills', ICT in Education, UNESCO, Bangkok, available at: www.unescobkk.org/education/ict/online-resources/databases/ict-in-education-database/item/article/innovative-teaching-and-learning-itl-research-a-global-look-at-pedagogies-for-21st-century-skills/ (accessed August 18, 2016).

United Nations Educational, Scientific and Cultural Organisation (UNESCO). (2011). ICT Competency Framework. Retrieved from http://unesdoc.unesco.org/images/0021/002134/213475e.pdf

United Nations Educational, Scientific and Cultural Organisation UNESCO (2015). Global Citizenship Education: Topics and Learning Objectives. Retrieved from http://unesdoc.unesco.org/images/0023/002329/232993e.pdf

Vieira, C., O. Parsons, and V. Byrd. 2018. 'Visual Learning Analytics of Educational Data: A Systematic Literature Review and Research Agenda.' Computers & Education 122: 119–135.

Vygotsky, L. S. (1978) Mind in Society. Cambridge, MA: Harvard University Press.

Ward, W. et al (2011), My science tutor: A conversational multimedia virtual tutor for elementary school science, ACM Transactions on Speech and Language Processing (TSLP). Vol., 7 Issue 4, August, Article No. 18

Waxman, H. C., & Huang, S. L. (1996). Classroom instruction differences by level of technology use in middle school mathematics. Journal of Educational Computing Research, 14(2), 157–169.

Whitelock, D. and Brasher, A. (2006) Roadmap for e-assessment. A JISC report. Available from: http://www.jiscinfonet.ac.uk/InfoKits/effective-use-of-VLEs/resources/roadmap-for-eassessment

Wiliam, D. (2008) International comparisons and sensitivity to instruction, in Assessment in Education: Principles, Policy & Practice, Vol. 15, No. 3, November 2008, pp. 253-258.

Wiliam, D. (2011). What is Assessment for Learning? in Studies in Educational Evaluation, Vol. 37, pp. 3-14.

Wiliam, D. (2017) Learning and assessment: a long and winding road?, in Assessment in Education: Principles, Policy & Practice, 24:3, 309-316.

Wright, K.B., Kandel-Cisco, B., Hodges, T.S., Metoyer, S., Boriack, A.W., Franco-Fuenmayor, S.E., Stillisano, J.R. & Waxman, H.C. (2014). Developing and assessing learners' collaboration in the IB programme. The Hague, Netherlands: International Baccalaureate Organization. Retrieved from https://www.ibo.org/globalassets/publications/ib-research/developingandassessinglearnercollaborationfinalreport.pdf

Yang, Min & Wang, Tianchong & Lim, Cher Ping. (2017). E-Portfolios as Digital Assessment Tools in Higher Education. 10.1007/978-3-319-17727-4_83-1.

Yarbro, J., McKnight, K., Elliott, S., Kurz, A., & Wardlow, L. (2016) Digital Instructional Strategies and Their Role in Classroom Learning, Journal of Research on Technology in Education, 48:4, 274-289.

Zosh, J. M., Hopkins, E. J., Jensen, H., Liu, C., Neale, D., Hirsh-Pasek, K., Solis, S. L., & Whitebread, D. (2017). Learning through play: a review of the evidence. Billund, Denmark: The LEGO Foundation. Retrieved from https://www.legofoundation.com/media/1063/learning-through-play_web.pdf