OECD DIGITAL EDUCATION OUTLOOK 2021

Pushing the frontiers with AI, blockchain, and robots

OECD

OECD Digital Education Outlook 2021

PUSHING THE FRONTIERS WITH ARTIFICIAL INTELLIGENCE, BLOCKCHAIN AND ROBOTS

OECD
BETTER POLICIES FOR BETTER LIVES

Please cite this publication as:
OECD (2021), *OECD Digital Education Outlook 2021: Pushing the Frontiers with Artificial Intelligence, Blockchain and Robots*, OECD Publishing, Paris, https://doi.org/10.1787/589b283f-en.

ISBN 978-92-64-64199-0 (print)
ISBN 978-92-64-90464-4 (pdf)
ISBN 978-92-64-87694-1 (HTML)
ISBN 978-92-64-38175-9 (epub)

OECD Digital Education Outlook
ISSN 2788-8568 (print)
ISSN 2788-8576 (online)

Revised version, January 2023
Details of revisions available at: https://www.oecd.org/about/publishing/Corrigendum_OECD-Digital-Education-Outlook-2021.pdf

Photo credits: Cover © Gerhard Richter 2021 (04052021).

Corrigenda to publications may be found on line at: www.oecd.org/about/publishing/corrigenda.htm.

Editorial

When countries closed down schools in early 2020 to deal with the Covid-19 pandemic, learning went digital. In the year since, teachers, students and administrators have done, what is effectively, a collective crash course on digital education. There have been many serious downsides to it, from screen fatigue and adaptation stress to the falling behind of those ill-equipped for digital learning or unprepared to learn on their own. But the experience has catapulted education systems, traditionally laggards when it comes to innovation, years ahead in what would have been a slow slouch towards smart schooling.

Remote classrooms, however, are not the same as smart ones. Rather, they have been a stop-gap measure that has kept learning going and conserved existing educational practice rather than transform it. As a result, remote classrooms have rarely been of the same calibre as physical ones, and, again, not for students without the wherewithal – technologically or mentally – to do online classes.

Two years before the pandemic, OECD's TALIS survey found that only half of teachers were letting students frequently use technology for projects or class work. But necessity is the mother of invention and, in the absence of physical classes, many teachers are catching the digital wave.

Now is the time for schools to dive into digital. Many have liked the "anytime-anywhere" capabilities of remote classes. More and more educators are getting ready, is the technology?

As it turns out, education technology is ready too. The OECD's Digital Education Outlook brings us up to speed on three areas of technology that are already being used in education systems: artificial intelligence (AI) or machine learning, robots and blockchain.

Software and social robots that are fed constant streams of data have the greatest disruption potential for teaching and learning: it's not just technology, it's teachology. While we study mathematics on a computer, the computer can now study how we study and then make our learning experience so much more granular, adaptive and interactive. Together with sensors and learning management systems, AI can give teachers a real sense of how different students learn differently, where students get interested and where they get bored, where they advance and where they get stuck. Teachology can help adapt learning to different student needs and give learners greater ownership over what they learn, how they learn, where they learn and when they learn.

AI can help teachers, especially novice ones, read the room better and slow down, speed up, or throw out a pop quiz question when there's a lull. Learning analytics can tell a teacher working out the next day's lesson plan who aced the homework assignment on carboxylic acid derivatives and who needs to review it still. A classroom robot can take Gabriel and Ishita to a corner of the classroom and have a 10-minute Spanish conversation while the rest of the class works on action verbs.

And of course, AI is helping assessment and exams make big leaps, whether these are assessments through simulations, hands-on assessments in vocational settings, or machine-learning algorithms scoring essays. One of the most consequential mistakes that education made as it industrialised learning over the past centuries was to divorce learning from assessment; that is, having students pile up lots of learning and then, sometimes years later, testing whether they could reproduce some narrow slices of that learning within a short window of time. Technology can now reintegrate learning and assessment, using real-time data and feedback to help students learn better, teachers teach better and education systems become more effective.

But for teachology to really take off it has to be user-friendly. Data on what percentage of the quadratic equations unit Hanzhou has mastered or whether Emilia has gotten bored with post-war social welfare history is no good if the teacher has to stop a classroom lesson to consult the data. Data needs to be intuitive. Education ministries can encourage developers to co-create with teachers and students digital tools that are relevant, affordable, interoperable and easy to use. Technology is unlikely to work for learning unless the teaching profession is part of the design of those tools.

Another constituent that can especially benefit from smart education tools are students with special needs. Smart software and robot tutors can adapt to Hector's learning needs or Farid's pace. They can help detect and diagnose problems that too often go under the radar, especially in primary school. And they can help plot curricular trajectories that fit each student's needs better.

That data-powered technology can help level the playing field (in the classroom!) and has applications beyond students with special needs to the situation now at hand. We face what is likely a catch-up period for young people who have struggled with school during the pandemic. Pinpointing where they need help and where they have excelled is something individualised Edtech can support.

Of the three areas of technology the report covers, blockchain is the most mature though applications, so far, are not in teaching and learning. Blockchain looks promising as a reliable, user-friendly credentialing system that can replace lumpy and expensive degrees, and help unbundle the institutional monopolies that often come with them. Authenticated certificates of completion from education and training programmes outside traditional academic institutions – like on-the-job training and massive open online courses (MOOCs) – are an important piece of the puzzle in bringing us closer to lifelong, life-wide learning. If everybody, independent of their jobs, can upskill and reskill and have blockchain-verified qualifications at their fingers, job-changing will be faster and more fluid, and much less anxiety-ridden.

But going back to teaching and learning, more technology here does not automatically translate into better learning outcomes. In fact, results from OECD's latest PISA assessment showed a persistent negative relationship between the intensity of technology use in classrooms and the digital reading, mathematics and science skills of 15-year-olds. Students who spent more time posting work on their school's website, playing simulations at school, using learning apps and websites or doing homework on a school computer tended to perform more poorly on the assessment.

Of course, there are many reasons that can explain the connection between higher tech use at school and lower cognitive performance. Perhaps lower-performing students simply spend more time on their homework. Or they spend more time on the computer because they are being directed to more practical digital assignments. It is also possible that the digital world helps develop knowledge and skills that are not easily captured by current assessments. But we should not discard the possibility that low-quality digital learning tools displace valuable instructional activities that could be better done without digital devices, or activities that teachers simply know better how to do in the analogue world. Not least, the demands that effective digital learning place on students' autonomy – their capacity for independent learning as well as their executive functioning and self-monitoring – are easily underestimated. These hypotheses are supported by the fact that the relationship between technology use and learning outcomes varies so widely across countries.

What is clear is that for robots, classware, predictive analytics and the like to work effectively will require reinventing the role of teachers. Technology and AI are not magic powers, they are just extraordinary amplifiers and accelerators that add speed and accuracy. AI will amplify good educational ideas and good practice in the same way it amplifies bad ideas and bad practice. AI can help remove bias and discrimination from educational practice in the same way it can spread and scale bias in educational practice. It can empower teachers to identify children at risk or disempower them from exercising human judgment. In so doing, AI can induce a paradigm shift from an education of consequences – with teachers helping their students understand who they are and who they want to become – to an education of correlations where all the technology does is to look back at what has happened with students with similar characteristics in the past. While technology is ethically neutral, it will always be in the hands of educators who are not neutral. The real risks do not come from AI but from the consequences of its application. When early warning systems flag a student in trouble, it should be a person who evaluates why and help get her/him/they back on track.

Humans have always been far better at inventing new tools than using them wisely. It is only by investing in teachers that technology can liberate them from routine administrative and instructional tasks, and provide them with the opportunity and support to become great coaches, mentors, role models, inspirers and leaders. Education will always work best when humans are kept in the loop, not left to their devices, whether their own or not.

The pandemic has picked up our education systems and hurtled them at light-speed from the 19th to the 21st century. From one-size-fits-all, factory-styled schools to more individualisable, free-range learning. In a way, the pandemic has revealed the enormous potential for innovation that has been dormant in education systems so often dominated by hierarchical structures that reward compliance.

But moving beyond the crisis will require a more level playing field for innovation in schools. Governments can help strengthen the professional autonomy of teachers and school leaders, and a collaborative culture where great ideas are refined and shared. Governments can also help fund incentives that raise the profile of, and demand for, what works. But governments alone can only do so much. Silicon Valley works because governments created the conditions for innovation, not because governments do the innovating. Similarly, governments cannot innovate in the classroom; they can only help by opening up systems so that there is an evidence-based, innovation-friendly climate where transformative ideas can bloom. That means encouraging innovation within the system and opening it up to creative ideas from the outside.

How do we know if education systems have opened up? When they communicate the need for change and build support for it. When they invest in capacity development and change-management skills. When they signal that teachers not passively implement technological and social innovations but get involved in designing them too. When they make it easier for innovators to take risks and bring out new ideas. And when they help innovators find more effective ways of scaling and disseminating their technologies.

We have learned many things during the pandemic. The trick is not to forget them when things return to 'normal'. Artificial intelligence, robots and blockchain are poised to transform how we teach, learn, and run schools. The technology is ready; are we? School closures forced us to dip our toe in the digital waters and for some students and teachers it wasn't so bad. With the quickly evolving smart education tools the report discusses, many of us may be ready to fully take the plunge.

Andreas Schleicher
Director for Education and Skills
Special Advisor on Education Policy
to the Secretary-General

Andreas Schleicher

Acknowledgments

Stéphan Vincent-Lancrin (Senior Analyst and Deputy Head of Division, OECD) edited the book.

Ryan S. Baker (Associate Professor, University of Pennsylvania) was a Scientific Advisor for the book, ensuring it embeds the latest scientific knowledge in the domain.

Some of the authors provided graphics, pictures and video materials to make these technology developments more visible and tangible to the readers.

The book is one of the outputs of the "Smart data and Digital Technology in Education: AI, Learning Analytics and Beyond" project of the Centre for Educational Research and Innovation (CERI) within the OECD Directorate for Education and Skills. Stéphan Vincent-Lancrin conceptualised and leads it.

Within the OECD, Andreas Schleicher (Director for Education and Skills and Advisor to the OECD Secretary-General on Education Policy) provided invaluable feedback, contributions and intellectual input to the report. Dirk Van Damme (Senior Counsellor and Acting Head of CERI, OECD) is also thanked for his support. Carlos González-Sancho (Analyst, OECD) played a key role in setting the foundations of the project before moving to other responsibilities within the OECD. Vivien Liu (Intern), Mathias Bouckaert, Cassie Hague, Koen Van Lieshout, Quentin Vidal, Reyer van der Vlies (Analysts, OECD) identified multiple country examples in the report and drafted some of the boxes in the report. Gwénaël Jacotin (Statistician), Tirso Martin and Yunze Xie (former Interns, OECD) also contributed in different ways, notably through analyses of the education technology industry. Madeleine Gereke provided excellent project assistance during the whole process, with special thanks for the completion efforts and work on the launch of the book. Cassandra Davis, Sophie Limoges, Della Shin, Henri Pearson, Jason Fallow and Alexis Cheney oversaw the publication and/or communication process. Clara Young copy-edited the book (with the help of Jennifer Allain for Chapter 5) – and sometimes more... Other colleagues are also gratefully acknowledged for their comments and suggestions: Glenda Quintini (OECD Directorate for Employment, Labour and Social Affairs) and Marieke Vandeweyer (OECD Skills Centre).

The CERI governing board is acknowledged for very helpful comments on the initial plan for the book and for their feedback on the final manuscript. Special thanks to Osama Rahman (United Kingdom), who made very helpful suggestions to clarify (and simplify) the framing of the analysis, and to Aune Vaulk (Estonia) for thorough comments on the initial project.

Colleagues outside the organisation were also invited to give feedback on the report. At Microsoft, Alexa Joyce and Maria Langworthy provided detailed and very helpful feedback on the manuscript.

Finally, a big thankyou to Gerhard Richter and his atelier, who generously permitted the use of his famous 1971 painting as the cover image of the book. Beyond evoking the binary and pixellated world of digitalisation, it shows how straitjacketed rigour can coexist with colour, fantasy, energy and beauty.

Table of Contents

CHAPTER 5. CLASSROOM ANALYTICS: ZOOMING OUT FROM A PUPIL TO A CLASSROOM
by Pierre Dillenbourg

CHAPTER 6. SERVING STUDENTS WITH SPECIAL NEEDS BETTER: HOW DIGITAL TECHNOLOGY CAN HELP
by Judith Good

CHAPTER 7. SOCIAL ROBOTS AS EDUCATORS
by Tony Belpaeme and Fumihide Tanaka

PART 2

CHAPTER 8. LEARNING ANALYTICS FOR SCHOOL AND SYSTEM MANAGEMENT
by Dirk Ifenthaler

CHAPTER 9. EARLY WARNING SYSTEMS AND INDICATORS OF DROPPING OUT OF UPPER SECONDARY SCHOOL: THE EMERGING ROLE OF DIGITAL TECHNOLOGIES
by Alex J. Bowers

CHAPTER 10. GAME-BASED ASSESSMENT FOR EDUCATION
by Jack Buckley, Laura Colosimo, Rebecca Kantar, Marty McCall, and Erica Snow

CHAPTER 11. BLOCKCHAIN FOR EDUCATION: A NEW CREDENTIALING ECOSYSTEM
by Natalie Smolenski

BOXES

FIGURES

TABLE

Executive Summary

Digitalisation opens up new possibilities for education. While education has always been rich in data such as grades or administrative information on students' absenteeism, the use of data to help students learn better and teachers to teach better, and to inform decision-making in educational administrations is recent. Education stakeholders have had a difficult relationship with technology, alternating between strong enthusiasm and scepticism. Might digital technology, and, notably, smart, technologies based on artificial intelligence (AI), learning analytics, robotics, and others, transform education in the same way they are transforming the rest of society? If so, how might this look? This book explores this question.

After an overview of the opportunities and challenges of digital technology (Chapter 1) and state-of-the-art smart technology solutions, including those not covered in-depth in the book (Chapter 2), the book focuses on how smart technologies can change education in the classroom and support the management of education organisations and systems.

Smart technologies in the classroom

Adaptive learning technology such as intelligent tutoring systems enable the personalisation of students' learning using similar approaches: they detect the knowledge (or knowledge gaps) of students; they diagnose the next appropriate steps for students' learning; they act by providing new exercises, new curriculum units, some form of instruction, or just notifying the teacher. This approach is now being expanded beyond mere knowledge acquisition and factoring in behavioural dimensions such as learning self-regulation or style (Chapter 3).

As keeping students engaged and motivated is key to learning effectiveness, a new domain of technology development focuses on measuring engagement and interventions to keep students engaged, both in digital and physical learning environments. Measuring engagement is difficult but a host of new automated approaches have been developed, from eye trackers to the monitoring and analysis of other facial features. Improving engagement typically takes two routes: proactive approaches try to stimulate engagement with incentives, gamification, etc.; reactive approaches do it in a more sophisticated way by continually monitoring engagement, detecting when engagement is waning, and adapting instruction to address periods of disengagement (Chapter 4).

While smart technologies focusing on personalising learning for individuals are probably the most pervasive, another approach is to consider the classroom or rather what happens in the classroom as the subject of the learning analytics. The objective is to support teachers in orchestrating the learning in their classroom and to propose rich and effective learning scenarios to their students. Some classroom analytics techniques provide teachers with real-time feedback to help manage transitions from one task to the next as their students work individually, in small groups or collectively, for example. They also give feedback to teachers on their classroom behaviour so they can reflect on and learn from their practice (Chapter 5).

Social robots are also being increasingly developed for learning uses. Usually powered by the personalisation systems mentioned above, they support teachers in different ways: as instructors or tutors for individuals or small groups, but also as peer learners allowing students to "teach" them. Telepresence robots also allow teachers or students to teach or study remotely and offer new possibilities for students who are ill and cannot physically attend class. They can also mobilise a remotely located teaching workforce, for example teachers from another country to teach foreign languages (Chapter 7).

Technology also enables students with special needs to participate in education and to make inclusive education a reality. With well-known applications such as speech-to-text, text-to-speech, and auto-captioning, etc., AI allows blind, visually impaired, deaf and hard-of-hearing students to participate in traditional educational settings and practices. Some smart technologies facilitate the diagnosis and remediation of some special needs (e.g. dysgraphia) and support the socio-emotional learning of students with autism so they can more easily participate in mainstream education (Chapter 6).

Those smart technologies usually assume and require a human-in-the-loop: a teacher. The level of automation of actions and decisions should be conceived of as a continuum between actions that are fully automated at one end and, at the other end, actions over which humans have full control. As of today, AI systems remain hybrid and request human intervention at a certain point in the process.

Smart technologies at the organisation and system levels

Smart technologies powered by AI and learning analytics also allow for the management of education organisations. They can be used for a variety of purposes; for example, to enhance an institution's curriculum based on an analysis of students' learning and study paths. While this is still a nascent trend, a whole-of-organisation adoption of learning analytics can transform educational institutions' culture (Chapter 8).

Early warning systems that identify students at risk of dropping out from high school are a good use of the administrative micro-data that are increasingly being collected by education systems and organisations. While identifying a good set of early warning indicators remains difficult, a few systems have shown a high level of accuracy and enriched thinking about the reasons students drop out. In order to avoid the risks of student profiling, open and transparent algorithms are important (Chapter 9).

Game-based standardised assessments also build on smart technologies and smart data analysis techniques to expand assessment to skills that cannot be easily measured by traditional (paper-and-pencil or computer-based) tests. These include higher-order skills (e.g. creativity) or emotional and behavioural skills (e.g. collaboration, behavioural strategy). Game-based tests may analyse eye-tracking data and audio recording, and process natural language and information such as time-on-task or use simulations (Chapter 10).

Finally, as a "verification infrastructure", blockchain technology opens new avenues for credentialing in education and training. Blockchain technology enables the validation of claims about an individual or institution, including their characteristics and qualifications, and to do this instantly and with a very high level of certainty. This helps eliminate diploma (and other records) fraud, facilitates the movement of learners and workers between institutions and geographies, and empowers individuals by giving them increased control over their own data. Many blockchain initiatives are underway across the world, which may transform how education and lifelong learning systems manage degrees and qualifications (Chapter 11).

Policy pointers

There are good reasons to believe that smart technologies can contribute to the effectiveness, equity and cost-efficiency of education systems. At the same times, there are a few important aspects of smart technologies to keep in mind to reap those benefits:

- Smart technologies are human-AI hybrid systems. Involving end users in their design, giving control to humans for important decisions, and negotiating their usage with society in a transparent way is key to making them both useful and socially acceptable.
- Smart technologies support humans in many different ways without being perfect. Transparency about how accurate they are at measuring, diagnosing or acting is an important requirement. However, their limits should be compared to the limits of human beings performing similar tasks.
- More evidence about effective pedagogical uses of smart technologies in and outside of the classroom as well as their uses for system management purposes should be funded without focusing on the technology exclusively. Criteria for this evidence to be produced quickly could also be developed.

- The adoption of smart technologies relies on robust data protection and privacy regulation based on risk assessment but also ethical considerations where regulation does not exist. For example, there is mounting concern about the fairness of algorithms, which could be verified through "open algorithms" verified by third parties.
- Smart technologies have a cost, and cost-benefit analysis should guide their adoption, acknowledging that their benefits go beyond pecuniary ones. In many cases, the identification of data patterns allows for better policy design and interventions that are more likely to improve equity or effectiveness. Policy makers should also encourage the development of technologies that are affordable and sustainable thanks to open standards and interoperability.

1 Frontiers of smart education technology: Opportunities and challenges

Stéphan Vincent-Lancrin
OECD

This chapter serves as an introduction to the book and presents some of its findings and policy implications. After highlighting the importance of digitalisation as a societal trend for education, it introduces the main focus of the book: exploring the frontiers of education technology. Artificial intelligence and learning analytics are transforming (or have the potential to transform) educational practices, and so have other smart or advanced technologies such as robotics and blockchain. How can they improve classroom instruction and the management of educational establishments and systems? After presenting the objectives and chapters of the book, the chapter highlights the opportunities of smart technologies for education systems and points to some emerging policy issues and dimensions to consider before making some forward-looking concluding remarks.

Smart data and digital technology in education

Digitalisation opens new possibilities for education. While education has always been a sector rich in data such as grades or administrative information, the use of this data to help students learn better, teachers to teach better, and inform decision-making in educational administrations is recent. That said, education stakeholders have always paid attention to new technologies and their potential to revolutionise education. This was true with the invention of radio, television, and more recently computers and the Internet. However, most uses of innovative technology have been to conserve existing educational practice and sometime enrich it, but rarely transform it. Might digital technology, and, notably, smart technologies based on artificial intelligence, learning analytics, robotics, and others, transform education in the same ways they are transforming the rest of society (OECD, 2019[1]; OECD, 2019[2])? If so, how might this look like?

There are two important aspects to the "digitalisation" discussion in education.

The first aspect relates to the changes that technology could induce in the delivery of education, from early childhood to adult learning. This is what this book explores. In this book we ask: how does and how could digitalisation transform education as a sector in the short, medium and long term? How does or may the rapid advances in artificial intelligence, learning analytics, robotics, etc., change how teachers and students teach and learn? What tasks do teachers perform that computers or robots might take over? Those technological advances can also translate into new work and management processes at the establishment or sector levels, sometimes in quest of cost-efficiency and productivity enhancement, sometimes to improve the effectiveness of the sector in reaching its traditional objectives (learning outcomes, equity, completion, etc.). Is digitalisation going to change schooling, higher education or lifelong learning? Will the educational processes be different, with more automated or computer-based tasks throughout the learning process? Will the digital infrastructure available to students, teachers, administrators and policy makers be different? Will an increased use of computers, data, smart devices, robots (and the technology that powers them) translate into better learning outcomes, more equity, more efficiency and productivity in education? What are the new possibilities, the opportunities and the challenges to be expected? Those questions have become more strategic for education policy makers in the past years. Between 2015 and 2019, 17 OECD countries published a digital strategy for education (and 16 others included an education chapter in their new national digital strategy) (van der Vlies, 2020[3]).

The second important question about digitalisation relates to how adequately education is responding to emerging societal and labour-market needs. This points to the 21st-century skills discussion in education circles, and the increasing importance of skills that are more difficult to automate and which foster innovation, such as creativity, critical thinking, communication and collaboration (Vincent-Lancrin et al., 2019[4]). The digitalisation of society and future shifts in labour market demand make the question of the content and nature of education more important: what are the knowledge, skills, attitudes and values people need in a highly digitalised, AI-inflected world? While this is not the primary focus of this book, some of the analysis will show how smart technology can also support the acquisition and assessment of those skills, for example through gamification or new forms of assessments.

After presenting the objectives and chapters of the book, the chapter highlights the opportunities of smart technologies for education systems and points to some emerging policy issues and dimensions to consider before making some forward-looking concluding remarks.

Current frontiers of digitalisation in education

As the objective is to get close to the "technology frontiers" of education technology and take stock of what technology in education can already do, this book limits itself to technology that is demonstrably possible and currently used in some jurisdictions, establishments or laboratories. Whenever possible, evidence about their effectiveness is provided.

The book is organised by educational objective or issue, rather than technology, thereby acknowledging that several technologies can be used to address similar issues (as alternatives or supplements). Roughly speaking, it covers three main types of technology fields (or families): artificial intelligence (in the all-encompassing meaning it currently has) and learning analytics; robotics (which adds a physical embodiment to artificial intelligence); and blockchain. Box 1.1 provides some initial definitions for those technologies. The book focuses on two areas where technology has (and will have) a transformative effect: teaching and learning in the classroom, and managing educational establishments and systems.

The chapters present how smart technologies are addressing (or could address) a number of educational issues, how they work, what they do well, what their shortcomings currently are, and what role they may play in the future in countries' education systems. The selection of applications was made in areas where technology is either sufficiently mature and its benefits appear as low-hanging fruits, or where recent breakthroughs may be less well-known by policy makers and a wider audience. The analyses focus on formal education, from primary education to higher education, and leave out all applications for informal and non-formal education, applications focusing on the teaching and learning of specific subjects (e.g. foreign language, mathematics, reading, etc.) as well as the teaching and learning of technology itself (coding, etc.).Chapter 2 by Ryan Baker (University of Pennsylvania, United States) provides a general overview of artificial intelligence in education. After clarifying different terms and definitions to help readers understand the relations and sometimes overlaps between different technological techniques and terms, Baker provides an overview of the technologies currently being used in education, their core applications and their potential to bring education forward. The overview introduces some of the core applications that are explored in more depth in the report, and which could transform teaching and learning (e.g. personalisation of learning). It also highlights the potential of smart technologies in other areas such as formative assessment, digital games and simulations or just the provision of data to inform pedagogy. Likewise, beyond highlighting some of the technology applications used in managing education establishments and systems (e.g. early warning systems), it also points to a vast array of other possible applications such as real-time reporting to parents, admission systems or proctoring systems.

Box 1.1 Description of digital technologies

Artificial intelligence (AI): An AI system is a machine-based system that can, for a given set of human-defined objectives, make predictions, recommendations, or decisions influencing real or virtual environments. AI systems are designed to operate with varying levels of autonomy. AI system lifecycle phases involve: i) 'design, data and models'; which is a context-dependent sequence encompassing planning and design, data collection and processing, as well as model building; ii) 'verification and validation'; iii) 'deployment'; and iv) 'operation and monitoring'. These phases often take place in an iterative manner and are not necessarily sequential. The decision to retire an AI system from operation may occur at any point during the operation and monitoring phase (OECD, 2019[1]).

Learning analytics: Learning analytics is one of the new young disciplines in data science. It studies how to employ data mining, machine learning, natural language processing, visualisation, and human-computer interaction approaches among others to provide educators and learners with insights that might improve learning processes and teaching practice.

Internet of Things (IoT) / Smart devices: The Internet of Things includes all devices and objects whose state can be altered via the Internet, with or without the active involvement of individuals. "Smart" devices, equipment, machines and infrastructure are creating opportunities for automation and for interaction in real time. Applications and services built for the Internet of Things, with insights provided by data analytics, are expected to become pervasive and educational establishments and classrooms may become "connected".

Robots: A robot is a physical machine with sensing, computing and actuating capabilities, able to carry actions automatically. Often robots can make autonomous decisions and can adapt these decisions based on prior knowledge and sensor input. In education, most robots used are "social robots" that interact with learners.

Blockchain: Fundamentally, blockchain is a combination of already existing technologies that together can create networks that secure trust between people or parties who otherwise have no reason to trust one another. The marriage of these technologies gives blockchain networks key characteristics that can remove the need for trust, and therefore enable a secure transfer of value and data directly between parties. Specifically, it utilises distributed ledger technology (DLT) to store information verified by cryptography among a group of users, which is agreed through a pre-defined network protocol, often without the control of a central authority. You can think of a ledger as a record book: it records and stores all transactions between users in chronological order. Instead of one authority controlling this ledger (like a bank), an identical copy of the ledger is held by all users on the network, called nodes. Along with its own hash, each block stores the hash of the block before it. A hash is a unique string of letters and numbers created from text using a mathematical formula. Blocks are therefore "chained" together, making the ledger (almost) immutable or unable to be changed (OECD, 2017[5]).

Source: ECHOES picutre and video database (reproduced with permission).

Part I of the book mainly focuses on the use of smart technologies *in the classroom.* It covers a variety of applications of smart technologies, from the most common ones (intelligent tutoring systems) to new developments (class orchestration, social robots, learning engagement).

Chapter 3 by Inge Molenaar (Radboud University, the Netherlands) discusses state-of-the-art personalisation of learning. Adaptive learning technology has arguably been one of the oldest lines of work for technology in education and is a mature field in the area. Existing learning technologies have a strong focus on diagnosing students' knowledge and adjusting feedback or problems at the task level (which task next?); at the step level (what is the next step in a given task?); or at the curriculum level (what topic or curriculum unit next?). The frontiers in this area constitute computers taking into account a broader range of learner characteristics such as self-regulation, motivation and emotion. The state of the art is described using a "6 levels of automation of personalised learning" model that articulates the different roles of AI, teachers and learners, and shows how hybrid human-AI solutions combine the strengths of human and artificial intelligence to implement personalised learning.

Chapter 4 by Sidney D'Mello (University of Colorado Boulder, United States) zeroes in on a frontier area of personalisation, which is learning engagement. It gives a broad overview of some promising avenues for measuring students' level of engagement during learning in an automated way with digital technologies. It also discusses how these technologies can be designed to improve engagement on the onset of learning or when learners' disengagement sets in. After discussing why engagement matters for learning, it presents different types of approaches to improve students' engagement and learning by using data and technology, for example, facial feature analysis, gaze analysis or eye tracking, sometimes in online but often in in-presence environments.

Chapter 5 by Pierre Dillenbourg (Ecole Polytechnique Fédérale de Lausanne, Switzerland) shifts the focus to the use of learning analytics and artificial intelligence to support teachers in orchestrating the teaching and learning of students in their classrooms. Here, the classroom as a whole and what is happening within it is the unit of analysis. Equipped with sensors, cameras or connected devices, classrooms become a hybrid physical-digital space in which computers analyse the behaviours of both students and teachers, and give teachers feedback on different parameters. Through different types of dashboards and displays, teachers get real-time information, for example, about when to move on to the next sequence of the lesson, or receive feedback after the class for their professional development or the planning of their next lessons.

Chapter 6 by Judith Good (University of Sussex, United Kingdom, and University of Amsterdam, the Netherlands) shows how technology provides specific services to the learning and engagement of students with physical impairments and mental health issues in education. Some of these technologies help bypass some of the obstacles to learning (e.g. text-to-speech or speech-to-text for blind and visually impaired students). Some simple applications assist adults in providing a first diagnosis of special needs, as is the case with dysgraphia. The chapter emphasises the importance of human-AI systems, both in the diagnosis and learning process. Because smart technologies are still limited compared to human beings, they sometimes help connect students with special needs with teachers or other people (e.g. the operator of a digital learning environment helping students with autism spectrum disorder). It also shows the value of engaging students with special needs in the development of the technologies designed to support them.

Chapter 7 by Tony Belpaeme (Ghent University, Belgium) and Fumihide Tanaka (Tsukuba University, Japan) presents possible roles for robots as educators. While it is based on a different type of technology (robotics), some of the smart technologies presented in the previous chapters can be embedded in the robots. Two main roles for robots are presented. Robots can be educators and tutors (typically one-on-one) or peer learners (with students teaching robots about something they are learning). Typically, social robots are designed for a learning environment supervised by teachers. Robots can also be telepresence devices, allowing teachers to teach a class (or students to attend class) remotely, offering more opportunities than videoconference systems. While research shows that robots are rather effective in the narrow tasks they perform, it seems unlikely that they will have the capability to replace teachers in the foreseeable future. Their cost is also a limit to their mainstreaming in education.

Part II of the book mainly focuses on the use of smart technologies for the *management of educational institutions and systems.*

Chapter 8 by Dirk Ifenthaler (University of Mannheim, Germany and Curtin University, Australia) starts with an overview of the different possible uses of learning analytics to manage higher education institutions and provide information to decision-makers on a variety of governance and organisational processes from forecasting future education to productivity enhancements. In spite of interesting cases of change management through learning analytics at the organisational (or faculty) level, few cases of holistic approach to learning analytics are documented. Some of the challenges at the organisational level apply to the system level as well and several general guidelines for a stronger adoption of learning analytics within organisations and system-wide are presented. Most of the research on learning analytics for the management of organisations concerns higher education institutions so that specific challenges for schools remain to be explored.

Chapter 9 by Alex Bowers (Columbia University, United States) shows how smart data and technology helps understand and tackle an important problem in most OECD countries: high-school dropouts. This is one of the most immediate and increasing uses of administrative data and the application of smart technologies here is relatively common, at least in some countries. A first important step is to have good predictors, and while the chapter notes how a variety of early warning indicators (and systems using them) fall short of their claimed objective, it also presents advanced data techniques that allow for a much more accurate identification of "at risk" students. Open data and open algorithms should allow third parties to verify the quality and fairness of algorithms. Recent analysis shows that students who drop out have different profiles, which require a greater variety of policy interventions than those currently provided to students considered "at risk" of dropping out.

Chapter 10 by Jack Buckley, Laura Colosimo, Rebecca Kantar, Marty McCall and Erica Snow (Imbellus Inc., United States) discusses how recent advancements in digital technology could lead to a new generation of game-based standardised assessments and provide education systems with feedback on students' higher-order or socio-emotional skills that are difficult to assess via traditional standardised tests, be they computerised or not. Game-based tests may analyse eye-tracking data and audio recording, and process natural language in addition to analysing information such as time on task or use simulations. Given their cost, the complexity of their development and also some of their intrinsic limitations, they will supplement rather than replace traditional standardised tests, which have their advantages for assessing certain knowledge and skills.

Chapter 11 by Natalie Smolenski (Hyland, United States) focuses on the use of blockchain technology to make the credentialing process more efficient – and also possibly some other educational administrative processes requiring verification. The chapter starts with a history of blockchain in cryptocurrencies to make the functionalities of the technology more easily understandable. Credentials are a good use case for blockchain technology and many blockchain initiatives around credentialing are underway worldwide. Blockchain allows for secure and transparent sharing of qualifications, credits and badges. At the national (but even more so at the international level), it could help stop fake degrees and certifications, facilitate the transfer of educational records as well as the credentialing of small units of learning such as MOOCs or corporate professional development provided by companies. The human and legal infrastructure for their full mainstreaming remains to be developed, including open standards and interoperability. Compared to the current processes in place, the chapter argues it would be a cost-efficient solution.

Key opportunities

Smart technologies can improve education systems and education delivery in different ways. They can enhance access to education, improve its quality for learners, and enhance its cost-efficiency for societies. This section highlights how smart technology contributes (or could contribute) to the achievement of these goals.

Effectiveness

Attending school or university does not always translate into as much academic learning as one would hope for. The OECD Programme for International Student Assessment (PISA) has shown that attending school may actually lead to very different levels of learning outcomes across countries. While there is no similar evidence at the higher education level as yet, this is arguably the same at that level too. One of the key promises of smart technologies is to enhance the effectiveness of teaching and learning for better student learning.

In the classroom, applications that directly support student learning show early promise. Personalised learning aims to provide all students with the appropriate curriculum or task, and scaffold them within a task, based on a diagnosis of their knowledge and knowledge gaps. This is not only done at the academic level, focusing on the "what", but increasingly takes into account how students learn and factors such as self-regulation, motivation or effort (Molenaar, 2021[8]). Engagement is key for learning, and solutions to keep students engaged within digital or physical learning environments are being developed to identify their affective states during learning and nudge them towards re-engagement when they seem to disengage (D'Mello, 2021[9]). Social robots perform similar tasks in different ways: they can use adaptive learning to tutor students with natural language, but they can also teach, or motivate them to learn by playing the role of a peer student. They support teachers by enabling the implementation of different types of teaching and learning strategies (Belpaeme and Fumihide, 2021[10]). Finally, smart technologies give students with impairments and special needs access to curriculum materials and allow those students to participate in learning activities to an extent that was not possible before, here again increasing the effectiveness of education (Good, 2021[11]).

Those solutions can be used and remain helpful outside of the classroom too, either for homework, as automated private tutoring or practice solutions, and for lifelong learning. In fact, the largest market for educational technology companies is the consumer market targeting students and parents directly, either for recreational learning activities or for tutoring or test preparation.

A second promise of learning effectiveness comes from classroom analytics that support teachers in providing more effective teaching. This is still a work in progress but many applications already show how a variety of solutions could support teachers in better using their time in class, for example, by suggesting when it is a good time to shift to the next teaching or learning activity, who would require their attention the most, how they could engage the whole class in collaborative learning activities. While classroom orchestration solutions can help teachers in real time, they also provide feedback on their own practice, for example, how much they talk, to whom, or how they divide their time between different types of activities (Dillenbourg, 2021[12]). Both real-time and post-hoc feedback are akin to personal professional learning opportunities for the teacher in question, and have the significant advantage of being about the specific teacher who was (digitally) observed rather than about theoretical or general teaching practice. In that sense, smart technology has real potential to improve the teaching practice of all individual teachers, and subsequently the learning outcomes of their students.

Box 1.2 Integrating AI and learning analytics in school: examples from China

Increasingly, school buildings will be equipped with sensors, cameras, and computers to fulfil certain administrative as well as teaching and learning functions. Some schools are already experimenting and developing innovative ways to integrate smart technologies in their operations. Here are a few examples from Shanghai (China).

The Luwan No 1 Central Primary School (Huangpu District, Shanghai) is a public school integrating AI in its school resource management as well as its teaching and learning – a digital model that may then be extended to other schools. The management of the campus, and the teaching and learning all rely on smart technologies. Using IoT sensing technology, the "digital campus" consists of collecting and analysing campus data to automatically control and manage environmental factors such as security, lighting, water quality and air quality, but also to collect campus activity data; for example, people density in corridors etc. Combined with wearable devices, the school also collects physiological data such as students' body temperature and heart rate as well as academic data and learning process data in order to support teachers and learners. The "digital students" application analyses student data to create a detailed, holistic portrait of students. The collection of data increases the understanding of students' learning status and growth, and provides teachers with data to tailor their teaching to their needs. The data cover discipline, academic level, physical and mental health, aesthetic taste and social practice. Socio-emotional aspects such as learning engagement and affective states are measured by voice and face-recognition technology. Finally, a "digital teaching" system «provides teachers with support on five aspects of teaching: lesson preparation, classroom orchestration, homework, tutoring and evaluation – with functionalities such as "classroom orchestration", "intelligent assessment" and "intelligent homework review". The intelligent tutoring system supports students directly in accessing resources, tools, pathways and personalised guidance. As of June 2021, this model has been studied and adopted by more than 250 schools in Shanghai, Qinghai, Shaanxi, Guizhou, etc.

Figure 1.1 Digitalisation at the Luwan No 1 Central Primary School in Shanghai

Note: The left panel shows a screenshot of the campus management system. The right panel shows a student using the classroom handwriting board collecting data about this progress.
Source: Courtesy of the municipal government of Shanghai

The demonstration high school affiliated to Tongji University (China) is also implementing a new "digital classroom" system in English, geography and biology. Students' learning data collected in the system is the basis for teaching and further learning. Before the lessons, teachers use digital interactive "practice" tools to assess students' learning; they also conduct short and concise in-class tests from time to time to obtain real-time student learning data. This allows them to change their teaching strategies during class, and to develop individualised after-class strategies. Based on this information, teachers will set online assignments, which are automatically marked by the system and provide the basis to the generation of personalised "knowledge analysis" reports (Figure 1.2). Based on these individual cards, the system proposes micro-tutoring video resources and exercises to meet individual learning needs, and teachers receive guidance to propose targeted after-class assignments and counselling and to customise their teaching to students' needs. The system also allows for collaborative learning with students and teachers being able to see and comment on students' work.

Figure 1.2 The "digital classroom" system at Tongji University's first demonstration high school

Note: The left picture shows the visualisation of a student's acquired knowledge in a chemistry curricula unit. The right picture shows how the system can be used for cooperative learning, with both students and teacher being able to view and assess how all students responded to a given assignment.
Source: Courtesy of the municipal government of Shanghai

Other demonstration schools focusing on digital technology explore other possible aspects of technology.

- Shanghai Xuhui Middle School has a traditional emphasis on science education and 22 engineering science and innovation labs (18 kinds). After developing 5G in the school campus, it developed a "holographic" science education model based on Mixed Reality in order to make difficult knowledge more directly understandable and intuitive, and to enhance students' attention and enthusiasm for the subjects. As of June 2021, two lessons ("Exploring the Mysteries of the Solar System" and "Understanding the Bones of the Human Body") were developed and delivered with real-time interaction with Yuanyang No. 1 Middle School (Honghe Prefecture, Yunnan Province).
- The Shanghai Industrial Technology School provides its students with advanced mixed reality and simulation technology to learn manufacturing. Simulation-based training projects are carried out in a 3D virtual environment, containing a series of workstations such as graphic drawing, workpiece handling and work units such as loading and unloading of computer numerical controlled (CNC) machine tools.

At the organisational and system levels, smart technologies also hold promise in making education more effective. While this remains relatively rare (Ifenthaler, 2021[13]), smart technologies can be integrated in most dimensions of school activities, providing administrators, teachers and learners with feedback to manage school resources as well as improve the effectiveness of teaching and learning (see Box 1.1). The rise of a new generation of assessments powered by AI also opens new avenues for recognising and evaluating competences that were hard to assess through paper and pencil tests. This could accompany most education systems in their shift towards emphasising skills (in addition to the traditional emphasis on knowledge). Game-based assessments and simulations allow assessments to be designed to be more realistic but also to assess skills such as complex problem solving, creativity or collaboration in new ways (Buckley et al., 2021[14]).

Finally, the emergence of longitudinal education data systems that follow students through the course of their studies also allows for more effective policy and organisational interventions and a better design of educational offerings. For example, in the United States, the analysis of community college graduation rates, the success of their student placement strategies in "remedial courses" and students' study patterns as part-time or full-time enrolment led to revisiting what the educational experience of community college students actually is and "redesigning" community colleges (Bailey, Smith-Jaggars and Jenkins, 2015[15]). As in other sectors (OECD, 2019[16]), the use of data supports policy design and interventions.

Equity

Smart technologies can help education systems provide more equitable learning opportunities. In this respect, smart technologies are more ambivalent. One the one hand, they clearly do or could help reduce inequity both by increasing access to learning opportunities for all and improving learning effectiveness for those who need it the most. On the other hand, without a widespread and equitable availability of smart technologies, inequity could also rise. They may also leave achievement gaps unchanged or even widened, depending on their differential impact on different learners.

Let us start with the difficulties. There are at least two reasons why technology may have a negative effect on equity. The first, obvious reason lies in the difference in access to devices and connectivity by students from different groups, notably students from lower socio-economic backgrounds. These students may not have the devices, the connectivity or the resources that allow accessing and using smart technologies either at the school they attend or at home. A second reason is that, if technology (e.g. personalised learning) works the same for everyone, those who start with stronger prior knowledge can maintain their advantage or even make faster progress than those with less prior knowledge. In spite of supporting students with less prior knowledge, it is thus possible that technology could help more advanced students more. This would widen rather than reduce the achievement gap.

There are also many reasons to believe that smart technologies can advance the equity agenda.

First, learning technology can expand access to learning opportunities. Educational platforms proposing open educational resources (Orr, Rimini and van Damme, 2015[17]) or massive open online course (MOOC) platforms are good examples. They allow learners to access learning materials with a quality that may be superior to what they can access locally. While many studies have shown that this increased access has not decreased inequity at scale given the low take-up and the fact that most users are already well educated, a recent systematic review of their effect on equity provides a more optimistic perspective, notably for non-English MOOCs or open educational resources (Lambert, 2020[18]).

As importantly, smart technologies can reduce inequity by facilitating the inclusion of students with special needs and by adapting learning to different learning styles. Technology has, for example, made it much easier to support the diagnosis of learning difficulties such as dysgraphia, and remedial digital responses have also been developed. A variety of smart technologies applied to learning solutions also make it easier for blind or visually impaired students as well as deaf or hard-of-hearing students to access learning materials and easily perform the educational tasks required from other students. Artificial intelligence enabling speech to text (and vice versa) or automatic subtitles are the most obvious examples. Learning technologies also tackle more difficult issues and support the socio-emotional (and thus the subsequent academic) learning of autistic children. They increasingly propose ways to help children with attention deficit hyperactivity disorder (ADHD) to self-regulate and better benefit from their schooling. One caveat here is that inclusion is not just about the individual "fitting in" but also for society to be more inclusive and open to differences (Good, 2021[11]). Technology encourages that by enabling students with special needs to study in a traditional (and inclusive) learning environment, which also changes peoples' view on disability and special needs.

Second, solutions such as early warning systems are entirely focused on reducing inequity by helping students at risk of dropping out from high school (or university) to graduate – students who drop out typically come from disadvantaged and minority backgrounds. Early warning systems also allow designing appropriate interventions by identifying the factors or indicators most likely to predict dropout (Bowers, 2021[19]). Some use of learning analytics within institutions, for example, to monitor student engagement or redesign study programmes, could also have the same effects, should the educational institution pay particular attention to inequity (Ifenthaler, 2021[13]).

Third, the use of learning analytics as exemplified by personalisation at the individual level, be it using intelligent tutoring systems or learning analytics to keep students engaged in learning, all hold promise in reducing inequity, notably by supporting students with less prior knowledge to learn at the right pace. Box 1.3 gives an example of an online solution that reduced the learning gap between the strongest and weakest students in mathematics at the beginning of the intervention. There is, however, little evidence that adaptive learning generally reduces achievement gaps between students. Classroom analytics can also give feedback to teachers on how they could improve their teaching; specifically how and when to pay more attention to different group of students in their class, based on their academic level, gender, ethnicity, etc. Adaptive learning technology can also help students practice and make progress at home, outside of the classroom, supported by intelligent tutoring systems. This may be particularly important for students coming from households where parents can support their students less effectively with their school work, be it directly or indirectly.

Box 1.3 Personalisation in maths homework can help reduce the achievement gap: a U.S. study

Few studies show that adaptive technology (or personalised learning) reduces the achievement gap between students with more and less prior academic knowledge. And yet, in order for intelligent tutoring systems to reduce achievement gaps, this would indeed be the objective. Evaluated through a randomised control trial, an intervention in the state of Maine (United States) showed that this may become the case (Murphy et al., 2020[20]). Teachers in the intervention were asked to use ASSISTments software to provide students with mathematics homework. The system provides feedback to students as they solve mathematics homework problems and automatically prepares reports for teachers about student performance on daily assignments. Teachers received training and coaching on formative assessment. The study found that students in the schools that were assigned to ASSISTments learned more compared to their peers in the control schools, with large effect sizes, and that the impact was greater for students with lower prior mathematics achievement. The evaluation confirms initial results by Roschelle et al. (2016[21]), which found both evidence of strong maths learning outcomes when using the platform and also a reduction of the achievement gap.

Efficiency

In most other sectors than education, smart technologies are used as a tool to enhance the cost-efficiency of operations, notably by automating a number of tasks and processes, making services faster and often cheaper (OECD, 2019[16]). While education might be behind most other sectors in this respect, digitalisation is also making many educational processes more efficient as interactions between stakeholders and educational institutions become increasingly automated. As noted above, also in teaching and learning, some degree of automation is gaining ground. To what extent will digitalisation allow for enhanced cost-efficiency and productivity in education?

Any discussion of cost-efficiency should keep in mind that technology incurs investment and maintenance costs, which have to be compared with the costs of current arrangements. Digital technology has not always delivered on its cost-efficiency promises in the past because one has sometimes forgotten that, beyond the initial investment, it needs to be continuously upgraded, maintained, etc.

Nevertheless, as in other sectors, there are good reasons to believe that smart technologies could increase cost-efficiency in education.

One example lies in the application (and admission) process to educational institutions. They are increasingly undertaken through digital platforms, especially in higher education, where a "matching" (or selection) process is often necessary. In cases of open admission institutions, when no selection is required, implementing seamless automated processes is even easier. The implementation of the National Education Information System (NEIS) in

Korea, an e-government system allowing, inter alie, for the digital transfer of students' academic records from one school to the other (as well as from school to university) was estimated to save USD 237 million a year in 2010 (KERIS, 2010[22]).

A second area where digitalisation could possibly lead to cost-efficiency is the provision of verifiable degrees and other credentials on blockchain. The gradual development of an infrastructure for digital credentials and the adoption of open standards will gradually lead to a different way of certifying and holding degrees, with individuals being able to manage their qualifications themselves. This is one of the strongest and most immediate arguments of cost-efficiency in the different digital solutions examined in this book.

A third area where cost-efficiency is underway is the collection of system-level statistical information. While statistical information often relied on the establishment of statistical panels (of representative samples of individuals or institutions) and often involved multiple handlings of the same data, the use of administrative data combined with the interoperability of diverse systems has made it much easier to get statistical information from operational services in almost real time (González-Sancho and Vincent-Lancrin, 2016[23]; n.d.[24]).

At the end of the day, a cost-benefit analysis comparing the benefits of smart technology, including the non-financial ones mentioned above, to that of an existing solution, will determine how cost-efficient it is for a given service (or educational goal).

Policy pointers

The emergence of smart education technologies or solutions powered by artificial intelligence, learning analytics, algorithms and other technologies presents many opportunities. At the same time, they raise a number of policy questions. How can governments best harness the benefits of technology in education while limiting its possible risks? This involves a good understanding of the opportunities and risks, both from a technical and political dimension. One success factor lies in the social willingness to adopt those technologies. This section will draw some key lessons from the book from a policy perspective and highlight some key features of smart technologies that matter for policy making and for the effective deployment of smart technologies in education.

Smart technologies as socio-technical systems

Whether they are already available or under development, most of the smart education technologies for teaching and learning covered in this book do not aim to replace teachers or human beings. They were in fact developed with the current education model in mind. One common thread of all analyses in the book is that most teaching and learning solutions are designed as hybrid human-AI systems and require teacher-student interactions and human oversight of the machine at different points. Molenaar (2021[8]) offers a model to better understand the continuum between fully automated and teacher-only education. Most advanced personalisation solutions require teacher intervention or alert teachers when they should intervene, for example, because students are still struggling or they need to move on to another step of their learning process. Most solutions to support classroom orchestration are also hybrid solutions that merely scaffold teachers in implementing rich learning scenarios for their students. As Dillenbourg (2021[12]) puts it, "there is a teacher in the loop" and classroom analytics are designed to support teachers in orchestrating the teaching and learning of students in the class and in providing them effectively with rich learning scenarios – not to replace them.

Contrary to how robots are often presented in other sectors, the social robots presented by Belpaeme and Tanaka (2021[10]) are not meant to replace teachers either, but to support their students for specific learning tasks, in the same spirit as personalisation tools. Not that it could not be possible one day, in a distant future. As of 2021, social robots are mainly effective in accomplishing narrowly defined tasks. They play the role of a teacher assistant, as computers do in their different way. As for telepresence robots, they enable human teachers to be present from a distance. Good (2021[11]) provides an excellent case showing how smart technologies for students with special needs may actually create new social relationships between learners and the humans in charge of providing them with appropriate learning tasks – rather than suppress them.

At the system and organisational level, the use of smart technologies follows the same pattern. Early warning systems help predict dropout, but they require a human intervention for "at risk" students not to drop out (Bowers, 2021[19]). Other types of learning analytics used within educational institutions to support decision-making also provide information that needs to be acted upon; they do not make final decisions in the place of administrators and teachers (Ifenthaler, 2021[13]).

This is not to say that smart technologies never make decisions or are not designed for full automation. Personalisation systems, classroom analytics, and early warning systems all make some decisions to enact their next step or recommend one to human beings. But they typically only provide input to decisions. Game-based standardised assessments do more than provide a suggestion: they automatically score the test-takers and assess their skills – as is already the case with traditional computer-based standardised assessments. Blockchain technology does not make decisions, it just records trustfully what a variety of (usually) human actors have done, building on different social processes: accrediting institutions, awarding a qualification or credential, storing the credential on some blockchain, sharing the credential with other parties, verifying the authenticity of the credential, etc. (Smolenski, 2021[25]). But both cases also highlight the relevance of thinking of smart technologies as socio-technical systems, that is, systems in which social and technical features interact and are shaped together.

One of the challenges of game-based standardised assessments will be to create social acceptance, if not full trust, as was the case for traditional standardised assessments. Buckley et al. (2021[14]) note that building valid, reliable, and fair game-based assessments is considerably more complex and challenging than designing traditional standardised tests. While it solves a problem with a clear level of efficacy, one challenge to the widespread use of blockchain for credentialing also relates to social change and legal adaptations: social processes to certify credentials already exist and not everyone may be willing to change those social habits – or cope with the uncertainties attached to any new solution.

There are different ways to acknowledge the fact that most smart technologies are hybrid AI-human systems or that, more generally speaking, smart technologies are better designed and understood as socio-technical systems. One is to clearly communicate that while technology could play a bigger role in the future, it currently needs to be supplemented and controlled by human actions in most cases. The ways to recognise those current realities would be as follows:

- Involving teachers, students and other end users as co-designers in the research and development process would ensure the usefulness and use of smart digital solutions. It also help the people involved in understanding and shaping the social context in which smart education technologies would best be used (the classroom, home, etc.). This should be an aim even when it is challenging to involve end users, for example, students with special needs.
- Public-private partnerships between government, technology researchers within universities and companies, and the education technology industry should be a key characteristic of most research and development projects in this area. This research should go beyond the technology functionalities to analyse how it is used in context and also work on the social and legal adjustments that would be required for their widespread adoption.

Algorithm accuracy

Smart technologies have made and are making very rapid progress, and this book illustrates their potential benefits in a wide range of educational areas both in terms of teaching and learning, and administrative efficacy. Smart technologies often outperform traditional data analysis and technology thanks to more powerful algorithms. Game-based assessments allow skills to be assessed that are difficult to assess through traditional computer-based or paper-and-pencil tests. Personalisation adapts to learner characteristics in ways that pre-existing personalisation methods could not permit – and they possibly do it as well as human teachers. The new algorithms in early warning systems outperform traditional regressions in terms of predictive power and bring visibility to dropout patterns that were not traditionally acknowledged by human administrators or teachers.

Nevertheless, many of the smart technologies presented in the book are not fully mature yet. For example, while some early warning systems now approach good predictive power, Bowers (2021[19]) shows that most early warning systems rely on predictors that are no better than a random guess. In the areas of student engagement, D'Mello (2021[9]) points to new approaches that are developed to better measure students' engagement in learning using facial image analysis and other ways but also notes the inaccuracy of many of the measures used in the field of learning engagement. In the area of classroom analytics, some solutions manage to identify whether learners are working individually or in groups with a very high level of accuracy (90%) but identifying the type of teaching and learning activity remains more challenging (67% of accuracy). Those are just three examples, which are encouraging as accuracy levels can be very high, but show that this is not guaranteed for any AI-powered education application.

In the current state of technology, one policy challenge is to ensure that the developed technology solutions perform their tasks with accuracy – or to get a clear sense of that level of accuracy. In spite of the very rapid advancement of smart technologies, computers and smart education technologies still remain imperfect – though not necessarily more imperfect than humans. Some of those imperfections are expected to be rapidly solved while it may take much longer for others. The contrary would be surprising when looking at technologies considered to be at the education technology frontier. But it is even possible that for some tasks, smart technologies will never become perfectly accurate and remain confounded by false positives, etc. However, the real question is how they compare to humans. After all, human beings perform those tasks with some levels of imperfection too (when they do) as the problems that they are addressing are usually complex. Whether full accuracy should be the expected standard is an open question. Probably this should depend on the social stakes related to the task. In many cases, an "accurate enough" diagnosis or decision should be sufficient.

Given that smart technologies are not fully mature yet and still have some intrinsic limitations, it is important for users and governments to remain aware of those limitations without preventing those technologies from continuing to improve and grow thanks to their actual use. Some possible policy pointers to mitigate the limits of smart technologies while embracing their potential are as follows:

- While smart education technologies can be useful before they are fully accurate (or even without being fully accurate), they should demonstrate a certain level of accuracy in their predictions and diagnoses when they support decision-making – or just be effective enough when they have a different role. Education technology companies could be asked to demonstrate the level of accuracy or effectiveness of their technology solutions with different accuracy requirements depending on the stakes of the supported decision (when they support a decision). Those accuracy requirements should ideally be compared with the current performance of human teachers and administrators.
- When still imperfect in terms of accuracy, smart education technologies should merely inform human decision-making and reflection rather than make fully automated decisions or support a decision process that will rarely deviate from their recommendations, especially for high-stakes solutions. Those technology requirements should have a risk-mitigation rather than no-risk policy, acknowledging that smart technologies can be beneficial even when they are not fully accurate. This may just imply keeping humans in control in the final stages when social stakes are high.

Designing for use

Sometimes, education technology solutions are designed and proposed because they are possible rather than because they are useful and provide clear benefits to end users in education. Most education technology products are mere educational derivatives of solutions designed for other sectors. Even when technology applications are useful and beneficial, some teachers, learners and users may have no interest in using them. Instances of lack of use and lack of usefulness of education technology have given rise to several critiques of education technology (Cuban, 1986[26]; Reich, 2020[27]) even though the increased use of technology in classroom instruction represents one of the biggest changes in classrooms of the 2010s (Vincent-Lancrin et al., 2019[28]).

How can one overcome this problem? Several chapters in this book discuss smart technology solutions that may not be sufficiently useful to be used at scale (given stakeholders' usual ways of working). One reason for this lack of use lies in the design of the smart technology solutions or in an insufficient understanding of how teachers can use them in their professional practice in ways that support rather than distract them. For example, classroom analytics are useful when they make visible to teachers what is either invisible or not easy to see (either in real time or after class) and when they provide information that they can act upon and interpret (Dillenbourg, 2021[12]).

One aspect of smart technologies that makes them more or less useful to their users relates to how they display the final information to end users. The interface between technology and humans is essential. For example, research shows that different types of dashboards can be more or less effective to support teachers and learners, or more appropriate in certain contexts than others (Molenaar, 2021[8]; Dillenbourg, 2021[12]). Dashboards typically display the final output of the analytics: they can take different forms (e.g. centralised, distributed, ambient) and use different display devices. While the mere appearance of social robots does not seem to matter much, social robots work better than virtual agents as users may relate to them in different ways than with a virtual image or a computer (Belpaeme and Fumihide, 2021[10]). The effectiveness of some solutions such as the ECHOES learning environment, which scaffolds autistic children's exploration and learning of social communication skills, partly lies in a setting that fosters communication between the autistic child and the adult monitoring the software and learning environment. This important dimension was actually discovered as the tool was tested and improved rather than as a preconceived use case, showcasing the importance in designing and adapting technology solutions with end users (Good, 2021[11]).

In some cases though, the usefulness of education technology may go beyond offering a technical solution to a specific problem. This is why usefulness and algorithmic accuracy are not always related: a solution with an algorithm accurately performing its task may not be so useful, whereas algorithms imperfectly performing their task may be. Changing the stakeholders' mindset or catalysing some broader change within an institution or an education system may be its usefulness. Generally speaking, innovation is a driver of professional learning and change (Avvisati et al., 2013[29]; Vincent-Lancrin, 2016[30]; Vincent-Lancrin et al., 2019[28]). Smart technologies play the same role.

Ifenthaler (2021[13]) shows that many universities and institutions introduce learning analytics at the institution-wide level in order to change their organisational culture or processes, and perhaps sometimes to foster new collaborations and ways of working between different stakeholders within the institution. Providing a solution to a specific problem or automating their processes may only be a secondary objective. Regardless of their effectiveness in reducing student dropout, one of the benefits that early warning systems (and the related research) has already delivered lies in a better and broader understanding of the circumstances leading students to drop out.

Bowers (2021[19]) shows that the traditional conception of students at risk of dropping out (as students with low and declining grades who do not like school) corresponds to only 38% of actual dropouts in the United States – so that traditional interventions miss out on the majority of students who actually drop out. Beyond providing real-time information, several of the functionalities of classroom analytics provide feedback to teachers on what happened in their class and one of their virtues can be to trigger professional reflection and learning, hopefully followed by behavioural change and improved teaching proficiency (Dillenbourg, 2021[12]).

Beyond design, including the final display of information to learners and other users, cost-benefit analysis is a final way to think about this "use" problem: if cheaper alternatives are available, in budget, time, cognitive load (or whatever relevant metric), new smart technology solutions may remain unattractive compared to existing human or older technology solutions. This is easily understandable when learning analytics predict, diagnose or act with little accuracy, but this can also be true for digital solutions with fully accurate or effective algorithms. Many of the digital tools that support teachers and administrators help them to solve very specific problems and sometimes the existing solutions may outperform the new ones. For example, at the system level, game-based assessments and simulations are likely to supplement rather than replace traditional standardised assessments based on batteries of questions: they are much more expensive to design, have less generalisability, and are only better suited to assess complex competences that are more difficult (or impossible) to assess through cheaper traditional alternatives (Buckley et al., 2021[14]). Smolenski (2021[25]) shows that blockchain has the potential to make the credentialing process more cost-efficient and simple for individuals. The technology is superior to some other technologies when fraud is possible, but not necessarily an adequate solution in all situations.

One last important point on usefulness relates to the affordability of smart technologies for public establishments and individuals. In the context of education systems, digital technology usually has to be very affordable to be bought – and thus be useful and have a chance to be actually used. As noted by Good (2021[11]) in the case of students with special needs, smart technologies should be designed to run on low-cost and widely available platforms (or devices). This is not always the case, but remains one of the conditions for them to be used – and to make their benefits widely accessible. Smolenski (2021[25]) also highlights the importance of open standards in the case of blockchain, partly as a way to make long term costs lower and ensure the solution is more sustainable for the end users (both institutions and individuals). This is true with many other technologies: open standards allow for greater interoperability, more sustainability over time, more competition among vendors, and often lower user costs. Technology solutions that run on widely available platforms are also more affordable and useable than specialised devices.

Several key messages emerge in terms of enhancing the usefulness and usage of smart technologies in education:

- Cost-benefit analysis should typically guide the design and adoption of smart digital solutions for different types of problems, acknowledging that benefits and costs are not just pecuniary.
- While the benefits of any solution may go beyond immediate student academic learning gains, and the costs, beyond merely financial ones, both the expected costs and benefits of smart technologies should be clearly identified and estimated either through research evidence when available and possible or a good theory of action (or theory of change).
- The display (or communication) of the information provided by learning analytics and other technology matters in making smart technologies useful to students, teachers and decision makers. More broadly speaking, the design of the interface between human and smart education is often a key aspect of the useability and impact of the digital solutions on learning or other targeted goals.
- Smart technology solutions should aim to be low cost and run on widely available platforms/devices to be as affordable as possible, possibly using open standards (and interoperability standards). Governments can support the development of those standards, preferably at the international level. Attention to affordability is essential to making smart technologies accessible to all so as not to reinforce the digital divide. Ensuring smart technology solutions can benefit all learners or institutions is key to equity and inclusiveness.

Smart technology and data governance: transparency, fairness and ethics

One key element of any socio-technical system is the broader social context in which the system operates, including its values and principles. Because they rely on large amounts of education data, including, sometimes, personal data such as biological markers, face recognition or expression, etc., or require a permanent monitoring and tracking of learners, classrooms or institutions, common concerns about the development and use of smart technologies relates to data protection and privacy, but also to ethical and political concerns. Could (or should) education establishments and systems become a new version of "Big Brother" for the sake of improved learning outcomes? Can governments and other parties be trusted to use this information for the mere sake of educational improvement – and to enforce strong data protection regimes? What could be the adverse consequences of this use, if not done properly, either in the present or in the future. Could data-rich education technologies, for example, perpetuate and reinforce biases and inequity? Embracing smart technologies implies some trust in how they are used, credible safeguards, and some level of understanding and acceptance of their processes and outputs.

Most OECD countries have strong data protection regulation that ensures that personal education data cannot be shared with (or used by) third parties beyond the educational processes for which they are collected unless certain privacy conditions are met. This is the case in the Europe Union with the General Data Protection Regulation (GDPR) and in the United States with the Family Educational Rights and Privacy Act (FERPA), which have both influenced many other data protection laws across countries. Much of the data concerns administrative micro-data (González-Sancho and Vincent-Lancrin, 2016[23]; n.d.[24]). The data protection regime usually extends to vendors providing technology solutions to schools and education administrations. It is noteworthy that the enforcement and implementation of data protection regulations can vary from one country to the other (or even one place to the other within a country). Given these arguably strong safeguards, the fact that privacy and data protection issues remain central in public discussions may point to a lack of trust in how the data are used (or could be used) within the education system and beyond.

Data protection is just one aspect of data governance though. One important question relates to the relationship between governments, the data subjects (who are usually the users of (sometimes mandatory) education services), and the private sector that usually develops the smart education technologies. This relates to questions of data ownership and of competition policy in the digital world: how should administrative and other education data be shared across education technology companies and public researchers to allow for progress and for a sufficient amount of competition in the sector of smart education technologies? How can the interests of learners and other individuals be best preserved in this context? The different solutions proposed for other sectors than education could probably be adapted to education (OECD, 2019[31]; 2019[1]).

Ethical discussions should normally concern what is not regulated within a country and thus aspects for which individuals and authorities have more freedom of action. This is how we interpret the question of "ethics" in educational AI.

The regulation of algorithms is usually not as strong as regulation for data protection. A major concern about algorithms is that they could be biased and have an undesirable social impact for some population groups (based on gender, ethnicity, socio-economic status, etc.) – but also that they could be flawed or just reinforce past human biases rather than reflect current societal values. A usual requirement is to ensure they are transparent and open, and that their "decisions" can be explained and challenged when automated. In the case of the EU General Data Protection Regulation, the regulatory language about algorithms is ambiguous (articles 13-15, 22, and recital 73) and lawyers are still debating what those articles imply in terms of "right of explanation" (transparency) and of the possibility to opt out of or challenge "automated decisions" for citizens. In the European Union, only the French and Hungarian laws have an explicit law about the "right of explanation" (with French law requiring both ex-ante and ex-post explanation in an intelligible way, and Hungarian law, some level of explanation) (Malgieri, 2019[32]). In the United States, there is no regulation about algorithms and their requirements as part of FERPA (or other regulation). Most OECD countries do not have clear regulatory requirements about them as of 2021.

Because algorithms based on machine learning are trained with historical data, many observers are concerned they will reproduce past biased (human) practices, as has apparently been the case in some countries in domains other than education (finance, justice, etc.) (O'Neil, 2016[33]). Several guidelines have been developed to avoid these pitfalls, which can happen at different steps of the process: measurement (data collection or labelling), model learning (when machine learning is involved), and action (when the algorithms detect, diagnose and act, for example). Different possible measures of fairness are also possible, which makes the issue even more complex (Kizilcec and Lee, 2020[34]; Baker and Hawn, 2021[35]). Ifenthaler (2021[13]) mentions several check lists of good practice and ethics for learning analytics. Bowers (2021[19]) points to the "open algorithm" movement in the area of early warning systems and notably to two overlapping sets of criteria to ensure transparency, verification and replicability of algorithms: the AAAA (accurate, accessible, actionable and accountable) and FAIR (findable, accessible, interoperable and reproducible) principles. Molenaar (2021[8]) also points to the importance of transparency to govern learning analytics and algorithms. As was shown in 2019 with some difficulties around exams and grade assignment for university admission, when it comes to high-stakes automated decisions, transparency is also about initiating an early dialogue about the criteria, expected social outcomes, relevance and acceptability of smart algorithms with diverse stakeholders from experts through to final users and other social bodies (Box 1.3). In some cases, the algorithms can be human-coded rather than involve AI techniques.

For example, given that accurate predictors for early warning systems can rely on minimal information that does not include information about gender, race and socio-economic status (Bowers, 2021[19]), it may not be considered ethical (or even necessary) to include these kinds of indicators to diagnose dropout in early warning systems unless they improve significantly the performance of the algorithms. On the other hand, results that do not include any personal information about learners may still lead to biased or socially/politically undesirable outcomes. Ethical concerns should therefore include a verification and discussion of the effects of smart technologies on different groups and ensure they are aligned with countries' social and political principles. As few people are, in practice, able to verify the effects and impacts of algorithms, some independent groups of stakeholders may be responsible for or even assigned this task. While anyone should be allowed to do it in the frame of an open algorithm culture (at least when algorithms lead to a decision or a quasi-decision), education researchers, non-governmental organisations, but also, possibly, independent governmental agencies, could play an enhanced role in this area.

In the case of the most advanced applications of learning analytics based on a continuous monitoring of individuals (e.g. engagement, self-regulation, classroom orchestration, game-based assessments), another question is whether stakeholders feel comfortable with some aspects of the applications even if they are legal. While the tracking and data collection necessary to power learning analytics focusing on student engagement, self-regulation or classroom orchestration have to comply with domestic data protection regulations (and algorithm regulation, if any), the question is how to make them compatible with the political values of the country where they are implemented. This may require some imagination in terms of data protection arrangements (such as deleting immediately the data once processed). As in the exam case mentioned above, this also requires a social negotiation with all stakeholders including transparency about the data collection and how they are used. This is not just a matter of regulation or even perceived ethics. Even within the same country, what can be acceptable in some communities may not be in others depending on how the smart technologies were introduced (Box 1.5).

Box 1.4 Two examples of controversies over predictive end-of-school grades during COVID-19

In summer 2020, the controversies that can arise around the use of algorithms in education were placed into sharp focus for both the International Baccalaureate Organization (IBO) and the Office of Qualifications and Examinations Regulation (Ofqual) in England. Lockdown measures and school building closures put in place to curb the spread of COVID-19 led to high-stakes end-of-high-school exams (English A level and International Baccalaureate) being cancelled. The results of these exams were used to allocate university admission places, creating the need to find ways of assigning grades to students for the same purpose.

Both IBO and Ofqual opted to develop algorithms to standardise grades on the basis of data from teacher assessments, previous performance, and a variety of other factors. In the case of IBO, historical assessment data from previous exam sessions as well as individual school data was used. In the case of England's A levels, teachers were asked to provide a centre assessment grade representing the grade students would have been most likely to achieve if teaching and learning had continued as usual rather than being disrupted by COVID-19. Teachers were also asked to provide a rank order of students for each student and grade. When centre assessment grades were considered, Ofqual found that their compound effect was likely to lead to an increase of up to 13% percentage points for some grade points compared to 2019. A statistical standardisation model was therefore developed and tested to produce grades based on the historical performance of the school or college in particular subjects as well as take into account factors such as changes in the prior attainment of candidates, centre assessment grades and rankings, and size of the cohort (Ofqual, 2020[36]).

In the case of both A levels and the IB, this process of standardisation led to substantial differences between predicted and assigned grades, meaning that offers of university places could be revoked, especially where they were conditional on students achieving a particular grade. When A-level results were released to students in England on 13 August 2020, for example, media reports suggested that around 40% of results were downgraded from teacher predictions meaning that many students did not meet the requirements of their offers for study at their first or second-choice universities (BBC, 2020[37]). Some alleged this disproportionately affected high-performing students in low-attaining schools, often in disadvantaged areas, because the algorithm had used the average of the previous performance of the school as part of the measures used to avoid grade inflation. This resulted in petitions, including from scientific associations such as the UK Royal Statistical Society, protests, and media articles complaining of a lack of transparency regarding the algorithms, models, and processes used and claiming compounded disadvantage for a cohort of students who were already having to cope with the effects of the pandemic (e.g. Studemann, 2020[38]; Adams and McIntyre, 2020[39]).

By mid-August 2020, it was announced that both A-level and IB results would be adjusted to reflect teacher estimates rather than results produced by the algorithms, ensuring in both cases that students would receive the highest grade of the two methods.

These examples highlight the socio-technical nature of smart technologies and the need for authorities to engage in a political dialogue with stakeholders to make the outcomes of smart technologies socially acceptable. In both cases, no one contested the accuracy of the algorithms, which did what they were coded to do, but rather the design parameters for predicting or adjusting grades, the observable or perceived outcome, especially for some sub-populations, and the lack of transparency of the process. Exams and grade assignment are social institutions that have been built over decades and centuries to become acceptable and part of the "meritocracy" construct (Sandel, 2020[40]). A big challenge ahead for the use of smart technologies will be to develop similar negotiated social acceptance. This is particularly true when algorithms lead to high-stakes decisions, as was the case in these two examples.

Box 1.5 Two examples of controversies related to social acceptability of smart technologies in school

Measuring and monitoring students' attention, behaviour or emotions in the classroom may help teachers to keep them engaged in learning. However, privacy protection and engagement with parents and other stakeholders are key dimensions for success. Two examples from China show the importance of social acceptability, stakeholder engagement and transparency in the deployment of such technologies.

In 2019, the Jinhua Xiaoshun Primary School (Zhejiang Province, China) trialled the FocusEDU headband. In combination with a software platform, these brainwave-tracking headbands used electroencephalography (EEG) technology to measure the extent to which students paid attention in class. Three hydrogel electrodes – one at the forehead and two behind the ears – detected electrical brain signals that an AI algorithm then converts into an attention score. The FocusEDU software provided teachers real-time access to individual attention levels in the class through a dashboard. In addition, lights on the headband's front showed different colours for different attention levels, signalling to teachers which students were identified as not paying attention. Local authorities suspended the trial in October 2019 due to privacy concerns.

Another example of monitoring students based on their behaviour and emotions was piloted in Hangzhou No. 11 Middle School (Zhejiang Province, China). Hikvision, a manufacturer of video surveillance equipment based in Hangzhou, developed cameras equipped with facial recognition technology that monitored students' in-class behaviour and facial expression under the name of "smart classroom behavioural management system". An AI algorithm classifies behaviour into six categories (reading, writing, listening, standing up, and lying on the desk) and distinguishes seven facial expressions (neutral, happy, sad, disappointed, angry, scared and surprised). An overall attention score was computed from these classifications that teachers could access in real-time through a screen. Following parents' concerns, the use of the technology to evaluate facial expressions was suspended in May 2018. Since then, Hangzhou No. 11 middle school mainly uses the facial recognition cameras for monitoring attendance and for on-campus payments.

Those two examples show the difficulties that come with some monitoring aspects of smart technologies. On the research side, one difficulty lies in the quality of the theoretical models used to identify emotions and link them to learning outcomes. (There is no published research about those pilots and the underlying models to measure engagement to our knowledge.) As the proposed solutions are not so different from those that are in use at some other Chinese schools (see Box 1.1), the interruption of both pilots points more to difficulties related to local acceptance and communication, which matter also where regulation on privacy and data protection might not be as restrictive as in some OECD countries.

Source: Focus EDU (Standaert, 2019[41]; Wang, Hong and Tai, 2019[42]); Hikvision (Li and Jourdan, 2018[43]; Yujie, 2019[44]; Lee, 2018[45])

Another pragmatic and ethical issue relates to the use of the information generated by data analytics about teachers and other staff. Typically, staff within an organisation do not benefit of the same data protection and privacy regulation as students and other users. While smart technologies and learning analytics have the potential to provide feedback and support to teachers and other education stakeholders to make better decisions and improve their professional practices, they could also be used against them and unintentionally lead to undesirable social behaviours. Classroom analytics can be used to monitor teachers' professional behaviour, and sometimes identify shortcoming in how they orchestrate learning in their classroom (Dillenbourg, 2021[12]). Should this information be used to sanction or support them? One possible intervention for students at risk of dropping out also consists of expelling or pushing them out of the school detecting this risk, which eventually enhances their likelihood to drop out eventually (Bowers, 2021[19]). Given their possible intrusive surveillance nature, the adoption of smart technologies relies on some level of trust in their positive, human-empowering ethical use. Should their voluntary use have adverse effects on teachers, school principals and decision-makers, they may appear as less acceptable and be resisted. Ethics about their use may require either full confidentiality about their results or a discussion and clear disclosure of how they can affect the staff using them. (This may also depend on their expected accuracy and effectiveness). As in the case of the information provided by longitudinal information systems, two different philosophies are possible. Some argue that their information should be used to reward and sanction stakeholders as an accountability mechanism, which is also a way to make them pay attention to the provided information. Others argue that the information should not be used to reward or sanction stakeholders as this may lead to their opposing the use of the information or just incentivise them to try to "game the system" – and thus lead to unethical behaviour. The jury is still out on what the best strategy is.

Building on the fact that smart technologies are socio-technical systems, their adoption and use will typically require some level of trust in teachers, schools, governments, and other stakeholders – and some attention paid to their possible adverse effects. Building trust should build both on regulation and the establishment of ethical practices, including different governance mechanisms:

- Regulation on data protection and privacy, as well as practical guidance on how to implement those provisions with different types of data and different levels of resources and competence;
- Regulation or good practice on data governance, and notably a strategy that makes data collected through smart technologies available to researchers and possibly to competitors to the vendors whose technology solution collects them;
- Regulation or guidelines about the transparency, openness and replicability of algorithms, as well as funding and support for the verification of the design and the final results of algorithms by independent parties;
- A risk-management approach to data protection and algorithm supervision that strikes an appropriate balance between risk-taking and data/usage sensitivity, acknowledging that regulations (or ethical guidance) that are too risk-averse will hinder the development of smart education technologies within a country or region and prevent reaping their possible benefits. The regulations above should be updated as smart technologies continue their rapid progress, and a close collaboration between researchers on data protection and on smart technologies may help co-design good solutions.

Infrastructure and public good

Smart technologies do usually require a strong Internet, computer and data infrastructure. Artificial intelligence systems, adaptive learning systems providing real-time information to teachers and learners, game-based assessments, blockchain, social robots, all require sufficient computer hardware in schools and universities (including students' hardware under "bring your own device" policies), but also increasingly at home, as well as a bandwidth and networking capacity capable of stable and acceptable data transfer speeds. This also implies some level of investment in IT staff to deal with the maintenance of the hardware in (or across) educational establishments. Hardware is a necessary basis, but there is more to what a digital learning infrastructure should be.

Policy makers need to think about infrastructure also in terms of digital resources, both content and tools, that should be publicly provided to citizens, students, and educational institutions – either directly or indirectly through subsidies or funding to institutions. Budget constraints can limit these aspirations, but it remains important to define what should be the core of digital resources for all, and what should be accessible privately. A new question is the extent to which smart technologies should be part of the core digital education infrastructure. Is it enough to make some educational resources available to the public or should they also have the personalised learning features that some intelligent tutoring systems propose?

A final aspect of a digital infrastructure lies in people's "digital skills", that is, the ability to use digital resources as part of one's professional practice. In the case of teachers and professors, digital skills are less about mastering the technology than about integrating technology tools, resources and outputs in their pedagogy. Unless fully automated, technology solutions are indeed mere tools for human teaching, learning or managing education systems. Professional learning opportunities for staff, both through training and organised continuous professional learning opportunities, are thus a final aspect that should be an integral part of a strong digital infrastructure.

While university researchers and researchers in public technology agencies may have the skills to develop some of those smart technologies, governments, schools and universities will typically rely on private education technology companies to provide and maintain them (hopefully through public-private partnerships, as mentioned above).

Beyond this investment in its digital infrastructure and connectivity for all, governments still have two important responsibilities in dealing with the private sector:

- Ensuring through its procurement policies and other incentives that publicly funded or purchased smart technologies are available to schools with affordable costs.
- Ensuring that some key techniques or discoveries of those smart education technologies become or remain a public, international good and allow more actors internationally to develop new interoperable solutions that will help improve education for all across the globe.
- Ensuring that staff have the learning opportunities to properly use the smart technologies and digital resources at their disposal.

Research and development

This book shows how smart technologies can be beneficial to education in various ways. At the same time, a common thread through the different chapters of the book is the insufficient amount of strong evidence about the effectiveness of the various uses of smart technologies. As noted above, there is a relative lack of evidence and transparency about the actual algorithmic accuracy of the technology solutions but more evidence on their effective use in real-life educational settings should also be produced. Social robots are one of the few areas where meta-analyses and some strong evidence results could be established about some specific use of technology to improve learning outcomes (Belpaeme and Fumihide, 2021[10]).

While education research and development is needed, new approaches may be warranted.

Policy makers and research funders should remain aware that technology is not to be tested in itself. As a mere tool that is part of a socio-technical system, the appropriate research question is rarely whether some specific technological application is "effective" but more about identifying and testing whether some specific uses yield positive outcomes. Usually, when it comes to instruction, the research question will be about pedagogy and how technology supports this pedagogy rather than about the technology itself. When it comes to administrative processes, the question may be about the interventions supported or induced by the smart technology rather than the technology itself.

Some specific technological applications of smart technology have been (and can be) researched through robust experimental research designs. However, as the development of smart technologies (and computing power) advances quickly, the evidence about their effectiveness may be quickly outdated and irrelevant as the technology gets upgraded.

As other digital technologies, smart technologies allow for new types of speedy research when they are done online: A/B research. Typically, this consists of trying two different designs (A and B) of a given technology with two different groups to identify quickly which one works best (when there are large numbers of use/exposure online). This approach can be extended to classrooms to evaluate the effectiveness in using different designs of some given smart technology or of different smart technologies pursuing similar objectives in a classroom or at the system level, factoring in the complex human behaviours that are attached to their use.

Should one wait for strong evidence before using technology in education? In practice, this may be too restrictive an approach. Requirements for strong evidence before the introduction of any innovation are often a way to protect the status quo of practices whose effectiveness is rarely proven. Disciplined innovation can happen in other ways too, notably through a proportionate use of effectiveness evidence. In the case of some smart technologies, usual evaluation processes may not be the most appropriate and sometimes it may be difficult to carry them out. It would, for example, be difficult to evaluate the effectiveness of early warning systems through a randomised control trial – one reason being that systems may have to be different from one site to the other to accommodate local circumstances. The speed of development of the technology will also typically mean that experimental results will either validate or invalidate technology solutions that are already outdated by improved versions of the solution (or new technologies altogether). Perhaps another approach would be to evaluate families of technology solutions used in certain ways.

The evidence standards required for different types of solutions should probably vary depending on the possible stakes and harm coming from the solution. Smart technologies supported by a good "theory of action" or good "underlying theory" may be acceptable for use even without a strong evidence base for low-stakes applications. When there is little underlying theory, no clear theory of action, no related basis of evidence, policy makers and stakeholders should be more reluctant to allow their use in public settings. Finally, in the case of high-stakes situations, the highest evidence standards should be in place, although algorithmic effectiveness should be compared to evidence of human effectiveness when performing similar tasks.

To further address this question, governments could:

- Invest in educational research about the use of technology in education in real life settings with a focus on pedagogy or administrative processes rather than the technology itself;
- Develop domestic and possibly international repositories of evidence on different types of uses of technology in education and different types of families of education technology;
- Develop common standards about good educational research and development in educational technology, acknowledging some of the specificities of the field.

Concluding remarks

Smart technologies that could transform education are already available. Some are more mature than others, but a range of solutions could and will make education systems and establishments operate differently in the future. Some of these tools are about the personalisation of learning (intelligent tutoring systems), about keeping students motivated and engaged in their learning, about allowing students with special needs to benefit fully from the curriculum. Technology can also support teachers. Smart technologies based on classroom analytics allow them to orchestrate teaching and learning in their class in real time and post-hoc while social robots can support them as teaching assistants, instructors and even peer students. Technology has also made big strides in supporting the management of education systems, with a host of solutions at the system and organisational level to manage budget, study paths, relationships with external stakeholders, etc. The development of early warning systems to help prevent students from dropping out of high school (or university) is a good example of this progress.

Some of the promises of smart technologies relate to the effectiveness of education. They can support students to achieve better learning outcomes, and teachers to teach (and also learn) better. Another promise lies in enhancing equity: technology helps to make education more inclusive and can provide additional learning opportunities to students from more disadvantaged groups – assuming that they are widely accessible and used. Cost-efficiency through automation is one aspect that digitalisation has brought to many sectors of society: this is also gradually happening in education. At the same time, developing and maintaining technology can be costly, and the public cost has to be balanced against its benefits.

Even if none of these promises of digitalisation materialised, digitalisation could still open new avenues for formal education and make it more convenient, more enjoyable, or just… different and aligned with modern life. Innovation is in itself a source of professional learning for teachers and also for students: it is a means to create new capacity within a system just because people have to adjust to the new requirements it promotes (Cohen and Moffitt, 2009[46]; Vincent-Lancrin, 2016[30]). Introducing digital tools in schools and universities may not have a narrow objective but be a tool to trigger change and improvement efforts. It is also a way for formal education to be part of its digital time. Should schools and universities resist or embrace digitalisation, which is gaining ground in all OECD societies, regardless of what happens in education? While formal education systems should empower everyone to enjoy, access and learn from all the knowledge and experiences that have been developed by humanity, education should be more than a museum.

Several scenarios are possible (and several have been developed: see OECD (2020[47]) for general schooling scenarios, and HolonIQ (2018[48]) for scenarios on digitalisation).

One scenario would be for education to change minimally and continue to have little adoption of technology and digital resources in teaching and learning. This may mean that most smart technologies would be available privately for out-of-school learning for those who can afford it. The education technology market would continue to mainly target its supply to the informal education market and corporate training. One long-term question in this scenario is whether education systems will remain relevant and whether out-of-school learning could become as important or more important than in-school learning.

A second scenario would be for education to look similar, on the surface, but become quite different, exactly as cars or planes look more or less the same as 40 years ago, but have become quite different now that they are fully equipped with sensors and computing devices. Education establishments may also become connected buildings with cameras, sensors and digital devices supporting students, teachers and administrators to make decisions to improve their teaching, learning and management practices. Technology may also become more prevalent for learning at home, with more intelligent tutoring resources etc. available for everyone to use.

A third scenario could be for education to build on smart technologies and other social trends related to digitalisation to reshape as a social institution. People may increasingly telework, more schoolwork may happen at home, sometimes with more involvement of parents and communities, and social time in school may become used mainly for individual tutoring and collective learning. For example, students can choose to go to school to do some tasks individually or perform them at home while other activities must be done at school with peers and under the guidance of their teachers.

The two latter scenarios would have implications for teachers and the main aspects of teaching, but also for what it means to be a student and how parents can support their children. Similar scenarios could be envisaged when it comes to the management of education systems and organisations. For example, many administrative processes could be fully automated, from assessments to the allocation of students to different educational institutions.

Of course, the future might hold completely different scenarios or any combination of them. But now is the time to think about what is possible and how digital technology can best support the improvement of education.

References

Adams, R. and N. McIntyre (2020), *England A-level downgrades hit pupils from disadvantaged areas hardest (The Guardian)*, https://www.theguardian.com/education/2020/aug/13/england-a-level-downgrades-hit-pupils-from-disadvantaged-areas-hardest (accessed on 19 May 2021). [39]

Avvisati, F., S. Hennessy, R.B. Kozma and S. Vincent-Lancrin (2013), "Review of the Italian Strategy for Digital Schools", *OECD Education Working Paper*, Vol. 90, http://www.oecd.org/education/ceri/Innovation%20Strategy%20Working%20Paper%2090.pdf. [29]

Bailey, T., S. Smith-Jaggars and D. Jenkins (2015), *Redesigning America's Community Colleges: A Clearer Path to Student Success*, Harvard University Press. [15]

Baker, R. and A. Hawn (2021), *Algorithmic Bias in Education*, Center for Open Science, http://dx.doi.org/10.35542/osf.io/pbmvz. [35]

BBC (2020), *A-levels and GCSEs: U-turn as teacher estimates to be used for exam results*, https://www.bbc.com/news/uk-53810655 (accessed on 19 May 2021). [37]

Belpaeme, T. and T. Fumihide (2021), "Social Robots as Educators", in *OECD Digital Education Outlook 2021: Pushing the frontiers with AI, blockchain, and robots*, OECD Publishing. [10]

Bowers, A. (2021), "Early Warning systems and indicators of dropping out of upper secondary school: the emerging role of digital technologies", in *OECD Digital Education Outlook 2021: Pushing the frontiers with AI, blockchain, and robots*, OECD Publishing. [19]

Buckley, J., L. Colosimo, R. Kantar, M. McCall and E. Snow (2021), "Game-based Assessment for Education", in *OECD Digital Education Outlook 2021: Pushing the frontiers with AI, blockchain, and robots*, OECD Publishing. [14]

Cohen, D. and S. Moffitt (2009), *The ordeal of equality: Did federal regulation fix the schools?*, Harvard University Press, Cambridge, MA. [46]

Cuban, L. (1986), *Teachers and Machines: The Classroom of Technology Since 1920*, Teachers College Press. [26]

D'Mello, S. (2021), "Improving Student Engagement in and with Digital Learning Technologies", in *OECD Education Outlook 2021: Pushing the frontiers with AI, blockchain, and robots*, OECD Publishing. [9]

Dillenbourg, P. (2021), "Classroom analytics: Zooming out from a pupil to a classroom", in *OECD Digital Education Outlook 2021: Pushing the frontiers with AI, blockchain, and robots*, OECD Publishing. [12]

González-Sancho, C. and S. Vincent-Lancrin (2016), "Transforming education by using a new generation of information systems", *Policy Futures in Education*, Vol. 14/6, pp. 741-758. [23]

González-Sancho, C. and S. Vincent-Lancrin (n.d.), *Transforming education by using a new generation of information systems.* [24]

Good, J. (2021), "Serving students with special needs better: How digital technology can help", in *OECD Digital Education Outlook 2021: Pushing the frontiers with AI, blockchain, and robots*, OECD Publishing. [11]

HolonIQ (2018), *Building Scenarios for Education in 2030*, https://www.holoniq.com/2030/building-scenarios-education-in-2030 (accessed on 19 May 2021). [48]

Ifenthaler, D. (2021), "Learning Analytics for School and System Management", in *OECD Digital Education Outlook 2021: Pushing the frontiers with AI, blockchain, and robots*, OECD Publishing. [13]

KERIS (2010), "The Future of Education: NEIS (National Education Information System)", *Presentation at OECD/SSRC/Stupski Workshop on Educational Information Systems for Innovation and Improvement*, https://www.oecd.org/education/ceri/46182781.ppt (accessed on 19 May 2021). [22]

Kizilcec, R. and H. Lee (2020), "Algorithmic Fairness in Education", in Holmes, W. and K. Porayska-Pomsta (eds.), *Ethics in Artificial Intelligence in Education (forthcoming)*, Taylor & Francis. ArXiv E-Prints, https://arxiv.org/abs/2007.05443. [34]

Lambert, S. (2020), "Do MOOCs contribute to student equity and social inclusion? A systematic review 2014–18", *Computers & Education*, Vol. 145, p. 103693, http://dx.doi.org/10.1016/j.compedu.2019.103693. [18]

Lang, C., G. Siemens, A. Wise and G. Gašević (eds.) (2017), *Handbook of Learning Analytics*, SOLAR, https://solaresearch.org/wp-content/uploads/2017/05/hla17.pdf. [7]

Lee, D. (2018), *Must Reads: At this Chinese school, Big Brother was watching students — and charting every smile or frown (Los Angeles Times)*, https://www.latimes.com/world/la-fg-china-face-surveillance-2018-story.html (accessed on 19 May 2021). [45]

Li, P. and A. Jourdan (2018), *Sleepy pupils in the picture at high-tech Chinese school (Reuters)*, https://www.reuters.com/article/us-china-surveillance-education/sleepy-pupils-in-the-picture-at-high-tech-chinese-school-idUSKCN1II123 (accessed on 19 May 2021). [43]

Malgieri, G. (2019), "Automated decision-making in the EU Member States: The right to explanation and other "suitable safeguards" in the national legislations", *Computer Law & Security Review*, Vol. 35/5, p. 105327, http://dx.doi.org/10.1016/j.clsr.2019.05.002. [32]

Molenaar, I. (2021), "Personalisation of learning: Towards hybrid human-AI learning technologies", in *OECD Digital Education Outlook 2021: Pushing the frontiers with AI, blockchain, and robots*, OECD Publishing. [8]

Murphy, R., J. Roschelle, M. Feng and C.A. Mason (2020), "Investigating Efficacy, Moderators and Mediators for an Online Mathematics Homework Intervention", *Journal of Research on Educational Effectiveness*, Vol. 13/2, pp. 235-270, http://dx.doi.org/10.1080/19345747.2019.1710885. [20]

O'Neil, C. (2016), *Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy*, Penguin Books. [33]

OECD (2020), *Back to the Future(s) of Education*, OECD, http://dx.doi.org/10.1787/178ef527-en. [47]

OECD (2019), *Artificial Intelligence in Society*, OECD Publishing, Paris, https://dx.doi.org/10.1787/eedfee77-en. [2]

OECD (2019), *Enhancing Access to and Sharing of Data: Reconciling Risks and Benefits for Data Re-use across Societies*, OECD Publishing, Paris, https://dx.doi.org/10.1787/276aaca8-en. [31]

OECD (2019), *Going Digital: Shaping Policies, Improving Lives*, OECD Publishing, Paris, https://dx.doi.org/10.1787/9789264312012-en. [1]

OECD (2019), "Using digital technologies to improve the design and enforcement of public policies", *OECD Digital Economy Papers*, No. 274, OECD Publishing, Paris, https://dx.doi.org/10.1787/99b9ba70-en. [16]

OECD (2018), "IoT measurement and applications", *OECD Digital Economy Papers*, No. 271, OECD Publishing, Paris, https://dx.doi.org/10.1787/35209dbf-en. [6]

OECD (2017), *OECD Blockchain Primer*, https://www.oecd.org/finance/OECD-Blockchain-Primer.pdf. [5]

Ofqual (2020), *Research and analysis: Awarding GCSE, AS & A levels in summer 2020: interim report*, https://www.gov.uk/government/publications/awarding-gcse-as-a-levels-in-summer-2020-interim-report (accessed on 19 May 2021). [36]

Orr, D., M. Rimini and D. van Damme (2015), *Open Educational Resources: A Catalyst for Innovation*, Educational Research and Innovation, OECD Publishing, Paris, https://dx.doi.org/10.1787/9789264247543-en. [17]

Reich, J. (2020), *Failure to Disrupt: Why Technology Alone Can't Transform Education*, Harvard University Press. [27]

Roschelle, J., M. Feng, R.F. Murphy and C.A. Masonet (2016), "Online Mathematics Homework Increases Student Achievement", *AERA Open*, Vol. 2/4, p. 233285841667396, http://dx.doi.org/10.1177/2332858416673968. [21]

Sandel, M. (2020), *The Tyranny of Merit: What's Become of the Common Good?*, Allen Lane. [40]

Smolenski, N. (2021), "Blockchain for education: a new credentialing ecosystem", in *OECD Digital Education Outlook 2021: Pushing the frontiers with AI, blockchain, and robots*, OECD Publishing. [25]

Standaert, M. (2019), *Chinese primary school halts trial of device that monitors pupils' brainwaves (The Guardian)*, https://www.theguardian.com/world/2019/nov/01/chinese-primary-school-halts-trial-of-device-that-monitors-pupils-brainwaves (accessed on 19 May 2021). [41]

Studemann, F. (2020), *A-level results chaos will have lasting impact on class of 2021 (Financial Times)*, https://www.ft.com/content/0e86d8c1-cb47-4965-bef2-6f694352235a (accessed on 19 May 2021). [38]

van der Vlies, R. (2020), "Digital strategies in education across OECD countries", *OECD Education Working Papers*, Organisation for Economic Co-Operation and Development (OECD), http://dx.doi.org/10.1787/33dd4c26-en. [3]

Vincent-Lancrin, S. (2016), "Innovation, Skills and Adult Learning: two or three things we know about them", *European Journal of Education*, Vol. 51/2, pp. 146-153. [30]

Vincent-Lancrin, S., C. González-Sancho, M. Bouckaert, F. de Luca, M. Fernández-Barrerra, G. Jacotin, J. Urgel and Q. Vidal (2019), *Fostering Students' Creativity and Critical Thinking: What it Means in School*, Educational Research and Innovation, OECD Publishing, Paris, https://dx.doi.org/10.1787/62212c37-en. [4]

Vincent-Lancrin, S., J. Urgel, S. Kar and G. Jacotin (2019), *Measuring Innovation in Education 2019: What Has Changed in the Classroom?*, Educational Research and Innovation, OECD Publishing, Paris, https://dx.doi.org/10.1787/9789264311671-en. [28]

Wang, J., S. Hong and C. Tai (2019), *China's Efforts to Lead the Way in AI Start in Its Classrooms (The Wall Street Journal)*, https://www.wsj.com/articles/chinas-efforts-to-lead-the-way-in-ai-start-in-its-classrooms-11571958181 (accessed on 19 May 2021). [42]

Yujie, X. (2019), *Camera Above the Classroom (Sixth tone)*, https://www.sixthtone.com/news/1003759/camera-above-the-classroom (accessed on 19 May 2021). [44]

2 Artificial intelligence in education: Bringing it all together

Ryan S. Baker

University of Pennsylvania, United States

Artificial intelligence has led to a generation of technologies in education – for use in classrooms and by school systems more broadly – with considerable potential to bring education forward. This chapter provides a broad overview of the technologies currently being used, their core applications, and their potential going forward. The chapter also provides definitions of some of the key terms that will be used throughout this book. It concludes with a discussion of the potentials that may be achieved if these technologies are integrated, the shifts in thinking about supporting learners through one-on-one learning experiences to influencing systems more broadly, and other key directions for R&D and policy in the future.

Introduction

For decades, educators and researchers have looked to computers as having the potential to revolutionise education. Today, much of the use of computers in education still falls short of revolutionary – a lot of learning still involves one instructor teaching many students simultaneously, and considerable computer-based learning takes place using curricula and technologies that replicate traditional practices such as drill and practice. However, the best practices of computers in education appear to go considerably beyond that. Millions of learners now use intelligent tutoring systems as part of their mathematics classes – systems that recognise student knowledge, implement mastery learning where students do not advance until they can demonstrate understanding of a topic, and have hints available on demand (VanLehn, 2011[1]). Millions of learners around the world watch lectures and complete exercises within massive online open courses, offering the potential to study thousands of topics that would not be available in colleges locally (Milligan and Littlejohn, 2017[2]). An increasing number of children and adults learn from (and are even assessed within) advanced online interactions such as simulations, games, virtual reality, and augmented reality (De Jong and Van Joolingen, 1998[3]; Rodrigo et al., 2015[4]; Shin, 2017[5]; De Freitas, 2018[6]). Perhaps none of these systems fully captures the sophistication dreamed of in early accounts of the potential of these systems (Carbonell, 1970[7]) (Stephenson, 1998[8]). On the other hand, their scale and degree integration into formal educational systems have gone beyond what seemed plausible even as recently as the turn of the century.

Increasingly, computer-based education has been artificially intelligent education. Advances in artificial intelligence (AI) in the 1980s, 1990s, and first decade of the new millennium have translated to new potentials for learning technologies in several areas. The core advances in AI in those decades led to advances in more specialised use of AI in education – the research and practice communities of learning analytics and educational data mining – from around 2004 to today. As the research field advanced, new methods filtered into systems used by learners at scale. AI is today used to recognise what students know (and their engagement and learning strategies) to predict their future trajectories, better assess learners along multiple dimensions, and – ultimately – help both humans and computers decide how to better support students.

As these technologies develop, as they mature, as they scale, it becomes worth asking the question: where are we going? And where could we be going? If we can understand the frontiers and potentials of artificial intelligence in education, we may be able to shape research and development through careful design of policy over the next decade to get there.

In the chapters of the book, the authors use their expertise in specific areas and challenges in artificial intelligence in education to explore the frontiers and potential of this area. What technology and teaching approaches are just becoming available in research classrooms that could soon be available to a much broader range of students? How can artificial intelligence shape educational systems more broadly, from academic advising to credentialing, to make them more adaptive to learner needs? Where could we be in 10 years with the right guidance and support for research and development? Where are the opportunities for incremental but positive impacts on learners? And where are the opportunities for radically transforming education and learner experiences?

In the remainder of this overview, I will clarify some terms and domains relevant to this book. Next, I will situate the chapters of this book in the context of broader trends and opportunities in the field (including some trends and opportunities that were not explicitly covered by the authors). The final section will discuss upcoming opportunities of a broader nature, cutting across types of artificially intelligent learning technologies that can be supported through shifts in policy.

Smart education technologies: Definitions and context

In this section, some definitions and context that are key to understanding smart technologies in education are presented.

Educational technology

Educational technology at its most obvious level refers simply to the use of any technology – any applied machinery or equipment – in education. Throughout the last hundred years, practitioners and researchers have sometimes become overly enthusiastic about finding applications for new technologies in education. See, for instance, reports by Cuban (1986[9]) of instructors teaching students with traditional lecture pedagogies but inside an early-generation airplane.

Today, most discussion of technology in education is focused on computers and digitalisation, though older technologies such as radio and television still play an important role – especially in many middle-income countries during the recent COVID-19 pandemic (OECD, n.d.[10]). Educational technologies can refer to a range of technologies. I provide a few examples here (others are given in the context of the chapters in this book).

- *Computer tutors or intelligent tutoring systems* provide students with a learning experience where the learning system adapts presentation based on some model or ongoing assessment of the student, a model of the subject area being learned, and a model of how to teach (Wenger, 1987[11]). Each of these models can be more sophisticated or more basic. Baker (2016[12]) notes that contemporary intelligent tutoring systems tend to be sophisticated in only one area (which differs between systems) and very simple in other areas.
- *Digital learning games* embed learning into a fun activity that resembles a game. The degree of gamification can vary from activities that embed learning into core gameplay and which may not even seem to be a learning activity (see, for instance, SimCity and Civilisation) to more obvious learning activities where the student gets rewards for successful performance (for instance, getting to throw a banana at a monkey after answering a math problem correctly in MathBlaster).
- *Simulations* are computerised imitations of a process or activity that would be difficult or costly to do in the real world as an educational activity. Increasing numbers of students today use virtual laboratories to conduct experiments that could be dangerous, expensive, or difficult, and also to receive feedback and learning support while completing these activities.

- *Virtual reality systems* embed learners in 3D depictions of real-world activities. Like simulations, they make it feasible to engage in activities from a home or computer lab that would be expensive, dangerous, or simply impossible to engage in otherwise. Augmented reality systems embed additional information and experiences into real-world activities, ranging from pop-up details that appear and ambient displays (information that is available in the environment without having to focus on it) to overlaying a different world on top of the current one. Both augmented reality and virtual reality often rely upon headsets to present visual information to learners.
- *Educational robots* have a physical presence and interact with students in real-world activities to support their learning. While robots as educational DIY kits have been available since the 1980s, a recent development sees robots take up the role of tutor.
- *Massive online open courses (MOOCs)* provide students with a basic learning experience, typically consisting of videos and quizzes. The innovation around MOOCs is not in the learning experience – it is typically a simplified version of a large lecture class – but, rather, in making materials developed by faculty at world-famous universities, often on highly specialised topics, accessible to learners around the world.

Educational data

Data are, quite simply, facts gathered together. Whereas a few facts gathered together do not enable us to reason about the relationships represented in that information, the accumulation of large quantities of information does and that is the modern power of big data. Educational data used to be dispersed, hard to collect, and small-scale. Individual teachers might keep a gradebook on paper; the school might keep disciplinary records in the basement; and curriculum developers would have a very limited idea of how their materials were being used and what students were struggling with. Today, educational data is gathered at a much larger scale. Gradebooks, disciplinary data, assessment data, absence data and more is stored centrally by local education agencies (or often by national or even trans-national vendors). Curriculum developers often gather extensive data on usage and learning. As of this writing, the regulations around handling, storage, and use of educational data vary considerably between countries, with some countries having very strict practices (particularly on the European continent), and other countries having less restrictive regulations. Each of these sources of data can be used to improve educational quality and support learning, supporting both artificial intelligence/machine learning (next definition) and human refinement of learning content and experiences.

Artificial intelligence and machine learning

Artificial intelligence is the capacity for computers to perform tasks traditionally thought to involve human intelligence or, more recently, tasks beyond the ability of human intelligence. Stemming from relatively simple, general-purpose systems in the 1960s, artificial intelligence today generally involves more specific-purpose systems that complete a specific task involving reasoning about data or the world, and then interaction with the world (more commonly through a phone or a computer interface than actual physical interaction). Machine learning (increasingly called data science, and also called both data mining and analytics) is a sub-area of artificial intelligence, present at a low level since the beginning of the field but becoming a particular emphasis in the 1990s through to today. Machine learning is when a system discovers patterns from data – becoming more effective at doing so when more data is available (and even more so, when more comprehensive or representative data is available). There is a broad range of machine-learning methods, classified mostly into supervised learning (attempting to predict or infer a specific known variable) and unsupervised learning (trying to discover the structure or relationships in a set of variables). There have roughly been two generations of machine learning: a first generation of relatively simple, interpretable methods and a second generation of much more complex, sophisticated, hard-to-interpret methods.

Artificial Intelligence in Education (AIED)

Artificial Intelligence in Education (AIED) arose as an interdisciplinary subfield in the early 1980s with a bi-annual (now annual) conference and peer-reviewed journal, although examples of this research area were present even before that. Much of the early work in artificial intelligence in education involved intelligent tutoring systems but the field has broadened over the years to include all of the types of educational systems/interactions defined above, and has expanded to include several independent conferences and journals. The revolution in machine learning and data mining impacted artificial intelligence in education as well, with a significant shift around 2010 – influenced by the emergence of a separate scientific conference, Educational Data Mining – towards much more use of this type of method. Today, AIED systems incorporate a range of functionality for identifying aspects of the learner, and a range of ways they can interact with and respond to learners.

Learning analytics

Learning Analytics, also referred to as Educational Data Mining, has emerged as a field since 2008 with two major international conferences and peer-reviewed journals. The goal of learning analytics is to use the increasing amounts of data coming from education to better understand and make inferences on learners and the contexts which they learn from. Learning analytics and educational data mining apply the methods of machine learning/ data science to education, with methods and problems emerging specific to education. Challenges such as inferring student knowledge in real-time and predicting future school dropout have seen particular interest, but there have been a range of other applications for these methods, from inferring prerequisite relationships in a domain such as mathematics to understanding the factors that lead to student boredom. A taxonomy of methods and applications for learning analytics is given in (Baker and Siemens, 2014[13]; DeFalco et al., 2017[14]). Learning analytics models are most frequently deployed in two types of technology: intelligence augmentation systems and personalised learning systems (discussed in the next section).

Intelligence augmentation systems, also called decision support systems, communicate information to stakeholders such as teachers and stakeholders in a way that supports decision-making. While they can simply provide raw data, they often provide information distilled through machine-learning models, predictions, or recommendations. Intelligence augmentation systems often leverage predictive analytics systems, which make predictions about students' potential future outcomes, and – ideally – also provide understandable reasons for these predictions. Predictive analytics systems are now used at scale to try to understand which students are at risk of dropping out of high school or failing to complete college, with an eye towards providing interventions which get students back on track. Intelligence augmentation systems often communicate information to stakeholders through dashboards, which communicate data through graphs and tables that allow the user to drill down for information about specific learners. Today, personalised learning systems and predictive analytics systems often use dashboards to communicate information to teachers, occasionally make dashboards available for school counselors, academic advisors, and school leaders, and rarely make dashboards available for parents. The quality of the data presented in dashboards can vary considerably from learning system to learning system.

The uses of artificial intelligence in classrooms and educational systems

This book focuses on two key areas: 1) New Educational Technologies and Approaches for the Classroom, and 2) New Educational Technologies and Approaches for Educational Systems. These new technologies often but not always involve artificial intelligence. Within this section, I will summarise work in each of these areas, including both the work discussed in the chapters in this report, but going beyond as well.

New educational technologies and approaches for the classroom

As computerised educational technologies become more commonly accessible to teachers and students, there is increasing awareness that the technology does not simply increase convenience for teachers or provide a fun alternative activity for students – it can promote new methods for teaching and learning.

Personalised Learning. One major trend within learning, driven by these technologies, is the move towards personalising learning to a greater degree. Personalisation of learning did not start with computerised technology – in a sense, it has been available since the first use of one-on-one tutoring, thousands of years ago (if not earlier). However, with the increase in systematised, standardised schooling and teaching over a hundred years ago, awareness increased that many students' learning needs were being poorly met by one-size-fits-all curriculum. Classroom approaches such as mastery learning (each student works on material until mastery and only then moves on to the next topic) were developed, but proved difficult to scale due to the demands on the teacher. Educational technologies provided a ready solution to this problem – the computer could manage some of the demands of personalising learning, identifying each individual student's degree of mastery and providing them with learning activities relevant to their current position within the curriculum.

The first dimension that educational technologies became effective at personalising for was a student's knowledge or state of learning. Molenaar (2021[15]) details efforts to develop better personalisation of learning for learners, providing a framework for the degree of automation in personalised learning systems. Her chapter discusses the shift from teacher-driven systems to computer-based technologies that can take a larger role in immediate decision-making, remaining within guidelines and goals specified by the teacher.

Next, educational technologies became more effective at personalising for differences in students' self-regulated learning – their ability to make good choices during learning that enhance their learning outcomes and efficiency. This topic is also discussed in Inge Molenaar's chapter (Molenaar, 2021[15]). Modern educational technologies in many cases have the ability to recognise when students are using ineffective or inefficient strategies, and to provide them recommendations or nudges to get back onto a more effective trajectory.

A contemporary trend, which is still primarily in research classrooms rather than wide-scale deployment, is the move towards also recognising and adapting to student engagement, affect, and emotion. Discussed by Sidney D'Mello (2021[16]), these systems recognise these aspects of a student's experience either from their interaction and behaviour within the system or from physical and physiological sensors. There are now several examples of educational technologies – particularly intelligent tutoring systems and games – which have been able to identify a student who is bored, frustrated, or gaming the system (trying to find strategies to complete materials without needing to learn) and re-engage them productively (e.g. DeFalco et al., 2017[14]).

Increasing research also looks at trying to personalise to increase broader motivation or interest. This work differs from the work on engagement and affect in terms of time-scale. Whereas engagement and affect often manifests in brief time periods – as short as a few seconds – motivation and interest are more long-term stable aspects of student experience. Work by Kizilcec and colleagues (Kizilcec et al., 2017[17]), for instance, has tried to connect student learning experiences with their values, leading to greater degrees of completion of online courses. Work by Walkington and colleagues (Walkington, 2013[18]; Walkington and Bernacki, 2019[19]) has modified the contents of learning systems to match student personal interests, leading students to work faster, become disengaged less often, and learn more.

New Pedagogies. Although the most obvious impact of artificially-intelligent educational technologies is through personalising learning directly, new pedagogies and teacher practices have also emerged. These pedagogies and practices enable teachers to support their students or provide their students with experiences in ways that were generally not feasible prior to the technology being developed.

Perhaps the largest shift has been in the information available to teachers. Dashboards provide teachers with data on a range of aspects of their students' performance and learning. This has produced a major shift in how homework is used. In the past, homework would need to be brought to class by students. It could be graded by the teacher after that (meaning that feedback and learning support would be delayed), or students could grade it with the teacher in a large group, which is not a very time-efficient approach. In contrast, data from homework technologies today can become available to teachers in real-time. This means that teachers can identify which students are struggling and which materials students struggled on in general before class even starts. This enables strategies where, for instance, teachers identify which students displayed common errors and can identify students who can demonstrate both incorrect and correct problem-solving strategies for whole-class discussion. It also enables teachers to message students who are behind in completing materials (or even in starting to work through materials), helping get the student back on track (Arnold and Pistilli, 2012[20]).

Similar uses are available for formative assessment systems, which are being increasingly used in contexts where students have high-stakes end-of-year examinations. These systems often go beyond teacher-designed homework in terms of their breadth and comprehensiveness of coverage of key skills and concepts. They are increasingly used by teachers to determine what topics to review with their classes as well as what types of supplemental supports to provide to specific students.

Box 2.1 Formative assessment systems

Formative assessment systems are increasingly used in K-12 education worldwide. The most widely used formative assessment systems, such as NWEA MAP (Finnerty, 2018[21]), present students with traditional multiple-choice items and measure straightforward mathematics and language arts competencies – essentially providing another test to students to complete, but one where their teachers will get useful data linked to competencies that will be seen on future standardised examination. A small number of emerging formative assessment systems assess more complicated constructs and/or embed assessment into more complex activities, such as games (Shute and Kim, 2014[22]).

Data from formative assessment systems can be used with platforms designed to provide lists of supplemental resources for specific skills, concepts, and topics. Especially post-COVID, both local education agencies, and regional and national governments have worked to develop platforms with supplemental learning resources for students and parents. However, right now, these platforms are generally not connected directly to formative assessment systems so the teacher or parent needs to look up the resources for a student struggling with a specific competency.

One concern about formative assessment systems is that time spent using a formative assessment system is time not spent learning – a loss of instructional time. For this reason, there has been a trend towards embedding formative assessment into personalised learning. Several widely-used personalised learning systems, such as MATHia, Mindspark, Reasoning Mind, and ASSISTments, provide teachers with formative assessment data on which competencies the student is missing (Feng and Heffernan, 2006[23]; Khachatryan et al., 2014[24]). This information is distilled from students' regular learning activity, avoiding a loss of instructional time.

There is also better information available to teachers on what is going on in their classes in real-time, an area discussed in detail by Pierre Dillenbourg (Dillenbourg, 2021[25]). Classroom analytics can provide the teacher with information on a range of aspects of class performance, from individual students' difficulties with material in real-time to the relative effectiveness of collaboration by different student groups. A teacher cannot watch every student (or every student group) at all times – better data can help them understand where to focus their efforts, and which students would benefit from a conversation right now.

Beyond just providing better data, it is possible to use technology to give students a range of experiences that were not feasible a generation ago. In their chapter, Tony Belpaeme and Fumihide Tanaka (Belpaeme and Tanaka, 2021[26]) discuss the new potentials of having robots interact with students in classrooms.

Using simulations and games in class can enable teachers to demonstrate complex and hard to understand systems to students. They can also allow students to explore and interact with these systems on their own. There seems to be particular educational benefit to the combination of a rich simulation or game experience that enables a student to develop informal, practical understanding, and then a teacher lecture or explanation that helps a student bridge from that informal understanding to more formal, academic conceptual understanding (Asbell-Clarke et al., 2020[27]). Modern technologies also offer new potentials for the use of collaborative learning, with systems that can scaffold effective collaboration strategies (Strauß and Rummel, 2020[28]), and systems that can provide rich experiences to collaborate around, such as interactive tabletops (Martinez Maldonado et al., 2012[29]).

Equity. New educational technologies are typically designed with the goal of improving student and teacher experiences and outcomes. However, the designers of these systems do not always consider how the full spectrum of learners are impacted. Often, systems are designed by members of specific demographic groups (typically higher socio-economic status, not identified as having special needs, and members of racial/ethnic/national majority groups) with members of their own groups in mind (not always intentionally). This can lead to lower educational effectiveness for members of other groups.

For example, Judith Good (2021[30]) discusses how there has been little effort to create educational technologies specifically designed for students with disabilities or special needs. She discusses examples of technologies that could support learners with autism, dysgraphia and visual impairment. The lack of attention to individuals with special needs by the scientific community and by developers of artificially intelligent educational technologies is a major source of inequity and a missed opportunity. Designing policies that facilitate developing systems to support learners with special needs (for instance, by developing approaches that improve access to data on disabilities while protecting student privacy) and the creation of incentives to develop for special needs populations may help to address this inequity.

Another key area of inequity is in support of historically underserved and underrepresented populations, including ethnic/racial minorities and linguistic minorities. Most educational technologies are developed by members of historically well-supported populations. They are often first piloted with members of historically well-supported populations. Testing for effectiveness with historically underrepresented populations often occurs only in later stages of development (or in final large-scale evaluations of efficacy) when it is too late to make major design changes. There is increasing evidence that both educational research design findings and algorithms obtained on majority populations can fail to apply or function more poorly for other populations of learners (Ocumpaugh et al., 2014[31]; Karumbaiah, Ocumpaugh and Baker, 2019[32]).

New educational technologies and approaches for educational systems

The benefits of modern educational technology – artificial intelligence and machine learning – goes beyond just supporting teaching and learning. There are a range of other ways that modern educational technology (not always artificially intelligent) benefits students and the schools supporting them in their educational journey. In a sense, this entire area is a focus of Dirk Ifenthaler's chapter (Ifenthaler, 2021[33]), but specific areas are also emphasised in other chapters.

Early Warning Systems. One of the major uses of predictive analytics in education is the creation of early warning systems. These systems, discussed in detail in Alex Bowers's chapter (Bowers, 2021[34]), attempt to predict in advance which students are at risk of a negative outcome – most frequently dropout and failure to graduate, but sometimes other outcomes such as failing a course as well. These predictions are often augmented with information on why a student is thought to be likely to have this negative outcome, such as poor grades in a specific class or an overly high number of disciplinary incidents.

The same types of data are also used in school-wide reporting systems for tracking student learning or disciplinary incidents. These school- (or district-) level dashboards give school and district leaders a broad picture of school climate and success. In the United States, these systems are increasingly provided to districts by a small number of vendors who combine expertise in ingesting school information system data (often in a disconnected set of different databases) with expertise in creating meaningful dashboards. Through these vendors, this type of artificially intelligent technology is made available for millions of students.

Reports for Parents. Many schools, school districts, and other local education agencies provide reports to parents on their students' progress. Often expanded over time from classic report cards, which simply provided a grade for each subject, these reports now provide a variety of information about learners to parents.

Admissions and School Allocations. Admissions and school allocations come down to the same process – determining if a student will be invited to attend a specific school or university – but from different perspectives. Admissions typically involve a decision made by a single school in an environment where students can be admitted to multiple schools; allocations typically involve a single centralised decision-making facility. Either way, algorithms are used to allocate limited resources (school/university placements) in line with institutional values, be they equity or selectivity. This is done in an increasing number of places – from charter school networks in the United States and public high-school systems in France to universities in Hungary.

Proctoring Systems. With intermittent (or continual) school closures occurring almost worldwide due to the COVID-19 pandemic, concerns about the security of examinations have emerged – for instance, that students will cheat by having another individual take a test for them, or that students will access unauthorised resources while taking a test. This has led to an explosion in proctoring systems where students will (for example) show a picture ID at the beginning of the test and keep their webcam on while taking it, and a proctor will watch a group of students' webcam feeds. Some proctoring systems also monitor other activity on students' computers during the duration of the test and what is going on in the room where students take the test, in many cases using artificial intelligence to supplement human proctors.

Box 2.2 Reports for parents

Increasingly, parents are provided with reports on their children's learning and activity in school. These reports for parents vary considerably in what information is provided. There is considerable variation in what data are provided to parents. Reports can range from very macro-scale reports (a student is at risk of dropping out of school or failing a course) to meso-scale reports (a student has 7 absences or is getting a C in mathematics) to micro-scale reports (the student answered "D" on problem 6, and here's why it was wrong).

These reports are provided in a variety of fashions. Many schools, school districts, and local education agencies still provide information to parents via physical letters. Text messages and phone calls are also used, particularly for warnings or reminders of various sorts. Some learning platforms and learning management systems provide web-based portals that parents can log into and look through. These platforms tend to provide relatively more data. For example, the ASSISTments platform (Broderick et al., 2011[35]) provides parents with data on which items the student has recently worked on, what their performance was, and what the correct answers were. The Edgenuity platform provides parents with data on how many minutes their student worked on each subject, and how much the student is behind pace or ahead of pace for the semester.

Though there is general agreement that providing data to parents is beneficial, there is debate as to how much parents look at the reports and data they have access to (Broderick et al., 2011[35]), and there has been research into developing reports that parents are more likely to use. When effectively designed, data reports can have positive impacts on parental engagement and student outcomes (Bergman and Chan, 2017[36]; Kraft and Monti-Nussbaum, 2017[37]).

Advances in Credentialing. Recent advances in computational technologies have led to advances in credentialing. Perhaps the most noteworthy advance in credentialing is the use in blockchain in education, discussed in Natalie Smolenski's chapter (Smolenski, 2021[38]). Blockchain offers a secure way to reduce credential fraud and streamline credential validation.

Recent major shifts in credentialing coming from the supply side (the availability of new credentials) make these new advances even more relevant. Increasing numbers of organisations offer certificates, such as the technical certificates available from Cisco, Microsoft, or CompTIA. In addition, massive online open course providers offer online courses and certificates developed in partnership with a range of universities. The ecosystem of massive online open course providers includes both large, international providers such as edX, Coursera, and FutureLearn, which partner with a range of universities worldwide to offer large numbers of courses. The ecosystem also includes a range of regional, national, or more specialised providers. This creates new use cases for credentials through blockchain.

Customer Relationship Management Systems. Customer relationship management systems, which originated outside of education, and were originally used for the purpose of sales, are now used in educational systems management as well. These systems track individuals' interactions with an educational institution over time – who they interacted with, and how they interacted. Some online universities and programmes integrate these systems with early warning systems to track how an at-risk student is supported. At universities where academic advisors regularly contact students, such as Southern New Hampshire University and Liberty University, an academic advisor might look to such a system to get a weekly update of the student's progress, look at past interactions between that student and their instructors or other academic advisors, and then track their own interaction with the student after calling them on the phone.

Resource Allocation and Planning. An increasing number of school districts and local education agencies now use algorithmic systems for estimating their future needs for equipment, staffing, and other resources. Systems of this nature are also used, often with support from consulting firms, to determine how resources such as government funding can be applied for and/or leveraged at the right time to fill future resource gaps.

Future potentials

Artificial intelligence has emerged as a powerful tool for improving education in recent years. The use of these technologies has expanded, albeit at different paces and in different ways for different technologies. Some technologies have expanded in use quickly, such as the explosion of early warning systems in the United States over the last few years, and some have expanded gradually, such as the slow and sometimes back-and-forth expansion of use of personalised learning technologies, class by class. Some technologies, particularly for studying classroom interaction and supporting classroom orchestration, have been slow to emerge from research classrooms and need greater support for the development of technology (and better ways to secure privacy) that enable their greater emergence.

This book discusses the many different applications and uses that have emerged for artificial intelligence in education. The text discusses them almost in isolation from each other – as separate and separable emerging trends. There is a reason for this: though they emerge from the same types of technology, they largely *have* been separate trends. They have emerged one by one, brought about by different stakeholders, with different goals, sometimes even in opposition to each other. For instance, the various forms of personalised learning technology have often competed with each other for sales and uptake, rather than trying to find ways to work in concert.

The result is a fragmented learning ecosystem. A school may use several different artificially intelligent technologies, but not together. A student may even use five or six different technologies within a single class, within the course of a semester. There are major costs to this lack of integration – multiple different learning technologies may each discover something about the student that the teacher already knows. Developing integration between AIED learning technologies, as called for in Baker (2019[39]), may reduce inefficiency and improve students' learning experiences.

Going further, if we can develop an ecosystem where various artificially intelligent technologies coordinate between each other and communicate information to teachers and other stakeholders, we can substantially improve student outcomes. Prediction of whether a student is at risk of dropping out will be facilitated through continual data on student use of personalised learning systems. Integrating formative assessments with classroom orchestration technologies will facilitate measurement of 21st-century skills while empowering teachers with real-time information on how they are developing. The possibilities are combinatorial – almost every possible pair of the technologies discussed in this chapter creates new opportunities when they are integrated together. There is the potential for the school of the future to move towards an integrated learning experience for students, where data is combined not just across learning platforms, but across every aspect of the learning experience. In this situation, teachers of different classes could coordinate to support each student's development of 21st-century skills, working in tandem with a variety of learning platforms to create an integrated, unified learning experience. A student struggling with seeking help, for instance, could be encouraged to do so (appropriately) in group activities in class within a personalised learning platform used by homework, and by an educational robot. Teachers could collaboratively review an integrated dashboard to understand the student's progress and its implications for his/her risk of dropping out of high school. The student's success at building this 21st-century skill could be assessed both formatively and summatively by assessment systems. Such a vision requires solving several challenges – perhaps the first is using policy to develop incentives to encourage the developers of these disparate systems to work together. Ultimately, the success of such a vision also requires the re-shaping of several systems – platform design, school practices, teacher professional development – to accommodate the opportunities that the new technology brings.

One driving force behind integrating information across technologies may be the increasing interest among teachers in high-quality data and reporting on student performance and progress. This trend was already building prior to 2020 but has amplified with the shift to home learning during the COVID-19 pandemic. Dashboards containing information on what students struggle with are already available for some advanced learning platforms, but are nowhere near universally deployed and available. This type of information becomes more essential to teachers when they cannot interact with students in person. As the demand for this type of functionality goes up, it creates opportunities to connect data sources and provide better information to teachers. It also creates a need for policy that increases support for teacher professional development in data-driven decision-making. Dashboards have generally been more widely-used by teachers with higher levels of data literacy. Fully capitalising on this opportunity will require improvements in the design of dashboards so that they are easier to understand by non-expert users, but also better support for teacher professional learning (both pre-service and in-service) in data literacy and learning to use data dashboards as part of instructional practice.

This shift matches broader shifts in the field's thinking on the uses of artificial intelligence and education as the technologies have matured. It is often hard to see how thinking has changed; the idea that AIED will completely replace teachers seems obviously wrong to most researchers and practitioners working in the field, and it can be difficult to believe that this view was once dominant. To illustrate, research on the design of teacher dashboards for AIED systems was rare as recently as 15 years ago (see Feng and Heffernan (2006[23])). This has been part of a broader shift from considering AIED systems as something a student uses to part of a broader ecosystem that also involves teachers, school leaders, and parents. Take, for example, the five visions for the future of artificially intelligent learning systems given 25 years ago by Shute and Psotka (1996[40]). Each of these five visions involved compelling, rich learning experiences. None of them involved teachers or parents (except as someone to say hello to on the way to a virtual reality cubicle). Increasingly, the field is aware that it does not have to choose between AI and humans as teachers; they can work together.

The next 20 years of changes to educational practice will be shaped in no small part by the greater uptake of artificial intelligence. For this shift to achieve its full potential, it will need to be driven not just by technology and research but in full partnership with teachers, school leaders, and the learners themselves. Putting the right policies in place can create a future for vision that matches – that exceeds – this book's optimistic vision (OECD, 2021[41]).

References

Arnold, K. and M. Pistilli (2012), "Course signals at Purdue", *Proceedings of the 2nd International Conference on Learning Analytics and Knowledge - LAK '12*, http://dx.doi.org/10.1145/2330601.2330666. [20]

Asbell-Clarke, J., E. Rowe, E. Bardar and T. Edwards (2020), "The Importance of Teacher Bridging in Game-Based Learning Classrooms", in *Global Perspectives on Gameful and Playful Teaching and Learning*, IGI Global. [27]

Baker, R. (2019), "Challenges for the Future of Educational Data Mining: The Baker Learning Analytics Prizes", *Journal of Educational Data Mining*, Vol. 11/1, pp. 1-17. [39]

Baker, R. (2016), "Stupid Tutoring Systems, Intelligent Humans", *International Journal of Artificial Intelligence in Education*, Vol. 26/2, pp. 600-614, http://dx.doi.org/10.1007/s40593-016-0105-0. [12]

Baker, R. and G. Siemens (2014), "Educational Data Mining and Learning Analytics", in Sawyer, R. (ed.), *The Cambridge Handbook of the Learning Sciences*, Cambridge University Press, Cambridge, 2nd Edition, http://dx.doi.org/10.1017/cbo9781139519526.016. [13]

Belpaeme, T. and F. Tanaka (2021), "Social Robots as educators", in *OECD Digital Education Outlook 2021: Pushing the frontiers with AI, blockchain, and robots*, OECD Publishing. [26]

Bergman, P. and E. Chan (2017), "Leveraging technology to engage parents at scale: Evidence from a randomized controlled trial", *Working paper.* [36]

Bowers, A. (2021), "Early warning systems and indicators of dropping out of upper secondary school: the emerging role of digital technologies", in *OECD Digital Education Outlook 2021: Pushing the frontiers with AI, blockchain, and robots*, OECD Publishing. [34]

Broderick, Z., C. O'Connor, C. Mulcahy and N.T. Heffernan (2011), "Increasing parent engagement in student learning using an intelligent tutoring system", *Journal of Interactive Learning Research*, Vol. 22/4, pp. 523-550. [35]

Carbonell, J. (1970), "AI in CAI: An artificial-intelligence approach to computer-assisted instruction", *IEEE transactions on man-machine systems*, Vol. 11/4, pp. 190-202. [7]

Cuban, L. (1986), *Teachers and Machines: The Classroom of Technology Since 1920*, Teachers College Press. [9]

D'Mello, S. (2021), "Improving Student Engagement in and with Digital Learning Technologies", in *OECD Digital Education Outlook 2021: Pushing the frontiers with AI, blockchain, and robots*, OECD Publishing. [16]

De Freitas, S. (2018), "Are games effective learning tools? A review of educational games", *Journal of Educational Technology & Society*, Vol. 21/2, pp. 74-84. [6]

De Jong, T. and W. Van Joolingen (1998), "Scientific Discovery Learning with Computer Simulations of Conceptual Domains", *Review of Educational Research*, Vol. 68/2, pp. 179-201, http://dx.doi.org/10.3102/00346543068002179. [3]

DeFalco, J., J.P. Rowe, L. Paquette, V. Georgoulas-Sherry, K. Brawner, B.W. Mott, R. Baker and J.C. Lester (2017), "Detecting and Addressing Frustration in a Serious Game for Military Training", I*nternational Journal of Artificial Intelligence in Education*, Vol. 28/2, pp. 152-193, http://dx.doi.org/10.1007/s40593-017-0152-1. [14]

Dillenbourg, P. (2021), "Classroom analytics: Zooming out from a pupil to a classroom", in *OECD Digital Education Outlook 2021: Pushing the frontiers with AI, blockchain, and robots*, OECD Publishing. [25]

Feng, M. and N. Heffernan (2006), "Informing teachers live about student learning: Reporting in the assistment system", *Technology Instruction Cognition and Learning*, Vol. 3/1/2, pp. 63. [23]

Finnerty, D. (2018), "The Predictive Utility and Longitudinal Student Growth of NWEA MAP Interim Assessments in Two Pennsylvania Schools", *Unpublished doctoral dissertation, Lehigh University.* [21]

Good, J. (2021), "Serving students with special needs better: How digital technology can help", in *OECD Digital Education Outlook 2021: Pushing the frontiers with AI, blockchain, and robots*, OECD Publishing. [30]

Ifenthaler, D. (2021), "Learning analytics for school and system management", in *OECD Digital Education Outlook 2021: Pushing the frontiers with AI, blockchain, and robots*, OECD Publishing. [33]

Karumbaiah, S., J. Ocumpaugh and R. Baker (2019), "The Influence of School Demographics on the Relationship Between Students' Help-Seeking Behavior and Performance and Motivational Measures", *Proceedings of the 12th International Conference on Educational Data Mining*, pp. 99-108. [32]

Khachatryan, G., A.V. Romashov, A.R. Khachatryan, S.J. Gaudino, J.M. Khachatryan, K.R. Guarian and N.V. Yufa (2014), "Reasoning Mind Genie 2: An intelligent tutoring system as a vehicle for international transfer of instructional methods in mathematics", *International Journal of Artificial Intelligence in Education*, Vol. 24/3, pp. 333-382. [24]

Kizilcec, R., A.J. Saltarelli, J. Reich and G.L. Cohen (2017), "Closing global achievement gaps in MOOCs", *Science*, Vol. 355/6322, pp. 251-252, http://dx.doi.org/10.1126/science.aag2063. [17]

Kraft, M. and M. Monti-Nussbaum (2017), "Can schools enable parents to prevent summer learning loss? A text-messaging field experiment to promote literacy skills", *The Annals of the American Academy of Political and Social Science*, Vol. 674/1, pp. 85-112. [37]

Martinez Maldonado, R., J. Kay, K. Yacef et B. Schwendimann (2012), "An Interactive Teacher's Dashboard for Monitoring Groups in a Multi-tabletop Learning Environment", in *Intelligent Tutoring Systems, Lecture Notes in Computer Science*, Springer Berlin Heidelberg, Berlin, Heidelberg, http://dx.doi.org/10.1007/978-3-642-30950-2_62. [29]

Milligan, C. and A. Littlejohn (2017), "Why Study on a MOOC? The Motives of Students and Professionals", *The International Review of Research in Open and Distributed Learning*, Vol. 18/2, pp. 92-102, http://dx.doi.org/10.19173/irrodl.v18i2.3033. [2]

Molenaar, I. (2021), "Personalisation of learning: Towards hybrid human-AI learning technologies", in *OECD Digital Education Outlook 2021: Pushing the frontiers with AI, blockchain, and robots*, OECD Publishing. [15]

Ocumpaugh, J., R. Baker, S. Gowda, N. Heffernan and C. Heffernan (2014), "Population validity for educational data mining models: A case study in affect detection", *British Journal of Educational Technology*, Vol. 45/3, pp. 487-501, http://dx.doi.org/10.1111/bjet.12156. [31]

OECD (2021), *Digital Education Outlook: Pushing the frontiers with AI, blockchain, and robots*, OECD Publishing. [41]

OECD (n.d.), *Education continuity stories from the coronavirus crisis*, https://oecdedutoday.com/coronavirus/continuity-stories/ (accessed on 24 May 2021). [10]

Rodrigo, M., N. Caluya and E. Cesar E. Vidal Jr (2015), "Igpaw: intramuros—design of an augmented reality game for philippine", *In: Proceedings of the 23rd International Conference on Computers in Education.* [4]

Shin, D. (2017), "The role of affordance in the experience of virtual reality learning: Technological and affective affordances in virtual reality", *Telematics and Informatics*, Vol. 34/8, pp. 1826-1836, http://dx.doi.org/10.1016/j.tele.2017.05.013. [5]

Shute, V. and Y. Kim (2014), "Formative and stealth assessment. In Handbook of research on educational communications and technology", in *Handbook of research on educational communications and technology*, Springer. [22]

Shute, V. and J. Psotka (1996), "Intelligent tutoring systems: Past, Present and Future", in *Handbook of Research on Educational Communications and Technology*, Scholastic Publications. [40]

Smolenski, N. (2021), "Blockchain for Education: A New Credentialing Ecosystem", in *OECD Digital Education Outlook 2021: Pushing the frontiers with AI, blockchain, and robots*, OECD Publishing. [38]

Stephenson, N. (1998), *The diamond age, Penguin: London*, UK. [8]

Strauß, S. and N. Rummel (2020), "Promoting interaction in online distance education: designing, implementing and supporting collaborative learning", *Information and Learning Sciences*, Vol. 121/5/6, pp. 251-260, http://dx.doi.org/10.1108/ils-04-2020-0090. [28]

VanLehn, K. (2011), "The Relative Effectiveness of Human Tutoring, Intelligent Tutoring Systems, and Other Tutoring Systems", *Educational Psychologist*, Vol. 46/4, pp. 197-221, http://dx.doi.org/10.1080/00461520.2011.611369. [1]

Walkington, C. (2013), "Using adaptive learning technologies to personalize instruction to student interests: The impact of relevant contexts on performance and learning outcomes.", *Journal of Educational Psychology*, Vol. 105/4, pp. 932-945, http://dx.doi.org/10.1037/a0031882. [18]

Walkington, C. and M. Bernacki (2019), "Personalizing Algebra to Students' Individual Interests in an Intelligent Tutoring System: Moderators of Impact", *International Journal of Artificial Intelligence in Education*, Vol. 29/1, pp. 58-88, http://dx.doi.org/10.1007/s40593-018-0168-1. [19]

Wenger, E. (1987), *Artificial intelligence and tutoring systems: computational and cognitive approaches to the communication of knowledge*, Morgan Kaufmann Publishers Inc. [11]

Part 1

3 Personalisation of learning: Towards hybrid human-AI learning technologies

Inge Molenaar
Faculty of Social Sciences, Radboud University, Netherlands

This chapter outlines the research and development of personalised learning in research labs and schools across OECD countries. The state of the art of personalised learning is described using a model of 6 levels of automation of personalised learning that articulates the roles of AI, teachers and learners. This describes how hybrid human-AI solutions combine the strengths of human and artificial intelligence to accomplish personalised learning. Existing learning technologies have a strong focus on diagnosing students' knowledge and adjusting feedback, tasks and/or the curriculum. Developmental frontiers lie in taking into account a broader range of learner characteristics such as self-regulation, motivation and emotion.

Introduction

There are multiple scenarios in which artificial intelligence (AI) can improve teaching and learning. A dialogue between researchers, entrepreneurs, policymakers and education professionals can make the most promising hybrid human-AI solutions available to the educational sector. This chapter will focus on how those solutions apply to the personalisation of learning. Personalised learning using technology refers to the trend in which education is increasingly adjusted to the needs of the individual learner (Aleven et al., 2016[1]). The underlying assumption is that each learner's talent can be optimised when the learning environment is adapted to the learner's needs (Corno, 2008[2]). Traditionally, in classrooms, all students follow the same curriculum, receive the same instruction, do the same tasks and largely receive similar feedback. This "industrial" model of education has been criticised widely (Robinson, 2010[3]) and technological developments have been proposed to support transformations towards more personalisation. However, even with the wide availability of learning technologies that can be adjusted based on a learner's need, such as Computer Assisted Instruction (CAI), Adaptive Learning Technologies (ALT), Intelligent Tutor Systems (ITS) and Educational Games, the uptake in schools has been slow (Tondeur et al., 2013[4]).

However, three recent developments have moved education systems closer to personalised learning over the last five years. First, the availability of one device per student in many contexts now allows the continuous use of technology in classrooms and further integration of technology into day-to-day school practices. Second, the power of data to support learning has become more articulated in the developing field of Learning Analytics (LA). Highlighting that data is not only useful within learning technologies, but also valuable when given directly to teachers and learners. Third, learning technologies that include Learning Analytics and Artificial Intelligence (AI) techniques have started to be used at scale in schools.

The current generation of learning technologies for personalisation mostly adapt to learners based on predictions of learner domain knowledge (Aleven et al., 2016[1]). Typically, these technologies adjust topics to study, problems to work on or feedback to answers given (Vanlehn, 2006[5]). However, in addition to personalisation based on predictions of learners' knowledge, a number of other learner characteristics such as emotion, motivation, metacognition and self-regulation can be used as input to attune to the individual learners' needs (Winne and Baker, 2013[6]) (D'Mello, 2021[7]).

In order to develop our thinking about the potential of learning analytics and AI in personalising and enriching education, this chapter applies 6 levels of automation defined by the car industry to the field of education. In this model, the transition of control between teacher and technology is articulated to build on the combined strength of human and artificial intelligence. This aligns with the hybrid intelligence perspective that emphasises the importance of human-AI interaction (Kamar, 2016[8]). The model positions the current state of the art with respect to personalisation of learning and supports the discussion of future AI and education scenarios. This is critical to envisioning future developments and articulating different levels of personalisation of learning with accompanying roles for AI, teachers and learners.

The chapter starts with an overview of the levels of automation of personalised learning. Based on this model, state-of-the-art personalisation of learning as developed in research labs across OECD countries is described. Next, real world use of learning technologies in schools is outlined with reference to cases of learning technologies used at scale in schools. Finally, the frontiers of personalisation of learning are discussed. In particular, applications using multiple data streams offer new ways to detect and diagnose this broader range of learner characteristics for advanced personalisation of learning. Despite the rapid influx of educational technologies used at scale, state-of-the-art technologies for advanced personalisation of learning are not often extensively used in schools. Technologies that fit in traditional school organisation models and uphold teachers' agency are implemented more often in schools. To accomplish broader deployment of state-of-the-art technology this chapter concludes with three recommendations for policy makers: *i) govern ethics by design, transparency and data; ii) improve learning technologies with public-private partnerships;* and *iii) engage teachers and education professionals in these transitions.*

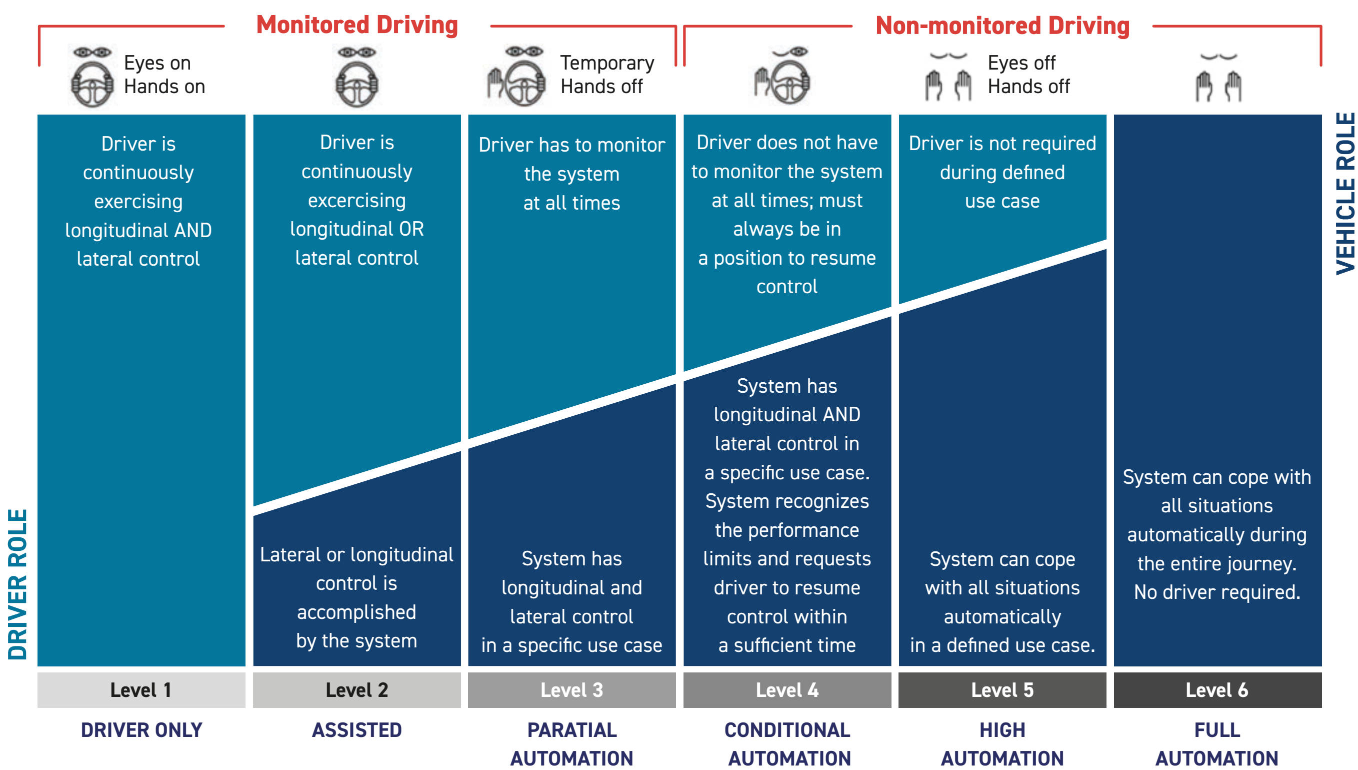
Hybrid human-AI systems: specifying the role of teachers and technology

With technologies increasingly gaining more data and intelligence, a new era of Human-AI interaction is emerging (Kamar, 2016[8]). Learning will increasingly be adjusted to individual learner characteristics. Many expect an ongoing trend in personalisation of learning (Holmes et al., 2018[9]). At the same time, there will be an ongoing fusion between human and artificial intelligence (AI) into so-called hybrid human-AI systems across many domains (Harari, 2018[10]). For example, even though self-driving cars are envisioned as eventually taking over driving from humans, currently they only assist human drivers (Awad et al., 2018[11]). Similarly, AI expert-systems support but do not replace doctors in medical decision-making (Topol, 2019[12]). A defining characteristic of a hybrid system is that the boundaries between AI and human decision-making fluctuate. Self-driving cars offload driving to the AI, but in situations that are too complex for the AI to navigate, control is transferred back to the human driver.

In order to distinguish the capabilities of self-driving cars needed for full automation, the Society of Automotive Engineers (SAE) (2016) has articulated 6 levels of automation of self-driving cars (Figure 3.1). These levels are based on earlier work that discusses different degrees of automation and consequences for human roles across different settings (Parasuraman, Sheridan and Wickens, 2000[13]). The 6 levels of automation highlight different stages in development towards an autonomous driving vehicle. At each level there is a further transition of control from the human driver to the self-driving technology. Moving up the levels, human control is reduced and the role of self-driving technology is increased until fully autonomous driving is achieved under all conditions. It is important to note that full automation may never be appropriate for particular domains such as education and medicine.

In the first three levels the human driver is in control, whereas in the last three levels control switches to the self-driving technology. In assisted driving (level 1), the self-driving technology provides supportive information to the driver. In partial automated driving (level 2), the self-driving technology controls driving in specific cases, for example on a highway under good weather conditions, but human drivers monitor the technology at all times. In contrast, in conditional automation (level 3), the self-driving technology takes over control but the driver should be ready to resume control at any time. The current state of the art for self-driving cars is between partial and conditional automation. In medicine this is estimated between assisted and partial automation (Topol, 2019[12]). Here, AI often supports medical decision-making, for example expert-systems detect tumours in x-rays, but human doctors still perform the final diagnosis of a patient and select the most appropriate treatment.

Figure 3.1 Six levels of automation towards the self-driving car



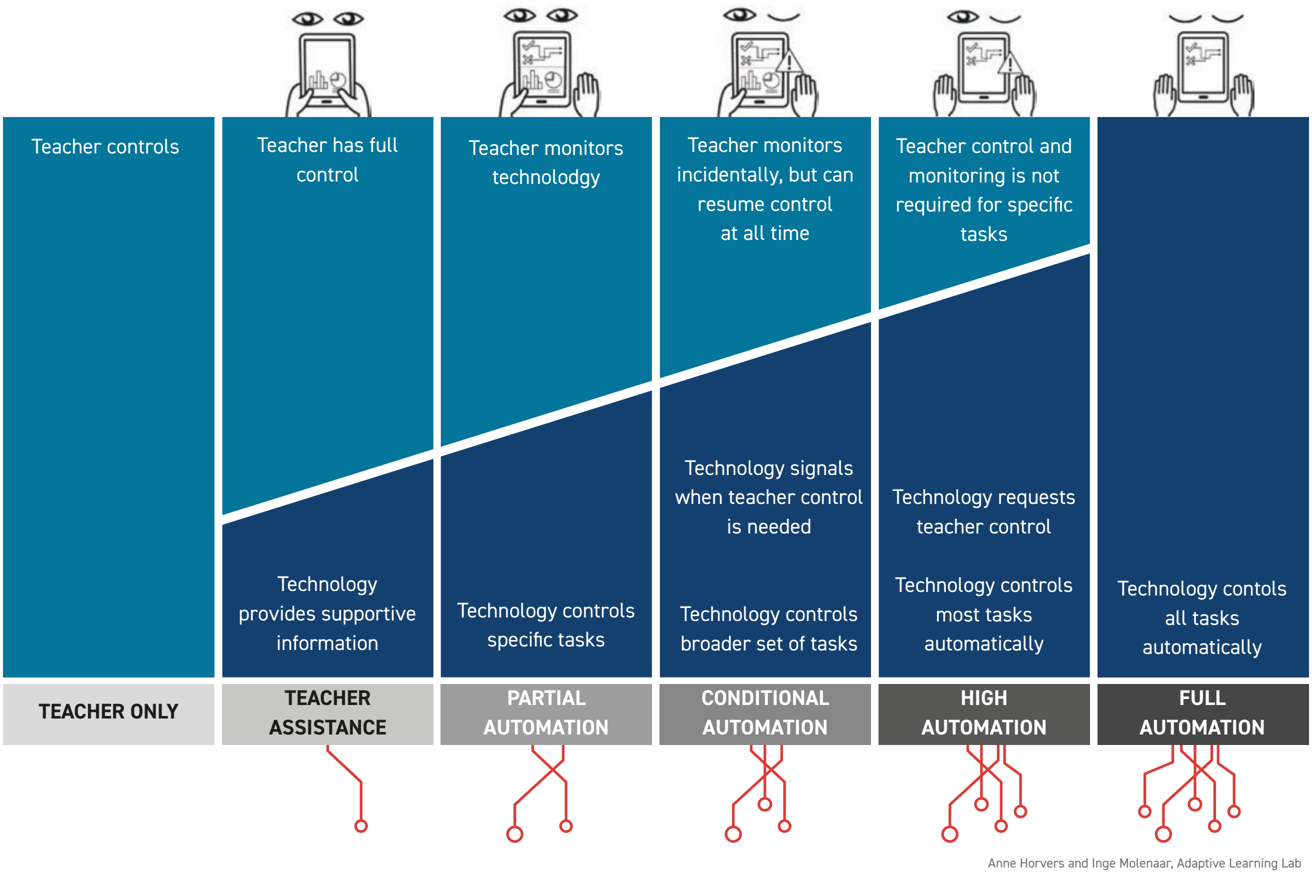
Source: Illustration - Mike Lemanski / ZF TRW; The Society of Automotive Engineers (https://www.sae.org/)

To the best of my knowledge, the levels of automation have not yet been translated into the field of education. This model helps to position current state-of-the-art learning technologies and how these are used in schools. The model may help us to understand the gap between state-of-the-art and day-to-day use of technologies in schools from the perspective of human control.

The six levels of automation applied to the field of educational technologies are outlined in Figure 3.2. The lines under the model represent the expectation that increasingly more data streams will be used in the transition towards full automation. These data streams can support more accurate detection and diagnosis of learners and their environment. At the top of the model, the level of human control is visualised across the levels. The hands on the tablet represent the level of teacher control. Full teacher control with two hands on the tablet, partial control with one hand and no hands symbolises no or incidental teacher control. The eyes represent the required level of teacher-monitoring, ranging from full, partial, incidental to no monitoring. The warning triangle indicates the ability of the AI to notify the teacher to resume control at critical moments. How the AI performs its actions using different types of data will be described in the next section.

In level 0 the teacher has full control over the learning environment and learning technologies have no organising function. This was the standard operating practice in most OECD countries until approximately 15 years ago.

Figure 3.2 Six levels of automation of personalised learning



Source: Illustration - Anne Horvers and Inge Molenaar, Adaptive Learning Lab
https://www.ru.nl/bsi/research/group-pages/adaptive-learning-lab-all/

In teacher assistance (level 1), teachers are in full control with the technology providing additional assistance to support the organisation of the learning environment. The technology provides access to learning materials and offers additional information about learners' activities. The technology does not control any aspects of the learning environment. For example, electronic learning environments and learning management systems that distribute learning materials are a form of teacher assistance. Teacher dashboards that provide teachers with concurrent information about students' activities, progress and performance are another form of teacher assistance (see also Dillenbour, 2021[14]). This information supports teachers to take pedagogical decisions before and during lessons (Bodily et al., 2018[15]). For example, a dashboard provides an overview of students' progress which helps teachers determine what instruction is appropriate in the next lesson. An overview of students' correct and incorrect answers helps teachers determine which students need additional feedback or extended instruction (Molenaar and Knoop-van Campen, 2019[16]). Other examples are systems that provide teachers with insights into student behaviours to help them judge which students need proactive remediation (Miller et al., 2015[17]). To sum up, the functions of learning technology in the teacher assistance level are to *support* teachers and *describe* and *mirror* learners' behaviour. Both control and monitoring are executed by the teacher. This is the current standard in OECD countries where IT solutions are increasingly being used in the classroom.

In partial automation (level 2), teachers give the technology control over specific organisational tasks. For example, Snappet (see Box 3.1) selects problems adjusted to the needs of individual students or provides feedback on a student's solution of a maths problem. When teachers allow the technology to take over these tasks, they can spend time on tasks that are beyond the reach of technology, such as providing elaborate feedback or helping students that need additional instruction. In partial automation the functions of learning technology are typified as *describe*, *diagnose*, *advise* and in specific cases *enact* actions. Hence at this level, teachers control most organisational tasks in the learning environment with a few exceptions where the technology takes over control. Teachers still completely monitor the functioning of the technology in which teacher dashboards often play an important role.

In level 3 conditional automation, technology takes control over a broader set of tasks in organising the learning environment. Teachers continue to hold a central position in organising the learning environment and they monitor how the technologies function. For example, Cognitive Tutor (see Box 3.2) selects problems and gives feedback on each problem-solving step as students' progress (Koedinger and Corbett, 2006[18]).The technology takes over both selection of problems and provision of feedback. It is important in this level that the technology recognises under what conditions it functions effectively and when teachers need to resume control. At these critical moments the technology notifies the teacher to take action. For example, when a student is not progressing at the anticipated speed, the technology notifies the teacher to step in Holstein, McLaren and Aleven (2017[19]). The teacher can then investigate why the student is not making sufficient progress and take remedial action. The functions of the learning technology in conditional automation are to enact a wider range of interventions in the learning environment and *notify* and *advise* if human action is needed.

In high automation (level 4), the technology takes full control over organising the learning environment within a specific domain. The technology deals effectively with most anticipated situations within the domain and it successfully addresses diversity among learners within the domain. Teachers do not generally need to be prepared to take over control and monitoring is not required. In specific specialised scenarios the technology reverts control or monitoring to teachers. This level may be reached by autonomous learning systems that are dedicated to supporting individual learners in domains such as maths, science and language learning. For example, MathSpring is an intelligent tutoring system that guides the learner in selecting learning goals and offers personalised instruction, practice opportunities and feedback driving towards the learning goals (Arroyo et al., 2014[20]). Within the scope of the technology the teacher does not execute any control or monitoring tasks. It is important to realise that such highly autonomous learning technologies remain uncommon and still only support a restricted part of a curriculum. The functions of the learning technology in high automation are to steer learners and only in exceptional cases *notify* other humans to take action.

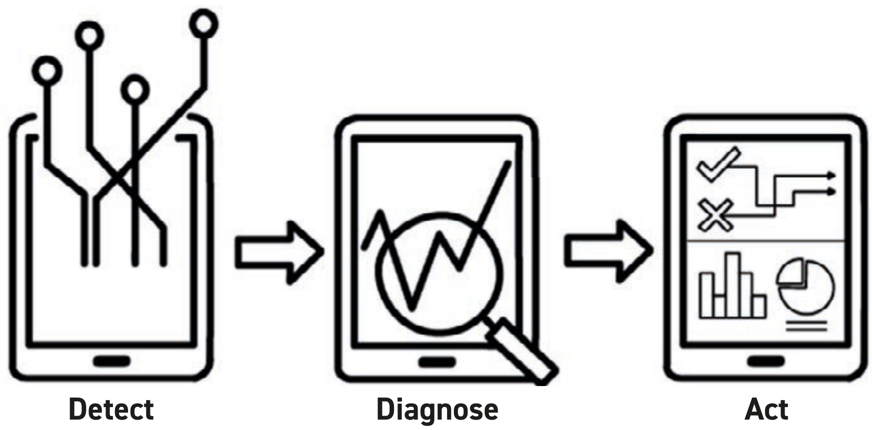
In level 5 full control, technology copes with learning across domains and situations automatically. The role of the teachers is completely taken over by the technology. For example, Alelo language-learning technology[1] that supports second language learning may evolve in that direction. It currently already uses natural language processing which analyses students' language usage, provides feedback, selects new learning goals and adjusts instruction and practice to improve the students' language level. These features could constitute the first steps towards fully automated learning. In conditions of full control, the technology would steer all learning of the learner. This level represents the paradigm for many learning solutions developed for the informal consumer learning market outside of the classroom and schools, from language learning through to music education or preparing a driving theory test. An important question is how viable the application of such technologies would be within the context of schools and what this would mean for teachers' responsibility and justification.

Speculations for the future: the ultimate role of AI

In addition to articulating the roles of teacher and AI, the 6 levels of automation help us to discuss expectations about the ultimate role of AI. It is generally accepted that even for self-driving cars, it is highly unlikely that full automation across all roads and conditions will ever be attained (Shladover, 2018[21]). Restrictions are found at three levels: i) tracking of the environment with sensors, ii) accurate diagnosis of risks and iii) determining appropriate actions. Especially with respect to the last step, research has highlighted the complexity of determining the appropriate action in sensitive situations (Awad et al., 2018[11]). Similarly, when it comes to the role of AI in medical decision-making, it is unlikely that developments will ever surpass the level of conditional automation (Topol, 2019[12]). Still, the ability to process massive sets of data quickly, accurately and inexpensively, plus the ability to detect and diagnose beyond human ability are put forward as a foundation for high-performance medicine.

These elements will also be foundational for high-performance education for similar reasons (see Figure 3.3). Traditionally the dominant template for AI in education has been full automation, that is, one student learning with a smart computer and teachers being replaced by technology (Blikstein, 2018[22]). This template emerged in the 1950s in Skinner's learning machines (Wingo, 1961[23]) and was further developed in intelligent tutor systems. Research showed that when a student is tutored individually by an expert tutor, learning performance can improve by two standard deviations over the situation where students learn under conventional classroom conditions (Bloom, 1984[24]). This means that 98% of students would perform better receiving one-to-one human tutoring than traditional classroom teaching. This "*two sigma problem*" has been used as a dominant argument for personalisation of learning with learning technologies ever since. Central to this analogy is the assumption that technology has the same ability to understand a learner as a human tutor. Although VanLehn's re-analysis of Bloom's data suggests a smaller benefit of human tutors (around 0.75 standard deviation), he also showed that well-developed Intelligent Tutoring Systems can be equally effective as human tutors (VanLEHN, 2011[25]). In specific cases technology does indeed resemble the abilities of good human tutors (VanLEHN, 2011[25]), but this is still limited to confined and specific domains. It is also important to note that the goals of education go well beyond the tutoring of domain knowledge. This raises the question of whether the aim is to optimise technology towards full automation or to optimise human-AI solutions in other levels of automation. Part of the answer lies in our abilities to detect, diagnose and act. Such advancements in personalisation of learning are, like advancements towards self-driving cars, dependent on three critical elements: i) our ability to follow learners and track their environment; ii) our ability to diagnose learners' current states and anticipate their development; and iii) our ability to consequently determine the most appropriate action to optimise learning. I will now discuss our current ability at each level.

Figure 3.3 Three challenges for high-performance education with AI



Source: Illustration - Anne Horvers and Inge Molenaar, Source: Adaptive Learning Lab, https://www.ru.nl/bsi/research/group-pages/adaptive-learning-lab-all

Detect: follow learners and track their environment

Our ability to track learners and their environment is improving progressively (Baker and Inventado, 2014[26]; Winne and Baker, 2013[6]). Learners differ widely and these differences are considered meaningful indicators for personalisation of learning (Azevedo and Gašević, D., 2019[27]). Traditionally there has been a strong focus in research on using log data from learning technologies to track learners. Although this is still an important focus in the field of learning personalisation, different data sources are increasingly used to understand important learner characteristics. These multimodal data sources can be conceptualised as physiological, behavioural and contextual data.

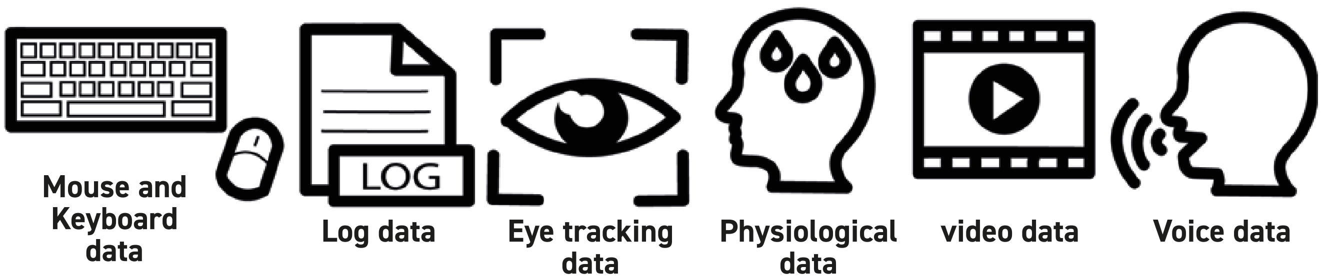
Physiological data indicate students' bodily responses during learning. Examples are heart rate (IBI), Electro Dermal Activity (EDA), Blood Volume Pulse (BVP), skin temperature and face capture software to assess learners' states.

Behavioural data detect aspects of students' behaviour during learning. One important source of behavioural data are log files. These data list the sequences of learner-technology interactions at the millisecond level leaving a trail of activities with learning technologies. Another source of behavioural data are mouse movement and keyboard entries. Eye movements indicate what students look at during learning and can be used to detect allocation of attention, viewing of multimedia sources and reading within learning technologies (Mudrick, Azevedo and Taub, 2019[28]). Wearable eye trackers also assess students' interaction with physical objects and social interaction during learning. In addition, specific eye-tracking data such as pupil dilation and blinking behaviour have been correlated to cognitive load and affective states.

Contextual data provide more holistic data from the learner's interactions with the learning technology, humans and resources in the learning environment. Both voice and video recordings hold rich data on how learners interact with their environment. Although these data allow for in-depth analysis of learning processes, they are mostly reliant on researchers' coding, scoring, annotation, understanding, and interpretation of the data. Progress in AI techniques may radically change this in the near future and first examples are available (Stewart et al., 2019[29]). Finally, self-reporting should not be disregarded as a valid source of data.

To sum up, multimodal data sources can be used to advance tracking of learners and their environment which is a critical step towards further automation.

Figure 3.4 Multimodal data source to track learners and their environment



Source: Illustration - Anne Horvers and Inge Molenaar, Source: Adaptive Learning Lab

Diagnosis: assessing learners' current states

The next step is to analyse the data in order to diagnose learners' current states and anticipate future development. Most work has been directed at diagnosing students' knowledge and the growth thereof during learning. Several models have been developed to assess learners' knowledge during learning based on students' problem-solving behaviour and answers to problems (Desmarais and Baker, 2011[30]). A detailed description follows in the next section. Increasingly, work is also being directed at assessing a broader range of learners' states such as motivation, metacognition, and emotion (Bosch et al., 2015[31]). Most of this work is done in structured domains such as maths and physics in which students give clear answers to well-specified problems. We also witness a growing development towards domain and context-specific diagnosis.

Advances in specific AI techniques allow the assessment of features that are critical in learning and developmental processes. For example, advances in automated speech recognition allow the ongoing detection of how young students learn to read . Analysis of a child's verbalisations during reading detected by automated speech recognition allows the extraction of specific features that aid the diagnosis of the child's reading ability, such as: the letters a child recognises correctly; the speed at which the child decodes words of different lengths; and the type of words that have been automated. Based on these extracted features, the development of a child's reading ability can be diagnosed accurately over time and can consequently be used to personalise teaching. In addition to speech recognition (text to speech), the first diagnostic reading software using eye-tracking data to diagnose learners' reading development is also available (Lexplore[3]).

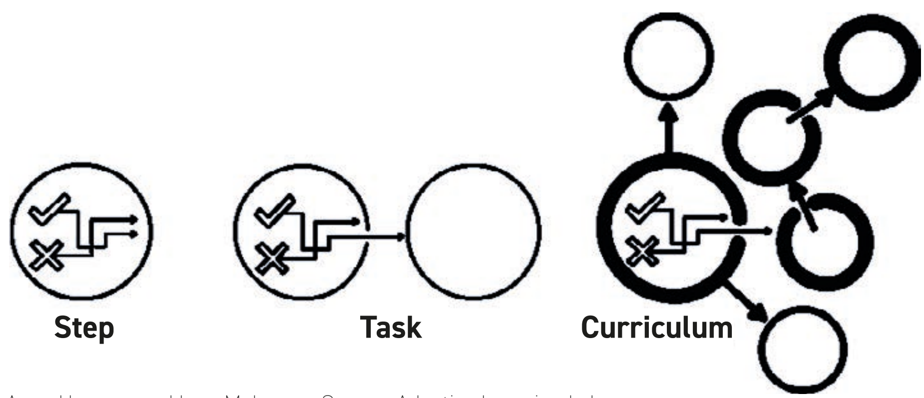
In a similar vein, the development of writing skills and even motor impairments such as dysgraphia can be diagnosed (Asselborn et al., 2018[32]). Asselborn's method measures a student's writing skills using a tablet computer and a digital pen. Important features related to the developmental trajectory of writing are detected when the child learns how to write, such as writing dynamics, pen pressure, and pen tilt. The method can extract as many as 56 features related to a student's writing skills which can consequently be translated into tutoring for appropriate writing practices. Other examples of domain-specific diagnosis mechanisms are diagnosis of dyslexia based on the type of mistakes a student makes on a specific problem-set (Dytective[4]), diagnosis of language development (Lingvist[5]) and functional writing skills (Letrus[6]).

To sum up, development of domain-generic and -specific diagnostic tools based on advanced AI techniques helps to further our ability to understand learner states and anticipate future developments. This is an important second step towards more advanced levels of automation in learning technologies.

Act: select appropriate actions

The last step is to interpret the diagnoses of learner states and translate them into actions that optimise learning (Shute and Zapata-Rivera, 2012). This translation of diagnosis into meaningful pedagogical actions is arguably the most complex step (Koedinger, Booth and Klahr, 2013[33]). As described by the Knowledge-Learning-Instruction Framework (Koedinger, Corbett and Perfetti, 2012[34]), there are endless possible response patterns and only limited evidence as to which interventions are most effective under particular circumstances. Even though advanced large-scale educational research into effective interventions is now possible within learning technologies, constraints on research funding prevent the field from taking full advantage of these opportunities. Acknowledgment of this complexity has also inspired the development of analytics that provide teachers with additional information to determine viable pedagogical responses. For example, providing teachers with dashboard with concurrent information about their learners activities, correct answers and progress helps teacher to improve feedback practices and learning outcomes (Knoop-van Campen and Molenaar, 2020[35]; Holstein, McLaren and Aleven, 2018[36]). Teacher dashboards have turned out to be an effective intervention in itself and can help to determine effective interventions for execution by learning technologies. The distinction between extracted and embedded analytics is an important step in developing more advanced response patterns that incorporate the pedagogical knowledge and skills of teachers and acknowledge the need to develop more advanced understanding of these relations within research (Wise, 2014[37]).

Figure 3.5 Three types of interventions in domain independent actions



Source: Illustration - Anne Horvers and Inge Molenaar, Source: Adaptive Learning Lab

When domain-generic adjustments are made within learning technologies they typically enact actions of three types, see Figure 3.5 (Vanlehn, 2006[5]). The first type is the step type, in which feedback is customised to the learner needs within problems. Learners receive elaborated feedback indicating how to proceed correctly within a particular step of a problem-solving task. The second type is the task type, in which the best next task or problem is selected based on a student's performance. Here students are provided with a problem fitting to their current knowledge base, based on their answers to the previous problem. The third type is the curriculum type, in which the organisation of instructional topics is adjusted to the learner. This entails in-depth selection of topics best suited to the developmental trajectory of students. In contrast, domain-specific adjustments follow a logic that is customised to the domain, such as the reading example described above where adjustments are driven by extensive knowledge about the development of reading.

To conclude, detection and diagnosis are enacted in different pedagogical actions with the learning technology, making adjustments of step, task or curriculum type. As well as direct enactment, there is also the option to feed the diagnosis back to the teacher to determine effective pedagogical interventions. This movement reflects the awareness that evidence-based interventions are essential to make progress towards more advanced automation in the field of education. As indicated above, most work in the field has focused on detection, diagnosis and actions based on students' knowledge. The next section will discuss the current state of the art in the field.

Personalisation based on students' knowledge

As personalisation based on learner knowledge and skills is the most dominant approach, I will continue to discuss detection, diagnosis and enactment of personalisation based on learners' knowledge and skills. Most research has been done on detection from log data, diagnosing students' knowledge and development of knowledge during learning, and translating this into pedagogical actions. As indicated above, based on the diagnosis of students' knowledge, adjustments are typically made at the step level, task level or curriculum level.

Adjusting problems at the task level

Historically, adjustments at the task level were made based on the assessment of students' knowledge prior to learning. For example, drill and practice solutions that are used for remedial purposes, i.e. helping students to overcome particular gaps in their knowledge. Although *Computer Assisted Instruction* (CAI) programs were found to improve learning by using performance on a formative test to determine a pre-set selection of problems that fit the learner's knowledge level, these programs lack the flexibility to adjust to changes in learners' knowledge as they learn (Dede, 1986[38]).

In order to resolve this limitation, research turned to detecting students' knowledge growth *during* learning. Technologies detect knowledge growth *during* learning using log data, such as correctness of answers and response time (Koedinger and Corbett, 2006[18]). This development was inspired by the theoretical assumptions of *mastery learning* that every student has individual needs with respect to instruction, timing of support and type of support (Bloom, 1984[24]). The first step was to adjust the time allowed to individual students to learn a particular topic (Corbett and Anderson, 1995[39]). The rationale is that before continuing to a new topic, the learner should learn the prerequisite topics. Based on a student's answers, technologies determine when a learner knows a certain topic well enough to proceed to the next topic. In this way, learning time is personalised for the individual learner to work on a topic long enough to learn it.

In addition to individualising learning time spent on a topic, learning technologies can also adjust each problem to a student's knowledge. For example, the adaptive learning technology *mathgarden* estimates a student's current knowledge to select the most appropriate next mathematical task for that learner (Klinkenberg, Straatemeier and van der Maas, 2011[40]). In this way, not only the time needed to master a topic but also the rate at which a student is learning it are personalised. These learning technologies match problems to a detected student knowledge level. This entails two elements: i) all problems are rated for their difficulty and ii) students' knowledge is assessed based on their answers to those problems. This allows the technology to follow a matching approach in which it selects problems at the same relative difficulty level for each student. These technologies are currently used at scale in a number of OECD countries (see Box 3.1).

Adjusting feedback at the step level

In addition to adjustment at the task level, a set of technologies makes adjustments at the step level. For example, complicated maths problems are solved in multiple steps. In structured domains such as maths, physics and chemistry, these steps can be related to specific topics they address. Based on this information, algorithms can not only detect students' current knowledge level, but can also analyse the type of errors students make (Baker, Corbett and Aleven, 2008[41]). Two types of errors are distinguished: i) slips when the student knows the answer but for instance reverses the numbers, and ii) mistakes when the students hold misconceptions that drive incorrect answers This distinction is important to determine which feedback is appropriate. Based on the type of mistake made, the technology can adjust feedback to the needs of the individual student or suggest problem-solving steps that can resolve the misunderstanding. For example, after finishing the first part of a calculation a student is given feedback on the correctness of the steps taken and the answers given. A number of advanced technologies, often referred to as *Intelligent Tutor Systems* (ITS), provide personalised feedback to students within a task or a problem (VanLEHN, 2011[25]). These systems investigate a student's response to the task to provide automated, elaborate feedback at each step the student takes. In this way the systems support the student in problem solving at the step level (see Box 3.2).

Box 3.1 Snappet: An example of Adaptive Learning Technology used at scale in the Netherlands

Detect:

- Students' answers to problems within a topic and with a certain difficulty level

Diagnose:

- Student knowledge growth during learning with the ELO algorithm
- Predictive analytics on student development over time

Act:

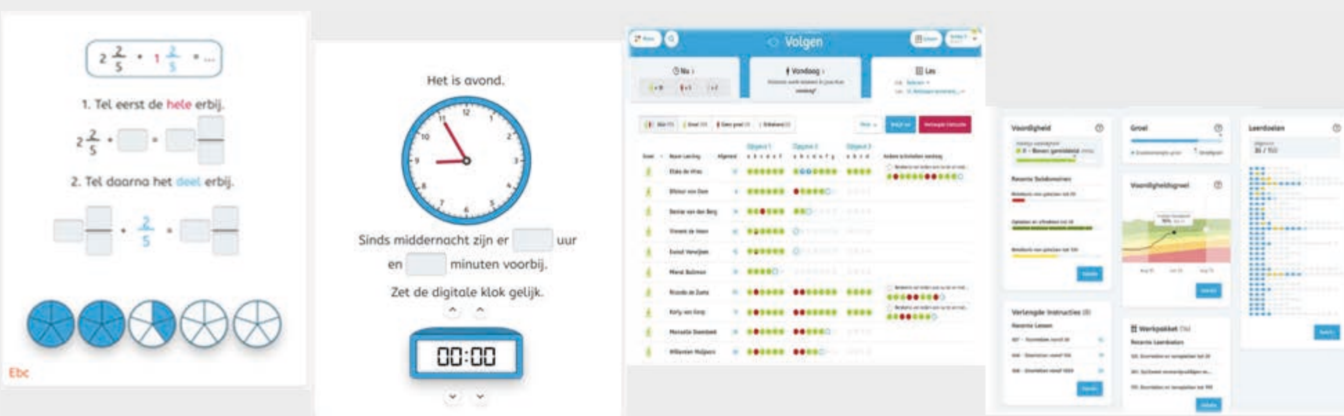
Partial automation:

- Provides direct feedback on student answers (step level)
- Selects problems adjusted to the knowledge of the student (task level)
- Determines when to continue on a next topic based on predictive analytics (curriculum level)

Teacher Assistance:

- Provides real-time teacher dashboards with progress and performance information at class level
- Provides real-time teacher dashboards with an overview of student development and predictive analytics
- Provides students with dashboards with progress information

Figure 3.6 Problems and dashboards in Snappet



Source: Illustration – Snappet (n.d.[42])

Usage information

Students: Primary education from grade 1 to 6

Subjects: Arithmetic, Mathematics, Dutch Spelling and Grammar, Reading comprehension

Scale: 2800 schools (45% of all primary schools in the Netherlands) and 1000 school in Spain

Effects: improved maths results after 6 months (Faber, Luyten and Visscher, 2017[43]), improved maths and marginally improved spelling after one year (Molenaar and van Campen, 2016[44]), diverse findings over multiple years of implementation (Molenaar, Knoop-van Campen and Hasselman, 2017[45]) in comparison to similar paper-based methods.

Box 3.2 Cognitive tutor, MATHia: an example of widely adopted adaptive learning technology in mathematics (from the United States)

Detect:

- Students' answers to problems and steps within maths problems

Diagnose:

- Student knowledge growth during learning with the Bayesian Knowledge Tracing algorithm
- Students errors on problems and steps within maths problems

Act:

Conditional automation:

- Provides direct elaborate automated feedback on student's answers to a problem (step level)
- Reveals next steps based on student's answers (step level)
- Determines when a student has mastered a topic (curriculum level)

Teacher Assistance:

- Reports to support shifts in pedagogy
- Provides planning reports for teachers and administrators as well as real-time dashboard (LiveLab) for teachers.

Figure 3.7 Problems and dashboards in MATHia

Source: Illustration - Carnegie Learning (n.d.[46])

Usage information

Students: students from grades 6-12

Subjects: Mathematics

Scale: 600k students

Effects: improved learning outcomes in maths (Koedinger and Corbett, 2006[18]; Pane et al., 2010[47]; Pane et al., 2014[48]; Ritter et al., 2007[49]) in comparison to both paper-based and different interventions in ITS.

Adjusting units at the curriculum level

Finally, optimisation at the curriculum level refers to the order in which a student works on different topics. The order is also called the learning pathway. Here the technology aim to build an overview of a learner's knowledge and skills to further inform future learning decisions (Falmagne et al., 2006[50]). These systems determine the order in which learners can best address a topic within a domain. Different pathways through the topics are interrelated based on the order in which topics are mastered by the students and in this way the order in which a student learns within a domain is personalised. An example of learning technology using this is ALEKS[7] which adjusts selections of units for students based on their previous performance.

Another important element that can be adjusted at the curriculum level is when a topic is restudied. The *spacing principle* indicates that there is an optimal time between learning particular topics. This time can be determined by the rate at which the learner forgets (Pavlik and Anderson, 2005[51]). Modelling this forgetting curve is a way to determine when a learner should repeat a particular topic (Pashler et al., 2007[52]). Especially for word learning this had been an important indictor to improve retention of new vocabulary. An example of such an application in a learning technology is WRTS[8] which uses an algorithm to determine the individual forgetting curves for a student when learning new vocabulary for a foreign language. Based on each student' individual forgetting curve, the system determines the interval at which words are repeated in practice problems. Both examples allow for customised pathways and repetition patterns for each student that would not be possible without the algorithm. Finally, answers of students can be used to diagnose misconceptions, which can be addressed at the step level (see Mathia example in Box 3.2), but also at the curriculum level. In that case the system makes an analysis of the errors during practice, which are translated into learning gaps that can be addressed in additional practice units. An example of such a system is Mindspark described in Box 3.3.

Below, I will discuss how these state-of-the-art technologies can be positioned in the levels of automation model.

Positioning current learning technologies in the levels of the automation model

As shown above, progress is being made in detecting and diagnosing students' knowledge and skills during learning, which is enacted most often at the task, problem and curriculum level. When we position existing learning technologies in the levels of the automation model, they mostly fall under the first three levels of automation. First, at the level of teacher assistance, technologies assist teachers with teacher dashboards. Second, at the level of partial automation, diagnosis of student knowledge and knowledge growth are used to make adjustments at the step, task or curriculum level. This results in automation of feedback, problem selection and optimisation of the curriculum. When these levels are applied separately this can be considered as examples of partial automation, but combined these solutions move towards conditional automation.

OECD schools are increasingly using technology in class to support the teaching and learning of maths, science and reading. This takes different forms, including the increasing use of computers to look for information, to practise skills in maths, science and reading, etc. (Vincent-Lancrin et al., 2019[54]). A range of adaptive learning technologies (ALT) and Intelligent Tutoring Systems (ITS) have become part of the daily routines in classrooms to personalise learning of foundation skills such as mathematics, grammar and spelling in many primary and secondary schools. Typically, young learners practise language or maths problems on a tablet or computer, while these technologies capture rich data on learners' performance. The current generation of technologies uses learners' data to adapt problems to learners' predicted knowledge levels and at the same time provide additional information on their progress in teachers' dashboards. This use of technology allows more efficient teaching of foundational skills (Faber, Luyten and Visscher, 2017[43]) and, in principle, frees up time to work on more complex skills such as problem-solving, self-regulation and creativity.

Most Adaptive Learning Technologies (ALT) only adjust at the task level and are therefore examples of partial automation. Intelligent Tutoring Systems (ITS) are an exception which can be considered as examples of conditional automation. These technologies take control over a broader set of elements in organising the learning environment. Most intelligent tutor systems enact adjustments at the task and step level but few solutions take over control at three levels. Moreover, while intelligent tutor systems are successfully applied within structured domains, such as maths and science, few solutions exist for unstructured domains (VanLehn, 2015[55]). Although intelligent tutor systems have been proven to increase learning outcomes and optimise learning (Kulik and Fletcher, 2016[56]), the uptake in schools has been limited (Baker, 2016[57]).

Box 3.3 Mindspark: An example of adaptive learning technology from India

Detect:

- Students take a diagnostic test to assess their knowledge and misconceptions on a topic
- Students answer questions designed to identify the type of misconception where knowledge gaps are detected within a topic

Diagnose:

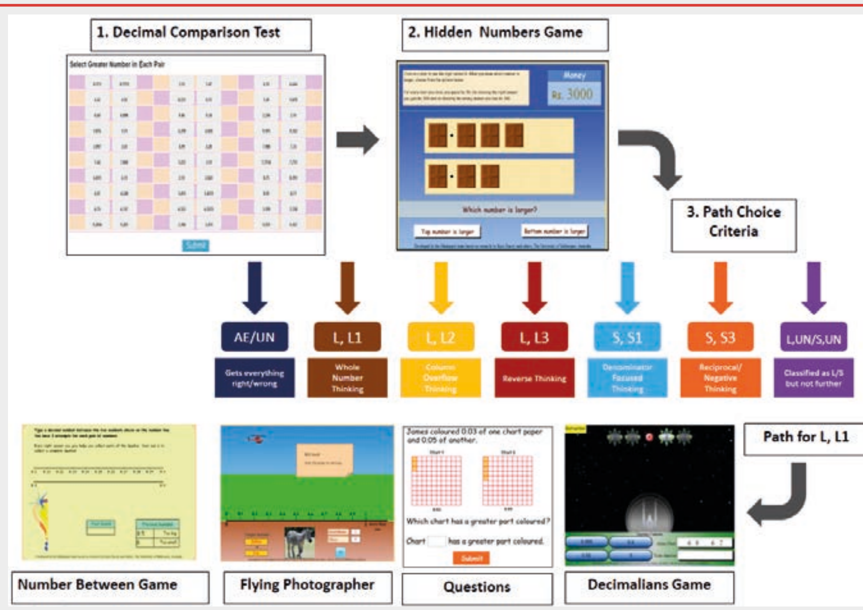
- Patterns of errors in student responses within a topic
- Assigns targeted remedial instruction based on conceptual bottleneck detected

Act:

Conditional automation:

- Provides diverse instructional materials (questions, interactive games, activities) created specifically to address misconceptions and learning gaps diagnosed (task level)
- Timely and specific feedback when a question is answered incorrectly (step level)
- Determines whether the student can move on to the next difficulty level (task or curriculum level)

Figure 3.8 Personalised remedial activities to learn decimals based on error patterns



Note: Illustration of adaptability based on error patterns: A child could be assigned to the "Decimal Comparison Test". If she gets most questions in that test correctly, she is assigned to the "Hidden Numbers Game", a slightly harder exercise. However, if she answers most questions incorrectly, she is assigned a remedial activity depending on her type of misconception. For example, if her errors stem from "Whole Number Thinking" (e.g. believing that 3.27 is greater than 3.3 because she compares the numbers to the right of the decimal point and concludes (incorrectly) that since 27 is greater than 3, 3.27 is greater than 3.3), she will be assigned the "Decimalians Game" designed to address this type of misconception. Developed by the Mindspark team based on research by Kaye Stacey and others, University of Melbourne.

Source: Muralidharan, Singh and Ganimian (2019[53]), Figure D.1

Usage information

Students: students from grades 1 to 10

Subjects: Mathematics, Hindi, English

Effects: after-school programme combining Mindspark and group instruction improved learning outcomes (Muralidharan, Singh and Ganimian, 2019[53]).

In this light, it is important to note two critical aspects for technology uptake by practice. First, teachers often feel that they are being pushed out-of-the-loop by these technologies (Holstein, McLaren and Aleven, 2019[58]) which may explain their limited take-up in practice (Baker, 2016[57]). An important step towards broader acceptance of these technologies would be to effectively address the limits of their functioning. Although dashboard functions have been developed in these systems for longer, there has been little research into how to make these dashboards useful for teachers. Only recently have dashboards been developed that aim to include teachers in the loop (Feng and Heffernan, 2005[59]; Holstein, McLaren and Aleven, 2017[19]). Although these developments are still in an early stage, first evidence shows this a powerful new direct to include teachers in the loop (Knoop-van Campen and Molenaar, 2020[35]).

Second, it is important to note that learning in most schools is organised around groups, and technologies directed at individuals do not fit easily into this context. Moreover, there are strong theoretical indications that learning is a social process in which students learn together. For this reason, solutions are developed that combine personalisation at the individual and class levels. In these contexts, personalisation of learning at the individual level is integrated into a class-paced curriculum. Technologies that are configured in this way can be implemented within the current context of schools without demanding extensive reorganisation. Uptake of these technologies may be easier than the uptake of solely individual-paced technologies.

Personalisation of learning based on self-regulated learning

In addition to personalisation based on student knowledge, there are a wide range of learner characteristics that could support more advanced forms of personalisation. Although these developments are still in progress, the first prototypes are being tested in labs across the OECD countries.

In the past few years, developments have shifted the focus to other learner characteristics beyond learner knowledge and skills, such as the ability to self-regulate learning, to apply metacognitive skills, to control and monitor learning activities, to motivate one-self to put in enough effort to learn and to regulate emotional responses. These are all considered potential input for personalisation (Bannert et al., 2017[60]; Järvelä and Bannert, 2019[61]). This wider focus on learner characteristics and behaviour during learning forms a natural fit with the central role of data to understanding learners (Azevedo and Gašević, D., 2019[27]). This move towards viewing the whole learner is based on research indicating that self-regulated learning (SRL), motivation and emotion play an important role during learning (Azevedo, 2009[62]).

Self-regulated learning theory defines learning as a goal-oriented process, in which students make conscious choices as they work towards learning goals, solve problems, reason about data and so forth across topics, domains, and contexts (Winne, 2017[63]; Winne and Hadwin, 1998[64]). Self-regulating learners use cognitive processes (read, practise, elaborate) to study a topic, and they use metacognitive processes (e.g. orientation, planning, goal-setting, monitoring and evaluation) to actively monitor and control their learning, reasoning, problem-solving processes and affective states (e.g. boredom, confusion, frustration) and to further motivate (e.g. increase task value and interest) themselves to engage in an appropriate level of effort (Greene and Azevedo, 2010[65]).

The relevance of supporting self-regulated learning is twofold. First, self-regulated learning skills are deemed essential for humans in the upcoming decades. Human intelligence will increasingly be augmented by artificial intelligence. Human agency is needed to take a leading role in these transitions and regulation skills are critical in these contexts where automation and robotisation of the economy increases (OECD, 2019[66]). Coupled with this shift, human intellectual capacity for solving society's most challenging problems will be in high demand (Luckin, 2017[67]). These human skills and competencies that AI cannot easily replicate will be necessary to succeed in a rapidly changing world (World Economic Forum, 2018[68]).The ability to self-regulate, that is, take initiative, set goals, and monitor self and others, is at the heart of these human skills and competences. Learners who can self-regulate learn more effectively and develop more elaborate mental structures to support the application of their knowledge in diverse situations (Paans et al., 2018[69]). Second, self-regulated learning skills are needed for effective lifelong learning (at school and in the workplace) to equip learners with agency, the feeling that one controls one's own life and to provide a means to adapt and regulate behaviour in challenging situations throughout life (e.g. family, hobbies, and work).

Including self-regulated learning in approaches to personalise education could therefore benefit both current and future learning (Molenaar, Horvers and Baker, 2019[70]). Yet, where researchers and developers have developed successful ways to measure students' knowledge during learning, the first hurdle to personalisation based on a broader range of learner characteristics has been found to lie in the measurement of self-regulated learning during learning. There has been a growing interest in capturing self-regulated learning processes with multimodal data which may offer unobtrusive and scalable measurement solutions. The different data streams introduced above are used in research to improve measurement but have only been used to a limited extent to support interventions. For example, Wayang Outpost uses students' affective states to adjust feedback to increase their motivation (Arroyo et al., 2014[20]) and AtGentive (Molenaar, Van Boxtel and Sleegers, 2011[71]) provides metacognitive scaffolds to learners based on their progress. MetaTutor students receive prompts from a strategy agent to set goals based on their navigation behaviour (Harley et al., 2015[72]). This measurement issue is discussed by D'Mello (D'Mello, 2021[7]) in the case of "learning engagement".

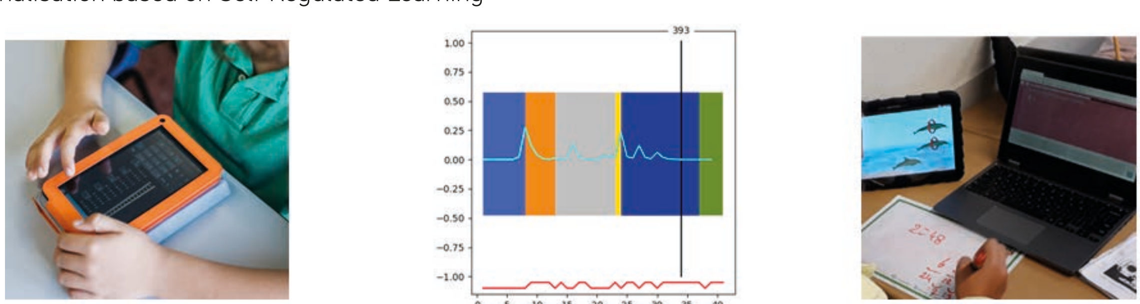
To illustrate this development, Figure 3.9 outlines an intervention that uses log data to detect self-regulated learning processes during learning based on data from an Adaptive Learning Technology used at scale. The moment-by-moment learning algorithm was developed to visualise the probability that a learner has learned a specific skill on a particular problem (Baker et al., 2013[73]). These visualisations show how much the learner is likely to have learned at each problem-solving opportunity, which is a representation of the learner's progress over time. Baker, Goldstein and Heffernan (2011[74]) found that spikiness in the graph showing the probability that a learner has just learned something at a specific moment was associated with sustained learning gain during their experiment. Moreover, the different visual patterns of the probability that the learner has just learned, the moment-by-moment curves, were related to different learning outcomes: for example an immediate peak was related to post-test results (Baker et al., 2013[73]).

These visual patterns also reflect learners' regulation (Molenaar, Horvers and Baker, 2019[75]). As such they provide insight into how learners regulate their accuracy and learning over time while learning with adaptive learning technologies (Molenaar, Horvers and Baker, 2019[70]). Personalised dashboards for students were developed based on this to provide feedback to learners. In this way, the role of learner-facing dashboards from discussing what learners learned to also incorporating how learners learned. Results of the study indicated that learners receiving dashboards improved regulation during learning, reached higher learning outcomes and achieved higher monitoring accuracy (Molenaar et al., 2020[76]). Overall, these findings indicate that the personalised dashboards did positively affect learners during learning.

This case (Figure 3.9) illustrates how widening the indicators that are tracked and increasing the scope of diagnosis can further the personalisation of learning and advance our ability to accurately understand a learner's current state and improve the prediction of future development. This supports better approaches towards the personalisation of learning that incorporate more diverse learner characteristics and a better understanding of the learner's environment.

Figure 3.9 The Moment-by-Moment Learning Algorithm

Personalisation based on Self-Regulated Learning



Detect **Diagnose** **Act**

Source: Illustrations - Inge Molenaar/Adaptive Learning Lab

Challenges for the future of personalisation of learning

I started this chapter by presenting the *6 levels of the automation of personalised learning* model to position the role of AI in education and to elicit a discussion about the envisioned level of automation for AI in education. Full automation may not be the optimal level of automation for AI and education in formal schooling settings. Hybrid systems in which human and artificial intelligence reinforce each other are put forward as new direction in these contexts. Notwithstanding the rapid influx of educational technologies used at scale, state-of-the-art technologies for advanced personalisation of learning are not often extensively used in schools. There still is a divide between learning technologies used at scale in schools and state-of-the-art technologies available in research labs. Most technologies used at scale are at the level of teacher assistance (digital materials) and partial automation. Even in state-of-the-art technologies, the major focus is on personalising based on students' knowledge. A focus on the whole student taking into account broader learner characteristics such as self-regulation, emotion and motivation could further enhance personalisation. In this quest multiple data streams offer new opportunities to detect and diagnose these diverse learner characteristics.

The development towards more advanced forms of personalisation is complex and requires international research and development (R&D) collaboration to ensure progress. Orchestration at the governmental level is needed to facilitate these technological innovations. In order to accomplish wider deployment of state-of-the-art technology this chapter concludes with three recommendations for policy makers: *i) govern ethics by design, transparency and data protection; ii) improve learning technologies with public-private partnerships;* and *iii) actively engage teachers and education professionals in these transitions.*

Recommendation 1. Govern ethics by design, transparency and data protection

Although this chapter does not aim to discuss the ethical issues related to AI in education, it is important that governments ensure that developments in AI and personalisation of learning continue to support the common good (Sharon, 2018[77]). Education is a basic human right and governments should take the necessary measures to safeguard access to open and unbiased educational opportunities for all (UNESCO, 2019[78]). Within the realm of above described developments, this requires in-depth reflection on the data infrastructure, data governance and legal framework needed to safeguard these basic human rights as well as learners' privacy, security and well-being. Governments should develop comprehensive data protection laws plus regulatory frameworks to guarantee ethical, non-discriminatory, equitable, transparent and auditable use and re-use of learners' data (UNESCO, 2019[78]). It is important to ensure that educational professionals responsibly use technologies, especially the data used for detection and algorithms used for diagnosis. Research plays a vital role in pre-forming critical analysis of AI solutions to ensure transparency in the sector.

Recommendation 2. Improving learning technologies with public-private partnerships

As indicated above, there still is a divide between the state of the art in research labs and the daily use of technologies in schools. Despite the sharp increase in use of technologies in classrooms over the last 10 years (Vincent-Lancrin et al., 2019[54]), there are no examples in which the full potential of learning technologies for personalisation of learning is being applied in schools. There are no advanced educational technology solutions that incorporate the three levels of adjustment (task, step and curriculum) and broad learner characteristics. Governments can encourage developments by coordinating public-private partnerships between research institutes and EdTech companies. Across sectors, large quantities of data combined with machine learning cause transformations and disrupt markets (Bughin et al., 2018[79]). Large datasets are required to realise evidence-based applications of machine learning in education. Partnerships between those that collect data with technologies used at scale (mostly private companies and public educational authorities), those who have the expertise to make the field progress (mostly researchers in universities and companies), and those who shape education (mostly teachers and educational leadership in schools) are crucial to ensure further personalisation of learning. Collaborations within this "golden triangle" (Cukurova, Luckin and Clark-Wilson, 2018[80]) have the potential to fast-forward personalisation of learning and, consequently, advance on the levels of automation in education. Good examples of such partnerships are the EDUCATE project of the University College London[9] and the Simon Initiative of Carnegie Mellon University[10].

Recommendation 3. Engaging teachers and education professionals in R&D

The 6 levels of automation model can also help teachers and education professionals to understand the role of AI in education. Traditional uptake of technologies has been low and resistance to full automation has been high (Tondeur et al., 2013[4]). Imagine if you were to step into a fully automated car tomorrow: How long would you continue to monitor or even control the cars functioning? Gradual transition through the levels of automation supports teachers' trust in AI and helps develop evidence for its effectiveness. Research shows that when teachers experience "teacher assistance" in level 1, this supports teacher empowerment and articulation of more advanced future education scenarios (Molenaar and Knoop-van Campen, 2019[16]). Teachers are responsible for their students' well-being; technologies should allow teachers to fulfil that responsibility. Interactive dialogues discussing the role of data, learning analytics and AI in education will help us achieve innovations beyond our current perception and truly learn how to grasp the endless possibilities AI offers education.

Notes

1. https://www.alelo.com/about-us/
2. http://hstrik.ruhosting.nl/DART/
3. https://www.lexplore.com/
4. https://www.changedyslexia.org/
5. https://lingvist.com/
6. https://www.letrus.com.br/
7. https://www.aleks.com/
8. https://wrts.nl/
9. https://www.ucl.ac.uk/ioe/departments-and-centres/centres/ucl-knowledge-lab/educate
10. https://www.cmu.edu/simon/

References

Aleven, V. and E.A. McLaughlin (2016), "Instruction Based on Adaptive Learning Technologies", in *Handbook of Research on Learning and Instruction*, Routledge Handbooks, http://dx.doi.org/10.4324/9781315736419. [1]

Arroyo, I., B. Park Woolf, W. Burelson, K. Muldner, D. Rai and M. Tai (2014), "A Multimedia Adaptive Tutoring System for Mathematics that Addresses Cognition, Metacognition and Affect", *International Journal of Artificial Intelligence in Education*, Vol. 24/4, pp. 387-426, http://dx.doi.org/10.1007/s40593-014-0023-y. [20]

Asselborn, T., T. Gargot, Ł. Kidziński, W. Johal, D. Cohen, C. Jolly and P. Dillenbourg (2018), "Automated human-level diagnosis of dysgraphia using a consumer tablet", npj Digital Medicine, Vol. 1/1, http://dx.doi.org/10.1038/s41746-018-0049-x. [32]

Awad, E., S. Dsouza, R. Kim, J. Schultz, J. Henrich, A. Shariff, J-F. Bonnefon and I. Rahwan (2018), "The Moral Machine experiment", Nature, Vol. 563/7729, pp. 59-64, http://dx.doi.org/10.1038/s41586-018-0637-6. [11]

Azevedo, R. (2009), "Theoretical, conceptual, methodological, and instructional issues in research on metacognition and self-regulated learning: A discussion", *Metacognition and Learning*, Vol. 4/1, pp. 87-95, http://dx.doi.org/10.1007/s11409-009-9035-7. [62]

Azevedo, R. and Gašević, D. (2019), "Analyzing Multimodal Multichannel Data about Self-Regulated Learning with Advanced Learning Technologies: Issues and Challenges", *Computers in Human Behavior*, Vol. 96, pp. 207-2010. [27]

Baker, R. (2016), "Stupid Tutoring Systems, Intelligent Humans", *International Journal of Artificial Intelligence in Education*, Vol. 26/2, pp. 600-614, http://dx.doi.org/10.1007/s40593-016-0105-0. [57]

Baker, R., A. Corbett and V. Aleven (2008), "More Accurate Student Modeling through Contextual Estimation of Slip and Guess Probabilities in Bayesian Knowledge Tracing", in I*ntelligent Tutoring Systems, Lecture Notes in Computer Science*, Springer Berlin Heidelberg, Berlin, Heidelberg, http://dx.doi.org/10.1007/978-3-540-69132-7_44. [41]

Baker, R., A. Goldstein and N. Heffernan (2011), "Detecting learning moment-by-moment", *International Journal of Artificial Intelligence in Education,* Vol. 21/1-2, pp. 5–25, https://doi.org/10.3233/JAI-2011-015. [74]

Baker, R., A. Hershkovitz, L.M. Rossi, A.B. Goldstein and S.M. Gowda (2013), "Predicting robust learning with the visual form of the moment-by-moment learning curve", *Journal of the Learning Sciences*, Vol. 22/4, pp. 639-666. [73]

Baker, R. and P. Inventado (2014), *Educational data mining and learning analytics*, Springer. [26]

Bannert, M., I. Molenaar, R. Azevedo, S. Järvelä and D. Gašević (2017), "Relevance of learning analytics to measure and support students' learning in adaptive educational technologies", *Proceedings of the Seventh International Learning Analytics & Knowledge Conference,* http://dx.doi.org/10.1145/3027385.3029463. [60]

Blikstein, P. (2018), *Time for hard choices in AIED*, Keynote at London Festival of Learning, https://vimeo.com/283023489. [22]

Bloom, B. (1984), "The 2 Sigma Problem: The Search for Methods of Group Instruction as Effective as One-to-One Tutoring", *Educational Researcher*, Vol. 13/6, pp. 4-16, http://dx.doi.org/10.3102/0013189x013006004. [24]

Bodily, R., J. Kay, V. Aleven, J. Jivet, D. Davis, F. Xhakaj and K. Verbert (2018), "Open learner models and learning analytics dashboards", *Proceedings of the 8th International Conference on Learning Analytics and Knowledge*, http://dx.doi.org/10.1145/3170358.3170409. [15]

Bosch, N., S. D'Mello, R. Baker, J. Ocumpaugh, V. Shute, M. Ventura, L. Wang and W. Zhao (2015), "Automatic Detection of Learning-Centered Affective States in the Wild", Proceedings of the 20th *International Conference on Intelligent User Interfaces - IUI '15*, http://dx.doi.org/10.1145/2678025.2701397. [31]

Bughin, J., J. Seong, J. Manyika, M. Chui and R. Joshi (2018), "Notes from the AI frontier: Modeling the global economic impact of AI", *McKinsey Global Institute: Vol. September*, http://www.mckinsey.com/mgi. [79]

Carnegie Learning (n.d.), Carnegie Learning, https://www.carnegielearning.com/products/software-platform/mathia-learning-software (accessed on 1 December 2020). [46]

Corbett, A. and J. Anderson (1995), "Knowledge tracing: Modeling the acquisition of procedural knowledge", *User Modelling and User-Adapted Interaction*, Vol. 4/4, pp. 253-278, http://dx.doi.org/10.1007/bf01099821. [39]

Corno, L. (2008), "On Teaching Adaptively", *Educational Psychologist*, Vol. 43/3, pp. 161-173, http://dx.doi.org/10.1080/00461520802178466. [2]

Cukurova, M., R. Luckin and A. Clark-Wilson (2018), "Creating the golden triangle of evidence-informed education technology with EDUCATE", *British Journal of Educational Technology*, Vol. 50/2, pp. 490-504, http://dx.doi.org/10.1111/bjet.12727. [80]

Dede, C. (1986), "A review and synthesis of recent research in intelligent computer-assisted instruction", *International Journal of Man-Machine Studies*, Vol. 24/4, pp. 329-353, http://dx.doi.org/10.1016/s0020-7373(86)80050-5. [38]

Desmarais, M. and R. Baker (2011), "A review of recent advances in learner and skill modeling in intelligent learning environments", *User Modeling and User-Adapted Interaction*, Vol. 22/1-2, pp. 9-38, http://dx.doi.org/10.1007/s11257-011-9106-8. [30]

Dillenbourg, P. (2021), "Classroom analytics: Zooming out from a pupil to a classroom", in OECD *Digital Education Outlook 2021: Pushing the frontiers with AI, blockchain, and robots.*, OECD Publishing. [14]

D'Mello, S. (2021), "Improving student engagement in and with digital learning technologies", in OECD *Digital Education Outlook 2021: Pushing the frontiers with AI, blockchain, and robots.*, OECD Publishing. [7]

Faber, J., H. Luyten and A. Visscher (2017), "The effects of a digital formative assessment tool on mathematics achievement and student motivation: Results of a randomized experiment", *Computers & Education*, Vol. 106, pp. 83-96, http://dx.doi.org/10.1016/j.compedu.2016.12.001. [43]

Falmagne, J., E. Cosyn, J-P. Doignon and N. Thiéry (2006), "The Assessment of Knowledge, in Theory and in Practice", in *Formal Concept Analysis, Lecture Notes in Computer Science*, Springer Berlin Heidelberg, Berlin, Heidelberg, http://dx.doi.org/10.1007/11671404_4. [50]

Feng, M. and N. Heffernan (2005), "Informing Teachers Live about Student Learning : Reporting in the Assistment System", *Tech., Inst., Cognition and Learning*, Vol. 3/508, pp. 1-14. [59]

Greene, J. and R. Azevedo (2010), "The Measurement of Learners' Self-Regulated Cognitive and Metacognitive Processes While Using Computer-Based Learning Environments", *Educational Psychologist*, Vol. 45/4, pp. 203-209, http://dx.doi.org/10.1080/00461520.2010.515935. [65]

Harari, Y. (2018), *21 Lessons for the 21st Century*, Random House. [10]

Harley, J., F. Bouchet, M. Sazzad Hussain, R. Azevedo and R. Calvo (2015), "A multi-componential analysis of emotions during complex learning with an intelligent multi-agent system", *Computers in Human Behavior*, Vol. 48, pp. 615-625. [72]

Holstein, K., B. McLaren and V. Aleven (2019), "Co-designing a real-time classroom orchestration tool to support teacher–AI complementarity", *Journal of Learning Analytics*, Vol. 6/2, pp. 27-52. [58]

Holstein, K., B. McLaren and V. Aleven (2018), "Student Learning Benefits of a Mixed-Reality Teacher Awareness Tool in AI-Enhanced Classrooms", in *Lecture Notes in Computer Science, Artificial Intelligence in Education*, Springer International Publishing, Cham, http://dx.doi.org/10.1007/978-3-319-93843-1_12. [36]

Holstein, K., B. McLaren and V. Aleven (2017), "Intelligent tutors as teachers' aides", *Proceedings of the Seventh International Learning Analytics & Knowledge Conference*, http://dx.doi.org/10.1145/3027385.3027451. [19]

Järvelä, S. and M. Bannert (2019), "Temporal and adaptive processes of regulated learning - What can multimodal data tell?", *Learning and Instruction*, p. 101268, http://dx.doi.org/10.1016/j.learninstruc.2019.101268. [61]

Kamar, E. (2016), "*Directions in Hybrid Intelligence: Complementing AI Systems with Human Intelligence*", IJCAI, pp. 4070–4073. [8]

Klinkenberg, S., M. Straatemeier and H. van der Maas (2011), "Computer adaptive practice of Maths ability using a new item response model for on the fly ability and difficulty estimation", *Computers & Education*, Vol. 57/2, pp. 1813-1824, http://dx.doi.org/10.1016/j.compedu.2011.02.003. [40]

Knoop-van Campen, C. and I. Molenaar (2020), "How Teachers integrate Dashboards into their Feedback Practices", *Frontline Learning Research*, pp. 37-51, http://dx.doi.org/10.14786/flr.v8i4.641. [35]

Koedinger, K., J. Booth and D. Klahr (2013), "Instructional Complexity and the Science to Constrain It", *Science*, Vol. 342/6161, pp. 935-937, http://dx.doi.org/10.1126/science.1238056. [33]

Koedinger, K. and A. Corbett (2006), "Cognitive Tutors: Technology Bringing Learning Sciences to the Classroom - Chapter 5", The Cambridge Handbook of: The Learning Sciences, pp. 61-78, https://www.academia.edu/download/39560171/koedingercorbett06.pdf. [18]

Koedinger, K., A. Corbett and C. Perfetti (2012), "The Knowledge-Learning-Instruction Framework: Bridging the Science-Practice Chasm to Enhance Robust Student Learning", *Cognitive Science*, Vol. 36/5, pp. 757-798, http://dx.doi.org/10.1111/j.1551-6709.2012.01245.x. [34]

Kulik, J. and J. Fletcher (2016), "Effectiveness of Intelligent Tutoring Systems", *Review of Educational Research*, Vol. 86/1, pp. 42-78, http://dx.doi.org/10.3102/0034654315581420. [56]

Luckin, R. (2017), "Towards artificial intelligence-based assessment systems", *Nature Human Behaviour*, Vol. 1/3, http://dx.doi.org/10.1038/s41562-016-0028. [67]

Miller, W., R. Baker, M.J. Labrum, K. Petsche, Y. Han Liu and A.Z. Wagner (2015), "Automated detection of proactive remediation by teachers in reasoning mind classrooms", *Proceedings of the Fifth International Conference on Learning Analytics And Knowledge - LAK '15*, http://dx.doi.org/10.1145/2723576.2723607. [17]

Molenaar, I., A. Horvers and R. Baker (2019), "Towards Hybrid Human-System Regulation", *Proceedings of the 9th International Conference on Learning Analytics & Knowledge*, http://dx.doi.org/10.1145/3303772.3303780. [70]

Molenaar, I., A. Horvers and R. Baker (2019), "What can moment-by-moment learning curves tell about students' self-regulated learning?", *Learning and Instruction*, p. 101206, http://dx.doi.org/10.1016/j.learninstruc.2019.05.003. [75]

Molenaar, I., A. Horvers, R. Dijkstra and R. Baker (2020), "Personalized visualizations to promote young learners' SRL", *Proceedings of the Tenth International Conference on Learning Analytics & Knowledge*, http://dx.doi.org/10.1145/3375462.3375465. [76]

Molenaar, I. and C. Knoop-van Campen (2019), "How Teachers Make Dashboard Information Actionable", *IEEE Transactions on Learning Technologies*, Vol. 12/3, pp. 347-355, http://dx.doi.org/10.1109/tlt.2018.2851585. [16]

Molenaar, I., C. Knoop-van Campen and F. Hasselman (2017), "The effects of a learning analytics empowered technology on students' arithmetic skill development", *Proceedings of the Seventh International Learning Analytics & Knowledge Conference*, http://dx.doi.org/10.1145/3027385.3029488. [45]

Molenaar, I., C. Van Boxtel and P. Sleegers (2011), "The Effect of Dynamic Computerized Scaffolding on Collaborative Discourse", in *Towards Ubiquitous Learning, Lecture Notes in Computer Science*, Springer Berlin Heidelberg, Berlin, Heidelberg, http://dx.doi.org/10.1007/978-3-642-23985-4_39. [71]

Molenaar, I. and C. van Campen (2016), "Learning analytics in practice", *Proceedings of the Sixth International Conference on Learning Analytics & Knowledge - LAK '16*, http://dx.doi.org/10.1145/2883851.2883892. [44]

Mudrick, N., R. Azevedo and M. Taub (2019), "Integrating metacognitive judgments and eye movements using sequential pattern mining to understand processes underlying multimedia learning", *Computers in Human Behavior*, Vol. 96, pp. 223–234. [28]

Muralidharan, K., A. Singh and A. Ganimian (2019), "Disrupting education? Experimental evidence on technology-aided instruction in India", *American Economic Review*, Vol. 109/4, pp. 1426-60. [53]

OECD (2019), *OECD Skills Outlook 2019 : Thriving in a Digital World*, OECD Publishing, Paris, https://dx.doi.org/10.1787/df80bc12-en. [66]

Paans, C., E. Segers, I. Molenaar and L. Verhoeven (2018), "The quality of the assignment matters in hypermedia learning", *Journal of Computer Assisted Learning*, Vol. 34/6, pp. 853-862, http://dx.doi.org/10.1111/jcal.12294. [69]

Pane, J. , B.A. Griffin, D.F. McCaffrey and R. Karam (2014), "Effectiveness of Cognitive Tutor Algebra I at Scale", *Educational Evaluation and Policy Analysis*, Vol. 36/2, pp. 127-144, http://dx.doi.org/10.3102/0162373713507480. [48]

Pane, J., D.F. McCaffrey, M.E. Slaughter, J.L. Steele and G.S. Ikemoto (2010), "An Experiment to Evaluate the Efficacy of Cognitive Tutor Geometry", *Journal of Research on Educational Effectiveness*, Vol. 3/3, pp. 254-281, http://dx.doi.org/10.1080/19345741003681189. [47]

Parasuraman, R., T. Sheridan and C. Wickens (2000), "A Model for Types and Levels of Human Interaction with Automation", *SYSTEMS AND HUMANS*, Vol. 30/3, https://ieeexplore.ieee.org/abstract/document/844354/. [13]

Pashler, H., D. Rohrer, N.J. Capeda and S.K. Carpenter (2007), "Enhancing learning and retarding forgetting: Choices and consequences", *Psychonomic Bulletin & Review*, Vol. 14/2, pp. 187-193, http://dx.doi.org/10.3758/bf03194050. [52]

Pavlik, P. and J. Anderson (2005), "Practice and Forgetting Effects on Vocabulary Memory: An Activation-Based Model of the Spacing Effect", *Cognitive Science*, Vol. 29/4, pp. 559-586, http://dx.doi.org/10.1207/s15516709cog0000_14. [51]

Ritter, S., J. Kulikowich, P. Lei, C. Mcguire and P. Morgan (2007), "What evidence matters? A randomized field trial of Cognitive Tutor Algebra I.", *Proceeding of the 15th International Conference on Computers in Education, ICCE 2007, November 5-9, 2007, Hiroshima, Japan.* [49]

Robinson, K. (2010), *Changing Edcuation Paradigms,* https://www.ted.com/talks/sir_ken_robinson_changing_education_paradigms (accessed on 26 November 2020). [3]

Sharon, T. (2018), "When digital health meets digital capitalism, how many common goods are at stake?", *Big Data & Society*, Vol. 5/2, p. 205395171881903, http://dx.doi.org/10.1177/2053951718819032. [77]

Shladover, S. (2018), "Connected and automated vehicle systems: Introduction and overview", *Journal of Intelligent Transportation Systems*, Vol. 22/3, pp. 190-200, http://dx.doi.org/10.1080/15472450.2017.1336053. [21]

Snappet (n.d.), *Snappet*, http://www.snappet.org (accessed on 1 December 2020). [42]

Stewart, A., H. Vrzakova, C. Sun, J. Yonehiro, C. Stone, N. Duran, V. Shute and S. D'Mello (2019), "I Say, You Say, We Say", *Proceedings of the ACM on Human-Computer Interaction*, Vol. 3/CSCW, pp. 1-19, http://dx.doi.org/10.1145/3359296. [29]

Stiftung, R. (ed.) (2018), *Technology-enhanced Personalised Learning: Untangling the Evidence Other How to cite*, http://www.studie-personalisiertes-lernen.de/en/ (accessed on 26 November 2020). [9]

Tondeur, J., L.H. Kershaw, R.R. Vanderlinde and J. van Braak (2013), "Getting inside the black box of technology integration in education: Teachers' stimulated recall of classroom observations", *Australasian Journal of Educational Technology 3*, https://ajet.org.au/index.php/AJET/article/view/16. [4]

Topol, E. (2019), "High-performance medicine: the convergence of human and artificial intelligence", *Nature Medicine*, Vol. 25/1, pp. 44-56, http://dx.doi.org/10.1038/s41591-018-0300-7. [12]

UNESCO (2019), *Beijing Consensus on Artificial Intelligence and Education*, https://unesdoc.unesco.org/ark:/48223/pf0000368303. [78]

VanLEHN, K. (2011), "The Relative Effectiveness of Human Tutoring, Intelligent Tutoring Systems, and Other Tutoring Systems", *Educational Psychologist*, Vol. 46/4, pp. 197-221, http://dx.doi.org/10.1080/00461520.2011.611369. [25]

VanLehn, K. (2015), "Regulative Loops, Step Loops and Task Loops", *International Journal of Artificial Intelligence in Education*, Vol. 26/1, pp. 107-112, http://dx.doi.org/10.1007/s40593-015-0056-x. [55]

Vanlehn, K. (2006), "The Behavior of Tutoring Systems", *International Journal of Artificial Intelligence in Education*, Vol. 16. [5]

Vincent-Lancrin, S., J. Urgel, S. Kar and G. Jacotin (2019), *Measuring Innovation in Education 2019: What Has Changed in the Classroom?*, Educational Research and Innovation, OECD Publishing, Paris, https://dx.doi.org/10.1787/9789264311671-en. [54]

Wingo, G. (1961), "Teaching Machines", *Journal of Teacher Education*, Vol. 12/2, pp. 248-249, http://dx.doi.org/10.1177/002248716101200225. [23]

Winne, P. (2017), "Learning Analytics for Self-Regulated Learning", in Handbook of *Learning Analytics*, Society for Learning Analytics Research (SoLAR), http://dx.doi.org/10.18608/hla17.021. [63]

Winne, P. and R. Baker (2013), "The Potentials of Educational Data Mining for Researching Metacognition, Motivation and Self-Regulated Learning", *JEDM - Journal of Educational Data Mining*, Vol. 5/1, pp. 1-8, https://doi.org/10.5281/zenodo.3554619. [6]

Winne, P. and A. Hadwin (1998), "Studying as self-regulated learning", *Metacognition in educational theory and practice.* [64]

Wise, A. (2014), "Designing pedagogical interventions to support student use of learning analytics", *Proceedins of the Fourth International Conference on Learning Analytics And Knowledge - LAK '14*, http://dx.doi.org/10.1145/2567574.2567588. [37]

World Economic Forum (2018), *The future of jobs report.* [68]

4 Improving student engagement in and with digital learning technologies

Sidney K. D'Mello
University of Colorado Boulder, United States

Engagement is critical to learning, but it has proved challenging to promote both meaningful engagement and deep learning. Can digital learning technologies help? This chapter provides a broad overview of some promising paths to measure students' level of engagement during learning with digital technologies and how these technologies can be designed to improve engagement on the onset of learning or when disengagement sets in. It discusses why engagement matters for learning, how to measure engagement with digital learning technologies, and presents different types of approaches in using data and technology to improve students' engagement and learning.

Introduction

Improving sustained student engagement in learning has become a key objective of education for at least two reasons: (1) engagement is a pre-requisite to meaningful learning; and (2) maintaining engagement involves cognitive and socio-emotional skills that are learning objectives in themselves. Progress in digital technologies, including sensors, advanced data analytics techniques, and innovative digital learning experiences are opening new avenues for measurement, theory development, and design of pedagogical interventions that can help students stay engaged in their learning activities.

Over the last two decades, researchers and developers have made considerable progress in designing learning technologies to promote both engagement and learning. The aim of this chapter is to provide a broad overview of this emerging area and discuss some promising paths for the next decade or so. It will mainly focus on digital learning environments rather than traditional classroom settings, even though some of the developed technologies may soon be adapted for the classroom. At this stage, most of the digital learning technologies that have been developed to measure and promote sustained engagement have been tested in the context of research labs, but studies in real-world settings are gradually increasing. Thus far, results have been mixed, which is par for the course in such a nascent area.

Engagement is an easy concept to comprehend, but difficult to define. Therefore, after emphasising the importance of engagement for learning, the chapter discusses the challenges of scientifically defining engagement. Next, it provides an overview of current approaches to measuring engagement and then discusses how digital technologies and advances in computing methods and technologies afford rapid improvements in measuring it. With definition and measurement issues in place, the paper then shows different ways in which learning technologies have attempted to improve engagement and learning through proactive strategies and reactive interactions. These approaches, illustrated by examples and case studies, will give some ideas of what could soon become possible if not already widely used in digital learning environments. The paper concludes with a discussion of important next steps for this research programme and how it could transform education practice in the future.

The importance of engagement

The importance of engagement to learning has been recognised and investigated for decades. The research broadly supports the following general conclusion: a student who is engaged is primed to learn; a student who is disengaged is not. In particular, boredom, which is somewhat the antithesis of engagement, is not merely an unpleasant feeling. Boredom is associated with difficulties in maintaining attention (Danckert and Merrifield, 2018[1]; Eastwood et al., 2012[2]; Hunter and Eastwood, 2018[3]) and is consistently negatively correlated with learning outcomes (Pekrun et al., 2014[4]; Putwain et al., 2018[5]). For example, a recent meta-analysis on 29 studies (N = 19 052 students) found an overall significant negative effect r = -.24 of boredom on academic outcomes (Tze, Daniels and Klassen, 2016[6]). It is of course not necessary, and even unrealistic, to expect a student to be engaged 100% of the time; some amount of disengagement is likely to always occur and should not be problematic in small doses. However, sustained disengagement is associated with a host of negative outcomes including lower levels of achievement, lower self-efficacy, diminished interest in educational activities, partaking in high-risk behaviours, and most importantly, increased attrition and dropout from school (Baker et al., 2010[7]; Csikszentmihalyi, 1975[8]; Daniels et al., 2009[9]; Farrell et al., 1988[10]; Finn and Voelkl, 1993[11]; Griffiths et al., 2012[12]; Mann and Robinson, 2009[13]; Patrick, Skinner and Connell, 1993[14]; Pekrun et al., 2010[15]; Perkins and Hill, 1985[16]; Wasson, 1981[17]). Put in a more positive light, being engaged in school is associated with several beneficial outcomes beyond academic achievement itself.

Much of the research on engagement has focused on traditional learning which occurs in the classroom and school settings. The Handbook of Research on Student Engagement (Christenson, Reschly and Wylie, 2012[18]) provides a detailed exposition of a range of issues pertaining to engagement in these learning contexts. However, with the advent of mobile devices, the Internet, and social media, much of learning occurs via digital media. This poses a challenge since it is particularly difficult to engage students when they interact with digital learning technologies, often in isolation. For example, Massive Open Online Courses (MOOCs) have achieved impressive outreach by opening up content to millions of learners worldwide. Yet, traditional MOOCs (xMOOCs), which mainly consist of watching videos, completing auto-graded assessments, and maybe participating in online discussions, have major problems with engagement and dropout (Yang et al., 2013[19]). Whereas a gifted human teacher or an expert tutor can design collaborative activities to increase engagement and can adapt the lesson when engagement appears to be waning, it is difficult for digital learning technologies to promote and sustain meaningful engagement for all learners. Even when a learning technology is successful at initially capturing students' attention, they have little recourse when novelty fades, the student gets stuck, or boredom eventually sets in.

Designing digital learning experiences that promote both engagement and deep conceptual learning is a challenging task because it requires bridging the gap between learning and liking, which are often at odds. Herein lies the conundrum. On one hand, it is very easy to entertain students via puzzles, games, comics, and other "edutainment" gimmicks. Undoubtedly, students find these experiences to be very engaging, but it is unclear if they learn anything meaningful, especially at deeper levels of comprehension (Charsky, 2010[20]; Papert, 1998[21]). Further, approaches to increase interest by manipulating the learning materials (e.g. cover stories, very realistic images) can actually harm learning, presumably by detracting attention and cognitive resources away from the learning content (Rey, 2012[22]). On the other hand, several decades of research into the cognitive science of learning has yielded principles of effective learning (e.g. Bransford, Brown and Cocking, 2000[23]; Karpicke and Blunt, 2011[24]; Roediger and Karpicke, 2006[25]) which can be implemented in learning technologies, such as intelligent tutoring systems (ITSs). However, despite the widely-documented learning benefits of intelligent tutoring systems (e.g. Steenbergen-Hu and Cooper, 2014[26]) and other technologies that implement deep learning principles (McNamara et al., 2006[27]), students find interacting with these technologies to be quite tedious (see Baker et al., 2010[7]; Craig et al., 2008[28]; but also see Rodrigo and Baker, 2011[29]) for an exception where intelligent tutoring systems were found to be more engaging than games. The central issue is that learning is hard because it takes considerable effort and practice to attain mastery (Ericsson, Krampe and Tesch-Römer, 1993[30]). It requires delaying gratification because it is much more rewarding (in the short term) to disengage from learning and engage in something more immediately rewarding, such as social media (Duckworth et al., 2019[31]). Of course, disengagement comes at a hefty price in the future as noted above.

Defining engagement

Unlike physical entities, such as temperature and mass, engagement in a conceptual entity (called a construct) that must be operationally (scientifically) defined. Consider, for example, four hypothetical middle-school students in a maths course.

- Keisha attends her maths class each day, listens attentively, asks good questions, and completes her homework. If asked, she will say that maths is her least favourite subject, but she recognises that it is important for her future.
- James frequently misses his maths class, tries to listen diligently when he does attend, but frequently catches himself zoning out. He completes about 50% of his maths homework and is ambivalent about maths and school in general.
- Rafael attends every maths class, sits silently, never speaks, takes detailed notes, and always completes his homework. He enjoys maths but is really passionate about science since he wants to become a biologist.
- Marcus is passionate about maths and plays maths games in his spare time. He attends every class but finds the material to be too easy, so he frequently gets bored, and attempts to find errors in what his teacher is saying as a way of passing the time. He tries to complete his maths homework every night but finds it too tedious and repetitive so he posts on social media instead.

Which of these students is engaged and which are not? Would a dictionary definition help? The Merriam-Webster dictionary defines engaged/engagement as "involved in activity", "greatly interested" and "emotional involvement or commitment." However, as Eccles and Wang (2012[32]) note, these generic, definitions make the construct more accessible for policy makers and the educated lay person, but less useful for scientific research where precise definitions are of greater value, especially when it comes to elucidating cause and effect relationships.

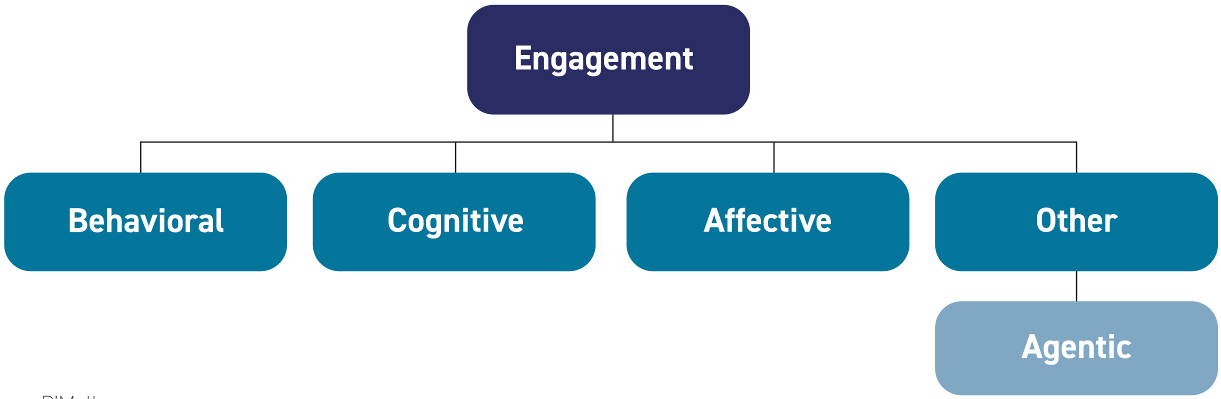
Unfortunately, a scientific definition of engagement remains elusive. Reschly and Christenson (2012[33]) note that the term *engagement* has been used to describe diverse behaviours, thoughts, perceptions, feelings, and attitudes, and at the same time, diverse terms have been used to refer to similar constructs. Engagement has been considered to be a broad and complex construct pertaining to diverse aspects of the educational experience (e.g. showing up, completing homework, feelings of belonging, getting a good time) and across multiple time scales (e.g. momentary periods of interest, stable dispositions such as general disengagement with school, and life-altering outcomes like dropping out of school). Thus, it may be more fruitful to study specific aspects of this complex construct rather than striving for an all-encompassing, but overly generic definition.

Researchers generally agree that engagement is a multidimensional construct, although the number and nature of the dimensions are unclear (Figure 4.1). Fredricks, Blumenfeld and Paris (2004[34]) proposed three components of engagement in a highly influential article. *Emotional engagement* encompasses feelings and attitudes about the learning task or learning context, such as feelings of interest towards a particular subject, teacher (Renninger and Bachrach, 2015[35]), or general satisfaction with school. *Behavioural engagement* broadly refers to learners' participation in learning, including effort, persistence, and concentration. *Cognitive engagement* pertains to learners' investment in the learning task, such as how they allocate effort towards learning, and their understanding and mastery of the material.

Reeve and Tseng (2011[36]) recently suggested a fourth dimension: *agentic engagement*, characterised by learners proactively contributing to the learning process. Alternatively, Pekrun and Linnenbrink-Garcia (2012[37]) proposed a five-component model that includes cognitive (e.g. attention and memory processes), motivational (e.g. intrinsic and extrinsic motivation), behavioural (e.g. effort and persistence), social-behavioural (e.g. participating with peers), and cognitive-behavioural (e.g. strategy use and self-regulation) aspects of engagement.

In addition to the components of engagement, the time course and the influence of contextual factors are also quite important. With respect to the time course, a fleeting interest in a topic that engages a student for a few minutes or a few hours is different from a student who sustains that engagement over months and years (Hidi and Renninger, 2006[38]). It is also widely acknowledged that the context in which an activity is situated has a major impact on the resultant patterns of engagement. Engagement is not an intrinsic property of individuals but emerges from interactions between individuals and their environments including their peers, teachers, family, and school structures (Christenson, Reschly and Wylie, 2012[18]).

Figure 4.1 Components of engagement



Source: D'Mello.

Accordingly, Sinatra, Heddy and Lombardi (2015[39]) proposed a framework that integrates the components, time course, and contextual influences of engagement (Figure 4.2). The framework conceptualises engagement along a continuum, anchored by person-oriented perspectives at one extreme, context-oriented at the other, and person-in-context in between. Person-oriented perspectives focus on the cognitive, affective, and motivational states of the student at the moment of learning and are best captured with fine-grained physiological and behavioural measures (e.g. response patterns, facial expressions). The context-oriented perspective emphasises the environmental context as the analytic unit. Here, the focus is on macro-level structures like teachers, classrooms, schools, and the community, rather than the individual student. The intermediate-grain size, person-in-context perspective focuses on the interaction between student and context (e.g. how students interact with each other or with technology). This level of analysis would examine, for example, whether particular classroom activities (e.g. small group work) are more engaging than others (e.g. a lecture).

When one focuses on digital learning, engagement can be operationally defined as a goal-directed state of active and focused involvement with a digital learning technology. This operationalisation aligns with the person-oriented level of analysis of Sinatra et al. (2015) in that the focus is on the state (not trait) of the affective, behavioural, and cognitive components of engagement across short time scales ranging from seconds to minutes. In most cases, boredom, drowsiness, and excessive zoning out are signs of disengagement whereas interest, curiosity, and the so-called "flow experience," where a learner is so engrossed in an activity that time and space disappear (Csikszentmihalyi, 1990[40]), would signal engagement. However, in some cases, the specific mental states (and

Figure 4.2 Time course of engagement and influence of context

Source: D'Mello.

their levels) associated with engagement vary based on the affordances of the learning technology. For example, despite being a negative emotion, frustration while playing a difficult educational game might indicate that the student is engaged as it is part and parcel of the "hard fun" of learning from games (Gee, 2003[41]), whereas frustration during a simple vocabulary learning task might signal something other than engaged learning.

Measuring engagement

The effectiveness of any approach to improve engagement depends on the validity of the measurement of engagement. So how has engagement been traditionally measured? Figure 4.3 provides an overview of the various measurement approaches.

Traditional measurement approaches

The most widely used measures of engagement in both traditional and digital education are self-report questionnaires (Fredricks and McColskey, 2012[42]; Greene, 2015[43]; Henrie, Halverson and Graham, 2015[44]). Here, students are asked to endorse items such as "When I am in class, I listen very carefully" (an item for behavioural engagement) or "I enjoy learning new things in class" (an item for emotional engagement). Although relatively inexpensive, easy to administer, and generally reliable, questionnaires have well-known limitations (Duckworth and Yeager, 2015[45]; Krosnick, 1999[46]). For instance, when endorsing items, respondents must compare the target (e.g. a student rating himself or herself) to some implicit standard, and standards may vary from respondent to respondent. To one student, "I am engaged in my maths class" may be exemplified by doing five hours of homework each day; for others, the same statement may be exemplified by simply showing up for class. Thus, biases that arise from heterogeneous frames of reference reduce the validity of self-report questionnaires (Heine et al., 2002[47]). Social desirability bias is another important limitation (Krosnick, 1999[46]), both when respondents aim to appear admirable to others and also when they inflate responses to preserve their own self-esteem. Likewise, memory recall limitations and acquiescence bias can influence self-report questionnaires (Podsakoff et al., 2003[48]).

Several non-questionnaire engagement measures have also been developed. Examples include experience-sampling methods (ESM) (Csikszentmihalyi and Larson, 1987[49]), day reconstruction (Kahneman et al., 2004[50]), and interviews (Turner and Meyer, 2000[51]). However, because they still rely on self-reports, they are subject to some of the same biases as questionnaires. For example, ESM is subject to social desirability biases but not to memory recall biases.

Observational methods are an attractive alternative to self-reports because they are more objective (Nystrand and Gamoran, 1991[52]; Pianta, Hamre and Allen, 2012[53]; Renninger and Bachrach, 2015[35]; Ryu and Lombardi, 2015[54]; Volpe et al., 2005[55]). Unfortunately, these methods entail considerable human effort, which poses a significant challenge for repeated long-term measurement at scale. Further, observations cannot be conducted in some learning contexts, such as students' homes. Some of the limitations may be addressed by combining automated data collection with semi-automated or manual data coding. For example, the Electronically Activated Recorder (EAR) randomly samples audio clips in naturalistic environments (Mehl et al., 2001[56]). Data collection with the EAR is efficient and cost-effective; however, the data still need to be transcribed and coded by humans, which increases cost and reduces scalability. Similarly, engagement can be coded from videos by researchers (Lehman et al., 2008[57]) or even teachers (D'Mello et al., 2008[58]), but video coding is a labour- and time-intensive effort. To address this limitation, there have also been some attempts to automatically analyse classroom video to infer aspects of student engagement (e.g. Aung, Ramakrishnan and Whitehill, 2018[59]; Bidwell and Fuchs, 2011[60]; Klein and Celik, 2017[61]; Raca, Kidzinski and Dillenbourg, 2015[62]), though research is still in its infancy partly due to the difficulty of recording classroom video and related privacy concerns.

Finally, engagement can be adduced from academic and behaviour records, such as homework completion, absences, achievement test scores, time spent on a digital learning platform, and so on (Arnold and Pistilli, 2012[63]; Lehr, Sinclair and Christenson, 2004[64]; Skinner and Belmont, 1993[65]), but these measures are limited in what they can reveal about the cognitive and affective components of engagement. For example, in the technology world, engagement is often equated with usage, and is quantified as the number of logins, clicks, and so on. This characterisation, which only captures a few overt behaviours, is clearly insufficient because it focuses on only one dimension (behavioural) of a multidimensional construct (see above).

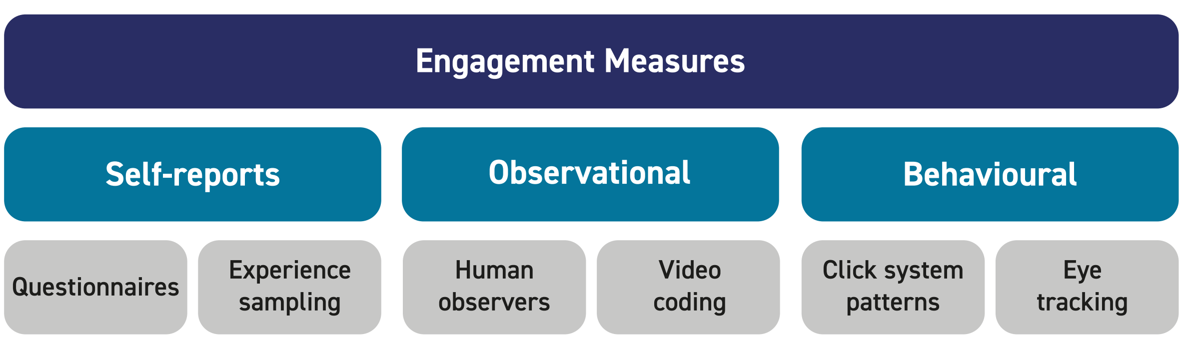
An automated digital measurement approach

Scientific advances on understanding engagement and interventions to increase it are limited by measurement approaches that are either costly or have known biases and limitations. Improving engagement in the context of digital learning is first and foremost a measurement and theoretical challenge that advances in sensors and digital technologies can help address. Among the most policy-, practice-, research-, and design-relevant finding from recent reports on engagement and other similar factors (Atkins-Burnett et al., 2012[66]; Shechtman et al., 2013[67]) is the overuse of and overreliance on student self-report measures, which limits current findings and theory even when produced by rigorous research designs and large sample sizes by the education research, design, policy, and practice communities. In the past and currently, student self-report data informs theory and programme interventions, which use student self-report measures to evaluate programmes and influence theory. Advanced digital technologies and data analytics methods allow one to break the cycle and advance our understanding of engagement systematically and significantly by going beyond the use of student self-report measures only. Reliable, valid, fair, and efficient measures collected from multiple sources and analysed by trained researchers applying methods and techniques appropriately will make the entire value of research studies and programme evaluation more productive.

D'Mello, Dieterle and Duckworth (2017[68]) recently proposed the advanced, analytic, automated (AAA) approach to measure engagement as an alternative method especially suited for interactions with digital learning technologies. This AAA-approach focuses on a person-oriented operationalisation of engagement as the momentary affective and cognitive states that arise throughout the learning process (i.e., it is consistent with the working definition provided above). AAA-based measures have several advantages over counterparts. For one, they are automated, which indicates that they can be applied widely and at scale. Second, they are more consistent because computers provide the measurements, thereby partially obviating reference, social desirability, acquiescence, and other biases associated with self- and observer- reports. These measures are also unaffected by momentary lapses in attention or by fatigue, as can occur with humans. They vastly reduce time and effort, which is a limitation of ESM, day reconstruction, video coding, and human observations.

The core idea of the AAA-approach (Figure 4.4) is to train machines to *infer* latent mental states associated with engagement (e.g. concentration, interest) from machine-readable signals and from aspects of the environmental context. The approach uses machine learning which requires *training* data to learn a computational model (computer programme) that can be applied on data collected in the future (*unseen data*). Accordingly, the AAA measurement approach, which begins with collection of training data.

Figure 4.3 Major categories of measures of engagement with examples (last row)



Source: D'Mello.

Step 1 consists of recording *signals* (video, physiology, log files, etc.) as students complete a learning activity within a particular learning context (Step 1a), followed by computing higher-order representations (called *features*) from the signals (Step 1b). A video is an example signal, from which features such as smiles, nods, frowns, etc., can be automatically extracted using computer vision techniques (Valstar et al., 2012[69]). Patterns of students' interactions (i.e., click-streams) with learning technologies provide another powerful signal into student engagement (Baker and Ocumpaugh, 2015[70]).

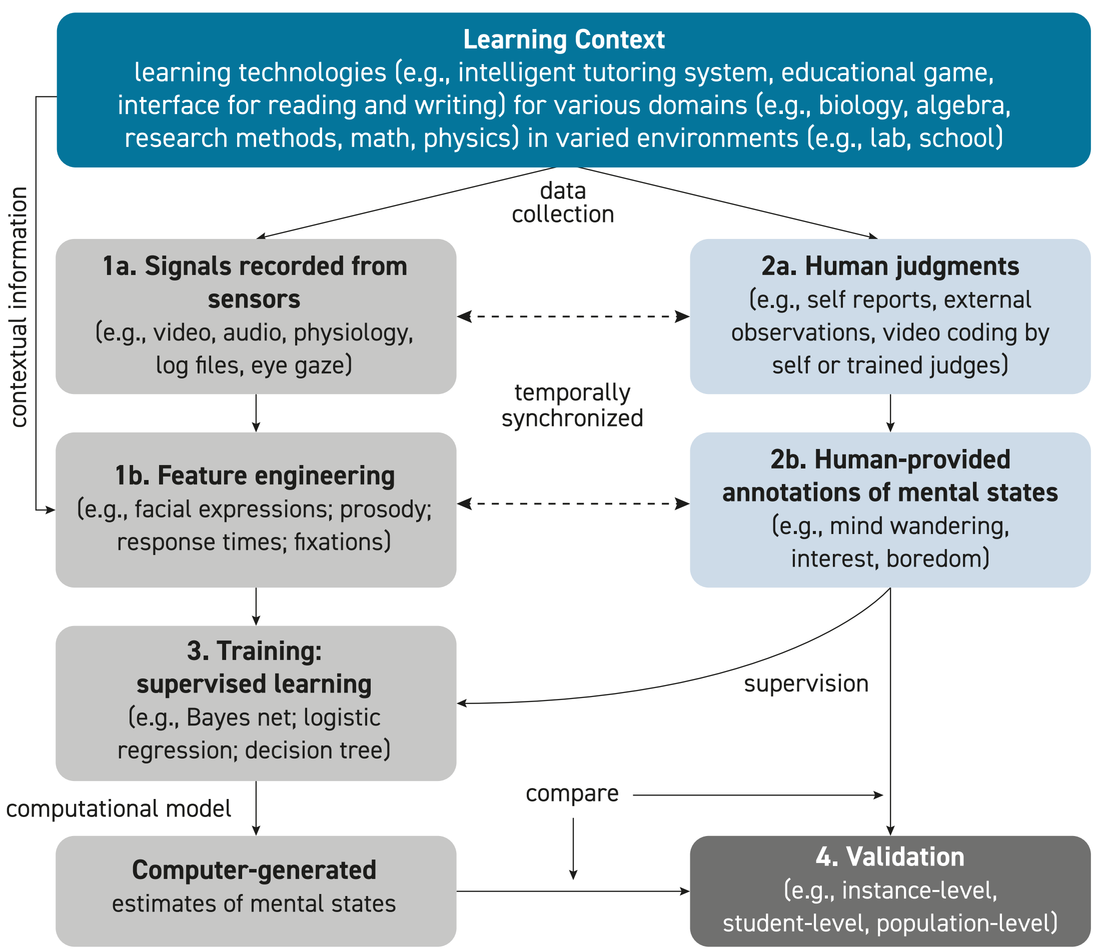
In Step 2, *annotations* of mental states reflecting various components of engagement are obtained from the students themselves, from external observers, or via some other method (see Porayska-Pomsta et al., 2013[71]) for a review). As an example, students can be asked to self-report their level of interest every five minutes (Sabourin and Lester, 2014[72]) or a trained human observer can watch the ensuing interaction and code students' perceived affect in five-minute intervals (Ocumpaugh, Baker and Rodrigo, 2015[73]). The features and annotations need to be temporally synchronised so that there is an association among the two. For example, the number of mouse clicks and smiles in each five-minute interval can be aligned to the self- or observer- reports of engagement in the same interval.

Step 3 in the AAA measurement approach involves supervised learning (a machine learning subfield), which learns to associate features (extracted from signals recorded by sensors as noted above) with temporally synchronised annotations of mental states (e.g. from self-reports or observer judgments) collected at multiple points in a learning session and/or from multiple students ideally in varied contexts. The output of supervised learning is a computational model (or a computer programme) which produces computer-generated annotations, which replaces the human-provided annotations. As summarised below:

1. features + human-provided annotations → *computational model*
2. features + computational model → *computer-generated annotations*

In Step 4, the computer-generated annotations are compared to the human-provided annotations to *validate* the model. The aim is for computer-generated annotations to match the human-provided annotations (accuracy) when applied to new or unseen data, for example, from a different set of students (generalisability). Once the model has been sufficiently validated, it can be deployed. Based on the sensor-collected data at some future time and/or from a new set of students, the model automatically generates estimates (annotations) of engagement. Figure 4.4 presents an overview of the approach.

Figure 4.4 Overview of the AAA-approach.



Source: Reproduced from D'Mello, Dieterle and Duckworth (2017[68])

It is important to consider privacy concerns for measures that use biometric signals (e.g. an image of a face, an audio sample). Some sensors can also inadvertently record sensitive information, as in the "WebcamGate" scandal where school authorities in a U.S. district remotely accessed students' school-issued laptops to take pictures of students in their own homes, their chat logs, and information on the websites they visited without informing students or parents (Martin, 2010[74]). An effective strategy to protect privacy is to only retain non-identifiable features from the signals while discarding the signals themselves as done in Bosch and D'Mello (2019[75]). In addition to privacy, there are important concerns pertaining to bias and fairness of the underlying models so it is important that training data is representative of different subgroups and the models demonstrate equivalent performance across subgroups (REFS) (Gardner, Brooks and Baker, 2019[76]; Jensen, 2019[77]). There are also ethical concerns with respect to how the measures are used. Their use for teacher or student evaluation is not recommended since the measures are imperfect and engagement is influenced by factors out of the control of teachers and students. Continual use is also not recommended as this can result in students rightly feeling monitored by these technologies – a particular concern for historically marginalised student populations. As we elaborate in the next section, these measures are best used in efforts to improve learning technologies, either by passively measuring periods of disengagement for retrospective review and refinement (e.g. Miller et al., 2014[78]) or by dynamically re-engaging disengaged students (D'Mello et al., 2016[79]; De Falco, Baker and D'Mello, 2014[80]). Their usage should be limited and ideally based on opt-in consent by students and caregivers.

Examples of engagement measures with the advanced, automated, analytic approach

D'Mello, Dieterle and Duckworth (2017[68]) provide examples of 15 studies that have utilised the AAA-approach to measure engagement across a range of learning technologies (e.g. intelligent tutoring system, educational game), learning domains (e.g. algebra, biology), student populations, operationalisations of engagement, methods used to obtain human annotations, supervised classification methods, and validation methods. They organised their review based on the sensor used for measurement. Sensor-free measures analyse digital traces recorded in log files (e.g. Bixler and D'Mello, 2013[81]; Gobert, Baker and Wixon, 2015[82]; Hutt, Grafsgaard and D'Mello, 2019[83]; Pardos et al., 2013[84]), while sensor-based measures use physical sensors. The latter can be further categorised as sensor-light if they use sensors that are readily available in contemporary digital devices, such as webcams and microphones (e.g. Bosch and D'Mello, 2019[75]; Bosch et al., 2016[85]; Forbes-Riley and Litman, 2011[86]; Monkaresi et al., 2017[87]; Pham and Wang, 2015[88]; Whitehill et al., 2014[89]) or sensor-heavy if they require non-standard sensors like eye trackers (e.g. Bixler and D'Mello, 2016[90]; Conati, Aleven and Mitrovic, 2013[91]; Hutt et al., 2019[92]), pressure pads (e.g. D'Mello and Graesser, 2009[93]; Mota and Picard, 2003[94]), and physiological sensors (e.g. Blanchard et al., 2014[95]; Dindar et al., 2017[96]; Mills et al., 2017[97]). Currently, sensor-free approaches are the most viable for learning technologies, but sensor-light approaches are gaining prominence and are expected to be major contenders in the next decade. Some research has also combined the two (e.g. Bosch et al., 2015[98]; D'Mello and Graesser, 2010[99]; Grafsgaard et al., 2014[100]; Kapoor and Picard, 2005[101]). Some examples of AAA-based engagement measures are listed below.

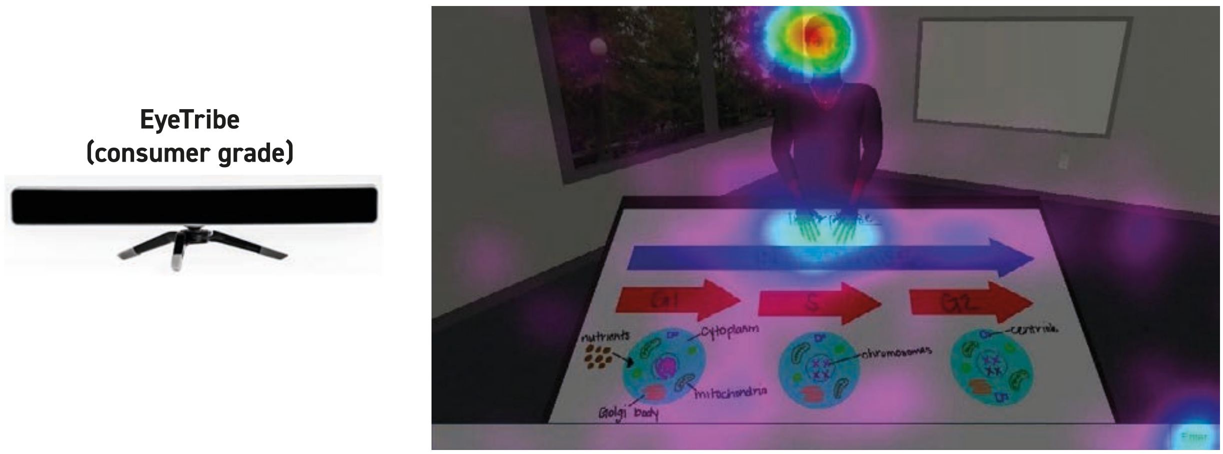
Interaction patterns to detect disengagement from task goals

Gobert, Baker and Wixon (2015[82]) developed an AAA-based measure for Inq-ITS, a computer-based learning environment with intelligent tutoring system to help students develop scientific inquiry skills. They focused on detecting if learners were disengaged from the task goal (DTG), defined as "engaging with the task, but in a fashion unrelated to the learning task's design goals or incentive structure" (p. 48). They collected training data from 144 middle-school U.S. students who used Inq-ITS as part of their science classes. Two humans coded disengagement from the task goal from human-readable excerpts (called clips) of Inq-ITS log files. Features used for supervised classification included the total number of actions, time between actions, duration of the longest pause, and number of simulations run. The researchers obtained moderate accuracies (about 41% above guessing) in discriminating disengagement from the task goal from clips with no disengagement from the task goal, highlighting the utility of the approach.

Eye-gaze to detect mind-wandering (zoning out)

Hutt et al. (2019[92]) used commercial off-the-shelf (COTS) eye trackers to track mind-wandering when high-school students completed tutorial sessions with a biology intelligent tutoring system called GuruTutor (Olney et al., 2012[102]) in their regular biology classroom (Figure 4.5). They asked students to self-report mind-wandering by responding to a question asking whether they were thinking about the learning context or something else entirely throughout the learning session. They computed eye gaze features (e.g. number of fixation, fixation durations, blinks) in 30-second windows preceding each probe and trained supervised classifiers to discriminate between positive and negative probe responses from the gaze features. The models were very successful in detecting mind-wandering, with accuracies more than double that of chance (guessing). Importantly, the models' predictions of mind-wandering negatively correlated with learning outcomes similar to self-reported mind-wandering. The researcher embedded the models into GuruTutor to provide real-time mind-wandering estimates for evaluation and to drive interventions (see reactive approach below).

Figure 4.5 Using consumer-grade eye tracker (left) to monitor visual attention while students interact with Guru (right) in classrooms



Source: Hutt et al. (2019[92])

Facial features, body moments, and interaction patterns to detect affect

Bosch et al. (2016[85]) developed an AAA-based measure of engagement while students played an educational game called Physics Playground (Shute, Ventura and Kim, 2013[103]); the game is described in more detail in the next section. They collected training data from 137 8th and 9th grade U.S. students in two 55-minute sessions across two days. Trained observers performed live annotations of boredom, engaged concentration, confusion, frustration, and delight using an observation protocol (Ocumpaugh, Baker and Rodrigo, 2015[73]), which were synchronised with videos of students' faces and upper bodies. The researchers extracted body movement and various facial expressions (e.g. smiles) from the videos and combined them with interaction features (e.g. number of restarts) extracted from the game log files. Supervised learning models trained to discriminate each affective state from the others (e.g. boredom vs. confusion, frustration, engaged concentration, and delight) yielded moderate accuracies (about 37% improvement over chance).

A sensor-free approach to measuring engagement during online learning

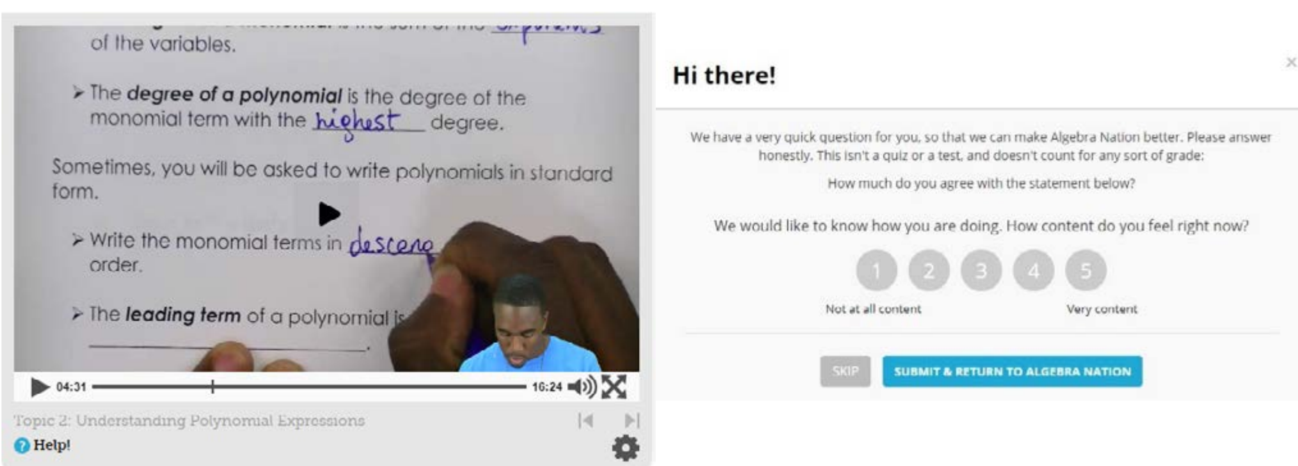
The above three examples highlight a variety of AAA-based engagement measures. However, these and all extant measures have been developed using data from a small number of students collected over one or two class sessions. As such, they serve as a proof-of-concept of the overall idea, but may not be sufficiently robust for real-world use. In contrast, Hutt, Grafsgaard and D'Mello (2019[83]) used the AAA-approach to develop a sensor-free measure of student engagement with an eye for scalability to tens of thousands of diverse students across extended time frames of an entire school year and beyond. The research was conducted in the context of Algebra Nation (Figure 4.6), an online maths learning platform that supports over 150 000 students studying Algebra 1, Algebra 2, and Geometry. For each topic, students can watch a video lecture from one of several human tutors. They can also use the Test Yourself quiz feature, which randomly selects 10 questions aligned with state standards. Students can review feedback on their answers or watch solution videos. Lastly, students can get more help through the Discussion Wall where they can interact with other students and study experts hired by Algebra Nation. Students can earn karma points by answering questions posted by other students.

The researchers collected a large-scale dataset of 69 174 students who used Algebra Nation as part of their regular maths classes for a semester. They used experience sampling to collect 133 966 self-reports (on a 1 to 5 scale) of 18 mental states (e.g. boredom, confusion, mind-wandering, curiosity, interest) related to engagement. They computed generic activity features (e.g. viewing a video, pausing a video taking a quiz) extracted from Algebra Nation log files; these features do not require specialised sensors and are domain- and (to a certain extent) system-independent. In all, they counted the occurrence of 22 such features in 5-minute windows prior to a self-report survey. They trained supervised learning models to predict each affective state from the features. Prediction accuracies, quantified with Spearman's rho (a correlation coefficient ranging from -1 to 1), were modest and ranged from .08 (for surprise) to .34 (for happiness), with a mean of .25.

The researchers tested the generalisability of the engagement models in several ways. First, they showed that the models trained on Algebra students generalised to a different data set of Geometry students (n = 28 458) on the same platform. They also studied the models' generalisability to clusters of students based on typical platform use and demographic features. They found that models trained on one group performed similarly well when tested on the other groups, although there was a small advantage of training individual subpopulation models compared to a general (all-population) model.

These results show the promise of scaling up sensor-free AAA-based engagement detection on the largest and most heterogeneous student sample to date using generic activity features that are not specific to a particular domain or system. The models have been embedded in the Algebra Learning platform, where they are being used as a component of a personalised system that recommends activities to students based on their ability and level of engagement (from the current models). The study is ongoing so the effectiveness of the approach awaits further research developments.

Figure 4.6 Sample video (left) in the Algebra Nation platform along with a self-report engagement questionnaire (right)



Source: Algebra Nation (n.d.[104])

Improving engagement

Learning technologies have traditionally mainly focused on promoting knowledge and skills. The unstated assumption underlying this approach is that cognition is all that matters, or at the very least, what really matters. Accordingly, emotion and motivation have been relegated to secondary design considerations, if considered at all. This was a reasonable assumption in the early days of learning technologies given that the dominant learning theories of the time put a premium on knowledge and skill acquisition (e.g. Anderson, 1982[105]; Brown and VanLehn, 1980[106]; Sleeman and Brown, 1982[107]). The benefit of hindsight tells us that this assumption is problematic since students need to be engaged to learn and it is very difficult to engage students with these traditional learning technologies. For example, intelligent tutoring systems (ITSs), which emulate one-on-one instruction of human-human tutoring, are effective in promoting learning for meta-analyses (see Ma et al., 2014[108]; Steenbergen-Hu and Cooper, 2013[109]; Steenbergen-Hu and Cooper, 2014[26]), yet students experience high levels of boredom while learning from intelligent tutoring systems (Craig et al., 2004[110]; D'Mello, 2013[111]; Hawkins, Heffernan and Baker, 2013[112]). Further, the basic functioning of the human cognitive system makes it difficult to sustain attention – a core component of engagement – even when one is motivated to do so. For example, students experience lapses in attention in the form of "zone outs" around 30% of the time during learning from technology (D'Mello, 2019[113]). Whereas some amount of zoning out and other forms of disengagement are normal occurrences in learning, higher amounts of zoning out is negatively related with learning outcomes (Randall, Oswald and Beier, 2014[114]; Risko et al., 2013[115]).

Can we design learning environments that are more sustainably engaging and thus enhance students' learning? It is only within the past two decades that researchers have made a concentrated effort to *design for engagement* (del Soldato and du Boulay, 1995[116]). Researchers have converged upon two main strategies to improve engagement in learning technologies: designs that promote engagement from the ground-up (proactive) or embedding mechanisms to monitor engagement in real-time and dynamically intervene when disengagement occurs or is imminent (reactive). It is also possible to combine the two approaches.

Proactive approaches

Proactive digital learning technologies are specifically designed to promote engagement and learning. Such systems aspire to increase the likelihood that the learner will experience cognitive and affective states which are generally positively associated with engagement (e.g. interest, curiosity, deep reasoning, critical thinking, being alert), while decreasing those which are usually negatively associated with engagement (e.g. boredom, zone outs, shallow processing).

It is important to distinguish between subtle or light-touch approaches that elicit mild engagement versus attempts to deeply engage students. For example, *gamification* – including elements of games into traditional learning technologies via points, challenges, badges, leaderboards and the like (Gibson et al., 2015[117]; Kapp, 2012[118]), see Box 4.1 as an example – can mildly increase engagement in the short term, but is unlikely to sustain deep engagement over extended time frames. Students might also resent the so-called "chocolate-covered broccoli" feel to some gamification attempts. Another subtle approach is *emotional design*, which involves altering digital learning materials to induce mild positive affect (e.g. adding facial anthropomorphisms to non-human graphical elements and/or adding pleasant colours in digital media) (Plass et al., 2014[119]; Um et al., 2012[120]). A recent meta-analysis (Brom, Stárková and D'Mello, 2018[121]) indicated that emotional design was surprisingly effective in increasing learning while also enhancing engagement as measured by intrinsic motivation, liking/enjoyment, and positive affect.

Promoting deep sustained engagement requires a fundamental reconceptualisation of the learning experience. The Interactive-Constructive-Active-Passive (ICAP) framework (Chi and Wylie, 2014[123]) provides a useful starting point. ICAP proposes four levels of engagement and learning based on the level of interactivity afforded by the learning activity. The levels, in decreasing order of expected engagement and learning, are Interactive > Constructive > Active > and Passive (Figure 4.7). An example of passive activity would include viewing a lecture or video with no overt behaviour, whereas taking verbatim notes without adding new ideas or organisation would be considered active. Generating a summary or self-explanation of lecture content by adding new ideas or reorganising old ideas would qualify as constructive. Interactive activities including some form of interaction or dialogue that accompany a constructive activity, for example, reciprocal peer tutoring, where students take turns tutoring each other. According to ICAP, one strategy to promote both engagement and learning is to incorporate more constructive and interactive experiences in learning technologies.

Educational games are one example of such learning technologies (Gee, 2003[41]). Well-designed educational games can enhance engagement and learning by turning work into play by minimising boredom, optimising engagement/flow, presenting challenges that are achievable with thought, creativity, and effort, and by engineering delight and pleasant surprises (Lepper and Henderlong, 2000[124]; Plass, Homer and Kinzer, 2015[125]; Ritterfeld, Cody and Vorderer, 2009[126]). Designing educational games can be quite challenging because game designers must balance a trade-off between game environments that are engaging but tangential to learning, and environments that promote deep learning but fail to foster engagement (Johnson and Mayer, 2010[127]). Well-designed games balance these goals by incorporating theoretically-grounded principles (problem solving, rules/constraints, challenge, control, ongoing feedback, and sensory stimulation (see Shute et al., 2014[128])) that make them intrinsically motivating and thus engaging (Fullerton, Swain and Hoffman, 2008[129]; Malone and Lepper, 1987[130]; Shute, Rieber and Van Eck, 2011[131]).

Box 4.1 Encouraging good behaviour and social learning with video games in Canada and the United States

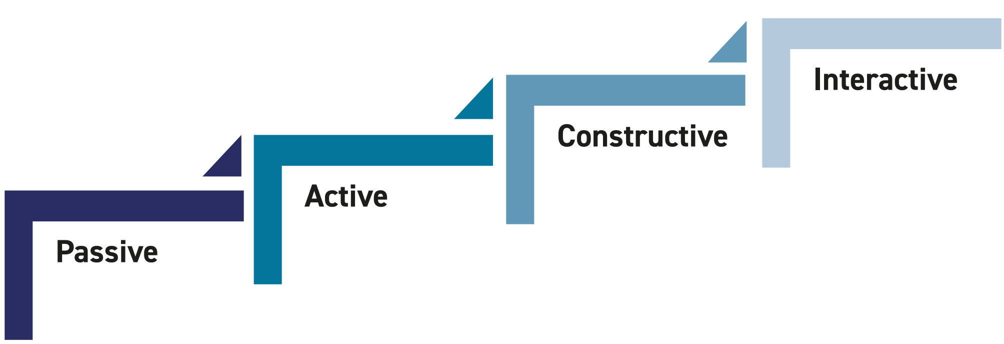
Fans of the Harry Potter series understand the significance of earning points for student behaviour. Teachers at Hogwarts school of witchcraft and wizardry award and take away points to reward or punish students for their behaviour. Students can earn points for their House by doing good deeds, answering questions correctly, and of course by winning Quidditch matches. By doing the opposite, students may lose points (and disappoint their fellow students in the same House). At the end of the school year, the House with most points wins the prestigious House Cup.

Real life may lack the savvy magic tricks of Hogwarts' teachers, but Québec- and New York-based company Classcraft tried to replicate engagement in learning with similar methods. Inspired by video games such as World of Warcraft, its cloud-based platform for secondary education was described as a "role-playing game for classroom management" (Sanchez, Young and Jouneau-Sion, 2016[122]).

Classcraft aims to help students learn about appropriate behaviour, like doing homework and not being late for class. The game is not related to any specific subject; the duration of a game can last from just one class to an entire school year. Classcraft is foremost a web application for use in class, although students can download the application on their mobile phones for use outside class too. Unlike most video games, Classcraft does not offer a 3D game world. Instead it offers a form of augmented reality: by adding a digital layer to the real world, the game creates new interactions between teachers and students.

Students play in teams of 4 to 6 students. At the beginning of the game, they each get to choose different avatars (healers, mages, warriors) with their own strengths and weaknesses. Their goal is to gain experience points and prevent their avatars from losing health points. By acting according to the criteria for appropriate behaviour, they earn points which they can for example use to upgrade their avatar or to help their teammates. Doing the opposite can of course also lead to losing points, which may affect the entire team. Teachers operate as "game masters": they distribute or remove points, and they can create certain random events in the game or certain dynamics in real life (like asking to fulfil specific tasks) which affect all students. Updates in the game appear in real-time in every player's application. Interaction with real life is at the core of the model: winning or losing virtual points can lead to teacher actions in real life, including praise, rewards, and perhaps even reprimands. Because the game also connects the health points of one student with other students in the same team, students are encouraged to work together. The use of Classcraft – as well as comparable software programmes for classroom management – may have a positive impact on student engagement, collaboration, and even on the school climate (Edweek).

Figure 4.7 Organisation of activities with respect to expected engagement and learning (from left to higher) as per the ICAP framework



Source: D'Mello.

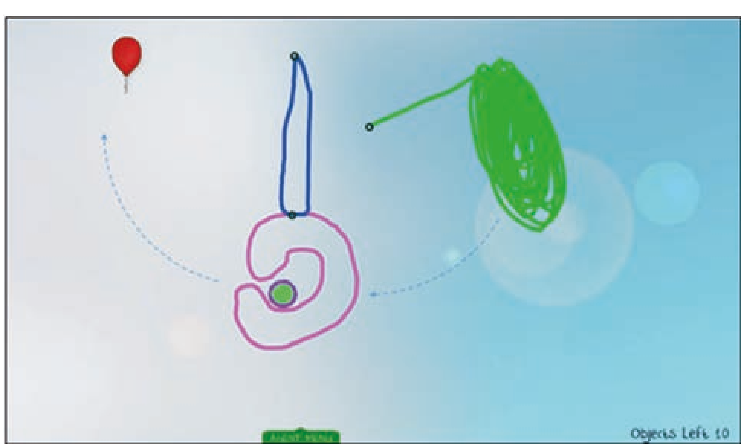
Well-designed games also directly embed meaningful learning experiences throughout gameplay. As such, playing well-designed games has been shown to be positively related to various learning competencies and outcomes such as visual-spatial abilities and attention (Green and Bavelier, 2012[132]; Green and Bavelier, 2007[133]; Shute, Ventura and Ke, 2015[134]), college grades (Skoric, Teo and Neo, 2009[135]; Ventura, Shute and Kim, 2012[136]), persistence (Ventura, Shute and Zhao, 2013[137]), creativity (Jackson et al., 2012[138]), and civic engagement (Ferguson and Garza, 2011[139]), as well as valuable academic content and skills (e.g. Coller and Scott, 2009[140]; DeRouin-Jessen, 2008[141]); for a review, see Tobias and Fletcher (2011[142]); Wilson et al. (2009[143]); Young et al. (2012[144]).

Two examples: Physics Playground and Crystal Island

As an example, consider Physics Playground (Shute, Ventura and Kim, 2013[103]), a highly engaging educational game for learning Newton's laws of force and motion, linear momentum, energy, and torque (Figure 4.8). The game obeys the basic rules of physics via a formal simulation of a virtual physics "world" and dynamically responds to players' interactions with the game. The primary goal is for players to guide a green ball to a red balloon, resulting in "solving" the level, and to do so, players must create *agents* – ramps, pendulums, levers, and springboards – that "come to life" on the screen. Game elements in Physics Playground include realism (accomplished via detailed formal simulation of a virtual physics "world"), ongoing feedback, interactive problem solving, and adaptive challenges. The game also gives players the freedom to try/fail by experimenting with a variety of solutions. Physics Playground is successful at increasing both engagement and learning by integrating physics concepts with authentic gameplay. Box 4.2 shows that the game also provides a good environment to research the relations between students' emotions and learning internationally.

Game-based learning can be further enhanced by increasing the immersion and realism of the experience. In traditional classroom settings, project-based learning (PBL) provides students or student teams with authentic, personally-meaningful, real-world problems over an extended period of time (Blumenfeld et al., 1991[145]). For example, student teams might take on the challenge of improving the water quality in their community; this requires an understanding of the water system, pollutants, purification, policy, chemistry, and so on. Thus, in addition to keeping students engaged through choice, realism, and collaboration, well-designed PBL curricula can help students acquire content knowledge along with real-world skills (e.g. inquiry, collaboration, scientific reasoning; see Schneider et al., 2020[146]). Unfortunately, implementing these learning experiences require extensive curricula, access to physical resources, and human scaffolding, and are difficult to realise via digital technologies alone. To address this challenge, researchers have explored the use of augmented and virtual reality to develop *immersive* digital learning experiences for students. Immersion is defined as the subjective perception of experiencing reality – akin to a suspension of disbelief – and is hypothesised to increase engagement and learning (Dede, 2009[147]). Examples include River City (Dede, 2009[147]), Quest Atlantis (Barab et al., 2005[148]), and Crystal Island (Sabourin and Lester, 2014[72]), which is discussed briefly below.

Figure 4.8 Example of a problem and pendulum solution in Physics Playground



Source: Shute, Ventura and Kim (2013[103])

Crystal Island (Rowe et al., 2009[149]; Spires et al., 2011[150]) is an immersive educational game that capitalises on the principle of narrativity, which posits that learners benefit most from educational games that weave a narrative theme into the game experience. In Crystal Island, the learner takes on the role of a protagonist named Alex who arrives on the island. Alex discovers that members of a research team are falling ill and is charged with identifying the source of an infectious disease. Alex proceeds with the investigation by visiting various areas of the island (dining hall, lab, inferminary, dorm), by interviewing other islanders, and by manipulating objects in the world. Through the generation of questions and hypotheses, and the collection and analysis of data, learners make gradual steps towards diagnosing the cause of the disease. Thus, Crystal Island embeds learning of microbiology knowledge and development of critical thinking and inquiry learning skills in an engaging narrative platform.

Evidence from learners' interactions with Crystal Island indicate that it is highly engaging and motivating and can improve learning outcomes (Rowe et al., 2009[149]; Rowe et al., 2010[151]). For example, Sabourin and Lester (2014[72]) conducted a study in which 450 8th grade students interacted with Crystal Island for 55 minutes. Students self-reported one of seven affective states (anxious, bored, confused, curious, excited, focused, and frustrated) at five-minute intervals during gameplay. The researchers found that self-reports of boredom (8%) were considerably lower than reports of excitement (13%) and curiosity (19%). Confusion (16%) and frustration (16%) were also quite frequent, suggesting that the game was deeply engaging students as these affective states are frequent in deep learning activities (D'Mello and Graesser, 2012[152]). In addition, students demonstrated significant learning gains of about 30%, assessed as the percent improvement of post-test over pre-test scores.

Recent advances in scalable augmented and virtual reality technologies have further closed the gap between the virtual and physical world (for a review, see Martín-Gutiérrez et al., 2017[154]), especially in terms of immersion and realism (Psotka, 1995[155]). Despite the enthusiasm for these technologies, and some initial promising results (e.g. Ibáñez et al. (2014[156])), there has yet to be rigorous research studies evaluating their learning effectiveness, particularly when it comes to conceptual rather than procedural or rote learning.

Box 4.2 Using digital environments to research learning emotions in the Philippines

Physics Playground was used in the Philippines to analyse the positive or negative relationships of different affective states to learning, including frustration, the affective state expressing annoyance and dissatisfaction. Using a standardised human-based codifying system for students' emotions while playing the game, and a different task (puzzle) than the one described in the main text, Banawan, Rodrigo and Andres (2015[153]) found that, after engaged concentration (79% of observed affective states), frustration (8%) was the second most experienced state among the 8 that were monitored, and the only one that had a statistical significant relationship with achievement – a negative relationship. The research shows that frustration does not behave like confusion, which can be positively related to achievement once solved. It also highlights the contextual nature of engagement, as Filipino students did not experience the same levels of enjoyment as the U.S. students using Physics Playground. This shows the importance of designing digital learning environments that remain sensitive to possibly different ways to deal with failure in learning and designing the right scaffoldings to help students cope with their negative feelings.

Reactive approaches

Consider the following hypothetical scenarios to motivate the reactive approach to improving engagement: "Imagine you are helping your niece prepare for an upcoming examination in evolutionary biology. Things started off quite well, but after a while, you realise that her mind is a million miles away. Although the plan is for the two of you to collaboratively model genetic frequency shifts in populations, you notice that her attention has drifted to unrelated thoughts of lunch, the football game, or an upcoming vacation. You might try to momentarily reorient her attention by asking a probing question. However, if her attentional focus continues to wane, you realise that you must adapt your instruction to better engage her by altering the course of the learning session. You shift the initiative from a collaborative discussion to a student-centred perspective by asking her to develop a strategy for tracking genetic changes in populations. This works and she appears to tackle this task with a renewed gusto and the session progresses quite smoothly. However, sometime later, you notice that she actually appears to be nodding off as you delve into the fundamentals of allele frequencies. So, you suggest switching topics or even taking a break, thereby giving her an opportunity to recharge."

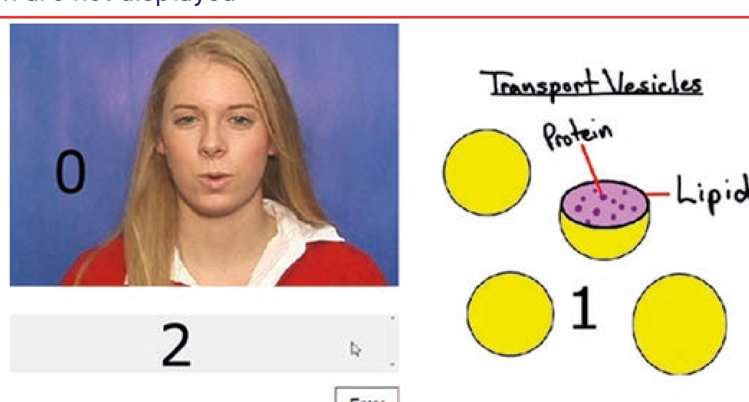
The example above provides a sense of what a reactive agent – a human in this case – can accomplish. Reactive approaches to increasing engagement focus on automatically detecting student engagement and dynamically *responding* when engagement appears to be declining, or providing motivating feedback when engagement is high (D'Mello and Graesser, 2015[157]). These approaches assume that engagement is a fluid, dynamical process that will wax and wane as the learning session progresses. Despite the best intentions of the technology designer to deliver an engaging experience, there will be individual differences in the extent to which a learning technology successfully engages a learner. Further, even if a learner is fully engaged at the onset of learning, engagement will eventually decrease over time as novelty fades and fatigue increases.

Reactive approaches are more sophisticated than proactive approaches. For one, the form of dynamic adaptivity exemplified above requires the ability to continually monitor engagement, to detect when engagement is waning, and to adapt instruction to address periods of disengagement. The technology could focus on a specific component of engagement, or could measure the construct more holistically. Engagement measurement can be achieved using the AAA-based approach discussed previously. A reactive learning technology must then alter its pedagogical/motivational strategies in response to the sensed engagement. It has multiple paths to pursue. It could do nothing if the learner is engaged and is on a positive learning trajectory. It could attempt to reorient attention if it detects that the learner is zoned out or distracted (D'Mello, 2016[158]). Hints and just-in-time explanations can be provided when confusion or frustration is detected (Forbes-Riley and Litman, 2011[86]). The system could provide choice, encourage breaks, or adjust the levels of challenge when it detects that a student is bored. It can also provide supportive messages which aim to motivate students to persevere in the learning (D'Mello and Graesser, 2012[159]; DeFalco et al., 2018[160]). If the technology is embodied in some form, such as with an animated pedagogical agent, it can utilise various social cues to boost engagement, for example by mirroring facial expressions and gestures (Burleson and Picard, 2007[161]) or looking unhappy when a disengaged behaviour is detected (Baker et al., 2006[162]). Examples of reactive learning technologies that implement some of these strategies are discussed below.

Responding to students' inattention with Gaze Tutor

D'Mello et al. (2012[163]) respond to behavioural engagement during multimedia learning. The interface consists of an animated conversational agent that provides explanations on biology concepts with synthesised speech that is synchronised with annotated image (Figure 4.9). The system uses an eye tracker to track when a student is not attending to the important parts of the interface (i.e., the tutor or image). Gaze Tutor simply assumes that learners are disengaged when their gaze is not on the tutor or image for at least five consecutive seconds. When this occurs, it attempts to reengage learners with "witty" statements that directs them to reorient their attention towards the agent or the image (e.g. "I'm over here you know" and "Snap out of it. Let's keep going.") Preliminary results suggest that these gaze-sensitive statements are successful in reorienting attention and improve learning outcomes.

Figure 4.9 GazeTutor interface with animated agent (0), image panel (1), and input box (2). Blank screen areas on the bottom are not displayed



Source: D'Mello et al. (2012[163])

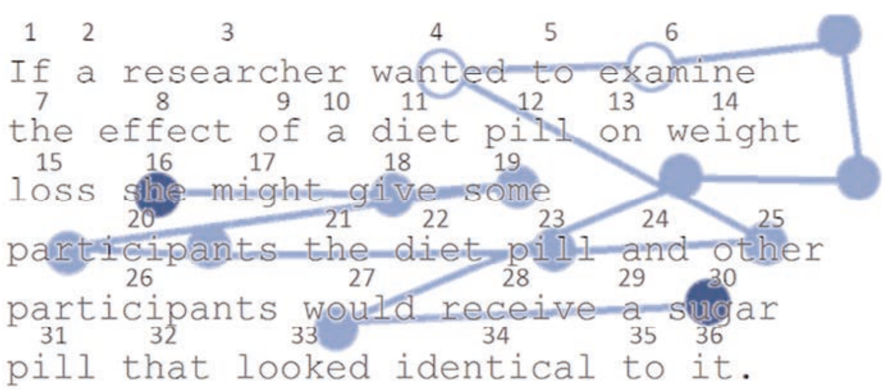
Responding to students' uncertainty with UNC-ITSPOKE

Forbes-Riley and Litman (2011[86]) was designed to examine whether automatic responses to learner uncertainty could improve learning outcomes while students engage with a spoken conversational intelligent tutoring system for physics. Uncertainty is related to confusion, an associated state that aligns with both the cognitive and affective components of engagement (D'Mello and Graesser, 2014[164]). UNC-ITSPOKE automatically detects learners' certainty/uncertainty (from speech and language using an AAA-approach) in addition to the correctness/incorrectness of a response. It provides an explanation-based sub dialogue when the student is correct but uncertain about a response as this suggests a metacognitive failure. A validation study indicated that this form of adaptivity achieved slightly (but not significantly) higher learning outcomes compared to control conditions.

Responding to students' mind-wandering with eye-mind reader

D'Mello et al. (2017[165]) and Mills et al. (2020[166]) respond to instances of mind-wandering (zone outs) during computerised reading (Figure 4.10). It uses an eye-gaze-based AAA-approach to detect mind-wandering (Faber, Bixler and D'Mello, 2018[167]) on a page-by-page basis (a page is a screen of text), and dynamically responds with comprehension assessments and re-reading opportunities. One initial intervention strategy consists of asking rote comprehension questions on the page where mind-wandering is detected and providing opportunities to re-read if the question is answered incorrectly. A validation study (D'Mello et al., 2017[165]) indicated that this intervention had the intended effect of reducing comprehension deficits attributable to mind-wandering in specific cases. However, since the interpolated questions were rote questions, this encouraged keyword spotting and a general shallow-level processing style. To address this, Mills et al. (2020[166]) replaced the rote multiple-choice questions with deeper-level questions that required learners to generate self-explanations in natural language. These were automatically scored and the system responded with feedback and provided opportunities to re-read and revise the explanations. Results suggest a positive effect for the intervention strategy compared to equivalent controls on retention (i.e. learning assessments completed a week later).

Figure 4.10 Eye tracking during reading. Filled circles display fixations (where eyes are focused) and lines display saccades (rapid eye movements between fixations)



Source: D'Mello et al. (2017[165])

Reactive technologies in the classroom

Most research on reactive learning technologies has mainly been done in in the lab, but this line of work is gradually moving to classrooms. Aslan et al. (2019[168]) developed Student Engagement Analytics Technology (SEAT) to assist teachers with the task of monitoring and responding to student real-time behavioural and emotional engagement. Here, the engagement measures provide assessments of student engagement to the teacher, who then can decide if and how to intervene. This research was done with students in Turkey who used a self-paced maths educational platform as part of their classes. An AAA-based engagement measure monitored facial features extracted from webcam video, patterns of interaction with the educational platform, and browser URL logs. It combined behavioural (on- or off-task) and emotional (bored, satisfied, confused) engagement assessments to yield an overall engagement score.

The engagement measure was then embedded in the SEAT interface which provided colour-coded (green, yellow, red) estimates of student engagement to teachers in real-time (Figure 4.11). Teachers used these data to intervene with individual students and their primary intervention strategies included verbal warnings, positive reinforcement, scaffolding (e.g. explaining the question, providing a hint), and close monitoring (watching the student's screen to ensure they were on-task).

The researchers tested SEAT in a 16-week research study using both qualitative and quantitative methods. Interviews indicated both teachers and students had positive experiences using SEAT. Teachers were proactive about using the interface and reported that it made it easier to monitor the needs of individual students in a large class. Students also reported receiving individual attention when they needed it and felt that they were more engaged with the learning session. Finally, students in a SEAT-enabled class had higher (but not significant) learning gains compared to a parallel class of students without SEAT. However, this result should be considered preliminary since the study used a quasi-experimental design and only 37 students were tested across the two classes.

Figure 4.11 SEAT interface displaying the overall class-view (left) and student-specific view (right).

Source: Reproduced from Asland et al. (2019[168])

Conclusions and future directions

Research on engagement during learning has exploded in past years. Not only is the science of engagement in education advancing, but so are technology designs that make engagement a central design feature. Shared among the scientific research and technology design communities is a perspective that recognises engagement as an outcome worthy of consideration in its own right. There is a risk, however, that each proceeds independently, thereby missing obvious benefits of the two progressing in tandem.

To illustrate, the last two decades have produced a flurry of theoretical and empirical research aimed at defining engagement, identifying its causes and effects, and devising interventions to increase it (Christenson, Reschly and Wylie, 2012[18]). Unfortunately, the science of measuring engagement has not kept up; most researchers exclusively rely on self-report questionnaires as the sole measure of engagement (Fredricks and McColskey, 2012[42]), partly because self-reports are convenient and familiar, and perhaps because accessing more advanced technologies is more costly and intimidating. At the same time, in many countries, the computer is becoming an indispensable component in everyday learning (see Vincent-Lancrin et al., 2019[169]). The increased bandwidth of available information in digital learning contexts also suggests that it may be unwise to exclusively rely on 20th-century methods to measure engagement in 21st-century digital learning contexts. Scientific research on engagement would benefit by considering digital learning experiences and incorporating associated methods into study designs.

On the flip side, the influx of technology in classrooms and the rise of online learning technologies produce huge volumes of data, which in principle, should lead to transformative advances in the measurement and facilitation of engagement. Unfortunately, this has not been the case yet. Despite the impressive volume, much of the data generated by these technologies is not very rich in depth, resulting in shallow measures of engagement that focus on very basic behavioural patterns, such as the number of logins, video plays, and so on. Further, despite many benefits, the proliferation of online learning has several side effects, most troublesome being a return to a transmission mode of education with its ineffectual passive learning strategies and lack of meaningful collaboration, which is one of the most engaging and productive ways to learn (Dillenbourg, 1999[170]; Stahl, Koschmann and Suthers, 2006[171]). Educational technology will benefit by incorporating principles from the scientific research on engagement into technology design.

Where will we be in the next 10 to 15 years? A pessimistic view is that research will proceed as usual, with scientific research on engagement chugging along in one direction and educational technology research proceeding in another, with rarely the two meeting. A more optimistic view is for the two to work hand in hand, resulting in learning technologies that meaningfully engage students and the emergence of more research on engagement with digital learning technologies. In this future, learning games – one of the best known experiences to foster deep engagement (as reviewed above) – also incorporate strategies such as scaffolds and learning supports that foster deep conceptual learning while also keeping learners engaged and motivated. And reactive-based approaches to improve engagement, which dynamically respond based on sensed engagement, will be deployed more readily as sensors become increasingly ubiquitous, wearable, and cost-effective; rigorous safeguards to protect privacy are established; efforts are made to ensure that the AI models underlying the technologies are unbiased and fair; and interdisciplinary collaborations between education and computer science researchers increase.

Importantly, these technologies should be developed with ethics, equity, and justice as core design considerations, not design after-thoughts, and technology designers seriously answer the question of what intelligent technologies should do rather than what they can do. There should be careful thought given to what data is collected and for how long it is stored, if at all, and end users (students, parents, teachers) should have the final say with respect to whether and when engagement is assessed. In this future, the ideas and research outlined in this paper will leave the confines of the research lab and small-scale studies in the real world to become a core component of digital learning that benefits millions of students worldwide. Here, opt-in measurement of engagement aimed for formative and non-evaluative purposes (while protecting privacy of students) coupled with intelligent strategies to enhance engagement will be essential elements of digital learning technologies. A future of efficient, effective, and engaging digital learning awaits us all.

References

Algebra Nation (n.d.), *Algebra Nation*, https://loveforquotes.com/t/algebra-nation (accessed on 29 April 2021). [104]

Anderson, J. (1982), "Acquisition of cognitive skill", *Psychological Review*, Vol. 89/4, pp. 369-406. [105]

Arnold, K. and M. Pistilli (2012), *Course signals at Purdue: Using learning analytics to increase student success.* [63]

Aslan, S., N. Alyuz, C. Tanriover, S. Mete, E. Okur, S. D'Mello and A. Esmeet (2019), "Investigating the Impact of a Real-time, Multimodal Student Engagement Analytics Technology in Authentic Classrooms", in *Proceedings of the ACM CHI Conference on Human Factors in Computing Systems (CHI 2019).*, ACM, New York. [168]

Atkins-Burnett, S., C. Fernandez, L. Akers, J. Jacobson and C. Smither-Wulsin (2012), *Landscape analysis of non-cognitive measures.* [66]

Aung, A., A. Ramakrishnan and J. Whitehill (2018), *Who Are They Looking At? Automatic Eye Gaze Following for Classroom Observation Video Analysis,* Paper presented at the International Conference on Educational Data Mining (EDM). [59]

Baker, R., A.T. Corbett, K.R. Koedinger, S. Evenson, I. Roll, A.Z. Wagner, M. Naim, J. Raspat, D.J. Baker and J.E. Beck (2006), Adapting to when students game an intelligent tutoring system, Paper presented at the Intelligent Tutoring Systems Conference. [162]

Baker, R., S. D'Mello, M.M. Rodrigo and A. Graesser (2010), "Better to be frustrated than bored: The incidence, persistence, and impact of learners' cognitive–affective states during interactions with three different computer-based learning environments", *International Journal of Human-Computer Studies,* Vol. 68/4, pp. 223-241. [7]

Baker, R. and J. Ocumpaugh (2015), "Interaction-based affect detection in educational software", in Calvo, R. et al. (eds.), *The Oxford Handbook of Affective Computing*, Oxford University Press, New York. [70]

Banawan, M., M. Rodrigo and J. Andres (2015), A*n investigation of frustration among students using physics playground. In Doctoral Student Consortium (DSC),* Proceedings of the 23rd International Conference on Computers in Education. [153]

Barab, S., M.K. Thomas, T. Dodge, R. Carteaux et H. Tuzun (2005), "Making learning fun: Quest Atlantis, a game without guns", *Educational technology research and development*, Vol. 53/1, pp. 86-107. [148]

Bidwell, J. and H. Fuchs (2011), "Classroom analytics: Measuring student engagement with automated gaze tracking", *Behav Res Methods*, Vol. 49, p. 113. [60]

Bixler, R. and S. D'Mello (2016), "Automatic gaze-based user-independent detection of mind wandering during computerized reading", *User Modeling & User-Adapted Interaction*, Vol. 26, pp. 33-68. [90]

Bixler, R. and S. D'Mello (2013), "Detecting engagement and boredom during writing with keystroke analysis, task appraisals, and stable traits", in *Proceedings of the 2013 International Conference on Intelligent User Interfaces (IUI 2013),* ACM, New York, NY. [81]

Blanchard, N., R. Bixler, T. Joyce and S. D'Mello (2014), "Automated physiological-based detection of mind wandering during learning", in Trausan-Matu, S. et al. (eds.), *Proceedings of the 12th International Conference on Intelligent Tutoring Systems (ITS 2014),* Springer International Publishing, Switzerland. [95]

Blumenfeld, P., E. Soloway, R.W. Marx, J.S. Krajcik, M. Guzdial and A.M. Palincsar (1991), "Motivating project-based learning: Sustaining the doing, supporting the learning", *Educational Psychologist,* Vol. 26/3-4, pp. 369-398. [145]

Bosch, N., H. Chen, S. D'Mello, R. Baker and V. Shute (2015), "Accuracy vs. Availability Heuristic in Multimodal Affect Detection in the Wild", in *Proceedings of the 17th ACM International Conference on Multimodal Interaction (ICMI 2015),* ACM, New York, NY. [98]

Bosch, N., S. D'Mello, J. Ocumpaugh, R. Baker and V. Shute (2016), "Using video to automatically detect learner affect in computer-enabled classrooms", *ACM Transactions on Interactive Intelligent Systems,* Vol. 6/2, pp. 17.1-17.31. [85]

Bosch, N. and S. D'Mello (2019), "Automatic Detection of Mind Wandering from Video in the Lab and in the Classroom", *IEEE Transactions on Affective Computing*, http://dx.doi.org/10.1109/TAFFC.2019.2908837. [75]

Bransford, J., A. Brown and R. Cocking (2000), *How people learn*, National Academy Press, Washington, DC. [23]

Brom, C., T. Stárková and S. D'Mello (2018), "How effective is emotional design? A meta-analysis on facial anthropomorphisms and pleasant colors during multimedia learning", *Educational Research Review,* Vol. 25, pp. 100-119. [121]

Brown, J. and K. VanLehn (1980), "Repair theory: A generative theory of bugs in procedural skills", *Cognitive Science,* Vol. 4/379-426, http://dx.doi.org/10.1016/S0364-0213(80)80010-3. [106]

Burleson, W. and R. Picard (2007), "Evidence for gender specific approaches to the development of emotionally intelligent learning companions", *IEEE Intelligent Systems*, Vol. 22/4, pp. 62-69. [161]

Charsky, D. (2010), "From edutainment to serious games: A change in the use of game characteristics", *Games and culture*, Vol. 5/2, pp. 177-198. [20]

Chi, M. and R. Wylie (2014), "The ICAP framework: Linking cognitive engagement to active learning outcomes", *Educational Psychologist*, Vol. 49/4, pp. 219-243. [123]

Christenson, S., A. Reschly and C. Wylie (2012), *Handbook of research on student engagement,* Springer, New York. [18]

Coller, B. and M. Scott (2009), "Effectiveness of using a video game to teach a course in mechanical engineering", *Computers & Education,* Vol. 53/3, pp. 900-912. [140]

Conati, C., V. Aleven and A. Mitrovic (2013), "Eye-Tracking for Student Modelling in Intelligent Tutoring Systems", in Sottilare, R. et al. (eds.), *Design Recommendations for Intelligent Tutoring Systems - Volume 1: Learner Modeling,* Army Research Laboratory, Orlando, FL. [91]

Craig, S., S. D'Mello, A. Witherspoon and A. Graesser (2008), "Emote aloud during learning with AutoTutor: Applying the Facial Action Coding System to cognitive–affective states during learning", *Cognition & Emotion,* Vol. 22/5, pp. 777-788, http://dx.doi.org/10.1080/02699930701516759. [28]

Craig, S., A. Graesser, J. Sullins and B. Gholson (2004), "Affect and learning: An exploratory look into the role of affect in learning", *Journal of Educational Media*, Vol. 29, pp. 241-250, http://dx.doi.org/10.1080/1358165042000283101. [110]

Csikszentmihalyi, M. (1990), *Flow: The psychology of optimal experience,* Harper and Row, New York. [40]

Csikszentmihalyi, M. (1975), *Beyond boredom and anxiety,* Jossey-Bass, San Francisco, CA. [8]

Csikszentmihalyi, M. and R. Larson (1987), "Validity and reliability of the experience-sampling method", *The Journal of nervous and mental disease,* Vol. 175/9, pp. 526-536. [49]

Danckert, J. and C. Merrifield (2018), "Boredom, sustained attention and the default mode network", *Experimental Brain Research*, Vol. 236/9, pp. 2507-2518. [1]

Daniels, L., R.H. Stupnisky, R. Pekrun, T.L. Haynes, R.P. Perry and N.E. Newall (2009), "A longitudinal analysis of achievement goals: From affective antecedents to emotional effects and achievement outcomes", *Journal of Educational Psychology*, Vol. 101/4, pp. 948-963, http://dx.doi.org/10.1037/a0016096. [9]

De Falco, J., R. Baker and S. D'Mello (2014), "Addressing behavioral disengagement in online learning", in Sottilare, R. et al. (eds.), *Design Recommendations for Adaptive Intelligent Tutoring Systems: Adaptive Instructional Strategies (Volume 2),* Army Research Laboratory, Orlando, FL. [80]

Dede, C. (2009), "Immersive interfaces for engagement and learning", *Science*, Vol. 323/5910, pp. 66-69. [147]

DeFalco, J., R. Baker and S. D'Mello (2018), "Detecting and addressing frustration in a serious game for military training", *International Journal of Artificial Intelligence in Education,* Vol. 28/2, pp. 152-193. [160]

del Soldato, T. and B. du Boulay (1995), "Implementation of motivational tactics in tutoring systems", *International Journal of Artificial Intelligence in Education*, Vol. 6, pp. 337-378. [116]

DeRouin-Jessen, R. (2008), *Game on: The Impact of Game Features in Computer-Based Training.* [141]

Dillenbourg, P. (1999), "What do you mean by collaborative learning", in Dillenbourg, P. (ed.), *Collaborative-learning: Cognitive and computational approaches.* [170]

Dindar, M., J. Malmberg, S. Järvelä, E. Haataja and P.A. Kirschner (2017), "Matching self-reports with electrodermal activity data: Investigating temporal changes in self-regulated learning", *Education and Information Technologies*, pp. 1-18. [96]

D'Mello, S. (2019), "What do we think about when we learn?", in Millis, K. et al. (eds.), *Deep Comprehension: Multi-Disciplinary Approaches to Understanding, Enhancing, and Measuring Comprehension,* Routledge, New York, NY. [113]

D'Mello, S. (2016), "Giving Eyesight to the Blind: Towards attention-aware AIED", *International Journal of Artificial Intelligence in Education,* Vol. 26/2, pp. 645-659. [158]

D'Mello, S. (2013), "A selective meta-analysis on the relative incidence of discrete affective states during learning with technology", *Journal of Educational Psychology,* Vol. 105/4, pp. 1082-1099. [111]

D'Mello, S., E. Dieterle and A. Duckworth (2017), "Advanced, Analytic, Automated (AAA) Measurement of Engagement during Learning", *Educational Psychologist,* Vol. 52/2, pp. 104-123. [68]

D'Mello, S. and A. Graesser (2015), "Feeling, thinking, and computing with affect-aware learning technologies", in Calvo, R. et al. (eds.), *The Oxford Handbook of Affective Computing, Oxford University Press*, New York. [157]

D'Mello, S. and A. Graesser (2014), "Confusion", in Pekrun, R. and L. Linnenbrink-Garcia (eds.), *International handbook of emotions in education, Routledge*, New York, NY. [164]

D'Mello, S. and A. Graesser (2012), "AutoTutor and Affective AutoTutor: Learning by talking with cognitively and emotionally intelligent computers that talk back", *ACM Transactions on Interactive Intelligent Systems,* Vol. 2/4, pp. 23:2-23:39. [159]

D'Mello, S. and A. Graesser (2012), "Dynamics of affective states during complex learning", *Learning and Instruction,* Vol. 22/2, pp. 145-157, http://dx.doi.org/10.1016/j.learninstruc.2011.10.001. [152]

D'Mello, S. and A. Graesser (2010), "Multimodal semi-automated affect detection from conversational cues, gross body language, and facial features", *User Modeling and User-adapted Interaction,* Vol. 20/2, pp. 147-187. [99]

D'Mello, S. and A. Graesser (2009), "Automatic detection of learners' affect from gross body language", *Applied Artificial Intelligence*, Vol. 23/2, pp. 123 - 150, http://www.informaworld.com/10.1080/08839510802631745. [93]

D'Mello, S., K. Kopp, R.E. Bixler and N. Bosch (2016), "Attending to attention: Detecting and combating mind wandering during computerized reading", in *Extended Abstracts of the ACM SIGCHI Conference on Human Factors in Computing Systems (CHI 2016)*, ACM, New York. [79]

D'Mello, S., C. Mills, R. Bixler and N. Bosch (2017), "Zone out no more: Mitigating mind wandering during computerized reading", in Hu, X. et al. (eds.), *Proceedings of the 10th International Conference on Educational Data Mining,* International Educational Data Mining Society. [165]

D'Mello, S., A. Olney, C. Williams and P. Hays (2012), "Gaze tutor: A gaze-reactive intelligent tutoring system", *International Journal of Human-Computer Studies,* Vol. 70/5, pp. 377-398. [163]

D'Mello, S., R. Taylor, K. Davidson and A. Graesser (2008), "Self versus teacher judgments of learner emotions during a tutoring session with AutoTutor", in Woolf, B. et al. (eds.), *Proceedings of the 9th international conference on Intelligent Tutoring Systems,* Springer-Verlag, Berlin, Heidelberg, http://dx.doi.org/10.1007/978-3-540-69132-7_6. [58]

Duckworth, A., J. Taxer, L. Eskreis-Winkler, B. Galla and J. Gross (2019), "Self-control and academic achievement", *Annual Review of Psychology,* Vol. 70, pp. 373-399. [31]

Duckworth, A. and D. Yeager (2015), "Measurement matters: Assessing personal qualities other than cognitive ability for educational purposes.", *Educational Researcher,* Vol. 44/4, pp. 237-251. [45]

Eastwood, J., A. Frischen, M. Frenske and D. Smilek (2012), "The unengaged mind: Defining boredom in terms of attention", *Perspectives on Psychological Science,* Vol. 7/5, pp. 482-495. [2]

Eccles, J. and M. Wang (2012), "Part I commentary: So what is student engagement anyway?", in Christenson, S., A. Reschly and C. Wylie (eds.), *Handbook of research on student engagement,* Springer, New York. [32]

Ericsson, K., R. Krampe and C. Tesch-Römer (1993), "The role of deliberate practice in the acquisition of expert performance", *Psychological Review,* Vol. 100/3, p. 363. [30]

Faber, M., R. Bixler and S. D'Mello (2018), "An automated behavioral measure of mind wandering during computerized reading", *Behavior Research Methods,* Vol. 50/1, pp. 134-150. [167]

Farrell, E., G. Peguero, R. Lindsey and R. White (1988), "Giving voice to high-school students: Pressure and boredom, ya know what I'm sayin'?", *American Educational Research Journal*, Vol. 25/4, pp. 489-502. [10]

Ferguson, C. and A. Garza (2011), "Call of (civic) duty: Action games and civic behavior in a large sample of youth", *Computers in Human behavior*, Vol. 27/2, pp. 770-775. [139]

Finn, J. and K. Voelkl (1993), "School characteristics related to student engagement", *The Journal of Negro Education*, Vol. 62/3, pp. 249-268. [11]

Forbes-Riley, K. and D. Litman (2011), "Benefits and challenges of real-time uncertainty detection and adaptation in a spoken dialogue computer tutor", *Speech Communication*, Vol. 53/9-10, pp. 1115-1136, http://dx.doi.org/10.1016/j.specom.2011.02.006. [86]

Fredricks, J., P. Blumenfeld and A. Paris (2004), "School engagement: Potential of the concept, state of the evidence", *Review of Educational Research*, Vol. 74/1, pp. 59-109. [34]

Fredricks, J. and W. McColskey (2012), "The measurement of student engagement: A comparative analysis of various methods and student self-report instruments", in Christenson, S., A. Reschly and C. Wylie (eds.), *Handbook of research on student engagement*, Springer, New York. [42]

Fullerton, T., C. Swain and S. Hoffman (2008), *Game design workshop: a playcentric approach to creating innovative games*, Morgan Kaufmann, Burlington, MA. [129]

Gardner, J., C. Brooks and R. Baker (2019), *Evaluating the fairness of predictive student models through slicing analysis.,* Paper presented at the Proceedings of the 9th International Conference on Learning Analytics & Knowledge. [76]

Gee, J. (2003), *What video games have to teach us about learning and literacy,* Palgrave Macmillan, New York. [41]

Gibson, D., N. Ostashewski, K. Flintoff, S. Grant and E. Knight (2015), "Digital badges in education", *Education and Information Technologies,* Vol. 20/2, pp. 403-410. [117]

Gobert, J., R. Baker and M. Wixon (2015), "Operationalizing and detecting disengagement within online science microworlds", *Educational Psychologist,* Vol. 50/1, pp. 43-57. [82]

Grafsgaard, J., J. Wiggins, K.E. Boyer, E. Wiebe and J.C. Lester (2014), "Predicting learning and affect from multimodal data streams in task-oriented tutorial dialogue", in Stamper, J. et al. (eds.), *Proceedings of the 7th International Conference on Educational Data Mining,* International Educational Data Mining Society. [100]

Green, C. and D. Bavelier (2012), "Learning, attentional control, and action video games", *Current Biology*, Vol. 22/6, pp. R197-R206. [132]

Green, C. and D. Bavelier (2007), "Action-video-game experience alters the spatial resolution of vision", *Psychological Science,* Vol. 18/1, pp. 88-94. [133]

Greene, B. (2015), "Measuring cognitive engagement with self-report scales: Reflections from over 20 years of research", *Educational Psychologist,* Vol. 50/1, pp. 1-17. [43]

Griffiths, A., E. Lilles, M.J. Furlong and J. Sidhwa (2012), "The relations of adolescent student engagement with troubling and high-risk behaviors", in Christenson, S., A. Reschly and C. Wylie (eds.), *Handbook of research on student engagement,* Springer, New York. [12]

Hawkins, W., N. Heffernan and R. Baker (2013), "Which is more responsible for boredom in intelligent tutoring systems: students (trait) or problems (state)?", in Nijholt, A., S. D'Mello and M. Pantic (eds.), *Proceedings of the 5th International Conference on Affective Computing and Intelligent Interaction (ACII 2013),* IEEE, Washington, DC. [112]

Heine, S., D. Lehman, K. Peng and J. Greenholtz (2002), "What's wrong with cross-cultural comparisons of subjective Likert scales?: The reference-group effect", *Journal of Personality and Social Psychology*, Vol. 82/6, pp. 903-918. [47]

Henrie, C., L. Halverson and C. Graham (2015), "Measuring student engagement in technology-mediated learning: A review", *Computers & Education,* Vol. 90, pp. 36-53. [44]

Hidi, S. and K. Renninger (2006), "The four-phase model of interest development", *Educational Psychologist,* Vol. 41/2, pp. 111-127, https://doi.org/10.1207/s15326985ep4102_4. [38]

Hunter, A. and J. Eastwood (2018), "Does state boredom cause failures of attention? Examining the relations between trait boredom, state boredom, and sustained attention", *Experimental Brain Research,* Vol. 236/9, pp. 2483-2492. [3]

Hutt, S., J. Grafsgaard and S. D'Mello (2019), "Time to Scale: Generalizable Affect Detection for Tens of Thousands of Students across An Entire Schoolyear", in *Proceedings of the ACM CHI Conference on Human Factors in Computing Systems (CHI 2019)*, ACM, New York. [83]

Hutt, S., K. Krasich, C. Mills, N. Bosch, S. White, J. Brockmole and S. D'Mello (2019), "Automated Gaze-Based Mind Wandering Detection during Computerized Learning in Classrooms", User Modeling & User-Adapted Interaction. [92]

Ibáñez, M., A. Di Serio, D. Villáran and C. Delgado Kloos (2014), "Experimenting with electromagnetism using augmented reality: Impact on flow student experience and educational effectiveness", *Computers & Education,* Vol. 71, pp. 1-13. [156]

Jackson, L., E. Witt, A. Games, H. Fitzgerald, A. von Eye and Y. Zhao (2012), "Information technology use and creativity: Findings from the Children and Technology Project", *Computers in Human behavior*, Vol. 28/2, pp. 370-376. [138]

Jensen, E., S. Hutt and S. D'Mello (2019), *Generalizability of Sensor-Free Affect Detection Models in a Longitudinal Dataset of Tens of Thousands of Students,* Proceedings of the 2019 Educational Data Mining Conference (EDM 2019). International Educational Data Mining Society [77]

Johnson, C. and R. Mayer (2010), "Applying the self-explanation principle to multimedia learning in a computer-based game-like environment", *Computers in Human behavior*, Vol. 26/6, pp. 1246-1252. [127]

Kahneman, D., A. Krueger, D. Schkade, N. Schwartz and A. Stone (2004), "A survey method for characterizing daily life experience: The day reconstruction method", *Science*, Vol. 306/5702, pp. 1776-1780. [50]

Kapoor, A. and R. Picard (2005), "Multimodal affect recognition in learning environments", in *Proceedings of the 13th annual ACM international conference on Multimedia*, ACM, New York, http://doi.acm.org/10.1145/1101149.1101300. [101]

Kapp, K. (2012), *The gamification of learning and instruction*, Wiley San Francisco. [118]

Karpicke, J. and J. Blunt (2011), "Retrieval practice produces more learning than elaborative studying with concept mapping", *Science*, Vol. 331/6018, pp. 772-775. [24]

Klein, R. and T. Celik (2017), *The Wits Intelligent Teaching System: Detecting student engagement during lectures using convolutional neural networks*, IEEE. [61]

Krosnick, J. (1999), "Survey research", *Annual Review of Psychology*, Vol. 50/1, pp. 537-567. [46]

Lehman, B., M. Matthews, S. D'Mello and N. Person (2008), "What Are You Feeling? Investigating Student Affective States During Expert Human Tutoring Sessions", in *Intelligent Tutoring Systems, Lecture Notes in Computer Science*, Springer Berlin Heidelberg, Berlin, Heidelberg, http://dx.doi.org/10.1007/978-3-540-69132-7_10. [57]

Lehr, C., M. Sinclair and S. Christenson (2004), "Addressing student engagement and truancy prevention during the elementary school years: A replication study of the check & connect model", *Journal of education for students placed at risk*, Vol. 9/3, pp. 279-301. [64]

Lepper, M. and J. Henderlong (2000), "Turning" play" into" work" and" work" into" play": 25 years of research on intrinsic versus extrinsic motivation", in Sansone, C. and J. Harackiewicz (eds.), *Intrinsic and extrinsic motivation: The search for optimal motivation and performance*, Academic Press, San Diego, CA. [124]

Malone, T. and M. Lepper (1987), "Making learning fun: A taxonomy of intrinsic motivations for learning", in Snow, R. and M. Farr (eds.), *Aptitude, learning, and instruction III: Cognitive and affective process analyses*, Erlbaum, Hilsdale, NJ. [130]

Replace with: Mann, S. and A. Robinson (2009), "Boredom in the lecture theatre: an investigation into the contributors, moderators and outcomes of boredom amongst university students", *British Educational Research Journal*, Vol. 35/2, pp. 243-258, https://doi.org/10.1080/01411920802042911. [13]

Martín-Gutiérrez, J. et al. (2017), "Virtual technologies trends in education", *EURASIA Journal of Mathematics Science and Technology Education*, Vol. 13/2, pp. 469-486. [154]

Martin, J. (2010), *Lower Merion district's laptop saga ends with $610,000 settlement*, http://articles.philly.com/2010-10-12/news/24981536_1_laptop-students-district-several-million-dollars. [74]

Ma, W., O. Olusola, J. Nesbit and Q. Liu (2014), "Intelligent tutoring systems and learning outcomes: A meta-analysis", *Journal of Educational Psychology*, Vol. 106/4, pp. 901-918. [108]

McNamara, D., T. O'Reilly, R. Best and Y. Ozuru (2006), "Improving adolescent students' reading comprehension with iSTART", *Journal of Educational Computing Research*, Vol. 34/147-171. [27]

Mehl, M., J.W. Pennebaker, D.M. Crow, J. Dabbs and J.H. Price (2001), "The Electronically Activated Recorder (EAR): A device for sampling naturalistic daily activities and conversations", *Behavior Research Methods, Instruments, & Computers*, Vol. 33/4, pp. 517-523. [56]

Miller, W., R. Baker, M. Labrum, K. Petsche and A.Z. Wagner (2014), *Boredom Across Activities, and Across the Year, within Reasoning Mind*, Paper presented at the Workshop on Data Mining for Educational Assessment and Feedback (ASSESS 2014), New York, NY. [78]

Mills, C., I. Fridman, W.V. Soussou, D. Waghray, A.M. Olney and S. D'Mello (2017), "Put Your Thinking Cap On: Detecting Cognitive Load using EEG during Learning", in Molenaar, I., O. X. and S. Dawson (eds.), *Proceedings of the 7th International Learning Analytics and Knowledge Conference (LAK'17)*, ACM, New York, NY. [97]

Mills, C., J. Gregg, R. Bixler and S. D'Mello (2020), "Eye-Mind Reader: An Intelligent Reading Interface that Promotes Long-term Comprehension by Detecting and Responding to Mind Wandering", *Human-Computer Interaction*, https://doi.org/10.1080/07370024.2020.1716762. [166]

Monkaresi, H., N. Bosch, R.A. Calvo and S. D'Mello (2017), "Automated detection of engagement using video-based estimation of facial expressions and heart rate", *IEEE Transactions on Affective Computing*, Vol. 8/1, pp. 15-28. [87]

Mota, S. and R. Picard (2003), *Automated posture analysis for detecting learner's interest level.* [94]

Nystrand, M. and A. Gamoran (1991), "Instructional discourse, student engagement, and literature achievement", *Research in the Teaching of English*, Vol. 25/3, pp. 261-290. [52]

Ocumpaugh, J., R. Baker and M. Rodrigo (2015), *Baker Rodrigo Ocumpaugh Monitoring Protocol (BROMP) 2.0 Technical and Training Manual.* [73]

Olney, A., S. D'Mello, N. Person, W. Cade, P. Hays, C. Williams, B. Lehman and A. Graesser (2012), "Guru: A computer tutor that models expert human tutors", in Cerri, S. et al. (eds.), *Proceedings of the 11th International Conference on Intelligent Tutoring Systems,* Springer-Verlag, Berlin/Heidelberg. [102]

Papert, S. (1998), "Does easy do it? Children, games, and learning", *Game Developer*, Vol. 5/6, p. 88. [21]

Pardos, Z., R. Baker, M. San Pedro, S. Gowda and S. Gowda (2013), "Affective states and state tests: Investigating how affect throughout the school year predicts end of year learning outcomes", in Suthers, D. et al. (eds.), *Proceedings of the 3rd International Conference on Learning Analytics and Knowledge*, ACM, New York, NY. [84]

Patrick, B., E. Skinner and J. Connell (1993), "What motivates children's behavior and emotion? Joint effects of perceived control and autonomy in the academic domain.", *Journal of Personality and Social Psychology,* Vol. 65/4, pp. 781-791, http://dx.doi.org/10.1037/0022-3514.65.4.781 [14]

Pekrun, R., T. Goetz, L. Daniels, R. Stupnisky and R. Perry (2010), "Boredom in achievement settings: Exploring control–value antecedents and performance outcomes of a neglected emotion", *Journal of Educational Psychology*, Vol. 102/3, pp. 531-549, http://dx.doi.org/10.1037/a0019243. [15]

Pekrun, R., N. Hall, R. Perry and T. Goetz (2014), "Boredom and academic achievement: Testing a model of reciprocal causation", *Journal of Educational Psychology,* Vol. 106/3, pp. 696-710. [4]

Pekrun, R. and L. Linnenbrink-Garcia (2012), "Academic emotions and student engagement", in Christenson, S., A. Reschly and C. Wylie (eds.), *Handbook of research on student engagement,* Springer, New York. [37]

Perkins, R. and A. Hill (1985), "Cognitive and affective aspects of boredom", *British Journal of Psychology,* Vol. 76/MAY, pp. 221-234. [16]

Pham, P. and J. Wang (2015), "AttentiveLearner: improving mobile MOOC learning via implicit heart rate tracking", in *International Conference on Artificial Intelligence in Education*, Springer, Berlin Heidelberg. [88]

Pianta, R., B. Hamre and J. Allen (2012), "Teacher-student relationships and engagement: Conceptualizing, measuring, and improving the capacity of classroom interactions", in Christenson, S., A. Reschly and C. Wylie (eds.), *Handbook of research on student engagement*, Springer, New York, NY. [53]

Plass, J., S. Heidig, E.O. Hayward, B. Homer and E. Um (2014), "Emotional design in multimedia learning: Effects of shape and color on affect and learning", *Learning and Instruction,* Vol. 29, pp. 128-140. [119]

Plass, J., B. Homer and C. Kinzer (2015), "Foundations of game-based learning", *Educational Psychologist,* Vol. 50/4, pp. 258-283. [125]

Podsakoff, P., S. MacKenzie, J-Y. Lee and N. Podsakof (2003), "Common method biases in behavioral research: A critical review of the literature and recommended remedies", *Journal of Applied Psychology*, Vol. 88/5, pp. 879-903. [48]

Porayska-Pomsta, K., M. Mavrikis, S. D'Mello, C. Conati and R. Baker (2013), "Knowledge elicitation methods for affect modelling in education", *International Journal of Artificial Intelligence in Education,* Vol. 22, pp. 107-140. [71]

Psotka, J. (1995), "Immersive training systems: Virtual reality and education and training", *Instructional science,* Vol. 23/5-6, pp. 405-431. [155]

Putwain, D., S. Baker, W. Symes and R. Pekrun (2018), "Reciprocal relations between students' academic enjoyment, boredom, and achievement over time", *Learning and Instruction,* Vol. 54, pp. 73-81. [5]

Raca, M., L. Kidzinski and P. Dillenbourg (2015), "Translating head motion into attention-towards processing of student's body language", in *Proceedings of the 8th International Conference on Educational Data Mining,* International Educational Data Mining Society. [62]

Randall, J., F. Oswald and M. Beier (2014), "Mind-wandering, cognition, and performance: A theory-driven meta-analysis of attention regulation", *Psychological Bulletin*, Vol. 140/6, pp. 1411-1431. [114]

Reeve, J. and C. Tseng (2011), "Agency as a fourth aspect of students' engagement during learning activities", *Contemporary Educational Psychology,* Vol. 36/4, pp. 257-267. [36]

Renninger, K. and J. Bachrach (2015), "Studying triggers for interest and engagement using observational methods", *Educational Psychologist*, Vol. 50/1, pp. 58-69. [35]

Reschly, A. and S. Christenson (2012), "Jingle, jangle, and conceptual haziness: Evolution and future directions of the engagement construct", in S. Christenson, A. Reschly and C. Wylie (eds.), *Handbook of research on student engagement,* Springer, Berlin. [33]

Rey, G. (2012), "A review of research and a meta-analysis of the seductive detail effect", *Educational Research Review*, Vol. 7/3, pp. 216-237. [22]

Risko, E., D. Buchanan, S. Medimorec and A. Kingstone (2013), "Everyday attention: mind wandering and computer use during lectures", *Computers & Education*, Vol. 68/1, pp. 275-283. [115]

Ritterfeld, U., M. Cody and P. Vorderer (2009), *Serious games: Mechanisms and effects, Routledge, New York and London.* [126]

Rodrigo, M. and R. Baker (2011), "Comparing Learners' Affect While Using an Intelligent Tutor and an Educational Game", *Research and Practice in Technology Enhanced Learning,* Vol. 6, pp. 43-66. [29]

Roediger, H. and J. Karpicke (2006), "Test-enhanced learning - Taking memory tests improves long-term retention", *Psychological Science,* Vol. 17/3, pp. 249-255. [25]

Rowe, J., B. Mott, S. McQuiggan, J. Robison, S. Lee and J. Lester (2009), *Crystal island: A narrative-centered learning environment for eighth grade microbiology,* Paper presented at the Workshop on Intelligent Educational Games at the 14th International Conference on Artificial Intelligence in Education, Brighton, UK. [149]

Rowe, J., L. Shores, B. Mott and J. Lester (2010), "Integrating learning and engagement in narrative-centered learning environments", in Kay, J. and V. Aleven (eds.), *Proceedings of the 10th International Conference on Intelligent Tutoring Systems,* Springer, Berlin / Heidelberg. [151]

Ryu, S. and D. Lombardi (2015), "Coding classroom interactions for collective and individual engagement", *Educational Psychologist*, Vol. 50/1, pp. 70-83. [54]

Sabourin, J. and J. Lester (2014), "Affect and Engagement in Game-Based Learning Environments", *IEEE Transactions on Affective Computing*, Vol. 5/1, pp. 45-56. [72]

Sanchez, E., S. Young and C. Jouneau-Sion (2016), "Classcraft: from gamification to ludicization of classroom management", *Education and Information Technologies,* Vol. 22/2, pp. 497-513, http://dx.doi.org/10.1007/s10639-016-9489-6. [122]

Schneider, B., J. Krajcik, J. Lavonen, K. Salmela-aro and M. Geller (2020), *Learning Science-the Value of Crafting Engagement in Science Environments,* New Haven. CT: Yale University Press. [146]

Shechtman, N., A. DeBarger, C. Dornsife, S. Rosier and L. Yarnal (2013), *Promoting grit, tenacity, and perseverance: Critical factors for success in the 21st century.* [67]

Shute, V., L. Rieber and R. Van Eck (2011), "Games... and... learning", in Reiser, R. and J. Dempsey (eds.), *Trends and issues in instructional design and technology,* Pearson Education, Inc. [131]

Shute, V., M. Ventura and F. Ke (2015), "The power of play: The effects of Portal 2 and Lumosity on cognitive and noncognitive skills", *Computers & Education,* Vol. 80, pp. 58-67. [134]

Shute, V., M. Ventura and Y. Kim (2013), "Assessment and learning of qualitative physics in Newton's playground", *The Journal of Educational Research,* Vol. 106/6, pp. 423-430. [103]

Shute, V. et al. (2014), "Video games and learning", in Tierney, W. et al. (eds.), *Postsecondary play: The role of games and social media in higher education,* John Hopkins Press, Baltimore, MD. [128]

Sinatra, G., B. Heddy and D. Lombardi (2015), "The challenges of defining and measuring student engagement in science", *Educational Psychologist*, Vol. 50/1, pp. 1-13. [39]

Skinner, E. and M. Belmont (1993), "Motivation in the classroom: Reciprocal effects of teacher behavior and student engagement across the school year", *Journal of Educational Psychology*, Vol. 85/4, pp. 571-581. [65]

Skoric, M., L. Teo and R. Neo (2009), "Children and video games: addiction, engagement, and scholastic achievement", *Cyberpsychology & behavior,* Vol. 12/5, pp. 567-572. [135]

Sleeman, D. and J. Brown (1982), *Intelligent tutoring systems*, Academic Press, New York. [107]

Spires, H., J. Rowe, B. Mott and J. Lester (2011), "Problem solving and game-based learning: Effects of middle grade students' hypothesis testing strategies on learning outcomes", *Journal of Educational Computing Research*, Vol. 44/4, pp. 453-472. [150]

Stahl, G., T. Koschmann and D. Suthers (2006), "Computer-supported collaborative learning: An historical perspective", in Sawyer, K. (ed.), *Cambridge handbook of the learning sciences*, Cambridge University Press, New York, NY. [171]

Steenbergen-Hu, S. and H. Cooper (2014), "A meta-analysis of the effectiveness of intelligent tutoring systems on college students' academic learning", *Journal of Educational Psychology*, Vol. 106/2, pp. 331-347. [26]

Steenbergen-Hu, S. and H. Cooper (2013), "A meta-analysis of the effectiveness of intelligent tutoring systems on K–12 students' mathematical learning", *Journal of Educational Psychology*, Vol. 105/4, pp. 970-987. [109]

Tobias, S. and J. Fletcher (2011), *Computer games and instruction*, Information Age Publishers, Charlotte, NC. [142]

Turner, J. and D. Meyer (2000), "Studying and understanding the instructional contexts of classrooms: Using our past to forge our future", *Educational Psychologist*, Vol. 35/2, pp. 69-85. [51]

Tze, V., L. Daniels and R. Klassen (2016), "Evaluating the relationship between boredom and academic outcomes: a meta-analysis", *Educational Psychology Review*, Vol. 28/1, pp. 119-144. [6]

Um, E., J. Plass, E. Hayward and B. Homer (2012), "Emotional design in multimedia learning", *Journal of Educational Psychology*, Vol. 104/2, pp. 485-498. [120]

Valstar, M., M. Mehu, B. Jiang, M. Pantic and K. Scherer (2012), "Meta-analysis of the first facial expression recognition challenge", *IEEE Transactions on Systems, Man, and Cybernetics, Part B: Cybernetics*, Vol. 42/4, pp. 966-979. [69]

Ventura, M., V. Shute and Y. Kim (2012), "Video gameplay, personality and academic performance", *Computers & Education*, Vol. 58/4, pp. 1260-1266. [136]

Ventura, M., V. Shute and W. Zhao (2013), "The relationship between video game use and a performance-based measure of persistence", *Computers & Education*, Vol. 60/1, pp. 52-58, http://dx.doi.org/10.1016/j.compedu.2012.07.003. [137]

Vincent-Lancrin, S., J. Urgel, S. Kar and G. Jacotin (2019), *Measuring Innovation in Education 2019: What Has Changed in the Classroom?*, Educational Research and Innovation, OECD Publishing, Paris, https://dx.doi.org/10.1787/9789264311671-en. [169]

Volpe, R., J. DiPerna, J. Hintze and E. Shapiro (2005), "Observing students in classroom settings: A review of seven coding schemes", *School Psychology Review*, Vol. 34/4, pp. 454-474. [55]

Wasson, A. (1981), "Susceptibility to Boredom and Deviant-Behavior at School", *Psychological Reports*, Vol. 48/3, pp. 901-902. [17]

Whitehill, J., Z. Serpell, Y. Lin, A. Foster and J. Movellan (2014), "The faces of engagement: Automatic recognition of student engagement from facial expressions", *IEEE Transactions on Affective Computing*, Vol. 5/1, pp. 86 - 98. [89]

Wilson, K., W. Bedwell, E. Lazzara, E. Salas, C. Burke, J. Estock, K. Orvis and C. Conkey (2009), "Relationships between game attributes and learning outcomes review and research proposals", *Simulation & gaming*, Vol. 40/2, pp. 217-266. [143]

Yang, D., T. Sinha, D. Adamson and C. Penstein Rose (2013), *Turn on, tune in, drop out: Anticipating student dropouts in massive open online courses*, Paper presented at the Proceedings of the 2013 NIPS Data-Driven Education Workshop. [19]

Young, M., S. Slota, A. Cutter, G. Jalette, G. Mullin, B. Lai, Z. Simeoni, M. Tran and M. Yukhymenko (2012), "Our Princess Is in Another Castle: A Review of Trends in Serious Gaming for Education", *Review of Educational Research*, Vol. 82/1, pp. 61-89. [144]

5 Classroom analytics: Zooming out from a pupil to a classroom

Pierre Dillenbourg
École polytechnique fédérale de Lausanne, Switzerland

The term "learning technologies" usually refers to the activities that one or a few learners perform on a digital device. How do technologies take into account the fact that a classroom has many more learners? Classroom analytics provide teachers with real-time support for class management: monitoring the engagement of learners in a class; deciding when and how to intervene in their learning activities; reusing the output from one activity in another one; forming student groups; integrating the learner's production in a lecture; deciding when to shift to another activity; or helping teachers to regulate their own behaviour. This chapter describes the classroom itself as a digital system. It speculates on how this vision can come true, and what can already be learned from it, namely the critical role of teachers in the success of digital education.

Introduction

Learning analytics aim at modelling the learning process, i.e. how knowledge, skills and competencies are developed by learners while performing some activities. When these activities are run on computers, these models enable the software to adapt the next learning activities to the learner's needs. What if 30 learners independently use this education software in the same classroom? Many interesting classroom events would then occur outside the software and be mostly ignored by analytics: the teacher's interventions, discussion among peers, etc. Do students engage more with the system when the teacher passes nearby than when he/she is further away? What if the individual activity on computers is only a part of the lesson, in addition to teamwork and lectures? These activities, invisible to the software, will not be taken into account by the analytics even though they do actually matter in terms of how much learners actually learn.

This chapter broadens the scope of learning analytics, from modelling the learner's interactions with the device to capturing anything that happens in this peculiar ecosystem called a classroom. Classroom analytics are multimodal, i.e. they collect relevant data with a variety of sensors in order to analyse conversation patterns, attention levels, body postures, etc. Some of these sensors are cameras, which immediately raises ethical issues. Balancing these risks with the potential benefits of classroom analytics is a core concern.

The benefits of classroom analytics should ideally occur in two steps. First, classroom analytics are designed to enhance the management of the classroom, for instance, by displaying a dashboard that shows the teacher which learners are struggling or that helps him/her to decide when to move on to the next activity. In the second step, improving classroom management is expected to lead to higher outcomes for the learners, as has been demonstrated by Do-Lenh et al. (2012[1]) as well Holstein, McLaren and Aleven (2018[2]). Another expected benefit of classroom analytics is to expand learning analytics to a rich diversity of learning activities such as lectures, teamwork, hands-on activities or collective discussions and even some activities outside the classroom, such as field trips. The management of such a rich pedagogical scenario, which includes individual, team and class-wide activities, some with and some without digital technology, is captured by the term "classroom orchestration". Classroom analytics do not make decisions in place of teachers. Rather, they provide teachers with information for them to interpret as teachers, generally, are aware of context: for instance, a learner detected as poorly performing may actually be ill, helping another student or experiencing slow connectivity. There should always be a teacher in the loop; classroom analytics empowers her/him to teach more effectively.

This chapter looks at how learning analytics may transform education in the next decade. I hypothesise that a 2030 classroom will be visually similar to a 2020 classroom, just as the 2020 one is similar to the 1920 one. If a person who drove a car in 1920 resurrected in 2020, they would have difficulties driving a new car. They would neither recognise an induction cooker as a stove nor a smartphone as a telephone, but they would know when they entered a "classroom". Education evolves slowly. Today's teachers live in the "classroom of the future" of teachers who lived in the 1980s. Therefore, analysing how digital education has evolved over the past 40 years affords linear projections for the next 10-20 years. Predicting that there will still be physical classrooms may sound conservative. Some would even argue classrooms will disappear. They forget an unpleasant reality: schools also fulfil a babysitting function, i.e. keeping kids busy while parents are at work. This chapter postulates that decision makers will still need to make choices about physical classrooms in 2030. Some initiatives entitled "classroom of the future" envision it as a cold space overloaded with individual screens and computers. This chapter proposes a different vision, in which classrooms remain rich social living places and not visually dominated by technological devices.

The vision: Classrooms as a digital system

Entering the classroom of tomorrow might be similar to getting into a car today. In both cases, one enters a physical space with doors, seats, windows, other people, etc., as well as many sensors, computers and actuators. Sitting in a car is much like being inside a digital system. It may sound scary but a digital system that sounds an alarm when the driver has nodded off has benefits that bypass the ethical concerns of video-monitoring the driver. Classroom systems face the same ethical trade-off.

A digital system captures (input), processes and communicates (output) information in a digital form. Since anything can be described as a system, why should we describe a classroom as a (digital) system? It emphasises the difference between the usual viewpoint which is that a classroom is a physical space where digital systems are introduced and the viewpoint whereby the classroom becomes the system.

As in any system, a classroom is composed of several subsystems and so forth. This set of subsystems can be described as a system if these subsystems collectively fulfil a function that none of them individually perform. A classroom may include several digital devices (e.g. software for mathematics, a spreadsheet, some educational robots, a teamwork platform, etc.) that each performs its specific function. Other cognitive functions are performed by the people in the room, the teacher and the learners, as well as by artefacts (e.g. a poster displaying the periodic table). The "classroom system" performs at a higher level function than these subsystems and one of these functions is classroom orchestration.

In personalised instruction systems, the input is the behaviour of learners, the function is adapting instruction for each learner and the output is the system's next action, for instance, the feedback given or the next activity proposed. In the classroom-as-system approach, the input is the analytics collected in the classroom, the output is the information given to the teacher or to the learners, for instance, dashboards, and the function is the classroom orchestration. Holstein, McLaren and Aleven (2017[3]) observed the behaviour of teachers when their students used an intelligent tutoring system and found that teachers spent on average 47% of their time either inactive or outside the classroom. They simply felt "out of the loop". In learning analytics, the vision is to keep one or more humans in the loop. We talk about the co-orchestration of the classroom by the teacher and the digital components of the system (Santos, 2012[4]).

How did we get there?

Looking at the evolution of learning technologies over the past 40 years, four trends led to the emergence of the concept of "classroom as a system".

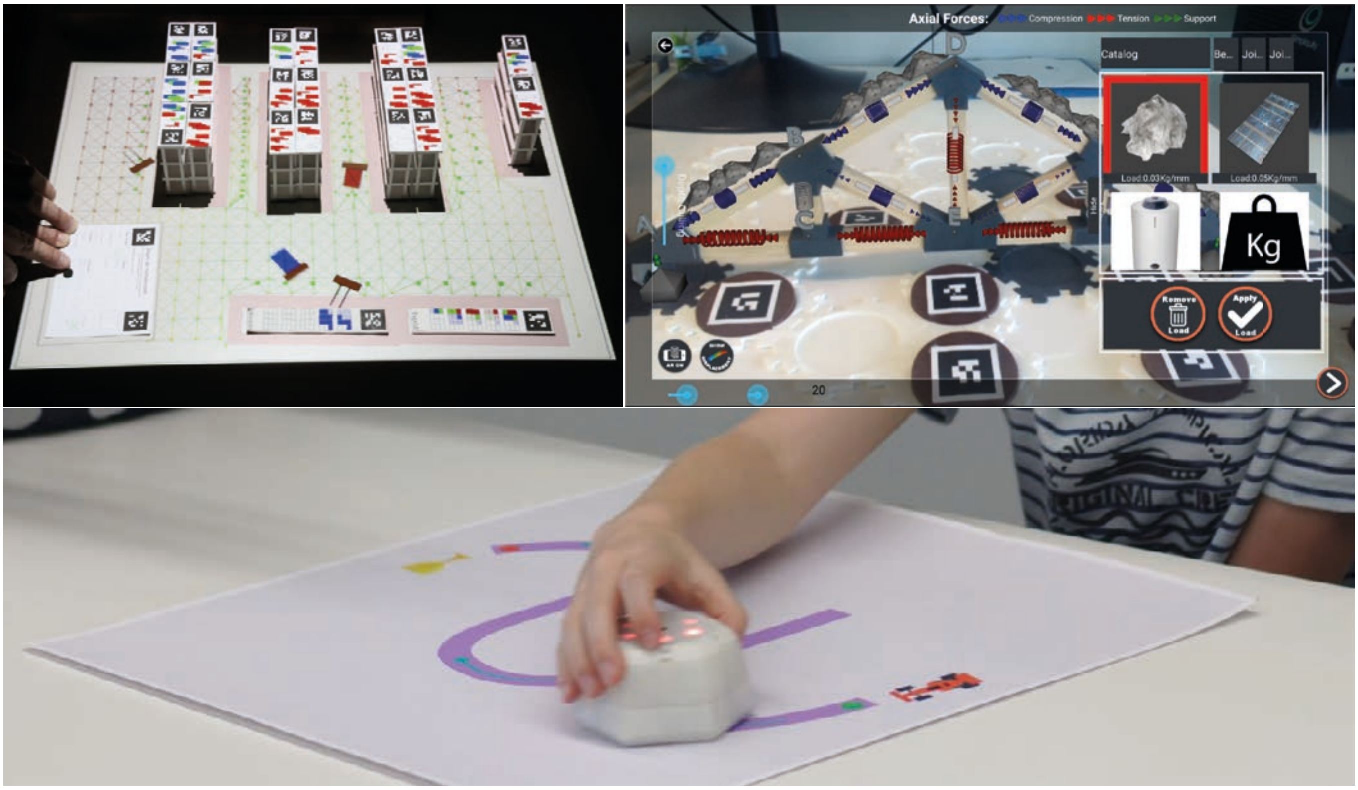
The first trend is the growing integration of pedagogical approaches that have for many years been considered as mutually exclusive. Many educational tools used in school are based on the "mastery learning" ideas (Bloom, 1968[5]): decomposing complex skills into simple ones by providing rapid feedback and incrementally practicing more complex skills. Another family of tools, called "micro-worlds" (Papert, 1987[6]), are some kind of digital sandboxes where learners acquire problem-solving skills by trial and error, reflecting constructivist theories. The same theory inspired inquiry-based learning tools, namely learning by running real or simulated experiments. "Instructionist" approaches inspired massive open online courses (MOOCs) and other environments where learners are mostly watching lectures or reading texts. These learning theories focus on individual learning, building on the ability to adapt instruction to the differences among learners. To the contrary, empirical studies revealed the benefits of collaborative learning, which gave rise to environments designed for learning in teams (Dillenbourg, Järvelä and Fischer, 2009[7]) based on social cognition theories (Vygotsky, 1964[8]). These oppositions are fading out. The human cognitive system is a social software running on an individual hardware, the brain. Why should a teacher bet on a single digital education approach when he/she can integrate several approaches where and when they are relevant?

The second trend is the growing compatibility between the technologies used in education. For many years, one could, for instance, not technically integrate a piece of software with math exercises and a video player. Web technologies contributed to the interoperability across almost any digital component. The possibility of technically integrating different tools converges with the first trend, the interest of integrating different pedagogical approaches. This evolution does not lead to the ultimate learning management system that offers all of the functions required for all learning activities, but rather to the development of ecosystems of digital tools, each having specific functions. Right now, interoperability among learning environments is still far from sufficient. Metadata standards (Duval, 2001[9]) have developed for exchanging digital contents (such as the Sharable Content Object Reference Model [SCORM] or Instructional Management System Learning Design [IMS LD]) and others, such as Learning Tools Interoperability (LTI), to foster interoperability, i.e. to exchange data about learners (Severance, Hanss and Hardin, 2010[10]). Today, the digital traces produced when learners use a tool are collected and aggregated into models specific to each tool they use. If they used multiple learning tools, no tool would have a comprehensive model of the learner. Recent projects (Mangaroska, Vesin and Giannakos, 2019[11]) aggregate data across applications in order to produce a synthetic account of the learners' learning paths. A standard for sharing records across applications, xAPI (Bakharia et al., 2016[12]) seems to be gaining momentum. Of course, producing cross-platform analytics multiplies the risks regarding data protection.

The third trend concerns the evolution of hardware. For many years, dedicated rooms were equipped with computers, so-called "computer labs", with rows of cumbersome boxes and vertical displays, making students hardly visible by teachers. Next, laptops started to enter genuine classrooms, then tablets and smartphones brought learning technologies to informal settings, from being seated on a sofa to walking in the forest. Nowadays, the diversity of devices that can be exploited in education has exploded with the integration of sensors and actuators in shoes, mugs, clothes, etc. – basically any object (the "Internet of Things"). Devices are becoming more present but less visible; they do not properly disappear, but are more in the background. The frontier between what is digital and what is not has been progressively blurred, as illustrated by the examples in Figure 5.1. Learners may physically manipulate tangible objects, tracked by sensors, combined with augmented reality. Physical-digital technologies expand the range of skills that can be practiced and assessed digitally, which is especially relevant for professional gestures taught in vocational education.

However, even if future classrooms become populated with digital components, they should not visually look like a NASA control room, fully packed with displays and devices. The classroom of the future may indeed appear close to being a technology-free room. Why not a room with wooden furniture and bright views of outside gardens? The more peripheral the digital tools, the less obtrusive they become for social interactions (dialogues, eye contact, etc.) and classroom orchestration.

Figure 5.1 The physical-digital blur



Note: Top left panel: Tangible interfaces enable students to interact by moving physical objects on a table. This simulation of a warehouse makes logistics principles more concrete for apprentices. Top right panel: Augmented reality combines real images of the environment with digital information displayed on a tablet or within head-mounted displays. This simulation of forces in roof structures provide future carpenters with an intuitive understanding of statics. Bottom panel: Educational robots are both physical and digital objects. This Cellulo robot provides learners with haptic feedback during various manipulative tasks.
Source: Do-Lenh et al. (2012[11]); Cuendet et al. (2015[13]).

The fourth trend is to pay more attention to the learning activity than to the learning technology. Consider the example of educational robots used to learn how to code. Some robots may be more appropriate than others, but the extent to which children learn depends less on the robot's features than on the activity that learners have to do with the robot. The same is true for MOOCs, augmented reality or virtual reality tools, and for any technology. The main variable of success is the ability to orchestrate rich activities in the classroom. Classroom orchestration refers to the real-time management of multiple activities under multiple constraints (Dillenbourg, 2013[14]). The multiplicity of activities refers to the integration of individual, team and class-wide activities with various digital environments as well as without any digital tool. The multiplicity of constraints emphasises the many practical aspects that shape teachers' decision making: managing learning time, coping with learners that missed previous lessons or joined late, taking into consideration the physical space – for instance, shifting from teamwork to lectures, maintaining a reasonable level of discipline, minimising the teacher's workload, etc. These constraints have somehow been neglected in scientific research but probably explain some difficulties in the adoption of learning technologies.

Classroom as the input

The rationale for expanding the input from usual keyboard and mouse to the whole classroom is that learner-software interaction traces only provide a partial (and limited) account of what is going on in a classroom. Even when learners are supposed to interact exclusively with a personal device, they actually often engage in "out-of-software" activities, some being on-task (e.g. asking the teacher for help), while others are off-task (e.g. chatting, surfing the web, daydreaming, etc.). Some will ask the teacher for help and conversely the teacher may intervene to nudge an inactive learner.

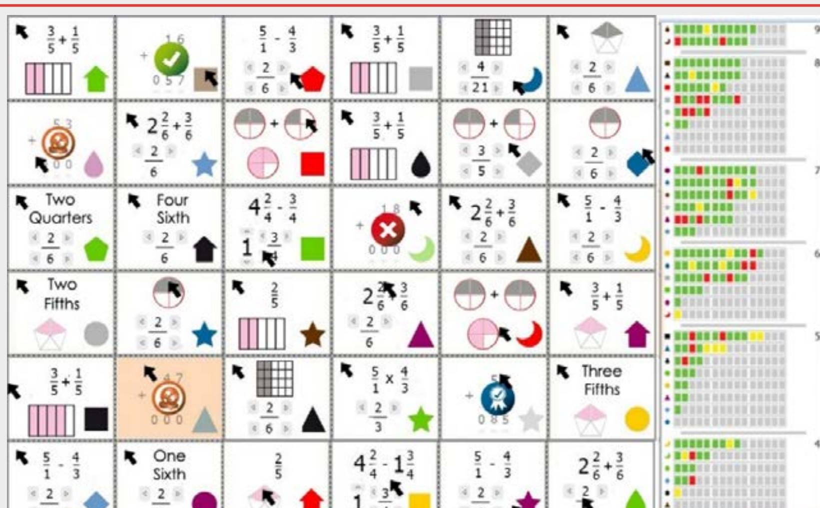
We propose the terms "*classroom analytics*" to emphasise that any event in the classroom may be captured and analysed for modelling the learning and teaching process. Holstein et al. (2017[3]) used a classroom replay tool for integrating these "out-of-software" interactions with the analytics produced by "in-software interactions". Data can indeed be collected for any classroom activity, including those with light technologies. For instance, the so-called "clickers" or "personal response systems" aim at increasing engagement during lectures, as well as collecting data: 1) the teacher interrupts her lecture and asks a multiple-choice question; 2) learners individually select a response on a personal device; 3) their answers are collected and visualised on the teacher's slides (output), enabling the teacher to give feedback and comment on frequent errors. Numerous variations of this scenario exist that allow for open questions, graphical questions, voting mechanisms, etc. In the peer instruction scenarios, between the individual answer phase (2) and the teacher's feedback phase (3), students are asked to compare their answer with their neighbour's answer and explain their choice. Fagen, Crouch and Mazur (2002[15]) collected robust evidence that this classroom scenario actually improves students' grades on university physics exams. Box 5.1 presents another example that was piloted in Chile.

Box 5.1 Single Display Groupware: A low-cost tool to make student activity visible to teachers (Chile)

Single Display Groupware (SDG) is an example of classroom analytics that is potentially very attractive for schools in developing countries because it only requires minimal equipment. SDG allows multiple people to work and interact over a single display, with everyone controlling their own input device. One application of SDG for the classroom is "One Mouse per Child for Basic Math", a software programme developed by researchers from Chile that only requires one PC, one projector, one mouse for the teacher and one mouse for each child participating in the activity (Alcoholado et al., 2012[16]). The programme integrates personalised learning in the classroom, which is particularly useful in contexts where a teacher has to manage a class with students at different levels.

As Figure 5.2 illustrates, the display – visible to everyone through the projector – contains individual cells, with the number of cells equal to the number of mice connected to the system. Students solve arithmetical exercises in their individual cells, moving to a new exercise when answering correctly and moving to the next difficulty level when solving multiple exercises correctly in a row. On the right of the screen, students are ranked based on the number of problems solved and difficulty level achieved. This informs them about their relative performance in the class, adding a competitive, fun element to the learning activity, and allows the teacher to identify students who are lagging behind. The teacher can freely move her cursor across the screen to help individual students, or switch the programme from practice to teaching mode to solve sample exercises and to explain math rules to the class. Results from an experimental evaluation in a primary school in Santiago using pre- and post-tests suggest that the programme has a positive impact on achievement, especially for low-achieving students.

Figure 5.2 One Mouse per Child for Basic Math

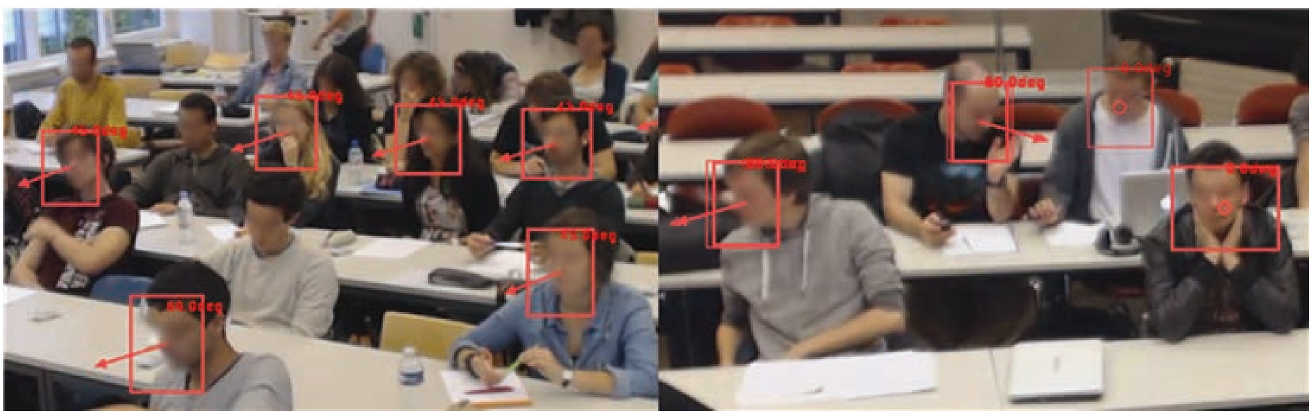


Note: Students solve math exercises in individual cells. Icons in the middle of the cells provide four types of feedback: correct, incorrect, pass to the next level, mouse inactivity. The panel on the right displays a ranking of students based on number of exercises solved and level achieved.

Source: Alcoholado et al. (2012[16]) (reproduced with permission)

Multimodal analytics (Ochoa and Worsley, 2016[17])) broaden the range of behaviours to be collected in classroom analytics. If a classroom is equipped with sensors, any gaze, gesture, body pause, stress level, etc. can be collected as input. If we consider that inside the smartphone of each students there are 15 20 sensors already, every classroom is potentially equipped with hundreds of sensors. Ahuja et al. (2019[18]) combined microphones and cameras in classrooms in order to detect which learners raised their hand, what posture and speech behaviour they displayed, and then correlated these features with lesson effectiveness. Scholars such as Yanga et al. (2018[19]) developed algorithms to identify emotions from facial images. The input data are not only behaviours (e.g. answers asking a question), but what could be called "behavioural dust", i.e. fragments of behaviour such as head rotation (Figure 5.3), a sigh or a gaze fixation. Taken individually, these fragments can hardly be interpreted; but aggregated over time or across learners, they eventually become meaningful. For instance, the approach taken by Raca, Kidzinski and Dillenbourg (2015[20]) was not to estimate the individual level of attention and then to average it over all students. Instead, they found a measure that could only be computed for a class: learners who pay attention to the lecture tend to rotate their head at the same time, simply because they are paying attention to a moving object, the teacher.

Figure 5.3 Estimating the average level of attention in a classroom by analysing head movements



Source: Raca, Kidzinski and Dillenbourg (2015[20])

Modelling an entire classroom is more complex than modelling interactions within a digital environment, in which correct and incorrect responses are often defined in advance. Some Bayesian Knowledge-Tracing variations integrated system-triggered instructional interventions (Lin and Chi, 2016[21]) into the model; one could also add the teacher's interventions. The more a learning environment is complex and open, the less accurate predictions can be. This lower accuracy is, however, not a concern in a classroom situation since the computational method does not aim to take autonomous decisions, but to inform a teacher who then takes decisions. Such a system, combining artificial and human intelligence, is often referred to as "a human in the loop". Classroom analytics *want to keep a teacher in the loop.*

Turning the classroom into an input device immediately raises an ethical red flag. In some projects, learners or teachers have been equipped with sensors (electroencephalography, skin conductivity, accelerometers, heart frequency, eye trackers, etc.). Placing cameras in a classroom is less intrusive, but does not comply with data protection principles. One way to be compliant would be that the system does not store images, but deletes them as soon as the relevant features have been extracted. Despite this solution, we believe that the risk of "Big Brotherisation" of schools remains high, and unacceptable in many cultures. As in many data protection debates, this risk has to be compared with the benefits, that is, the value of the output. According to Raca, Kidzinski and Dillenbourg (2015[20]), signalling to a teacher the sudden loss of attention can be useful for novice teachers or those who fail to keep their audience's interest. However, the benefits are not always obvious. Many scholars such as Yanga et al. (2018[19]) developed algorithms to infer emotions from facial images. What should the system do if it detects a learner's frustration? Strong negative feelings may hamper the learner's motivation, but some moderate level of confusion may indeed motivate them to try harder (D'Mello et al., 2014[22]). Classroom input should be restricted to what can actually provide a clear added-value to learning and teaching, and based on theories that have ideally sufficient empirical evidence or very plausible theories of action.

Classroom as the output

In adaptive personalisation systems, the output of the learning analytics is usually a decision to adapt instruction to the needs of an individual learner. In classroom analytics, the output is some information given to the humans in the loop – teachers and learners – who may then take a decision. This information often takes the form of a teaching dashboard, i.e. a visualisation of the state of learners or the progress of learning in the classroom, beamed onto the classroom walls or presented in a display (usually a screen).

The design of these dashboards involves a usability challenge: providing teachers with information without increasing their cognitive load. Most dashboards developed so far are indeed overwhelming teachers with too many details. One solution is to develop "zoomable" interfaces, i.e. providing a global picture, with minimal information per learner, but allowing the teacher to get more detailed information for any learner. Moreover, the dashboard should not reduce the visual attention that teachers pay to the classroom. Various solutions have pros and cons: to show the dashboards on the classroom display provides teachers with permanent access to its information (see Figure 5.4), but the information – including personal difficulties – is also made public to the whole class; to display the dashboard on the teacher's desktop preserves privacy but requires her to return to her desk; to display the dashboard on a tablet gives permanent and private access, but may be cumbersome; to display the dashboard on a head-up display (e.g. glasses) (Holstein et al., 2018[23]) provides both information while maintaining visual contact with learners and keeping the teachers' hands free, but is not very natural. Other design dimensions concern the nature of the displayed data (e.g. the response contents vs. the score), the social level (e.g. individual, teams, class), etc. One design choice that matters for classroom orchestration in particular is the spatial mapping of the given information (does the position of John on the dashboard correspond to his physical location in the classroom?).

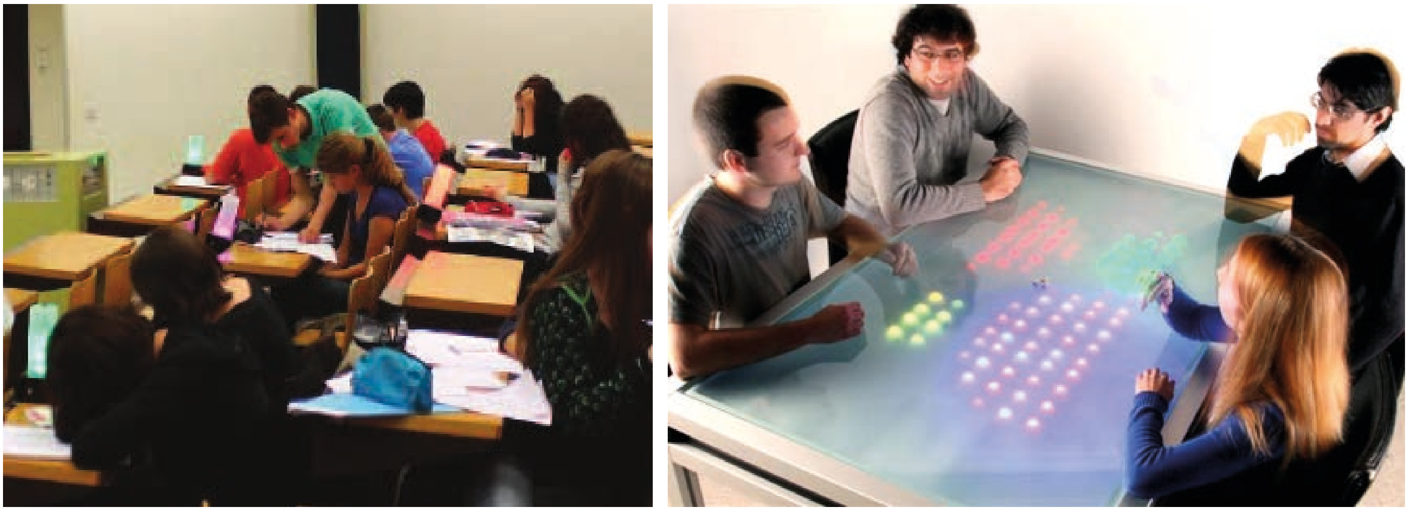
Classroom dashboards can be centralised (on a display), distributed (through several displays in the room) or ambient (just providing minimal information to teachers through distributed or centralised hints). This is another design choice that is important. Dashboards are generally centralised, but distributed dashboards also have their advantages for the orchestration of teaching and learning. Figure 5.4 shows a set of Lantern devices spread within the classroom: they constitute a distributed dashboard. Alavi and Dillenbourg (2012[24]) compared it with a centralised dashboard also visible to all, showing exactly the same information, and found that the centralised one tended to trigger competition among male students while the distributed one triggered some interactions among neighbouring teams.

Since the teacher's visual attention is saturated by the elements he/she needs to monitor, one may exploit peripheral vision and provide teachers with an "ambient" dashboard. For instance, Tomitsch, Grechenig and Mayrhofer (2007[25]) displayed information on the ceiling. The teacher is, of course, not expected to look at the ceiling, but if the colour of the classroom ceiling suddenly darkens, she will notice it. Gellersen, Schmidt and Beigl (1999[26]) conveyed information by changing the intensity of various lights or by controlling the pumps of a table fountain. Peripheral vision does not convey precise information, such as a numerical value, but a global impression. The term "ambient computing" describes technologies that do not require focal attention but change some contextual or background components. Today, ambient computing is alien to education stakeholders, but it has a great potential to turn the entire classroom into a display. It relates to "modest computing" (Dillenbourg et al., 2011[27]), which emphasises that the design of these displays deliberately degrades the accuracy of information: if the average score of learners in the classroom is 75%, it can be conveyed by setting the colour of the wall at the back of the classroom (which teachers often face) with a nuance of blue that is not as accurate as displaying the number 75, but is permanently visible to the teacher. On the Lantern device (Figure 5.4, left panel), the teacher perceives, for instance, which team has been waiting more than another team, without knowing exactly how much. Similarly, on the Reflect table (Figure 5.4, right panel), the colour of the table area in front of each learner approximates their amount of speech, but does not provide an exact count. It may happen that a participant keeps the floor for a while simply because an overall introduction or a long explanation is required. It does also occur that some participants game the system by deliberately and meaninglessly over-speaking. In both cases, the participants are aware of the conversation that happened, they know what the table display corresponds to. This justifies the "human in the loop" approach that we previously emphasised: knowing the context allows humans to interpret the feedback (while computers could misinterpret it).

Figure 5.4 Three types of interventions in domain independent actions

A. Regulating teachers' attention using Lantern devices

B. Regulating dialogues in teamwork using a Reflect table



Note: Left panel: Regulating the assistant's attention. During recitation sections, students communicate their progress to the teaching assistant by using a small device called a Lantern. The colour indicates which exercises they are working on, the number of LEDs indicate how long they have been working on this exercise. To ask for help, they push on the Lantern, which starts blinking slowly, then faster, revealing their waiting time. Right panel: Regulating dialogues in teamwork. A microphone array placed on the middle of the Reflect table captures the angle of the voices and thereby identifies who is speaking. The more someone speaks, the more the LEDs in front of him/her light up.
Source: Left: Alavi and Dillenbourg (2012[24]), Right: Bachour, Kaplan and Dillenbourg (2010[28])

While the previous examples rely on visual perception, Moher et al. (2010[29]) also exploited sound for a classroom simulation on seismology. Over a period of 6 weeks, 21 earthquakes were simulated in the classroom. A low-frequency rumbling sound was generated by a subwoofer and displays located in different parts of the room simulated seismographs, showing a continuously running strip chart recorder of ground vibrations. Then, during lessons dedicated to earthquakes, students analysed the seismograph waves in order to locate the earthquake epicentre inside the classroom and to mark it by hanging a Styrofoam ball from the ceiling whose colour indicated the magnitude of the earthquake. While science simulations in schools usually run inside computers, this simulation was embedded in physical space, i.e. the classroom was the output.

System functions

A digital system processes data between the input and the output, a typical example being to aggregate data over time or across learners. These processes implement the functions expected from the system. The overarching function of classroom analytics, classroom orchestration, is fulfilled by implementing some specific functions. Seven more specific functions help to understand the current possibilities offered by this approach: monitoring and intervention, data propagation, team formation, debriefing, timing transitions, teacher self-regulation and, orchestration as a whole.

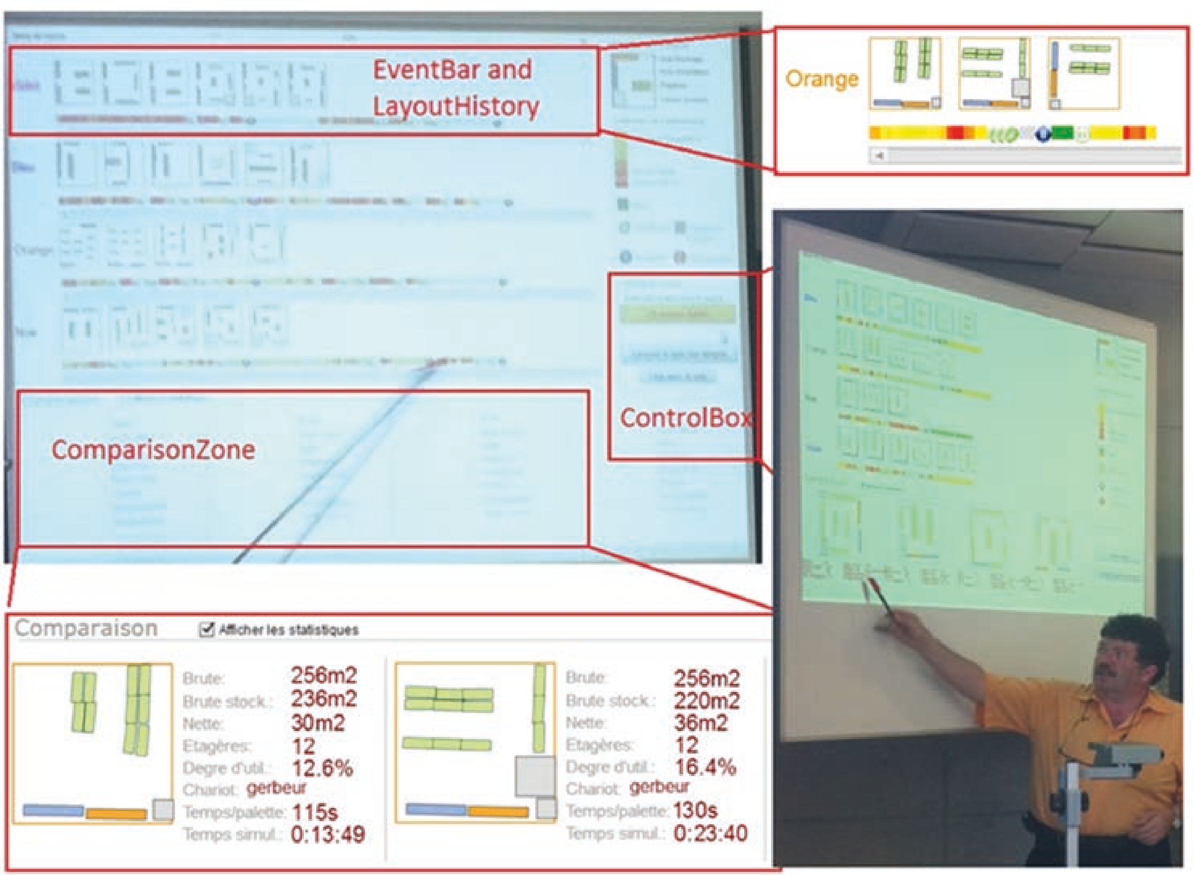
Monitoring and intervention

The main function of classroom dashboards is to monitor the state of the learners in order to detect which learner is inactive or struggling, which teams do not collaborate well, which learner could help another one, etc. Why would a teacher benefit from this information when he/she can, in a glance, see what the learners located in the classroom are doing? There are several answers to this question: when the number of learners is very high; when the learners' activities are not easily seen by the teacher (e.g. working on laptops); when the student's activity cannot be assessed at a glance (e.g. when they are writing complex code); when direct observation is intractable (e.g. monitoring 15 teams of 2 learners); when what matters is not only the current state, but what learners have done since the lesson outset, etc. In a nutshell, the key functionality of system is to make visible what is invisible, e.g. how long a learner has been silent, how much a learner dominates his teammates in a group discussion, etc.

Figure 5.5 illustrates this principle. Four teams are using the tangible logistics simulation tools shown in Figure 5.1. The four lines in the top panel show the history warehouse layouts designed by each team, which helps the teacher to perceive their strategy. Experiments carried out by Do Lenh et al. (2012[1]) showed that the pairs who modified the warehouse layout without much reflection and frequently ran the simulation did not learn much. Therefore, the dashboard includes the colour bar below the history which records the frequency of the warehouse manipulations, from yellow to red (too many manipulations). Students moving plastic shelves on the table is visible, but the variations of frequency of these movements for four teams is not visible.

For this function, the data processes are aggregation and evaluation. Aggregation consists of cumulating answers or behaviours over time and over teams to provide teachers with visualisations (timelines, histograms, etc.) that restore the "at a glance" effect. Simple evaluation processes compare the aggregated data to some thresholds (e.g. more than 5 minutes' idle time; less than 30% correct responses) or use a colour code from least to most desirable as in Figure 5.5. More sophisticated evaluation methods are, for instance, code synthesis (e.g. highlighting incorrect lines in the code written by learners) and text processing (e.g. finding similarities between texts). The goal is not to intervene in place of the teacher, but to trigger an alert for the teacher, an invitation to pay attention to someone or something. For instance, in Figure 5.5, when the colour bar includes many red periods, the teacher may pause the simulation and ask teams to predict the outcomes of the next simulation before restarting it. This triggers reflective effort. As we previously observed, teams with fewer reflection phases achieved lower learning gains.

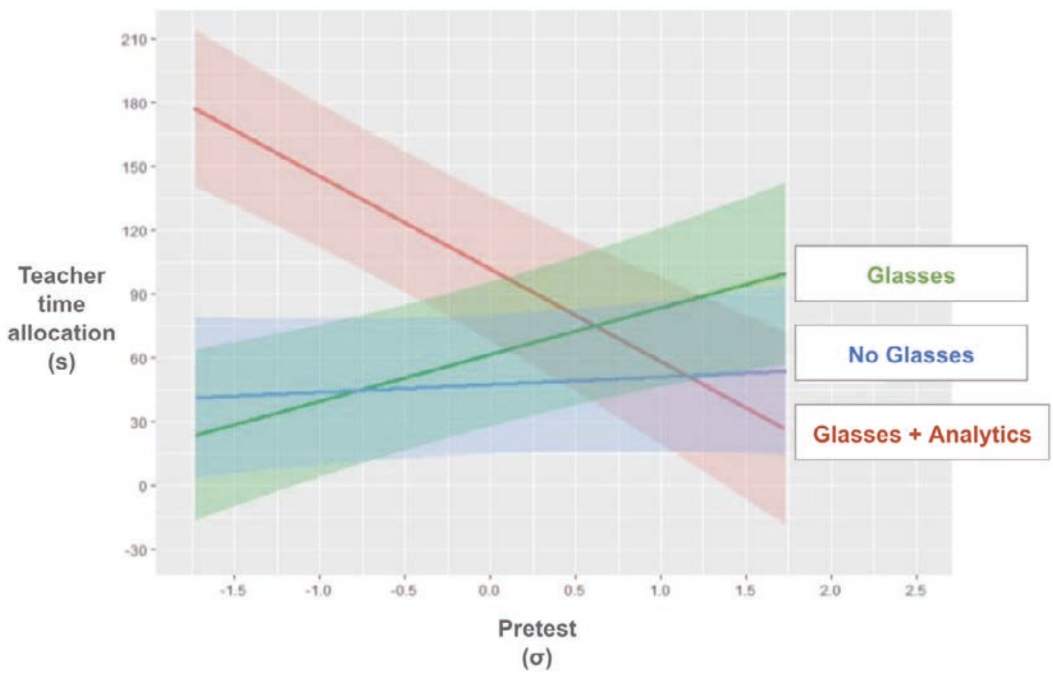
Figure 5.5 Teaching dashboard of a logistics lesson



Source: Do-Lenh et al. (2012[1]).

A plausible hypothesis is that, by supporting teachers' monitoring and intervention for learning tasks and processes that they might otherwise not perceive, dashboards would lead to higher learning outcomes. Evidence supporting this assumption is not abundant at present. In the previous example, Do Lenh et al. (2012[1]) showed that using the dashboard actually led to higher learning gains, but as the dashboard was combined with other orchestration tools, they may not be due to the dashboard itself. The state of research on teaching dashboards still has to mature. Schwendimann et al. (2017[30]) analysed 55 publications on this topic; only 15 included an evaluation in authentic contexts, the majority being based on questionnaires to teachers or learners. Only four of these papers actually measured the effects on learning. A more robust piece of evidence came from an experiment with 286 middle-school students: Holstein, Mc Laren and Aleven (2018[2]) showed that head-up display dashboards actually led to better orchestration which, in turn, increased learning gains for students using an intelligent tutoring system in mathematics. It is very interesting to zoom in on the relationship between providing a dashboard and the increase of learning gains. Holstein, McLaren and Aleven (2019[31]) observed (Figure 5.6) that the use of the dashboards led teachers to change their time allocation and pay more attention to weaker students, while it was the other way around without the dashboard.

Figure 5.6 Use of a dashboard to increase learning gains



Note: The teachers with the dashboard displayed on their glasses (red line) pay more visual attention (vertical axis) to weaker students, i.e. low pretest score (left end of the horizontal axis), than teachers without glasses (control group) or with inactive glasses (placebo group). Shaded regions indicate standard errors.
Source: Holstein, McLaren and Aleven (2019[31]).

Data propagation

Another orchestration function supported by classroom dashboards and analytics is to feed an activity with the data produced into a different activity. Some examples include:

1. During the first activity, Teams A and B each invent a small math problem. In the next activity, A solves B's problem and vice versa. The data process is simply to rotate the problem statements across the teams.

2. First, students are invited to enter the country of birth of their grandparents. In the next activity, the teacher shows a map that visualises the migration flows over two generations. The data process is to aggregate and visualise the individual data.

3. Learners collect pictures of mushrooms during a trip to the forest. In the next activity, they work in teams to classify the set of pictures collected by the class. The data process is to aggregate all of the pictures, but it could also include automatic picture annotation from existing mushroom libraries.

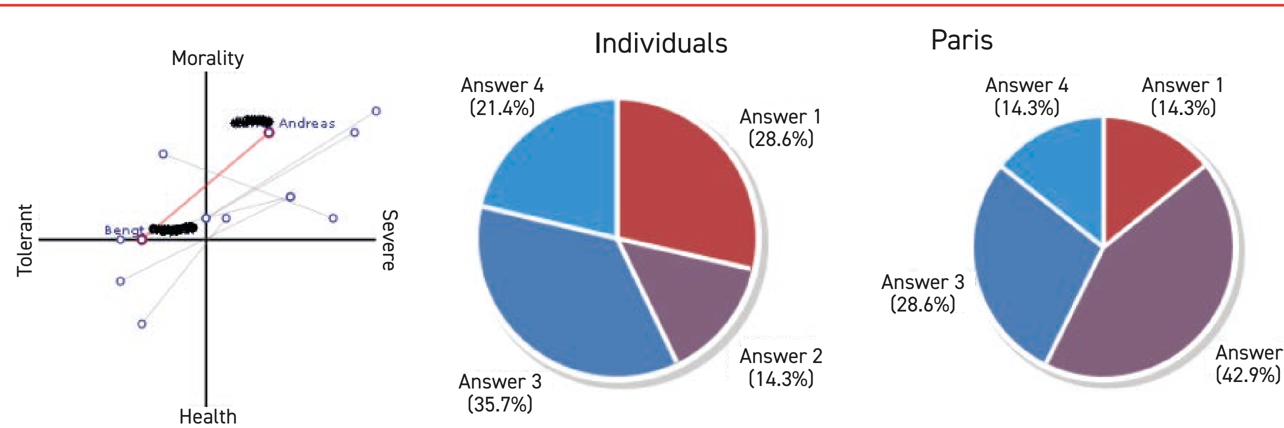
The list of examples is infinite and only bound by teachers' imagination. The "data propagation" function refers to a learning situation where an activity produces objects (or data) that are processed by an operator to feed into a subsequent activity. For physical objects, the operator is physical: in the first example, Teams A and B could simply exchange a sheet of paper. For digital objects, Dillenbourg (2015[32]) proposed a taxonomy of 26 operators that connect 2 or more learning activities. This flow of data across activities, referred to as a workflow, enables rich pedagogical scenarios. It can also create some rigidity, for instance if one team drops out. One challenge is to develop flexible workflows that enable teachers to fix on-the-fly the unexpected events that inevitably populate classroom life.

The two next subsections highlight two specific cases of data propagation that are especially relevant for classroom orchestration: team formation and debriefing.

Team formation

A specific function of classroom analytics is to process the data produced by learners in one activity in order to form dynamic teams for a subsequent activity. This function can be illustrated with an often tested pedagogical scenario that scaffolds cognitive conflict between peers (Dillenbourg and Jermann, 2007[33]). It is inspired by socio-constructivist theories that predict that the interactions necessary to overcome a cognitive conflict enhance learning (Doise, Mugny and Perret-Clermont, 1975[34]). In the first activity, each student responds to an online multiple-choice questionnaire. The questions do not have right or wrong answers, but reflect different viewpoints. For each answer, the students have to write a few words justifying their choice. In the second activity, the system produces a specific dashboard, a map of opinions (Figure 5.7, left panel): every answer in the first activity has been associated with an x,y value on the map. The teacher discusses this dashboard with students, who often comment on their positions. The system forms pairs of students in a way that maximises their distance on the map; that is, it finds students whose responses reveal opposite opinions. In the third activity, pairs are asked to answer the same online questionnaire as in the first activity. The environment provides them with the answers and justifications provided individually. In the fourth activity, the teacher uses another dashboard (Figure 5.7, right panel) for the debriefing activity (see Debriefing).

Figure 5.7 Using dashboards to form dynamic teams



Note: Left panel: The map represents the learners' opinions and the lines indicate how individuals with different opinions have been paired. Right panel: the pie charts indicate how they change opinions between the individual and the collaborative activities, which the teacher exploits during the debriefing activity.
Source: Dillenbourg and Jermann (2007[33])

The data operator used for this function consists of maximising differences within teams. This is also the case in the example of Gijlers and De Jong (2005[35]) on learning from simulations: they form teams of individuals who previously expressed an opposite hypothesis in order to reduce the natural bias of designing experiments that confirm one's own hypothesis.

Another pedagogical scenario could use an operator that minimises the difference among team members, e.g. making teams of learners who made the same error in the previous exercises. Group formation is an example of data propagation in which the data fed into an activity are not the object of the activity, but its social organisation.

Debriefing

The dashboard presented in Figure 5.5 includes a tool (bottom part) that enables the teacher to select two warehouse layouts designed by teams and to compare them in terms of capacity and performance. This allows the teacher to debrief the exploration activities of his/her students. Debriefing means reflecting on what has been done in order to extract concepts or principles to be taught. By comparing warehouses, the teacher will, for instance, illustrate the trade-off between capacity and performance. The dashboard on the right in Figure 5.7 is used by the teacher to push students to explain why they change opinions between the individual and collaborative phases in order to later connect their explanations to the scientific debate.

The role of debriefing activities is a critical orchestration phase in constructivist learning scenarios, based on discovery or open problem-solving activities. These approaches have been criticised by some for being unproductive per se. However, Schwartz and Bransford (1998[36]) or Kapur (2015[37]) showed that, if this phase of exploration is followed by direct instruction, the lesson is actually more effective than the opposite sequence, i.e. if direct instruction is followed by application exercises. The reason is that, during the exploration activity, learners rarely shout "eureka". More commonly, they get some intuition, some vague ideas, on the basis of which the concepts can later be clarified. They ask themselves questions that will give meaning to the teachers' lecture. As Bransford and Schwartz (1998[36]) put it, there is a "time for telling", but this instruction phase has to build on what learners have done during the exploration phase. It should not be a standard lecture disconnected from their experience. This is a very demanding activity for teachers, as it includes improvisation. The "debriefing" function of classroom analytics aims to support this task by collecting the learners' productions, comparing them, annotating them, etc., and facilitating their exploitation by the teacher.

Timing transitions

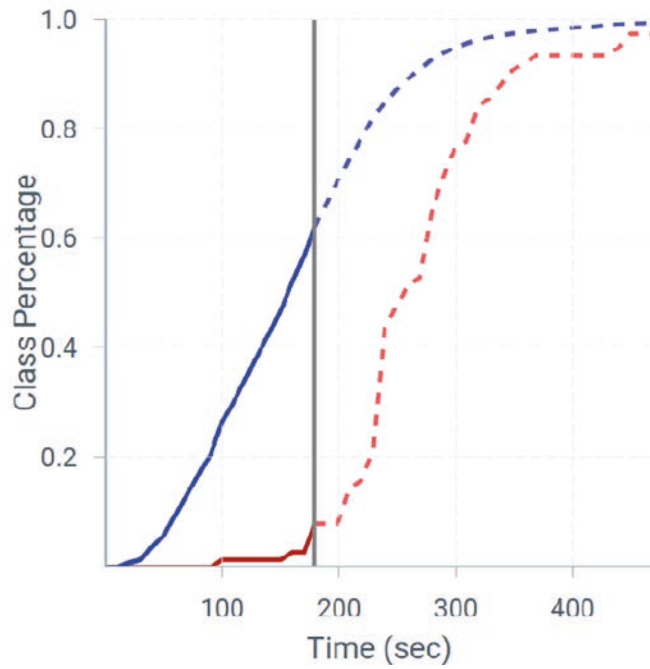
Orchestrating learning within classrooms is a very time-constrained process. Teachers permanently compare the remaining class time to the pedagogical activities to be done. In addition, orchestrating rich learning scenarios is difficult while transitioning between activities involving different social levels: individual activities, team activities or class activities, referred to as "social planes". A typical trade-off is as follows: the teacher had planned to devote 15 minutes to an individual exercise to practice some skill and this skill is required for the next activity, conducted in teams. After 15 minutes, the teacher realises that some learners have not finished their individual exercises. If he/she decides to start the next activity anyway, these late students and their teammates will be penalised. If he/she gives late individuals five more minutes, he/she will have to reduce the time budget of the next activity, and moreover, will have a majority of students wasting time and engaging in off-task interactions. Similar constraints arise when, as in the examples given in section Team formation, individual answers need to be provided in order to make automatic group formations.

One example of classroom analytics that addresses this issue is the "time extension gain", i.e. the percentage of additional students who would complete the activity if its duration was extended by one single time unit. On the progression chart presented in Figure 5.8, the "time extension gain" corresponds to the slope of the curve (Faucon et al., 2020[38]). When the curves become flat, it is time to move to the next activity. This chart has been used in real time for orchestrating a variety of activities in lecture theatres. Future dashboards are expected to provide more instances of similar time prediction tools, thus supporting teachers to introduce short activities without wasting time.

Teacher self-regulation

Since the input of classroom analytics can be any event in the classroom, classroom analytics can also capture the teacher's behaviour. So far, learning analytics have not often included the teacher in their analysis since they typically analyse behaviours in a learning environment where the teacher has no or few possibilities of intervening. On the contrary, modelling classroom processes requires modelling the teacher's behaviour, since the teacher plays a critical role in the teaching and learning processes: how much does the teacher speak, did they distribute verbal attention to all learners, how do they decide to whom they will ask a question, did they walk across the classroom, how varied is the tone of their voice?

Figure 5.8 Progression chart of the time extension gain



Note: When students are engaged in an activity composed of multiple steps, this dashboard display shows the percentage of students who are doing the activity (blue) and the percentage of students who have completed the activity (red). The vertical bar is the moment when the snapshot was taken. The plain lines on the left are based on observed data while the dotted lines on the right are predictions based on data collected in previous experiments. In this example, the teacher could move on after approximately 300 seconds, when about 95% of the class has completed the activity. Three additional minutes (stopping at 480 seconds) would allow virtually all students to complete the activity, but 95% of the class would have wasted time.
Source: Faucon et al. (2020[38])

For instance, classroom analytics may make a teacher realise that he/she has been speaking much longer than they planned to or that he/she has been neglecting some learners. This enables real-time self regulation, what Schön (2017[39]) called "reflection-in-action", which is cognitively demanding. More sophisticated analytics can also (or alternatively) be provided after the lesson, supporting "reflection-on-action" that enables the teacher to reflect later on in order to improve his/her teaching over time – a powerful form of professional development.

Prieto, Sharma and Dillenbourg (2015[40]) combined eye-tracking measures and personal questionnaires to estimate the orchestration load of teachers in a classroom. They found that high-load episodes occur when teachers provide explanations or questions to the entire class, often looking at students' faces in an attempt to assess their progress and understanding, which confirms the relevance of dashboards that provide such information. On the other hand, low-load episodes tended to correspond to individual or small group feedback, during which the teacher often focuses on the students' worksheets or laptops. By combining eye-tracking measures with other sensors (electroencephalogram, accelerometers, etc.), Prieto et al. (2016[41]) applied machine-learning methods (random forest and gradient-boosted decision trees) automatically characterise the ongoing learning activities. The lesson was following an orchestration graph composed of only two planes: team and class activities. In Figure 5.9, the colour represents the teacher's activity. The algorithm identified the plane of interactions with an accuracy of 90% but was less accurate, only 67%, for identifying the teacher's activity based on those digital observations.

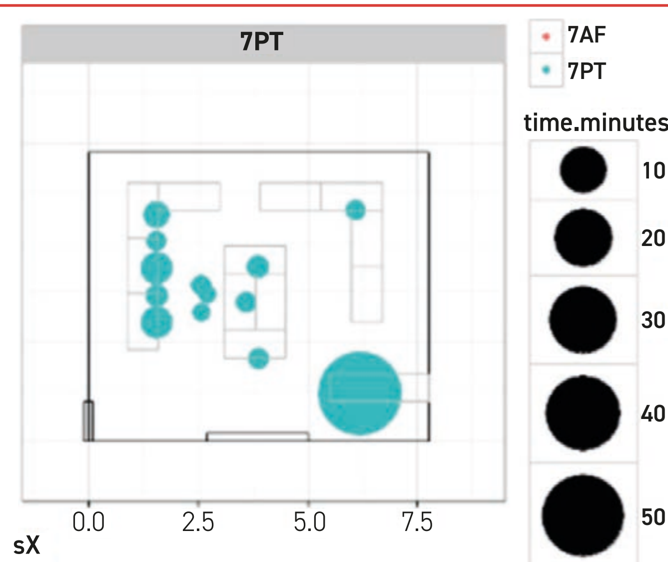
Now, any observation tool designed for teacher professional development may quickly drift into a teacher control or evaluation tool. The recommendation for this other touchy ethical issue is to strive for minimalism (capturing only information that may improve teaching) and to trust regulation and education stakeholders' self-regulation (showing data to the teachers only and not to the school principals, other people in their hierarchy, parents or students).

Figure 5.9 Progression chart of the time extension gain

Note: These simplified orchestration graphs represent the activities that actually happened during the lesson (upper panel) and those identified by the teaching analytics algorithm (lower panel).
Source: Prieto et al. (2016[41])

As for learning analytics, classroom analytics focusing on teaching are useful if they help teachers reflect on their practice in order to improve it. In another study, Prieto et al. (2017[42]) showed teachers their location in the classroom. One teacher, whose classroom locations during a class are represented in Figure 5.10, was indeed surprised to discover that she neglected the right-hand part of the classroom and offered her support mainly to the left-hand side and middle tables (when not at her desk). This behaviour is not problematic in itself: perhaps students sitting at these tables required more support than those on the right-hand side of the classroom. However, this is a good illustration of how classroom analytics can help teachers identify some characteristics or possible issues in their teaching practice – for example, if they show her that she mainly supports students who are strong academically rather than those who struggle, that she ignores her female students during science lessons, or students with minority or underprivileged backgrounds, etc.

Figure 5.10 Showing teachers their location in the classroom



Source: Prieto et al. (2017[42])

Orchestration

Classroom analytics aim to facilitate the orchestration of learning activities, which is not a single function, but an umbrella concept that includes many functions. Six of them have been described in the previous subsections. Basically, classroom analytics are designed to empower teachers in the demanding task of conducting rich scenarios; with individual, classroom and class-wide activities; with or without digital tools; considering all of the practical constraints of daily classroom life. The idea of empowering teachers to better orchestrate learning in their classroom can be viewed as provocative in times when many scholars define the teacher's role as being a facilitator or a guide on the side. Empowering teachers for a better orchestration of learning does not mean increasing lecturing time; it is about supporting them in steering rich learning scenarios, whatever their actual components. Implementing constructivist learning scenarios with 30 learners requires the teacher to feel comfortable with driving the variety of pedagogical activities they include.

Technology can help in different ways. Or not. For example, a typical mistake is to suddenly distribute a tablet to every child in a classroom. This may destroy teachers' established orchestration habits. What should the learner do? How can the teacher get their attention? Most initiatives to massively introduce technologies in education have failed because the hardware availability was not the bottleneck. The key is to provide teachers with scenarios describing which learning activities they can ask their students to do with technology. The answer is not one or two specific activities, but scenarios that include multiple activities, with or without technology, and embrace the whole classroom life. This is the educational proposition of classroom analytics – a new way to empower teachers in their classroom.

Perspectives

This chapter presents a new vision according to which a future classroom could be viewed as a digital system. It proposed the term "classware" to describe such a digital system which captures and supports classroom processes. This is a concept for the years to come. Classrooms are not digital systems yet, and very few classroom analytics today could be described as "classware". The chapter developed a graphical language to model-integrated pedagogical scenarios, called "orchestration graphs" (Dillenbourg, 2015[32]), as a first step to modelling the flow of data required by classroom orchestration.

Coming back to the car example – or perhaps the "connected home" idea, classrooms within schools could look apparently very much the same as today, but they could be equipped with sensors feeding learning analytics and digital tools that would not only help teachers to orchestrate rich learning scenarios during class, but would also give them feedback on their teaching in real time, and food for thought for improving their teaching, and, as a result, students' learning outcomes. Before this new schooling model emerges, research and development on classroom analytics and on better understanding the types of dashboards that would make the display of information the most helpful to teachers has to continue. The ethical and privacy issues these developments may raise also need to be addressed.

There are several immediate implications of this vision though.

First, this vision is a thinking tool for decision makers when designing or assessing educational projects. They should favour projects that do not bet on a single pedagogical approach or a single technology. Rather than projects that focus, for instance, on teamwork but neglect the need for individual practice, they should support projects that integrate individual activities, team activities and class-wide activities into consistent pedagogical scenarios. Good development projects cannot be defined by a technology only either, e.g. "virtual reality for X" or "3D printers for Y". The belief that technologies have intrinsic effects have been proven wrong many times. A promising school or education research project should be defined by pedagogical goals that drive the necessary sequence of classroom activities, whatever technology is supporting these activities and their orchestration at this point only. Technology can help individual learning, but it can also help and empower teachers to mix learning scenarios including activities with no technology, activities carried out with or without technology, or fully digital activities. All of those scenarios are possible, and none is intrinsically superior to the others.

Second, this proposition implies that the design of learning technologies should embed these classroom orchestration functions. Schools will fully exploit the potential of digital technologies only when teachers feel empowered and confident in using them. This will not come rapidly, but could come with the help of technology.

The education technology (EdTech) market is currently structured by competition between various tools, while the integration of digital tools is the condition to build digital ecosystems – and a sustainable EdTech market.

Third, all teacher training programme courses should include some learning on digital learning technologies. Currently, they often include one or two courses on that topic, but digital technologies can and should support any type of teaching.

Finally, policy makers and other education stakeholders have to address regulatory as well as ethical issues regarding classroom analytics. Everything that can be done within a regulatory framework is not necessarily desirable. Greater collaboration between researchers in learning sciences and in data protection should occur and inform regulation and ethical practice.

The road for learning technology and classroom analytics to reach their maturity is still long. But we may see the light at the end of the tunnel much sooner than many expect.

Acknowledgements

The presented technologies were developed by many current or former lab members: Jennifer Olsen, Patrick Jermann, Stian Haklev, Louis Faucon, Luis Prieto Santos, Son Do Lenh, Sébastien Cuendet, Guillaume Zufferey, Hamed Alavi, Khaled Bachour and Fredéric Kaplan. The development of a research community around orchestration benefited from interactions with Frank Fischer, Miguel Nussbaum, Yannis Dimitriadis, Manu Kapur, Nikol Rummel, Vincent Aleven and Chee Kit Looi. The author would also like to thank the colleagues who reviewed this chapter (Ryan Baker and Stéphan Vincent-Lancrin).

References

Ahuja, K., D. Kim, F. Xhakaj, V. Varga, A. Xie, S. Zhang, J.E. Townsend, C. Harrison, A. Ogan and Y. Agarwal (2019), "EduSense", *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies,* Vol. 3/3, pp. 1-26, http://dx.doi.org/10.1145/3351229. [18]

Alavi, H. and P. Dillenbourg (2012), "An Ambient Awareness Tool for Supporting Supervised Collaborative Problem Solving", *IEEE Transactions on Learning Technologies*, Vol. 5/3, pp. 264-274, http://dx.doi.org/10.1109/tlt.2012.7. [24]

Alcoholado, C., N. Nussbaum, A. Tagle, F. Gomez, F. Denardin, H. Susaeta, M. Villalta and K. Toyama (2012), "One Mouse per Child: interpersonal computer for individual arithmetic practice", *Journal of Computer Assisted Learning,* Vol. 28/4, pp. 295-309, http://dx.doi.org/10.1111/j.1365-2729.2011.00438.x. [16]

Bachour, K., F. Kaplan and P. Dillenbourg (2010), "An Interactive Table for Supporting Participation Balance in Face-to-Face Collaborative Learning", *IEEE Transactions on Learning Technologies,* Vol. 3/3, pp. 203-213, http://dx.doi.org/10.1109/tlt.2010.18. [28]

Bakharia, A., K. Kitto, A. Pardo, D. Gašević and S. Dawson (2016), "Recipe for success", *Proceedings of the Sixth International Conference on Learning Analytics & Knowledge - LAK '16,* http://dx.doi.org/10.1145/2883851.2883882. [12]

Bloom, B. (1968), "Learning for mastery: Instruction and curriculum", *Regional Education Laboratory for the Carolinas and Virginia. Topical Papers and Reprints.,* Vol. 1, https://files.eric.ed.gov/fulltext/ED053419.pdf. [5]

Cuendet, S., J. Dehler-Zufferey, G. Ortoleva and P. Dillenbourg (2015), "An integrated way of using a tangible user interface in a classroom", *International Journal of Computer-Supported Collaborative Learning*, Vol. 10/2, pp. 183-208, http://dx.doi.org/10.1007/s11412-015-9213-3. [13]

D'Mello, S., B. Lehman, R. Pekrun and A. Graesser (2014), "Confusion can be beneficial for learning", *Learning and Instruction*, Vol. 29, pp. 153-170, http://dx.doi.org/10.1016/j.learninstruc.2012.05.003. [22]

Dillenbourg, P. (2015), *Orchestration Graphs*, EPFL. [32]

Dillenbourg, P. (2013), "Design for classroom orchestration", *Computers & Education*, Vol. 69, pp. 485-492, http://dx.doi.org/10.1016/j.compedu.2013.04.013. [14]

Dillenbourg, P., S. Järvelä and F. Fischer (2009), "The Evolution of Research on Computer-Supported Collaborative Learning", in *Technology-Enhanced Learning,* Springer Netherlands, Dordrecht, http://dx.doi.org/10.1007/978-1-4020-9827-7_1. [7]

Dillenbourg, P. and P. Jermann (2007), "Designing Integrative Scripts", in *Scripting Computer-Supported Collaborative Learning*, Springer US, Boston, MA, http://dx.doi.org/10.1007/978-0-387-36949-5_16. [33]

Dillenbourg, P., G. Zufferey, H. Alavi, P. Jermann, S. Do-Lenh and Q. Bonnard (2011), *Classroom orchestration: The third circle of usability,* Proceedings of the 9th Computer-Supported Collaborative Learning Conference, Hong-Kong. July 4-8, 2011. [27]

Doise, W., G. Mugny and A. Perret-Clermont (1975), "Social interaction and the development of cognitive operations", *European Journal of Social Psychology*, Vol. 5/3, pp. 367-383, http://dx.doi.org/10.1002/ejsp.2420050309. [34]

Do-Lenh, S., P. Jermann, A. Legge, G. Zufferey and P. Dillenbourg (2012), "TinkerLamp 2.0: Designing and Evaluating Orchestration Technologies for the Classroom", in *Lecture Notes in Computer Science, 21st Century Learning for 21st Century Skills,* Springer Berlin Heidelberg, Berlin, Heidelberg, http://dx.doi.org/10.1007/978-3-642-33263-0_6. [1]

Duval, E. (2001), "Metadata standards: What, who & why", *Journal of Universal Computer Science*, Vol. 7/7, pp. 591-601. [9]

Fagen, A., C. Crouch and E. Mazur (2002), "Peer Instruction: Results from a Range of Classrooms", *The Physics Teacher,* Vol. 40/4, pp. 206-209, http://dx.doi.org/10.1119/1.1474140. [15]

Faucon, L., J.K. Olsen, S. Haklev and P. Dillenbourg (2020), "Real-Time Prediction of Students' Activity Progress and Completion Rates", *Journal of Learning Analytics,* Vol. 7/2, pp. 18-44, http://dx.doi.org/10.18608/jla.2020.72.2. [38]

Gellersen, H., A. Schmidt and M. Beigl (1999), "Ambient media for peripheral information display", *Personal Technologies,* Vol. 3/4, pp. 199-208, http://dx.doi.org/10.1007/bf01540553. [26]

Gijlers, H. and T. de Jong (2005), *Confronting ideas in collaborative scientific discovery learning,* Paper presented at AERA 2005. [35]

Holstein, K., G. Hong, M. Tegene, B. McLaren and V. Aleven (2018), "The classroom as a dashboard", P*roceedings of the 8th International Conference on Learning Analytics and Knowledge,* http://dx.doi.org/10.1145/3170358.3170377. [23]

Holstein, K., B. McLaren and V. Aleven (2019), "Co-Designing a Real-Time Classroom Orchestration Tool to Support Teacher–AI Complementarity", *Journal of Learning Analytics*, Vol. 6/2, http://dx.doi.org/10.18608/jla.2019.62.3. [31]

Holstein, K., B. McLaren and V. Aleven (2018), "Student Learning Benefits of a Mixed-Reality Teacher Awareness Tool in AI-Enhanced Classrooms", in *Lecture Notes in Computer Science, Artificial Intelligence in Education*, Springer International Publishing, Cham, http://dx.doi.org/10.1007/978-3-319-93843-1_12. [2]

Holstein, K., B. McLaren and V. Aleven (2017), "SPACLE", *Proceedings of the Seventh International Learning Analytics & Knowledge Conference*, http://dx.doi.org/10.1145/3027385.3027450. [3]

Kapur, M. (2015), "The preparatory effects of problem solving versus problem posing on learning from instruction", *Learning and Instruction*, Vol. 39, pp. 23-31, http://dx.doi.org/10.1016/j.learninstruc.2015.05.004. [37]

Lin, C. and M. Chi (2016), "Intervention-BKT: Incorporating Instructional Interventions into Bayesian Knowledge Tracing", in *Intelligent Tutoring Systems, Lecture Notes in Computer Science*, Springer International Publishing, Cham, http://dx.doi.org/10.1007/978-3-319-39583-8_20. [21]

Mangaroska, K., B. Vesin and M. Giannakos (2019), "Cross-Platform Analytics", *Proceedings of the 9th International Conference on Learning Analytics & Knowledge*, http://dx.doi.org/10.1145/3303772.3303825. [11]

Moher, T., J. Wiley, A. Jaeger, B. Lopez Silva, F. Novellis and D. Lyman Kilb (2010), "Spatial and temporal embedding for science inquiry: An empirical study of student learning", *Proceedings of the 9th International Conference of the Learning Sciences*, Vol. 1, pp. 826-833. [29]

Ochoa, X. and M. Worsley (2016), "Editorial: Augmenting Learning Analytics with Multimodal Sensory Data", *Journal of Learning Analytics*, Vol. 3/2, pp. 213-219, http://dx.doi.org/10.18608/jla.2016.32.10. [17]

Papert, S. (1987), "Microworlds: Transforming education", *Artificial Intelligence and Education*, Vol. 1: Learning Environments and Systems, pp. 79-94. [6]

Prieto, L., P. Magnuson, P. Dillenbourg and M. Saar (2017), *Reflection for Action: Designing Tools to Support Teacher Reflection on Everyday Evidence*, Center for Open Science, http://dx.doi.org/10.31219/osf.io/bj2rp. [42]

Prieto, L., K. Sharma and P. Dillenbourg (2015), "Studying Teacher Orchestration Load in Technology-Enhanced Classrooms", in *Design for Teaching and Learning in a Networked World, Lecture Notes in Computer Science*, Springer International Publishing, Cham, http://dx.doi.org/10.1007/978-3-319-24258-3_20. [40]

Prieto, L., K. Sharma, P. Dillenbourg and M. Jesús (2016), "Teaching analytics", *Proceedings of the Sixth International Conference on Learning Analytics & Knowledge - LAK '16*, http://dx.doi.org/10.1145/2883851.2883927. [41]

Raca, M., L. Kidzinski and P. Dillenbourg (2015), *Translating head motion into attention-towards processing of student's body language*, Proceedings of the 8th International Conference on Educational Data Mining, https://files.eric.ed.gov/fulltext/ED560534.pdf. [20]

Santos, P. (2012), *Supporting orchestration of blended CSCL scenarios in distributed learning environments*, Doctoral dissertation - Universidad de Valladolid, http://dx.doi.org/10.35376/10324/1794. [4]

Schön, D. (2017), *The Reflective Practitioner: How Professionals Think In Action.*, Routledge. [39]

Schwartz, D. and J. Bransford (1998), "A time for telling", *Cognition and Instruction*, Vol. 16/4, pp. 475-522, http://www.jstor.org/stable/3233709. [36]

Schwendimann, B., M. Jesús, A. Vozniuk, L. Prieto, M. Shrivani Boroujeni, A. Holzer, D. Gillet and P. Dillenbourg (2017), "Perceiving Learning at a Glance: A Systematic Literature Review of Learning Dashboard Research", *IEEE Transactions on Learning Technologies*, Vol. 10/1, pp. 30-41, http://dx.doi.org/10.1109/tlt.2016.2599522. [30]

Severance, C., T. Hanss and J. Hardin (2010), "Ims learning tools interoperability: Enabling a mash-up approach to teaching and learning tools", Technology, *Instruction, Cognition and Learning,* Vol. 7/3-4, pp. 245-262. [10]

Tomitsch, M., T. Grechenig and S. Mayrhofer (2007), "Mobility and emotional distance: exploring the ceiling as an ambient display to provide remote awareness", *3rd IET International Conference on Intelligent Environments (IE 07),* http://dx.doi.org/10.1049/cp:20070362. [25]

Vygotsky, L. (1964), "Thought and language", *Bulletin of the Orton Society,* Vol. 14/1, pp. 97-98, http://dx.doi.org/10.1007/bf02928399. [8]

Yang, D., A. Alsadoon, P.W.C. Prasad, A.K. Singh and A. Elchouemic (2018), "An Emotion Recognition Model Based on Facial Recognition in Virtual Learning Environment", *Procedia Computer Science*, Vol. 125, pp. 2-10, http://dx.doi.org/10.1016/j.procs.2017.12.003. [19]

6 Serving students with special needs better: How digital technology can help

Judith Good

University of Amsterdam, Netherlands; formerly University of Sussex, United Kingdom.

This chapter explores the role of technology in supporting students with special needs. The support ranges from helping disabled students to access the curriculum to providing disability specific support so that students can participate in inclusive school settings. After highlighting the importance of supporting students with special needs, the chapter shows how technology can support a variety of special needs. It then focusses on three cutting-edge technologies which aim to: 1) support the development of autistic children's social skills, 2) diagnose and support students with dysgraphia and 3) provide access to graphical materials for blind and visually impaired students. The examples highlight the importance of involving students and stakeholders in the design of the solutions and the need for developers to consider affordability as a key element of their development.

Introduction

The role of technology in providing support to students with special needs is widely recognised, and there is evidence for the effectiveness of both particular hardware platforms, such as mobile devices (Chelkowski, Yan and Asaro-Saddler, 2019[1]; Ok and Kim, 2017[2]) as well as numerous specialised software and apps. Educational technologies are likely to play an increasingly pervasive role for students with special needs, with educators called to keep abreast of technology developments so as to make informed decisions on the use of such technologies in the classroom (McLeskey et al., 2017[3]).

Interestingly however, despite the plethora of educational technologies for special needs (see, e.g. Cheng and Lai, 2020[4]; Erdem, 2017[5] for recent reviews), few could be considered to be "smart". In parallel with this, there is a significant and longstanding body of work in the area of artificial intelligence and education (for recent reviews, see Alkhatlan and Kalita, 2018[6]; Chen, Chen and Lin, 2020[7]) as well as clear evidence of the effectiveness of intelligent tutoring systems above and beyond other forms of computer-based learning (Kulik and Fletcher, 2016[8]). However, there is a dearth of educational artificial intelligence (AI) systems which target students with special needs. Indeed, a review of articles published in the *International Journal of Artificial Intelligence in Education* in the past five years did not uncover any which focused on inclusion, accessibility or special educational needs. As noted by Kazimzade et al. (2019[9]), although adaptive educational technologies and inclusion, in the broadest sense, are two key considerations in the current educational landscape, they intersect more seldom than one would expect.

It is difficult to say why there is so little overlap between these two areas. When considering educational provision for children with special needs, the World Health Organization (2011[10]) has highlighted the need to adopt more learner-centred approaches, which recognise differences in the way that people learn, and are able to flexibly adapt to individual learners. As such, this seems like the ideal opportunity to explore the ways in which already established and newly emerging methods and approaches in the field of AI and education could be extended and adapted to provide support for children with special needs.

The potential of smart technologies to serve and support students with special needs is particularly important in light of the likely increase in students identified with such needs. In 2000, the OECD estimated that at some point during their schooling, approximately 15 to 20% of young people would be considered to have a special educational need (OECD, 2000[11]). Twenty years on, this figure is likely to be even higher, given that the recognition of disability among children has been increasing steadily year on year (Houtrow et al., 2014[12]). Although the rate of physical disabilities has decreased over time, there have been significant increases in the rate of developmental disabilities (Zablotsky et al., 2019[13]) with the latter now estimated to affect 17.8% of children in the United States (Zablotsky and Black, 2020[14]).

There are a number of possible reasons for this increase, including changing definitions of what constitutes a particular disability (in the case of autism, Volkmar and McPartland (2014[15]) provide a detailed account of changing conceptualisations since its first formal description in 1943) as well as improved access to diagnostic services. An in-depth consideration of these phenomena is beyond the scope of this report, however, it is important to highlight two issues. Firstly, given that more than one in six children are now considered to have a developmental disability, it is extremely likely that any mainstream classroom will have at least one student, and probably more than one, who will require additional resources to support their learning. Secondly, the steady increase in diagnostic rates may continue as we uncover additional forms of disability and become more accurate at identifying those we already know, further increasing the number of children who will require additional support.

In an educational context, disabled children are at a disadvantage compared to their typically developing peers. According to the World Health Organization (2011, p. 208[10]), "despite improvements in recent decades, children and youth with disabilities are less likely to start school or attend school than other children. They also have lower transition rates to higher levels of education", a trend which continues (UNESCO Institute for Statistics, 2018[16]). This, in turn, has long-term negative impacts on people's futures, potentially affecting their integration within society, and career prospects. For example, in the UK, only 16% of autistic adults are in full-time paid employment, despite the fact that 77% have expressed a desire to work (National Autistic Society, 2016[17]). Furthermore, of the small minority who are in employment, over half feel that their jobs do not make use of the skills they actually possess.

A final important consideration is that, given this increasing rate of identified developmental needs, supporting students with special needs increasingly intersects with the broader equity agenda. The development of technologies that help diagnose and address student disabilities (e.g. dyslexia, dysgraphia, dyscalculia, some hearing or visual impairments) will help close the achievement gap and improve learning outcomes within countries.

In this report, I consider the near future of developing "smart" technologies for students with additional needs, by focusing on three case studies and drawing conclusions for future work. Before doing so, I consider definitions of disability and special needs, and their intersections with education and with technology, in more detail below.

Education, technology and special needs

Broadly speaking, special needs support in an educational context refers to support for needs that a disabled child may have that are different to those of their typically developing peers (see Box 6.1 for terminology). Providing effective support for special needs is complex, and requires careful thought and planning. Students' needs change over time due to various factors (their individual developmental trajectory, previous support, etc.). Their needs may become more or less pronounced, requiring an ongoing assessment of what support is appropriate at any given time. Co-morbidity, defined as having more than one disability (also termed 'multi-morbidity' depending on the source), is another complicating factor.

Box 6.1 Conceptions of disability and special need support

There are no agreed definitions of disability or special needs and furthermore, the relationship between the two is not always straightforward. Definitions vary by country, and are categorised and classified in different ways. Diagnostic processes and pathways also vary, both within and across countries, and change over time. However, it is important to have a basic understanding of differing perspectives on the nature of disability, as they have significant implications for learning and education.

As our understanding of disability changes and develops, so too does the terminology used to describe it, which has led to different models of disability (Marks, 1997[18]). Traditional models, such as the medical model, focus on "impairment", and locate the source of that impairment within the individual, often with a view to trying to provide a "cure". In contrast, social models look at the intersection between individuals and their environments and, in particular, at the ways in which a particular environment might act to produce an impairment. For example, a wheelchair user within the medical model might be considered to have a prima facie impairment, whereas within the social model, the impairment could be considered to arise from the fact that a given building does not have ramps or lifts, rather than being an intrinsic characteristic of the individual.

Furthermore, certain terms can have negative connotations, such as "disabled", which implicitly suggests that most people are "abled". This can lead to stigma and exclusion (Sayce, 1998[19]). Proponents of the term "neurodiversity" (originally coined by Singer (1999[20]) in reference to autism, but now used more broadly for a number of conditions including attention deficit hyperactivity disorder (ADHD) and dyslexia consider these conditions to be neurological variations which have both positive and negative aspects. They reject attempts at "normalisation" and instead argue for a deeper understanding of these different ways of being in, and interacting with, the world.

It is important to recognise the tensions inherent in these different views of disability, and the ways in which they shape our perspectives on the types of support that are necessary, and the types of technologies that are designed as a result.

The additional resources needed to provide this support can take a number of forms, including financial resources, personnel resources (e.g. additional teachers, or teaching assistants), or material resources. This chapter focuses on the latter type of resource, considering how technologies, and smart technologies in particular, contribute to support students with additional needs.

Although there are numerous ways of further categorising this support, it can be useful to consider the *aim* of the support along a continuum. At one end are technologies designed to facilitate access to the curriculum and allow disabled children to participate in typical classroom learning activities. In this case, the technology allows children to access the same curricular content as their typically developing peers. As an example, providing blind or visually impaired (BVI) students with technologies that have text-to-speech capabilities will allow them to access (at least some of) the curricular materials used by their peers, making it easier for them to learn in an inclusive school setting.

At the other end of the continuum are technologies designed to address and provide support for issues related to the child's disability. In this case, the content of the intervention is typically *not* part of the standard school curriculum. An example of this type of technology would be interventions for autistic students which are designed to support the development of their social and communication skills. Technologies at this end of the continuum tend to be more contentious: as mentioned above, differing perspectives on disability can lead to debates around the types of interventions and technologies that might best support disabled children. Often these views are implicit, but nonetheless drive the development of educational technology, influencing decisions about what types of support are needed, and why.

As an example, a recent review of technologies for autistic children found that the majority focus on social skills training (Spiel et al., 2019[21]), which implicitly suggests that this is the area of most concern for parents and educators (even though it may not be). The authors maintain that many of these technologies require children to

> *"learn the modes of interaction that are deemed as appropriate by neurotypical adults without the adults having to learn how the autistic children might want to engage…" (Spiel et al., 2019, p. 18[21]).*

At the same time, the authors do acknowledge that having such skills might provide autistic children with strategies for coping in a neurotypical world. It may be that increased coping skills might, in turn, lead to improvements in mental well-being, and indeed, this seems to be the case, with interventions targeting social skills also reducing both depression and anxiety (Rumney and MacMahon, 2017[22]).

Finally, it is important to note that many of the technologies designed primarily to provide support for the specific aspects of a child's disability may well have a secondary effect of improving access to the standard curriculum. For example, providing support for social and communication skills for autistic children (support for special needs) may well help them to participate more easily in curricular activities which involve group work and collaboration (access to curriculum). Similarly, technologies designed to support children with ADHD with self-regulation skills, such as the ones described in Box 6.2 (support for special needs), may well lead to them being able to engage with more of the topics being taught, or to engage with them in more depth (access to curriculum).

Box 6.2 Technologies supporting students with attention deficit hyperactivity disorder (ADHD)

Technologies designed for students with ADHD focus on different aspects of the condition, for example self-regulation (i.e. learning to manage one's thoughts, behaviours and emotions). The technologies described below are not yet widely used, but are in test phases in various countries.

A key component of self-regulation is emotional regulation, which involves learning to recognise one's emotions and manage them in a way that is appropriate to the situation. One way of emotionally regulating and reducing stress is through breathing exercises. However, children may not find the exercises engaging or motivating. ChillFish, a Danish solution, is a biofeedback game in which children use a breath-based game controller to control the movements of a fish on the screen. The aim is to help the fish to collect as many starfishes as possible, achieved through slow, continuous breathing patterns (Sonne and Jensen, 2016[23]; Sonne and Jensen, 2016[24]). ChillFish's impact was measured using Heart Rate Variability (HRV) and electrodermal activity and was found to be as calming as regular relaxation exercises for ADHD students.

Researchers in the United States are currently developing a smartwatch/smartphone application called CoolCraig (Doan et al., 2020[25]) to support co-regulation (where parents and teachers provide support such as helping redirect a child's attention, helping them initiate tasks, giving praise, etc.) of children with ADHD. Parents and teachers can use the CoolCraig app on their phone to set goals for the child, who can then select a goal from his/her smartphone. Once the goal is achieved, the parent or teacher receives a notification, and the child receives a set number of points determined by the adult, which can later be exchanged for a reward. CoolCraig also helps with emotional regulation by asking children to report on their current emotional state (using a colour-based system). The system can offer appropriate suggestions (e.g. "take a deep breath") and also allows children and adult to see a visualisation of their moods over time, which can encourage reflection.

Although the researchers feel that the approach behind CoolCraig may be of benefit, their preliminary design work with children with ADHD also uncovered a number of challenges (Cibrian et al., 2020[26]). Among them were children's hesitations about whether an app would be the best way of receiving support (rather than through a parent or teacher), their concern that the support may actually be a distracting factor, their fear of stigma and potential embarrassment (not wanting to receiving alerts or notifications in front of friends), and their desire for privacy (not wanting to be obliged to share their personal information with their parents and/or teachers). Such issues make it crucial to listen to children and understand their lived experiences when designing technology to support them, otherwise the technology runs the risk of not having the desired impact.

Some examples of learner-centred approaches to smart technologies

How can technology support the need for flexible, adaptable and learner-centred approaches for children with special needs (World Health Organization, 2011[10])? In this section, I describe three such approaches, focusing on autism, dysgraphia and visual impairment respectively.

The ECHOES environment

ECHOES (Porayska-Pomsta et al., 2018[27]) is a technology-enhanced learning environment designed to scaffold autistic children's exploration and learning of social communication skills through a series of playful learning activities, some of which involve a virtual character with which the child can interact. The target group for ECHOES is children with a developmental age of between 4 and 7 years (note that in the case of autistic children, their chronological age may be much higher as a result of having additional learning difficulties).

The ECHOES environment (Figure 6.1) was designed to run on a large multi-touch screen with sound output. Children can sit or stand in front of the screen, and physically interact with the system by dragging, tapping and shaking objects.

Interactions are set within a "magic garden" where the objects within the garden have unusual properties designed to spark curiosity and encourage exploration. For example, touching and dragging a flower head detaches it from its stem and transforms it into a bouncy ball. The magic garden is also home to Andy, an intelligent agent, with whom the child can play and interact. Andy acts both as a guide to the child, explaining activities and providing support, and also as a peer, taking turns with the child in activities such as a sorting task.

Autism is a lifelong, neurodevelopmental condition that affects the way in which a person interacts and communicates with others, and the manner in which they experience the world around them (National Autistic Society, 2016[28]). Social and communication difficulties are one of the hallmarks of autism and, given that autism is a spectrum condition, these difficulties will manifest in different ways at different points on the spectrum (e.g. they might range from difficulties in understanding the give and take of a typical conversation to not initiating communication at all). There can also be marked differences between individuals considered to be at the same point on the spectrum, and even within individuals at different times (for example, in situations of stress or anxiety, these difficulties are likely to be exacerbated).

Figure 6.1 The ECHOES environment

Source: ECHOES picutre and video database (reproduced with permission).

Social and communication difficulties can have a profound and sustained impact on an individual's social and emotional well-being, leading to issues with making and maintaining friendships (Kuo et al., 2013[29]), resulting in loneliness (Locke et al., 2010[30]), isolation (Chamberlain, Kasari and Rotheram-Fuller, 2007[31]) and a significantly increased likelihood of being bullied (Cappadocia, Weiss and Pepler, 2012[32]). Over time, these difficulties can have a profoundly negative effect on children's mental health (Whitehouse et al., 2009[33]), and their sense of self-worth and self-esteem (Bauminger, Shulman and Agam, 2004[34]). Furthermore, these difficulties persist throughout an individual's life, with many autistic adults reporting a marked sense of isolation, despite their desire to be more engaged with others (Müller, Schuler and Yates, 2008[35]).

ECHOES is based on the SCERTS model (Social Communication, Emotional Regulation, Transactional Support) (Prizant et al., 2006[36]). One of the overall goals of SCERTS is to support autistic children in developing competence and confidence in social activities. A particularly unique aspect of SCERTS is that it aims to identify children's strengths and build on those in developing further skills.

Another interesting aspect of SCERTS is what it terms "transactional support", which considers the role of the child's environment, including the people in that environment, in supporting the development of the skills in question. Children will be more successful in developing social competence when the environment can adapt to their particular needs so as best to support them (and this includes social partners: children's social competence increases when they are surrounded by partners who are understanding, supportive and enjoy interacting with them).

In terms of social communication, the SCERTS model focuses on two key foundational skills, namely, joint attention, and symbol use. Joint attention refers to the ability to share attention, emotion and intention with partners, engage in turn-taking and reciprocal social interactions. Symbol use includes using objects, pictures, words or signs to represent things and share intentions, and the ability to use objects in play.

In designing the virtual environment, the ECHOES team chose a participatory approach involving the widest range of stakeholders, including parents, carers, practitioners, teachers and, most importantly, autistic children (see, e.g. Frauenberger, Good and Keay-Bright, 2011[37]; Frauenberger et al., 2013[38]).

The aim of ECHOES was to create an environment in which children's strengths and abilities can be discovered and built upon, and it comprised both exploratory and task-focused activities. The exploratory activities were designed to give children a sense of autonomy and agency within the environment, while the task-focused activities provided opportunities for the agent to model initiation and response behaviours to the child. An example of the former would be taking turns with the virtual agent, Andy, to sort a fixed number of differently coloured balls into the appropriately coloured boxes. In the case of exploratory activities, there was no such fixed end point: examples include taking turns with the agent to shake clouds, which in turn causes it to rain, and for flowers to grow (see Figure 6.2). In these cases, the practitioner can use child interest and engagement in order to decide how long the activity should last, and when to move to a different activity.

The actions and behaviours of the intelligent virtual character, Andy, are underpinned by an autonomous planning-based agent. The planner works at both the reactive level and the deliberative level. The deliberative level is concerned with longer term plans related to a particular learning activity. For example, if the long-term goal within the activity is to encourage the child to pick up a basket, the deliberative level will focus on the set of actions that Andy needs to execute in order to encourage this to happen. By contrast, the reactive level refers to generating agent reactions to the child's immediate interface actions (e.g. what the child is touching, for how long, whether it is the right object, etc.).

Although the system can perceive a child's touch behaviours, and respond accordingly, it cannot detect other aspects of the child's interaction with the system or general behaviours and, as such, is unable to determine when it might be appropriate to repeat a particular activity (because the child is finding it soothing or enjoyable), move to another activity (because the child is bored or frustrated), or stop the session.

Instead, these decisions are made by the practitioner accompanying the child by using a specially designed practitioner interface. The practitioner interface is accessed on a separate screen from the main ECHOES touch screen so as not to distract the child. The interface allows the practitioner to control the choice, duration and sequencing of the learning activities. The practitioner can also use this screen to prompt the agent to repeat a behaviour, or to move on, where necessary. Thus, action within the ECHOES environment is driven by a combination of practitioner/researcher expertise and the system's intelligent planning.

Figure 6.2 Making it rain in ECHOES

Source: ECHOES picutre and video database (reproduced with permission)

As noted above, the team used a participatory design process, with the aim of involving end users in the design of the system as much as possible. One way of involving autistic children, where standard methods of gathering feedback such as focus groups and interviews would be inappropriate, was to quickly design prototypes and observe how children used them. In one such case, this led the team to completely reconceptualise the role of ECHOES within the broader context. It was initially envisaged that the primary social interactions would occur between the child and the social agent, and the study setup was arranged accordingly, with the researcher in retreat, and out of the line of sight of the child. However, from the initial prototype testing, the research team noticed that children often turned to the researcher or practitioner to share affect with them, or to initiate a conversation, typically about something happening in the ECHOES environment. This led to the reframing of ECHOES in a broader context, recognising the importance of the adult as another social partner, and arranging the setup to make it easier for the child to interact with the adult(s) in the room (see Figure 6.3).

Figure 6.3 A child interacting with ECHOES shares his joy with one of the researchers

Source: Reproduced with permission from Alcorn (2016[39])

ECHOES: Findings

The natural exchange of social interaction includes verbal and non-verbal initiations (e.g. asking a question, making an observation, pointing to something to draw another person's attention to it) and responses to those initiations (e.g. answering a question, nodding, following the person's pointing gesture to the object of interest). Although autistic children typically experience difficulties with both types of social interaction, initiations are typically more difficult than responses.

In evaluating the ECHOES environment, the team were interested in determining whether, when using ECHOES, autistic children showed an increase in initiations and responses, and whether there were differences in whether these social interactions involved either the agent or the human practitioner. Another research interest lay in understanding whether any increases transferred outside the ECHOES environment (in this case, in a free play session with a practitioner).

Interestingly, when children interacted with ECHOES, their initiations increased over time, and to both the human and the agent (although this increase was not statistically significant). Children's responses to the human partner also significantly increased, but their responses to the intelligent agent decreased. These increases in initiation and response behaviours did not, however, transfer to the free play session.

These are interesting findings, with particularly interesting implications for the use of technology in an educational context. The increase in initiations is positive, and it is also interesting to note that, over time, children responded more to the human partner, but less to the agent. This suggests that they were aware of the limits of the technology, realising that Andy could only detect their response if it was expressed through touch, unlike the human partner.

Above and beyond the changes in initiations and responses, one of the interesting findings emerging from ECHOES was the fact that the children seemed to genuinely enjoy interacting with it. As noted above, although the team initially imagined that children would interact primarily with the virtual agent, they noticed that children often wanted to share their experience with someone in the room (as in the case of the child in Figure 6.3 expressing their enjoyment with one of the researchers).

This likely explains two things with respect to the results: firstly, social communication may have increased within the ECHOES setting because the on-screen interactions provided the child with things they felt were "worth communicating about" (Alcorn, Pain and Good, 2014[40]). This would also explain the decrease in social communication behaviours once they were no longer using ECHOES. This suggests, firstly, thinking about how we can design technologies which, rather than taking a drill and practice approach to skills learning, aim to provide experiences that are engaging and motivating. Ideally, where possible, we should be thinking about how to make these experiences enjoyable, where learning is embedded in, for example, environments which allow for playful, self-directed exploration (Mora-Guiard et al., 2016[41]) or playful multisensory learning experiences (Gelsomini et al., 2019[42]).

Secondly, rather than considering technology as a discrete entity from which the child transfers skills into the "real world", it makes more sense to think about how children's experiences with learning technologies can effectively be embedded and integrated within the broader educational environment. As such, it is important to understand human-computer systems as forms of social systems, and consider them in the round.

Dysgraphia: diagnosis and beyond

As noted above, the field of special needs and disability is vast and wide-ranging, encompassing both physical disabilities as well as cognitive and neurodevelopmental disabilities. The diagnosis of any disability will require specialist input and assessment. However, certain disabilities, such as dyslexia or dysgraphia, for example, may only become apparent in an educational setting, therefore, the presumption of a potential disability may begin with teacher concern.

The process of obtaining a diagnosis for a special need is typically extremely time-consuming, lengthy and stressful for both children and their families. At the same time, early intervention, specifically targeted to an individual's needs, is often key to helping the child develop and progress. As such, any tools that might help teachers to recognise the early signs of a potential disability and support the child and their family in seeking a specialist diagnosis could have a huge impact on a child's education, and their future. This is not to suggest in any way that teachers themselves should be carrying out a diagnosis. However, in the course of the diagnostic process, the teacher's perspective is often sought as one input into the diagnostic process, usually in the form of a report, questionnaire, etc. As such, if tools were available to allow teachers to more quickly recognise differences in a child's development within an educational setting, it could potentially allow families to initiate the diagnosis process more quickly. Furthermore, providing specialists with more detailed information than would otherwise be the case might also allow diagnoses to be made in a more timely fashion.

A team at the EPFL in Lausanne has developed a system for detecting dysgraphia in children which has shown very promising results (Asselborn et al., 2018[43]; Asselborn, Chapatte and Dillenbourg, 2020[44]; Zolna et al., 2019[45]). Dysgraphia refers to difficulties with handwriting, that can manifest as distorted writing and difficulties forming the letters correctly, with letters sometimes written backwards and/or out of order. Associated problems with spelling may also be present.

Dysgraphia is typically diagnosed using one of a number of standardised tests. Although there are variations across the tests, all involve the child copying some text, which is then assessed by an expert in order to determine *legibility* (by measuring it against a set of criteria) and *efficiency* (by counting the amount of text which is produced within a given period of time) (Biotteau et al., 2019[46]).

The disadvantages of such tests are that they are subjective and expensive. Furthermore, their focus is primarily on the output, i.e. the written text, rather than on the processes used to arrive at that output. Asselborn and colleagues have developed a machine-learning algorithm which can detect dysgraphia, and which runs on a standard, commercially available tablet. In order to develop this tool, they first collected data from almost 300 children (both typically developing and those diagnosed with dysgraphia), who were asked to copy text onto a tablet with a sheet of paper overlaid on the surface (to mimic typical writing practices). They then used part of this data to train a machine-learning classifier to detect dysgraphia, and the remaining data to test the accuracy of the classifier. They were able to detect dysgraphia with a high degree of accuracy (approximately 96% in the study reported in Asselborn et al. (2018[43])).

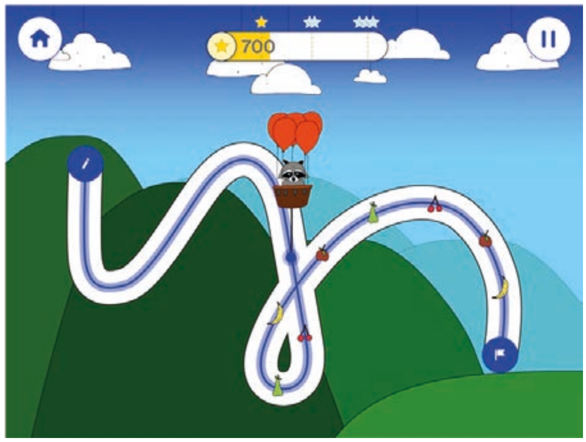
In the process, they were able to extract 53 features which describe various aspects of a child's handwriting, such as pen tilt, amount of pressure, speed and changes in speed. They were then able to determine which of these features were most discriminative, in other words, which ones distinguish handwriting from children with dysgraphia as compared to typically developing children.

One of the real advantages of the system is the fact that it does not require any specialist hardware: it works on a commercial graphics tablet, meaning that it is low cost, and potentially usable by non-specialists. Furthermore, compared to traditional methods of diagnosing dysgraphia, the system can analyse the process of writing, rather than just the product. The features identified above, such as pen pressure, pen tilt, etc., can provide a more finely grained analysis of the difficulties the child is experiencing, rather than simply identifying whether or not a child is dysgraphic. This then means that the child's support needs can be addressed in a much more specific and targeted way.

These findings have been taken forward in the form of Dynamico (www.dynamico.ch/), a tablet-based app which will soon be commercially available. Designed for an iPad and an Apple Pencil, Dynamico supports children with handwriting difficulties in multiple settings, and can be used at home, in the classroom and with therapists. The app includes tools which can analyse children's writing in 30 seconds, and then allow therapists to create a personalised remediation programme for the child, based on the results of the analysis. Teachers are also able to use the tool to create individualised sequences of learning activities for each child in the classroom. The tool can also be used at home, with teachers and/or therapists able to monitor the child's progress remotely. From the child's perspective, the activities take the form of games which have been designed to be playful and engaging. Figure 6.4 shows a screenshot from the app where children use the stylus to practice their letter tracing skills: doing so moves the raccoon from start to finish, and rewards for precise tracing can be collected along the way.

In addition to dysgraphia, technologies are also being designed to support students with dyslexia and dyscalculia: Box 6.3 provides some examples.

Figure 6.4 A screenshot from the Dynamico app



Source: Courtesy of Thibault Asselborn

Box 6.3 Smart technologies for dyslexia and dyscalculia

There are a number of technologies which can provide support for students with dyslexia, ranging from generic tools which can be used within an educational context, to those that are designed specifically for educational use. A good example of the former are web browser plugins which aim to facilitate reading by giving users the ability to change aspects of the web page such as background colour, font size, word spacing, etc. One such example, Help me read! (Berton et al., 2020[47]), also provides an "easy reading mode", which highlights and magnifies a single word of text at a time, allowing users to focus on each word, and move to the next at their own pace.

The use of smart technologies to diagnose dyslexia is also gaining traction. Interestingly, the diagnosis of dyslexia can be more or less difficult depending on the language. Languages differ in terms of the consistency of the relationship between grapheme (letter) and phoneme (sound). In languages with less consistent relationships, such as English, children with dyslexia may struggle more when learning to read, whereas in languages with a more consistent relationship, such as Spanish, dyslexia may not be picked up until much later, thus reducing the possibility for early intervention. Researchers in Spain and the United States have developed Dytective, an online game using machine-learning algorithms which was able to correctly diagnose dyslexia in Spanish speakers with over 80% accuracy (Rello et al., 2016[48]), thus increasing the potential that children can get the support they need as early as possible.

In addition to detection and diagnosis, smart technologies can also support the development of skills in children with dyslexia. Such technologies include PhonoBlocks (Fan et al., 2017[49]), a system incorporating 3D tangible letters which children can physically manipulate to spell words. The letters have the ability to change colour depending on the sound that they make in a given word. For example, the colour of the letter A is yellow in the word "fad", but changes to red in the word "fade", thus providing support for the child to better understand the relationship between the letter and the sound that it makes, as well as the ways in which these relationships can differ.

Contrary to dyslexia, there are very few new technologies designed to support students with dyscalculia (a difficulty with the comprehension of numbers and mathematical calculations). However, a research team in Germany has recently developed a system called Hands-On Math, which aims to support children in learning to use their fingers to represent numbers and perform simple calculations (Erfurt et al., 2019[50]). Typically, this is taught in one-on-one sessions with a trained practitioner, meaning that children have limited access to such support. The system developed calls out numbers or simple mathematical calculations which the child needs to represent using their fingers. By wearing gloves with markers attached to each finger, the system can use a camera to track the child's calculations and determine if they are correct. Hands-On Math is currently in development and has received very positive reviews.

Visual impairment and interactive tactile graphics

Background

Similar to other disabilities, pupils who are blind or visually impaired (BVI) face gaps in educational attainment, decreased access to further and higher education, high unemployment rates, and lack of access to certain careers.

In many OECD countries, the overwhelming majority of BVI students attend or are encouraged to attend a mainstream school. This presents a number of challenges to ensuring their participation in everyday classroom activities. Metatla (2017[51]) provides an excellent overview of some of these challenges, such as facilitating collaborative work when BVI children and sighted children are using different materials which they cannot share between them, or the fact that classrooms walls are typically covered in visual materials, such as posters, charts, etc., which are deemed important for learning, but which are inaccessible to BVI students. In this section, I will focus exclusively on access to educational material, specifically graphics.

Most educational materials have a significant visual component, whether they take the form of text or graphics (or typically, some combination of the two). There are various options for accessing written text. Although not without usability issues, screen readers take written digital material and use text-to-speech engines to translate the written text into audio form. Screen readers are widely available for both PC and mobile devices (e.g. JAWS or NVDA for Windows, or VoiceOver for Mac and iOS). In situations where a BVI person prefers to read the text rather than listen to it, refreshable Braille displays offer a solution. These displays are a separate piece of hardware which take input from the screen reader and translate it into Braille output (using moveable pins) which can then be read by the user (note that these displays are also input devices, and allow BVI users to input text into a computer) (see Box 6.4). Many of these devices can, however, be quite costly.

Box 6.4 Technology supporting blind and visually impaired students

A number of technologies support the educational needs of blind and visually impaired (BVI) students.

The first type is hardware which could help BVI students in both note-taking and reading. Such solutions are used in both low- and high-income countries. Portable refreshable Braille display devices featuring refreshable six- or eight-dot Braille cells allow BVI students to have access to written learning materials and books in Braille. This technology works for many languages and reads texts in Braille (mere reader) as well as translates other text formats from a variety of applications (Braille translator) – and allows for printing on Braille printers. They can work with other devices using Bluetooth or USB and thus allow teachers to interact with their BVI students using compatible apps on their smartphone, computer or tablet – providing teachers with real-time text translation of the Braille being read or written by the students on the device (or vice versa). Braille Me, BrailleRing, and Orbit Reader are examples of such devices. For example, eKatibu, EdTech Hub and the Leonard Cheshire foundation used the Orbit Reader 20 in the Kenyan Nyanza region during the COVID-19 school closures in conjunction with a teacher-training programme to ensure that blind and visually impaired students continued to learn (see https://edtechhub.org/2021/01/08/using-innovative-methods-to-train-teachers-of-blind-children-what-we-learned/). Another type of device based on AI technology and designed to support BVI students (and people) is the Finger Reader, a wearable device worn as a ring around the index finger that reads aloud whatever one points it to thanks to its camera, text recognition and text-to-speech algorithms (Shilkrot et al., 2015[52]).

Other forms of technology are software-based and use regular computers to boost the independence of visually impaired students. For example, SuperNova allows users to use magnification, announce punctuation, replace difficult colours, increase verbosity, hear webpages, and turn on Braille. This system allows the user to hear characters and words aloud as they type, and reads aloud web pages, applications, documents and emails to the user. Programmes that support reading in this manner are a key way for visually impaired students to be more self-sufficient in a regular classroom setting and are used throughout the world for such purposes. Beyond SuperNova, examples of such programmes are Jaws, Microsoft Narrator, NVDA, Orca or Window Eyes. Besides aiding blind or visually impaired students, this technology can also support students with dyslexia or related forms of learning difficulties.

However, both text-to-speech capabilities and refreshable Braille displays are designed to work with text-based materials. There is no analogous low-cost solution for graphical representations, meaning that access to graphical content for BVI individuals remains a challenge. On web pages, alt-text (alternative text) can be used to provide a textual description of an image, which is then read out by the screen reader. However, these alt-text descriptions are sometimes missing (as they are reliant on the content creator of the web page to include them in the page's source code) and often vary in quality (as the content creator has free rein in deciding how best to describe the image).

Recent research, although not education specific, but relevant and certainly timely, suggests a further reason for such difficulties. In a study on access to public health information about the COVID-19 pandemic, Holloway et al. (2020[53]) found that over 70% of the websites they surveyed used graphical representations to convey information about the pandemic (for example, visualisations of real-time statistical data). However, less than a quarter included alt-text information for these graphics, not because the web page creator did not include them, but because graphics which are interactive, or automatically updated, do not support alt-text. This means that important information which is present in these visualisations is inaccessible to BVI users.

This gap in access to graphical information seems to be widening unfortunately, as graphics are increasingly being favoured over text as a way of exchanging information (Gorlewicz et al., 2018[54]), and providing information in a digital format is both easier and more cost effective as compared to paper-based alternatives. Paradoxically, novel interactive technologies for learning may further decrease potential access to information for BVI children, given their reliance on visual content and interactions such as drag and drop (Metatla et al., 2018[55]). Particular curriculum subjects also present challenges. In some subjects, graphical information might be purely illustrative, or supplement textual information, however, in other subjects, it is difficult to present some types of information in any other way. This is particularly the case for STEM subjects, with their heavy reliance on representations such as charts and graphs.

One approach to graphical access for BVI relies on touch, similarly to Braille. Tactical graphics can be created by using an embosser, which raises the elements of the graphic so that they can be perceived through touch, or swell paper, which uses specialised paper and heating machines to cause the printed parts of an image to swell, again, becoming perceivable through touch. However, the output of such systems is static, meaning that a new graphic would need to be created if there were any changes or updates. Given the importance of dynamic graphics in education (e.g. understanding how changes to *m* affect the slope of a line in the equation $y = mx+b$), it's clear that these types of graphics represent only a partial solution.

In order to create dynamic graphical displays, similar technologies to refreshable Braille displays have been developed using pin arrays which allow BVI users to explore the graphic through touch. However, these solutions rely on bespoke hardware and are extremely expensive. Even newly released technologies, e.g. Graphiti (https://www.orbitresearch.com/product/graphiti/), which uses an array of pins which can be set to varying heights to convey topographical information, still take the form of specialised hardware. Although the aim is to eventually be able to reduce the price of purchasing Graphiti to USD 5 000 as a result of bulk orders, this still represents a significant expense for a school, particularly for a piece of hardware that can only be used for one purpose.

In summary, providing access to dynamic graphics which does not require specialised hardware is a major and yet incredibly important challenge. It could be transformative not just in the educational sector, but in almost every aspect of the daily life of a BVI individual.

Promising approaches

In trying to address the issue of graphical access in a way which benefits the greatest number of users, Gorlewicz et al. (2018[54]) make the case for using touchscreen-based smart devices such as phones or tablets as the hardware platform.

There are a number of advantages to such an approach. Firstly, the hardware platform is low in cost, readily available, and is already widely used by a significant proportion of the intended user group.

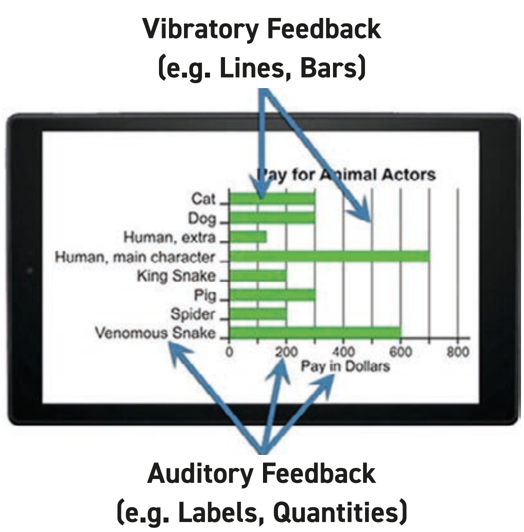
Furthermore, these devices already have the inbuilt capacity to provide information through multiple modalities, including visual, auditory and tactile modalities. In the case of BVI users, the fact that almost all touchscreen devices include a sound card and speaker, and text-to-speech capabilities, means that they can provide auditory access to information. Furthermore, the fact that many touchscreen displays also have vibratory capabilities gives them the capacity to provide vibrotactile feedback. Although this type of feedback is not typically used as a primary form of interaction with a device, there is no reason it could not be and, in the case of BVI users, it offers an additional modality through which to provide information.

As such, this approach offers advantages over current solutions such as tactile graphics displays, which only present information through a single modality: touch. And the fact that these features are already present in most touchscreen devices potentially eliminates the need to create expensive, specialised hardware which can only be used for a single purpose.

Gorlewicz et al. (2018[54]) envisage a display such as the one shown in Figure 6.5. In this case, textual information in the bar graph can be converted to auditory information, while the graphical, spatial elements of the graph (in this case, the bars) can be conveyed through vibratory feedback. This represents an obvious strength over systems which can only provide both types of information inherent in a graphic through a single modality.

In this example, although the technology exists to provide a potential solution to a key issue, further research is needed before such a system can become a reality. In particular, Gorlewicz et al. (2018[54]) note the need for research into understanding how people encode, interpret and represent graphical information which is presented via non-visual channels. Although they note that previous research has considered these issues in relation to tangible (i.e. raised) graphics, the results may not be applicable as the touchscreen does not offer a similar tangible experience and is therefore likely to trigger different sensory receptors. At the same time, a recent study comparing the use of touchscreen-based graphics to embossed graphics showed that there was no significant difference in performance (Gorlewicz et al., 2020[56]), suggesting that this is a promising area to continue to explore.

Figure 6.5 Using a combination of auditory and vibratory feedback to convey graphical information



Source: Courtesy of Thibault Asselborn

Looking to the future

The research described in this paper suggests three promising avenues in terms of developing and deploying smart technologies that could have the most potential impact in the lives of learners with special educational needs. The first is a more holistic development of smart systems, considering the need, the user, and the context of use. The second is creating smart systems using technologies which have the highest chance of being available to the intended users ("smart systems for all"). Finally, systems which incorporate a blend of human and artificial intelligence have great promise, and it is important to consider how to go about achieving the optimum blend. I consider these in turn below.

Holistic smart systems

If we want to see a real impact of smart technologies to support children with special needs in the short to medium term, we should prioritise the development of "holistic smart systems" by which I mean technologies that 1) address a real need, and are designed with an understanding of both 2) the end users and 3) the context of use. Neglecting any of these aspects is likely to lead to technologies that are either not adopted in the first place, or are quickly abandoned.

Address a real need

As we have seen above, there is a pressing need to support learners with special needs. There are vast pockets of need, and some types of disabilities receive less of a research and development focus than others (e.g. dyscalculia). Although it is clear that technology plays, and will continue to play, an increasing role in providing support for learning with special needs, it is important to take the time to understand what is really needed, rather than rely on the carers' or vendors' sense of what is needed. Equally, it is important to consider where the greatest potential impact lies: the case of access to graphical content for BVI users is one such example where providing a readily available and low-cost solution could have a hugely substantial impact, both within and beyond the classroom.

Design for users

The World Health Organization (2011[10]) highlights the importance of ensuring that the voices of disabled children are heard, while acknowledging that this is frequently not the case unfortunately, and the same holds true for the design of new technologies.

Involving children with disabilities in design activities leading to smart technologies supporting their education may present additional challenges (e.g. how to ensure that the needs and desires of non-verbal children can be expressed during the design process), however, the fact that their needs and desires are likely to be very different from the adults actually designing the technology makes it all the more important to do so. Fortunately, there is an extensive body of work which has considered participatory design with children with various types of disabilities: Benton and Johnson (2015[57]) provide a good overview.

In addition to the voices of the children, it is important to involve teachers in the design of any educational technology. For a start, they have unique insights and expertise in supporting children with special educational needs, and will be able to provide their views on what is likely to work, and what isn't (see e.g. Alcorn et al., 2019[58]).

Design for context

In addition to designing for users, it is important to design with an understanding of the context of use. Any type of learning technology exists in a wider pedagogical ecosystem (typically comprising the classroom and the school), and introducing advanced technologies into such a setting requires an awareness of the constraints within which the technology must operate. These include a number of practical considerations such as school budgets and funding priorities, fit with the curriculum, the robustness of the technology, its cost, and potential integration with existing technologies. Ease of use (for both teachers and children) is also a major concern, as are issues of maintenance and support.

Ease of use is something that can be addressed through the design process (ideally, by keeping end users involved throughout), however, the initial choice of hardware is a decision that will have a critical influence on whether the technology will actually be used in the first place.

When considering the potential uptake of technology for learning, there are two extremes (and a continuum in between). At one extreme are technologies being developed which are unlikely to make it into the classroom, either because the technology is too expensive, or the hardware or software is too difficult for non-specialists to set up, use, and/or maintain (such as ECHOES). At the other are systems which use existing, low-cost technologies, and which are designed to be easily usable by non-specialists (such as the Dynamico app and, once they are developed, systems which make use of touchscreen displays for vibrotactile graphics).

Although not usable in classrooms, ECHOES provided us with a much deeper understanding of how to build environments with the potential to provide playful, motivating interactions. These insights were then taken forward in a system which aimed to be both technologically simpler while also allowing non-specialists to author their own content (Porayska-Pomsta et al., 2013[59]). The ECHOES project also challenged accepted conceptions about the nature of autism, in particular, the purported "need for sameness", leading to further research and exploration which made an important theoretical contribution to the field (Alcorn, 2016[39]). We need to research and develop both types of system but, at the same time, researchers need to be clear about where on this continuum their systems are situated.

Smart systems for all

Part of the issue in designing for the context of use involves understanding issues around cost and availability. Many assistive technologies are prohibitively expensive for public schools and require specialist hardware that, in many cases, is limited to a single purpose. This means that many children are not able to access the technologies which could support their learning. If we really want to see the potential positive impact of smart technologies for special educational needs in the short term, then we need to think about how best to bring the latest advances in AI into schools in ways that are affordable and readily available.

One very promising way of doing so is likely to be through the use of touchscreen devices, i.e. smartphones and tablets, for two reasons. Firstly, these devices are commercially available, reasonably low cost, and are multi-functional. This gives them obvious advantages over specialised, bespoke hardware systems which are developed for a single use and are typically very expensive (such as many of the solutions for BVI students). Secondly, above and beyond their processing power, modern touchscreen devices incorporate numerous inbuilt sensors which offer multiple possibilities for multimodal input and output. This opens up opportunities for creating novel and innovative learning experiences for students with special needs, while still ensuring that they remain relatively low in cost and are widely available.

Two of the case studies presented in this report (for dysgraphia and access to graphical information respectively) offer excellent examples of how this can be achieved. Dynamico builds on robust scientific research and makes use of complex AI algorithms which can run on commercially available tablets. And it uses the tablet's inbuilt sensors (e.g. pressure) to be able to analyse a child's handwriting and detect dysgraphia. Although at a much earlier stage of development, the researchers investigating access to graphical information were taking much the same approach. They are currently carrying out foundational research in order to better understand how BVI students understand and use graphics presented on a multisensory tablet (Hahn, Mueller and Gorlewicz, 2019[60]), and stress that much more research will be needed before a commercial product can be envisaged. However, their ultimate aim is to use the existing capabilities of touchscreen devices in innovative ways in order to provide multimodal output to BVI students (in this case, a combination of auditory and vibrotactile feedback).

Investigating how the sophisticated capabilities of modern touchscreen devices could be leveraged to support students with a wider range of special needs, and via new and innovative forms of input and output, seems like a very promising way forward.

Blending human and artificial intelligence

Baker (2016[61]) points out that, initially, the grand vision of intelligent tutoring systems was to develop intelligent tutors that were as skilled as human tutors. They would be able to use the same strategies as expert human tutors, incorporating knowledge about the domain, and how to teach it. And while there are now a number of intelligent tutoring systems that are being used at scale, with hundreds of thousands of students benefiting from them (for example, Cognitive Tutor, ALEKS, Mindspark, and Alef), these systems are rather more simplified versions of this initial vision.

As Baker notes, one issue with automated interventions is that they are brittle, in the sense that if an intervention is not working, it is difficult for the system itself to recognise this and react accordingly. And these breakdowns in interaction do not go unnoticed by learners. In a study of children's perceptions of a humanoid empathic robot tutor, Serholt (2019[62]) found that children who had interacted with the robot previously were more critical of the concept of emotion recognition in robots than those who had not. One child graciously allowed that perhaps social robots "may become more useful in the future, when they [researchers/developers] have had a few more years to develop them and make them more human" (Serholt, 2019, p. 95[62]). This was similar to the ECHOES experience: although the children were not able to verbalise this, the fact that they stopped responding to Andy quite as much suggests that they were aware of the limits of his interaction skills. Similarly, a number of their initiations to Andy were concerned with offering him help when, due to planner issues, he did not behave as expected (for example, incorrect moves in the sorting task, walking off screen).

On a related note, some aspects of teaching are simply easier for humans to perform, at least currently. This point is particularly relevant to students with special needs. Supporting their learning is a subtle and individual art, where skilled teachers and teaching assistants will have spent many hours and considerable effort becoming attuned to each child's particular skills, needs, and ways of communicating.

This point can be illustrated using the ECHOES environment. The original conception of ECHOES was much more complex, aiming to be able to recognise the children' emotional state using facial emotion recognition, and to track their gaze and respond appropriately. In reality, there were problems with both. Children moved around, quite energetically in some cases, meaning that the tracking system could not function correctly. However, it seemed counter intuitive, in a system designed to encourage playful exploration, to require the child to maintain a fixed position. Furthermore, research suggests that in the same way that autistic individuals can struggle to understand the facial expressions of others, neurotypical individuals often struggle to understand autistic facial expressions (Brewer et al., 2016[63]). Therefore, it seemed unwise to rely on an automated system, which would necessarily be built from a neurotypical perspective, to detect the child's emotional state. This was also important more broadly in the sense that if the child became particularly distressed, they would be likely to move away from the screen, where they could no longer be detected, even though appropriate intervention would be required immediately. Therefore, it was important to ensure that anything included in the environment did not have the potential to lead to distress as a result of a breakdown in the interaction.

The research team therefore decided to use the intelligence of the ECHOES system to provide the playful, motivating interaction sequences with an intelligent agent, which children seemed to appreciate, while the system relied on human support to interpret the meaning of children's behaviours and ensure their overall well-being during the sessions.

This situation corresponds to a new vision of intelligent educational systems, where the best of human and artificial intelligence are combined in a way that most effectively supports learners.

Artificial intelligence, in its current form, excels at finding patterns in data, something humans are rather less good at. However, access to this data has the potential to improve teaching and learning. This is a vision which Baker (2016[61]) offers, where intelligent educational systems are used to provide data to humans, who can use this data to guide their pedagogical decisions. The use of AI in this way was evidenced in Dynamico, which provides practitioners with access to previously unavailable data, thus leading, in this case, to improvements in the way that dysgraphia can be diagnosed, and to the ways in which learners with dysgraphia can best be supported.

On the other hand, human intelligence is better suited to situations where the data might not be so clear-cut. For example, teachers working with autistic children have very specialised knowledge of autism, and will have devoted many hours to developing a nuanced and in-depth understanding of each individual child over time. They are able to interpret cues from the child's behaviours in ways which are likely to be impossible for people who do not know the child. This might include understanding particular triggers that may lead to a child becoming distressed. Teachers will probably also be sensitive to the specific gestures, behaviours or utterances that indicate that the child is becoming distressed. And they are also likely to know how to intervene in a way which de-escalates the situation, and provides support and reassurance to the child. However, although teachers might be best placed to provide this type of support, that is not to say that artificial intelligence could not provide insight into the child's unique way of interacting, thus allowing the human practitioner to intervene in the most effective way.

In line with this view, a positive view of smart technologies for learning could be one in which smart technologies are used to enhance the complex orchestration involved in working with a child with special educational needs. Although teachers would be responsible for this orchestration, smart technologies could support them in at least three ways:

1. by offering support for recognising the needs in the first place (as in the dysgraphia example discussed earlier);
2. by providing teachers and teaching assistants with additional knowledge and insight about the child, which may help them support the child in a more appropriate way (as discussed above);
3. by providing adaptable support at multiple levels (described in more detail below).

In-depth adaptivity and personalisation

One of the promises of the use of artificial intelligence in the field of education is that of adaptivity, both being able to adapt to the student's current level of knowledge and/or skill, but also, in some cases, to their current level of motivation and/or affective disposition. In most cases, this adaptivity works at the level of the individual learner.

In addition to adaptivity at the individual level, supporting learners with special needs offers a unique space in which to consider additional types of customisation and personalisation, which could be achieved through a combination of human and artificial intelligence, as explained below.

When describing ECHOES earlier in the report, it was explained that the system planner operated at the level of the individual learning activities within the system, while the practitioner worked at a higher level to structure the overall session, determining the choice of particular activities, including their duration, and sequencing. Although this worked well, it was sometimes necessary to customise the environment prior to being able to use it in a given school. Andy, the AI agent, used speech and gesture to communicate with children. However, different schools used different key phrases and gestures, for example, to signal to children the end of an activity and prepare them, emotionally and cognitively, to move on to the next. Therefore, the agent's speech and gestures needed to be changed prior to use in each school. The research team was able to do this as system developers, however, it would be better if schools had these options themselves.

In thinking about personalisation and adaptation in the context of support for disabled learners, it is likely that it could happen not only at an individual level, which is typically the case for intelligent systems, but also at the level of the disability and the particular school context. I describe these three levels below (using examples from ECHOES to illustrate):

1. *Disability level customisation*: this level involves customising and adapting the interaction based on what we know about working with particular types of disabilities. In the case of ECHOES, designed for autistic children, this meant thinking about the pacing of the interaction (slowing it down to give children time to reflect), the language used for instructions (i.e. using direct, simple language, not rephrasing instructions), and deciding how much time to allow for a child to process an instruction before re-issuing it. However, these decisions resulted in fixed parameters, built into the environment, and it would be better if these could be customised by the practitioner.
2. *School level customisation*: as explained above, customisation to the school's particular use of language, symbols and signs.
3. *Child level customisation*: for example, do not use particular sounds, turn off agent facial expressions, remove certain keywords or phrases that may be triggering.

Being able to provide more finely grained customisation and adaptation that can be activated through a combination of human and artificial intelligence, and that encapsulates extant knowledge of a particular disability, the broader school context and the particular child increases the chances that the environment will be truly useful, and in a range of situations.

To conclude, smart technologies offer real promise in offering targeted and more sophisticated support for learners with special needs. By embedding the latest advances in artificial intelligence, as well as the most up-to-date understanding of special needs and disability, in readily available, low-cost technologies, there is a real opportunity to make a difference to learners across the globe.

Acknowledgements

The ECHOES system, presented here, results from a large, inter-institutional and interdisciplinary project. In addition to the author, the team included Kaska Porayska-Pomsta (project leader), Alyssa Alcorn, Katerina Avramides, Sandra Beale, Sara Bernardini, Mary Ellen Foster, Christopher Frauenberger, Karen Guldberg, Wendy Keay-Bright, Lila Kossyvaki, Oliver Lemon, Marilena Mademtzi,

Rachel Menzies, Helen Pain, Gnanathusharan Rajendran, Tim Smith, and Annalu Waller. The ECHOES project was funded by the ESRC/EPSRC, TRLP TEL programme grant number: RES-139-25-0395.

References

Alcorn, A. (2016), *Embedding novel and surprising elements in touch-screen games for children with autism: creating experiences "worth communicating about"*, PhD thesis, The University of Edinburgh. [39]

Alcorn, A., E. Ainger, V. Carisi, S. Martinioti, S. Petrovic, B.R. Schadenberg, T. Tavassoli and E. Pellicano (2019), "Educators' Views on Using Humanoid Robots With Autistic Learners in Special Education Settings in England", *Frontiers in Robotics and AI*, Vol. 6, http://dx.doi.org/10.3389/frobt.2019.00107. [58]

Alcorn, A., H. Pain and J. Good (2014), "Motivating children's initiations with novelty and surprise", *Proceedings of the 2014 conference on Interaction design and children*, http://dx.doi.org/10.1145/2593968.2610458. [40]

Alkhatlan, A. and J. Kalita (2018), *Intelligent Tutoring Systems: A Comprehensive Historical Survey with Recent Developments.*, rXiv preprint arXiv:1812.09628., https://arxiv.org/abs/1812.09628 (accessed on 26 February 2021). [6]

Asselborn, T., M. Chapatte and P. Dillenbourg (2020), "Extending the Spectrum of Dysgraphia: A Data Driven Strategy to Estimate Handwriting Quality", *Scientific Reports*, Vol. 10/1, http://dx.doi.org/10.1038/s41598-020-60011-8. [44]

Asselborn, T., T. Gargot, Ł. Kidziński, W. Johal, D. Cohen, C. Jolly and P. Dillenbourg (2018), "Automated human-level diagnosis of dysgraphia using a consumer tablet", *npj Digital Medicine*, Vol. 1/1, http://dx.doi.org/10.1038/s41746-018-0049-x. [43]

Baker, R. (2016), "Stupid Tutoring Systems, *Intelligent Humans", International Journal of Artificial Intelligence in Education*, Vol. 26/2, pp. 600-614, http://dx.doi.org/10.1007/s40593-016-0105-0. [61]

Bauminger, N., C. Shulman and G. Agam (2004), "The Link Between Perceptions of Self and of Social Relationships in High-Functioning Children with Autism", *Journal of Developmental and Physical Disabilities*, Vol. 16/2, pp. 193-214, http://dx.doi.org/10.1023/b:jodd.0000026616.24896.c8. [34]

Benton, L. and H. Johnson (2015), "Widening participation in technology design: A review of the involvement of children with special educational needs and disabilities", *International Journal of Child-Computer Interaction*, Vol. 3-4, pp. 23-40, http://dx.doi.org/10.1016/j.ijcci.2015.07.001. [57]

Berton, R., A. Kolasinska, O. Gaggi, C. Palazzi and G. Quadrio (2020), "A Chrome extension to help people with dyslexia", *Proceedings of the International Conference on Advanced Visual Interfaces*, http://dx.doi.org/10.1145/3399715.3399843. [47]

Biotteau, M., J. Danna, E. Baudou, F. Puyjarinet, J. Velay, J. Albaret and Y. Chaix (2019), "<p>Developmental coordination disorder and dysgraphia: signs and symptoms, diagnosis, and rehabilitation</p>", *Neuropsychiatric Disease and Treatment*, Vol. Volume 15, pp. 1873-1885, http://dx.doi.org/10.2147/ndt.s120514. [46]

Brewer, R., F. Biotti, C. Catmur, C. Press, F. Happé, R. Cook and G. Bird (2016), "Can Neurotypical Individuals Read Autistic Facial Expressions? Atypical Production of Emotional Facial Expressions in Autism Spectrum Disorders", *Autism Research*, Vol. 9/2, pp. 262-271, http://dx.doi.org/10.1002/aur.1508. [63]

Cappadocia, M., J. Weiss and D. Pepler (2012), "Bullying Experiences Among Children and Youth with Autism Spectrum Disorders", *Journal of Autism and Developmental Disorders*, Vol. 42/2, pp. 266-277, http://dx.doi.org/10.1007/s10803-011-1241-x. [32]

Chamberlain, B., C. Kasari and E. Rotheram-Fuller (2007), "Involvement or Isolation? The Social Networks of Children with Autism in Regular Classrooms", *Journal of Autism and Developmental Disorders*, Vol. 37/2, pp. 230-242, *http://dx.doi.org/10.1007/s10803-006-0164-4*. [31]

Chelkowski, L., Z. Yan and K. Asaro-Saddler (2019), "The use of mobile devices with students with disabilities: a literature review", *Preventing School Failure: Alternative Education for Children and Youth*, Vol. 63/3, pp. 277-295, http://dx.doi.org/10.1080/1045988x.2019.1591336. [1]

Cheng, S. and C. Lai (2020), "Facilitating learning for students with special needs: a review of technology-supported special education studies", *Journal of Computers in Education*, Vol. 7/2, pp. 131-153, http://dx.doi.org/10.1007/s40692-019-00150-8. [4]

Chen, L., P. Chen and Z. Lin (2020), "Artificial Intelligence in Education: A Review", *IEEE Access*, Vol. 8, pp. 75264-75278, http://dx.doi.org/10.1109/access.2020.2988510. [7]

Cibrian, F., K. Lakes, A. Tavakoulnia, K. Guzman, S. Schuck and G. Hayes (2020), "Supporting Self-Regulation of Children with ADHD Using Wearables", *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, http://dx.doi.org/10.1145/3313831.3376837. [26]

Doan, M., F. Cibrian, A. Jang, N. Khare, S. Chang, A. Li, S. Schuck, K. Lakes and G. Hayes (2020), "CoolCraig: A Smart Watch/Phone Application Supporting Co-Regulation of Children with ADHD", *Extended Abstracts of the 2020 CHI Conference on Human Factors in Computing Systems*, http://dx.doi.org/10.1145/3334480.3382991. [25]

Erdem, R. (2017), "Students with special educational needs and assistive technologies: A literature review.", *Turkish Online Journal of Educational Technology-TOJET*, Vol. 16/1, pp. 128–146. [5]

Erfurt, G., E. Hornecker, J. Ehlers and S. Plasckies (2019), "Hands-On Math", *Extended Abstracts of the 2019 CHI Conference on Human Factors in Computing Systems*, http://dx.doi.org/10.1145/3290607.3313012. [50]

Fan, M., A. Antle, M. Hoskyn, C. Neustaedter and E. Cramer (2017), "Why Tangibility Matters", *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*, http://dx.doi.org/10.1145/3025453.3026048. [49]

Frauenberger, C., J. Good, A. Alcorn and H. Pain (2013), "Conversing through and about technologies: Design critique as an opportunity to engage children with autism and broaden research(er) perspectives", *International Journal of Child-Computer Interaction*, Vol. 1/2, pp. 38-49, http://dx.doi.org/10.1016/j.ijcci.2013.02.001. [38]

Frauenberger, C., J. Good and W. Keay-Bright (2011), "Designing technology for children with special needs: bridging perspectives through participatory design", *CoDesign*, Vol. 7/1, pp. 1-28, http://dx.doi.org/10.1080/15710882.2011.587013. [37]

Gelsomini, M., G. Cosentino, M. Spitale, M. Gianotti, D. Fisicaro, G. Leonardi, F. Riccardi, A. Piselli, E. Beccaluva, B. Bonadies, L. Di Terlizzi, M. Zinzone, S. Alberti, C. Rebourg, M. Carulli, F. Garzotto, V. Arquilla, M. Bisson, B. Del Curto and M. Bordegoni (2019), "Magika, a Multisensory Environment for Play, Education and Inclusion", *Extended Abstracts of the 2019 CHI Conference on Human Factors in Computing Systems*, http://dx.doi.org/10.1145/3290607.3312753. [42]

Gorlewicz, J., J. Tennison, H. Palani and N. Giudice (2018), "The Graphical Access Challenge for People with Visual Impairments: Positions and Pathways Forward", in *Interactive Multimedia - Multimedia Production and Digital Storytelling, IntechOpen*, http://dx.doi.org/10.5772/intechopen.82289. [54]

Gorlewicz, J., J. Tennison, P. Merlin Uesbeck, M. Richard, H.P. Palani, A. Stefik, D. Smith and N. Guidice (2020), "Design Guidelines and Recommendations for Multimodal, Touchscreen-based Graphics", *ACM Transactions on Accessible Computing*, Vol. 13/3, pp. 1-30, http://dx.doi.org/10.1145/3403933. [56]

Hahn, M., C. Mueller and J. Gorlewicz (2019), "The Comprehension of STEM Graphics via a Multisensory Tablet Electronic Device by Students with Visual Impairments", *Journal of Visual Impairment & Blindness*, Vol. 113/5, pp. 404-418, http://dx.doi.org/10.1177/0145482x19876463. [60]

Holloway, L., M. Butler, S. Reinders and K. Marriott (2020), "Non-visual access to graphical information on COVID-19", *The 22nd International ACM SIGACCESS Conference on Computers and Accessibility*, http://dx.doi.org/10.1145/3373625.3418015. [53]

Houtrow, A., K. Larson, L. Olson, P. Newacheck et N. Halfon (2014), "Changing Trends of Childhood Disability, 2001-2011", *PEDIATRICS*, Vol. 134/3, pp. 530-538, http://dx.doi.org/10.1542/peds.2014-0594. [12]

Kazimzade, G., Y. Patzer and N. Pinkwart (2019), "Artificial Intelligence in Education Meets Inclusive Educational Technology—The Technical State-of-the-Art and Possible Directions", in A*rtificial Intelligence and Inclusive Education, Perspectives on Rethinking and Reforming Education*, Springer Singapore, Singapore, http://dx.doi.org/10.1007/978-981-13-8161-4_4. [9]

Kulik, J. and J. Fletcher (2016), "Effectiveness of Intelligent Tutoring Systems", *Review of Educational Research*, Vol. 86/1, pp. 42-78, http://dx.doi.org/10.3102/0034654315581420. [8]

Kuo, M., G. Orsmond, E. Cohn and W. Coster (2013), "Friendship characteristics and activity patterns of adolescents with an autism spectrum disorder.", *Autism*, Vol. 17/4, pp. 481–500, https://doi.org/10.1177/1362361311416380. [29]

Locke, J., E. Ishijima, C. Kasari and N. London (2010), "Loneliness, friendship quality and the social networks of adolescents with high-functioning autism in an inclusive school setting", *Journal of Research in Special Educational Needs*, Vol. 10/2, pp. 74-81, http://dx.doi.org/10.1111/j.1471-3802.2010.01148.x. [30]

Marks, D. (1997), "Models of disability", *Disability and Rehabilitation*, Vol. 19/3, pp. 85-91, http://dx.doi.org/10.3109/09638289709166831. [18]

McLeskey, J., M. Barringer, B. Billingsley, M. Brownell, D. Jackson, M. Kennedy, T. Lewis, L. Maheady, J. Rodriguez, M. Scheeler, J. Winn and D. Ziegler (2017), *High-leverage practices in special education.*, Arlington, VA: Council for Exceptional Children & CEEDAR Center. [3]

Metatla, O. (2017), "Uncovering Challenges and Opportunities of Including Children with Visual Impairments in Mainstream Schools", http://dx.doi.org/10.14236/ewic/hci2017.102. [51]

Metatla, O., M. Serrano, C. Jouffrais, A. Thieme, S. Kane, S. Branham, E. Brulé and C. Bennett (2018), "Inclusive Education Technologies", *Extended Abstracts of the 2018 CHI Conference on Human Factors in Computing Systems*, http://dx.doi.org/10.1145/3170427.3170633. [55]

Mora-Guiard, J., C. Crowell, N. Pares and P. Heaton (2016), "Lands of Fog", *Proceedings of the The 15th International Conference on Interaction Design and Children*, http://dx.doi.org/10.1145/2930674.2930695. [41]

Müller, E., A. Schuler and G. Yates (2008), "Social challenges and supports from the perspective of individuals with Asperger syndrome and other autism spectrum disabilities", *Autism*, Vol. 12/2, pp. 173-190, http://dx.doi.org/10.1177/1362361307086664. [35]

National Autistic Society (2016), *The autism employment gap: Too Much Information in the workplace.*, https://www.basw.co.uk/resources/autism-employment-gap-too-much-information-workplace. [17]

National Autistic Society (2016), *What is autism?*, https://www.autism.org.uk/advice-and-guidance/what-is-autism (accessed on 29 July 2020). [28]

OECD (2000), *Inclusive Education at Work: Students with Disabilities in Mainstream Schools*, OECD Publishing, Paris, https://dx.doi.org/10.1787/9789264180383-en. [11]

Ok, M. and W. Kim (2017), "Use of iPads and iPods for Academic Performance and Engagement of PreK–12 Students with Disabilities: A Research Synthesis", *Exceptionality*, Vol. 25/1, pp. 54-75, http://dx.doi.org/10.1080/09362835.2016.1196446. [2]

Porayska-Pomsta, K., A. Alcorn, K. Avramides, S. Beale, S. Bernardini, M.E. Foster, C. Fraunberger, J. Good, K. Guldberg, W. Keay-Bright, L. Kossyvaki, O. Lemon, M. Mademtzi, R. Menzies, H. Pain, G. Rajendran, A. Waller, S. Wass and T.J. Smith (2018), "Blending Human and Artificial Intelligence to Support Autistic Children's Social Communication Skills", *ACM Transactions on Computer-Human Interaction*, Vol. 25/6, pp. 1-35, http://dx.doi.org/10.1145/3271484. [27]

Porayska-Pomsta, K., K. Anderson, S. Bernardini, K. Guldberg, T. Smith, L. Kossivaki, S. Hodgins and I. Lowe (2013), "Building an Intelligent, Authorable Serious Game for Autistic Children and Their Carers", in *Lecture Notes in Computer Science, Advances in Computer Entertainment,* Springer International Publishing, Cham, http://dx.doi.org/10.1007/978-3-319-03161-3_34. [59]

Prizant, B., A. Wetherby, E. Rubin, A. Laurent and P. Rydell (2006), *The SCERTS model: A comprehensive educational approach for children with autism spectrum disorders. Vol. 1.,* Paul H Brookes Publishing. [36]

Rello, L., M. Ballesteros, A. Ali, M. Serra, D. Alarcon Sanchez and J. Bigham (2016), "Dytective: Diagnosing Risk of Dyslexia with a Game", *Proceedings of the 10th EAI International Conference on Pervasive Computing Technologies for Healthcare*, http://dx.doi.org/10.4108/eai.16-5-2016.2263338. [48]

Rumney, H. and K. MacMahon (2017), "Do social skills interventions positively influence mood in children and young people with autism? A systematic review", *Mental Health & Prevention*, Vol. 5, pp. 12-20, http://dx.doi.org/10.1016/j.mhp.2016.12.001. [22]

Sayce, L. (1998), "Stigma, discrimination and social exclusion: What's in a word?", *Journal of Mental Health,* Vol. 7/4, pp. 331-343, http://dx.doi.org/10.1080/09638239817932. [19]

Serholt, S. (2019), "Interactions with an Empathic Robot Tutor in Education: Students' Perceptions Three Years Later", in *Artificial Intelligence and Inclusive Education, Perspectives on Rethinking and Reforming Education*, Springer Singapore, Singapore, http://dx.doi.org/10.1007/978-981-13-8161-4_5. [62]

Shilkrot, R., J. Huber, W.M. Ee, P. Maes and S. Chandima Nanayakkara (2015), "FingerReader", *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems,* http://dx.doi.org/10.1145/2702123.2702421. [52]

Singer, J. (1999), "Why can't you be normal for once in your life? From a problem with no name to the emergence of a new category of difference.", *Disability Discourse,* pp. 59-70. [20]

Sonne, T. and M. Jensen (2016), "ChillFish: A Respiration Game for Children with ADHD", *Proceedings of the TEI '16: Tenth International Conference on Tangible, Embedded, and Embodied Interaction*, http://dx.doi.org/10.1145/2839462.2839480. [23]

Sonne, T. and M. Jensen (2016), "Evaluating the ChillFish Biofeedback Game with Children with ADHD", *Proceedings of the The 15th International Conference on Interaction Design and Children,* http://dx.doi.org/10.1145/2930674.2935981. [24]

Spiel, K., C. Frauenberger, O. Keyes and G. Fitzpatrick (2019), "Agency of Autistic Children in Technology Research—A Critical Literature Review", *ACM Transactions on Computer-Human Interaction,* Vol. 26/6, pp. 1-40, http://dx.doi.org/10.1145/3344919. [21]

UNESCO Institute for Statistics (2018), *Education and disability: Analysis of data from 49 countries. Information Paper No. 49,* http://uis.unesco.org/sites/default/files/documents/ip49-education-disability-2018-en.pdf. [16]

Volkmar, F. and J. McPartland (2014), "From Kanner to DSM-5: Autism as an Evolving Diagnostic Concept", *Annual Review of Clinical Psychology*, Vol. 10/1, pp. 193-212, http://dx.doi.org/10.1146/annurev-clinpsy-032813-153710. [15]

Whitehouse, A., K. Durkin and E. Jaquet (2009), "Friendship, loneliness and depression in adolescents with Asperger's Syndrome", *Journal of Adolescence,* Vol. 32/2, pp. 309-322, http://dx.doi.org/10.1016/j.adolescence.2008.03.004. [33]

World Health Organization (2011), *World report on disability 2011,* https://www.who.int/disabilities/world_report/2011/report.pdf. [10]

Zablotsky, B. and L. Black (2020), *Prevalence of children aged 3–17 years with developmental disabilities, by urbanicity: United States, 2015–2018.,* Natl Health Stat Report. 2020 Feb;(139):1-7. PMID: 32510313., https://pubmed.ncbi.nlm.nih.gov/32510313/. [14]

Zablotsky, B., L. Black, M. Maenner, L. Scieve, M. Danielson, R. Bitsko, S. Blumberg, M. Kogan and C. Boyle (2019), "Prevalence and Trends of Developmental Disabilities among Children in the United States: 2009–2017", *Pediatrics*, Vol. 144/4, p. e20190811, http://dx.doi.org/10.1542/peds.2019-0811. [13]

Zolna, K., T. Asselborn, C. Jolly, L. Casteran, M.A. Nguyen-Morel, W. Johal and P. Dillenbourg (2019), "The dynamics of handwriting improves the automated diagnosis of dysgraphia", http://arXiv preprint arXiv:1906.07576. [45]

7 Social Robots as Educators

Tony Belpaeme

Ghent University, Belgium and University of Plymouth, United Kingdom

Fumihide Tanaka

University of Tsukuba, Japan

Robots in education largely fall into two categories: robots that are used to teach and enthuse children about STEM subjects, and the more recent application of robots as teachers. While the pedagogical potential of robots for STEM education has been extensively explored since the 1970s, robot teachers form a new technology, driven by new developments in artificial intelligence and robotics, which is currently the subject of research and proof-of-concept trials. These robots assist teachers in their pedagogical task by offering specific tutoring experiences to students. Their potential stems mainly from their ability to provide one-to-one tutoring and a physical presence, with the latter missing in traditional computer-based learning. While there are no commercial solutions yet aimed at formal education, research suggests that social robots do offer benefits which computer-based solutions do not. Their physical nature lends them to real-world interactions with learners, and they have an increased social presence, which enhances learning outcomes. There are, however, considerable technical, economical and logistical challenges to rolling out social robots in classrooms.

Introduction

Robots in education are typically used as a means to teach STEM subjects. Originally seen as a way to introduce students to programming and computational thinking, robots are now also used as a tool in the teaching of electronics, mechanical design, computational thinking, and even arts, and to also practice soft skills such as collaborative work and negotiation skills (Alnajjar et al., 2021[1]). The use of robots for STEM education has been well explored for over 50 years and is known to be particularly effective as a catalyst to teach a range of subjects (Benitti, 2012[2]). As a result, robots as learning tools are adopted with varying success in primary, secondary and tertiary education.

In recent years a new application has emerged of robots in education. Fuelled by advances in robotics and artificial intelligence, social robots are now being explored as teaching assistants (Belpaeme et al., 2018[3]). Social robots are robots that interact with people using the same interaction channels used in human-to-human communication. They use speech, facial expressions or body language to communicate. They are often designed to have visual appeal and their software is tailored to keep social interaction flowing. While these robots are still fairly limited in terms of their interaction abilities, they do well in restricted and contained interactions (Bartneck et al., 2020[4]; Breazeal, 2004[5]; VanLehn, 2011[6]). In recent years the potential of social robots has been explored in education, with a body of research showing that robots have significant potential in formal and home education.

In this chapter, we will present the promise of robots as educators and also the current limitations in this area. We will mainly focus on two types of robots: social robots that are designed to operate autonomously and aid teachers to carry out some tasks; and telepresence robots based on a more hybrid model that are operated remotely by teachers and embody the teacher in the classroom. After presenting some of the current learning domains of application and giving some explanations on their technology, we will highlight some difficulties for their wide adoption and stress that, while they can be useful supplements to teachers, they are unlikely to replace teachers in the near future. Their cost may also remain too high for a massive presence in education systems.

Why robots as educators?

Social robots are attractive in a number of ways. Their lifelike behaviour and their social responsiveness speaks to us, and this has not gone unnoticed by educators. In the most basic sense, robots appeal to a large audience and can be used to make teaching more engaging. But beyond this short-term attraction, perhaps the largest appeal of robots is that they offer the potential for taking over some of the tasks of the teacher. While financial resources put a limit on the amount of time teachers can spend on pupils, robots are relatively cheap and could potentially be used to teach and tutor small groups, with, in the ideal case, each learner being assigned their own robot tutor. Tutoring, that is, the teaching of only one learner or a small group of learners, is known to be one of the most effective forms of teaching. VanLehn (2011[6]) found that human tutoring has a mean effect size of $d = 0.79$ when compared to class-based teaching. Computer-based teaching, and specifically step-based Intelligent Tutoring Systems (ITS) which provide fine-grained customised instruction and feedback to the learner, can achieve a similar outcome of $d = 0.76$ for certain subjects (VanLehn, 2011[6]). The expectation is that robot as educators will achieve similar outcomes, with a recent review showing that early robot prototypes achieved learning gains of $d = 0.70$ when instructing subjects varying from language tutoring to teaching about the dynamics of energy use in a city (Belpaeme et al., 2018[3]).

Key to assigning an educational role to robots is the social interaction that these robots can support. Through their appearance and software, social robots are optimised to engage humans in a manner that comes naturally. Most often these robots will have a friendly design, with face-like features such as a head, eyes and a mouth evoking the ability of the robot to see, hear and speak. Through their artificial intelligence they are able to interact with people: they use face detection and recognition to detect and identify people, use speech recognition to extract words from spoken interaction, and will use dialogue models and speech synthesis to converse (Bartneck et al., 2020[4]). The natural interface makes them suitable for a large range of tutoring activities, from tutoring pre-literate children to tutoring languages using natural interaction.

While software-based solutions for learning, such as Intelligent Tutoring Systems, can offer individual and personalised teaching, a robot adds social and physical presence which is missing from a typical Intelligent Tutoring System. It has been shown that the physical and social presence of a robot induces behaviours which are conducive to learning (Bainbridge et al., 2011[7]; Li, 2015[8]). Attention is higher, compliance increases, and motivation is more persistent when social robots are used. This is likely due to the human brain responding strongly to social stimuli, and while on-screen avatars to some extent also offer a social presence, the tangible nature of a robot amplifies this and any associated effects. This may sound paradoxical as a criticism of computer-based learning assumes the inability of humans to engage emotionally with machines: while interactions with robots are of course different than with other humans, they still trigger social responses and create some form of connection (Belpaeme et al., 2012[9]).

The different teaching roles of robots

Social robots can play different roles to teach or support students during their learning journey. In education, social robots are typically designed and programmed to play one (or more) of the following roles: tutor, teacher (or teaching assistant), and peer learner.

The most promising and pragmatic role for social robots in education is as a *tutor* (e.g. Kennedy, Baxter and Belpaeme, 2015[10]; Leyzberg et al., 2012[11]; Saerbeck et al., 2010[12]). In this role the robot provides support to a single learner or to a small group of learners. The robot can offer individual attention in a manner that can augment the capabilities of a teacher in a typical classroom setup. It can offer help to children who fall behind, or can challenge children who are ahead, without disrupting the usual classroom activities. It has infinite patience and can rehearse subjects as long as the teacher allows. Furthermore, the robot is often seen as non-judgemental by the learner (Bhakta, Savin-Baden and Tombs, 2014[13]; Catlin, 2014[14]), thereby removing the anxiety often associated with answering questions with a human tutor or teacher. Box 7.1 presents an example in language learning.

Box 7.1 Language learning with a robot tutor in the Netherlands

Language learning, and specifically second language tutoring, has been identified as a particularly promising application of robot tutoring. A child's first language is acquired through interacting with parents, siblings and peers, but quite often a second language is learned in a formal education setting and the process of learning is radically different from that of acquiring a mother tongue. No longer is the language learnt through interaction, but instead through the rote learning of vocabulary lists and grammatical rules. This stark contrast in mode of learning is to a large extent due to resource limitations. The teacher is not in a position to interact in the target language with individual children in the classroom, and instead is forced to resort to class-based teaching. In addition, the teacher himself might not be confident enough to speak the target language, or does not master all aspects of the target language. For example, native English speaking teachers will often struggle with French pronunciation – and vice versa.

This is where a robot tutor can make a valuable contribution (van den Berghe et al., 2019[15]). The robot can support the child in learning a second language through not just tutoring, but through genuine interaction in the target language. Not only can the robot offer language lessons, but the robot is likely to have a better accent in the target language than the teacher, as modern computer voices are now almost indistinguishable from human voices. Moreover, while speaking a new language many people suffer from foreign language anxiety, something which is alleviated when speaking to a robot as learners do not feel judged by the robot.

In a recent study Dutch children aged 5 were taught English as a second language with the help of a social robot. The young learner and the robot sat around a tablet computer, on which short stories were displayed. The robot described the stories in both Dutch and English and taught the children a range of words, from nouns to mathematical concepts, but also introduced grammar through play rather than through formal instruction (Vogt et al., 2019[16]). The study wished to see how effective robots were for early second-language learning and also tested if gestures by the robot, such as the robot mimicking a little jog when teaching the verb "running", would speed up learning. The children were able to acquire and retain English taught by the robot tutor to a similar extent as when they are taught by a tablet application. However, learning overall was slow. When learning with the robot in which they met the robot for seven 20-minute sessions, the children's score on a comprehension test of English only increased from 3.47 to 7.69 out of a maximum of 34.

Figure 7.1 English-as-a-second-language teaching with the help of a social robot

Source: Vogt et al. (2019[17])

The robot can also be used as a *teacher* or a *teaching assistant*: here the robot will substitute for the teacher by, for example, delivering a lecture or by providing assistance during teaching. In this case the robot addresses the class, rather than a single learner. The robot can assist with administrative tasks, such as registering students, and can take over narrow teaching tasks, such as announcing today's topic, checking prerequisite knowledge, setting a learning task, asking multiple-choice questions, summarising responses and providing feedback. In doing so, the robot frees up time of the teacher. The value of the robot not only lies in freeing up the teacher from class-based interaction, instead allowing the teacher to provide individual attention to learners, but providing experiences which the teacher might find difficult to do, such as native pronunciation of foreign languages. However, more often than not, robots are just used to spice up a lecture or classroom experience; for example, by acting as a sidekick to the teacher (see Box 7.2).

A novel and particularly promising role for a robot is as a *peer learner* (Tanaka and Kimura, 2009[18]; Tanaka and Matsuzoe, 2012[19]; Hood, Lemaignan and Dillenbourg, 2015[20]). In this the robot is presented as a learner and children are invited to learn together with the robot or are asked to instruct the robot. This relies on the protégé or learning-by-teaching effect, the idea that explaining study material to others reinforces a student's understanding: the robot plays the role of a "teachable agent". In this model, children have been shown to spend more time and effort on learning activities and to learn more (Chase et al., 2009[21]). This has been shown to be effective in learning subjects as diverse as handwriting (Lemaignan et al., 2016[22]) and second-language learning (Tanaka and Matsuzoe, 2012[19]). This effect is more pronounced for weaker students, and is probably linked to the robot inducing more confidence in the student. With the peer-like robot in the classroom, the weakest learners are no longer the weakest students, the robot instead is weaker and its reliance on being taught elevates the status of the children.

Box 7.2 Robots used as a sidekick during language class in Iran

Iranian secondary school students have been using a small humanoid robot, nicknamed Nima, to support them during their English classes. The robot acts as a sidekick to the teacher, with both the teacher and the robot standing at the front of the classroom. The robot helps the all-female class of 12- to 13-year olds to practice their English (Alemi, Meghdari and Ghazisaedy, 2014[25]) through for example giving feedback on exercises or demonstrating the correct pronunciation of English words and phrases. Over a period of five weeks, the teacher taught the official curriculum to a group of 30 students together with a robot assistant. When compared to a balanced group of 16 students who were taught only by a teacher, the group with the robot assistant showed greater vocabulary gain and retention (the students' mean score on a pre-test was 13.45, after five weeks the robot group had a mean score of 39.76, the control group a mean score of 30.50). At the same time, the group with the robot assistant enjoyed the subject more, offering a possible explanation for their performance.

Figure 7.2 A class with a robot assistant

Source: Alemi, Meghdari and Ghazisaedy (2014[25])

Besides teaching, robots can also be used to provide *socio-emotional support*. As robots are generally considered to be non-judgemental and neutral, people will often share information with a robot which they would be reluctant to share with others. This can be used to discuss personal aspects and offer advice on how to address problems. The robot can, if given permission by the learner, share selected information with teaching professionals or support staff. This has for example been effectively in education to address bullying, the study showed that children disclosed more about the occurrence of bullying at school than they when reporting bullying using an anonymised form (Bethel, Stevenson and Scassellati, 2011[23]; Bethel et al., 2016[24]).

Robots as telepresence devices for teaching and learning

Social robots in education are usually designed to perform certain learning tasks by themselves. While most of the time they are still meant for a learning environment controlled by teachers, they are often autonomous to perform specific teaching (or learning) tasks. Robots can also be used in a different way, as a telepresence device, with varying functionalities. The telepresence robot becomes the teachers' avatar in the classroom.

In this scenario, a remote teacher controls the robot and engages the learners. As the artificial intelligence required to run fully fledged teaching is still unavailable, telepresence robots can fill this technological gap. As the robot is controlled by a human teacher, it can respond to the sometimes wide-ranging responses and needs of the children, something which an AI-controlled robot, for the time being, cannot do. It also allows learners to have access to expertise and skills which their local teacher cannot offer, or draws in expertise from people who are not locally present or who wish to contribute to teaching on an occasional basis.

Telepresence robots are robots that are remotely operated by a human operator and are capable of embodying the operator's presence as a robot avatar. These robots can provide several benefits by leveraging the embodied properties of the avatar in the field of education. This topic has been gaining particular attention under the risk of infectious diseases such as COVID-19.

One instance of use of telepresence robots in education is when a human teacher employs a robot to conduct lessons remotely, providing the teacher with a richer perception of the classroom than with the traditional video-conferencing. This is because the teacher can arbitrarily control the locations of the robots and of the sensors (cameras, microphones, etc.) that are installed on the avatar robot, whereas the sensor locations are mostly fixed in the traditional video-conferencing.

In addition, students in the classroom can experience the presence of the remote teacher better when a robot avatar is present in the classroom. Typically, the teacher's face is projected on the head of the robot. A field trial conducted in a public elementary school in Japan revealed that students felt some sort of tension (i.e. the classroom atmosphere under control) in the presence of an avatar robot in the classroom (Okamura and Tanaka, 2020[26]). The students were mostly in favour of this tension because it prevented distractions in the classroom in the absence of the human teacher.

Another instance of the benefit of telepresence robots is the situation in which a student operates the avatar robot. For example, a student can participate in the classroom activity of a remote school (even in a different country), which offers good opportunities for both learning various languages and experiencing different cultures. In a field trial (Tanaka et al., 2013[27]), two elementary school classrooms, one in Australia and the other in Japan, were connected by a telepresence robot, and students of one school participated in the classroom activities of the other school (Figure 7.3). The classroom teachers and school managers who observed the trial greatly appreciated the use of this technology for learning languages and cultures. This functionality is also often used to allow students who have a long-term illness to maintain the contact with their school (Box 7.4).

Furthermore, a study reported that a telepresence robot can be used to facilitate second-language learning in a face-to-face manner (Tanaka et al., 2014[32]). Here, learners participated in a private lesson conducted remotely by a native speaker of a language via traditional video-conferencing. This situation posed a challenge for some learners; in fact, learners often froze owing to their non-fluency in using the second language. However, when they participated in the lesson by using a telepresence robot, they could communicate with the remote teacher not only verbally but also physically (i.e. by utilising the body of the avatar robot), which facilitated the second-language learning (Figure 7.4). With gesturing and hand-to-hand interactions, both learners and teachers can relax, resulting in responsive learning between them (Tanaka et al., 2014[32]). In Korea, the Engkey robot developed by the Korea

Institute of Science and Technology was designed to support English learning in primary schools: a remote teacher, whose face appears on the robot screen, controls the robot using a computer. Results from a field study in the city of Daegu involving 29 classrooms suggest that the telepresence robot controlled by a native English speaker improved students' achievement, especially in speaking (Yun et al., 2011[33]).

Box 7.3 Attending class via a telepresence robot in Norway and France

Many telepresence robots were developed with the objective of allowing students with long-term illness to maintain a connection with their school.

For example, the AV1 robot developed by the Norwegian start-up No Isolation is a student-operated telepresence robot. When a student cannot attend class due to illness, AV1 can take the student's place in the classroom. The robot is equipped with a camera, speaker, microphone, and Internet connection, enabling remote students to listen, see and speak in class (by using an app on a smartphone or tablet). They can look around the room, raise their hand to speak in class, change the robot's eye expression to convey their emotions (e.g. confusion) and even whisper to a neighbouring classmate. The use of the robot can of course also allow them to stay in touch with friends by attending birthdays and other gatherings remotely.

In France, "my connected schoolbag" is a telepresence robot also designed to take the place of a student in the classroom when ill. It is equipped with the same hardware devices as the Norwegian AV1 robot (rotating camera, speaker, microphone, Internet connection, plus a tablet app). Developed as part of a non-profit initiative, this telepresence robot takes the form of a traditional schoolbag, with the objectives of allowing students to attend class and be present among their classmates, while avoiding substitution mechanisms that anthropomorphic robots might generate.

Many similar robots were developed for similar uses in other countries, such as Inbot Technology Ltd.'s PadBot in China; Axyn Robotique's Ubbo in France; FuutureRobot's FURo-i in Korea; Wicron Robotics' Webot and R.bot's Swan Synergy in the Russian Federation; Giraff Technologies' Giraff in Sweden; Xandex's Kubi, Blue Ocean Robotics' Beam, Orbis Robotics' Carl and Teleporter, Double Robotics' Double 3 in the United States.

Source: AV1: Belton (2018[28]); and Anthony (2017[29]); Mon cartable connecté (n.d.[30]); Telepresencerobots (n.d.[31])

Figure 7.3 Classrooms in Australia and Japan were connected by a telepresence robot in real time

Source: Tanaka et al. (2013[27]).

As a supplement to remote and virtual laboratories in science, telepresence robots are being tested to carry out science experiments in a real lab. Researchers in Canada studied the use of a specifically-designed telepresence robot in a mock-up smart laboratory as a way to enable students to conduct lab or field work remotely. They built an affordable prototype (about USD 350) with a two-degree of freedom arm; online users could easily operate it and results of the small pilot suggested increased engagement of online students (Tan et al., 2019[34]).

Considering the benefits explained so far, one may imagine students participating in a classroom by using a telepresence robot from his/her home, enabling them to safely attend classes in dangerous situations involving the spread of infectious diseases, for example. Currently, because of the high cost this would represent, this idea is not feasible in education. However, a single robot avatar can be controlled by multiple operators (shared-control), and therefore, fewer robot avatars than the number of students is required. It will be interesting to test whether a few (e.g. 2-3) robot avatars and a teacher can conduct a meaningful class for more (e.g. 4-6 or more) students.

Figure 7.4 A telepresence robot facilitated the second-language learning of students

Source: Tanaka et al. (2014[32])

With the help of technologies supporting remote education, we may be able to secure new teachers. For example, retired senior people who have the necessary skill and knowledge can teach school children from their homes (Figure 7.5) (Okamura and Tanaka, 2016[35]). By introducing AI features such as keyword-spotting and dialogue generation to facilitate intergenerational conversation between school children and these senior teachers, a telepresence robot can act as an intelligent interface that assists the senior teachers in grasping the status of the remote students and giving a lecture effectively.

Figure 7.5 Senior people can give a lecture from their homes by making use of a telepresence robot

Source: Okamura and Tanaka (2016[35]).

Robot effectiveness: age and learning domains

Robots seem to be effective for a wide range of ages. While most research focuses on children between 6 and 12, robots have been shown to be effective for other ages as well. They have been used for language tutoring in pre-school classrooms (Gordon et al., 2016[36]; Vogt et al., 2019[16]), but have equally been shown to be effective in higher education (Weber and Zeaiter, 2018[37]). While it was believed that robots would be most effective at the age where children still show "suspension of disbelief" it is now increasingly clear that social robots are convincing at any age. An appropriate design of the robot and an appropriate interaction experience for the age of the learner suffices. However, building robots for older students has proven to be more technically challenging, as older learners put higher demands on the robot's abilities (Beran et al., 2011[38]). Younger children will accept that the robot leads the interaction and keeps the interaction on track, and will still believe the robot to be social. Older children are more likely to want a more diverse and unscripted interaction, which in turn is technically challenging to deliver.

Robots as educators work best for topics that are relatively restricted and where the input to the robot is well-described. One reason for this is that social signal processing, a sub-branch of artificial intelligence which attempts to interpret the social environment (Vinciarelli, Pantic and Bourlard, 2009[39]), can for now only handle rather explicit social signals, such as strong facial expressions or gestures. The technology to transcribe and understand spoken language and non-verbal social signals, to understand the intents and beliefs of people, and come up with appropriate responses, while impressive, often struggles to interpret social interaction in context. But well-functioning social signal processing is necessary to offer a full interactive experience similar to that offered by a human teacher. This means, for example, that robots for now are not able to engage in unconstrained dialogue: while they can transcribe spoken language into written language, the artificial intelligence struggles to access the meaning of what is being said and without it cannot formulate an appropriate response. This is the reason why voice-operated technology still requires us to use short, structured commands and for now struggles with long and unconstrained spoken language. However, when appropriate constraints on the learning context are in place, robots can provide support in a wide range of subjects. Often, the robot is presented together with a screen, which is not only used to display educational content, but which also serves as an input device where the learner can enter responses or select exercises shortcutting the need for the robot to understand verbal input.

Content-based subjects, where rote learning is important, such as geography, vocabulary, or low-level science, lend themselves well to robot tutoring. The repetitive nature of this type of learning and the relative ease with which topics can be taught and tested make them suitable for computer-based and robot-based instruction. As these topics are well-structured, the level of the learner is easily assessed through formative assessment or quizzes, upon which the robot can adapt its tutoring.

Skill-based subjects such as reading or mathematics also lend themselves relatively well to tutoring by robots. Here, often the skills are displayed on a computer screen, as they often have a strong visual component, and the robot offers encouragement and support during learning. The robot adapts to the learner's profile and progress and will offer hints, encouragement and help at just the right time. As with Intelligent Tutoring Systems, it is important to dose support in an appropriate manner. Too much support often leads to learners' overreliance on the system's help function (Aleven et al., 2003[40]), but robots can actively shape help-seeking behaviours of the learner (Ramachandran, Litoiu and Scassellati, 2016[41]).

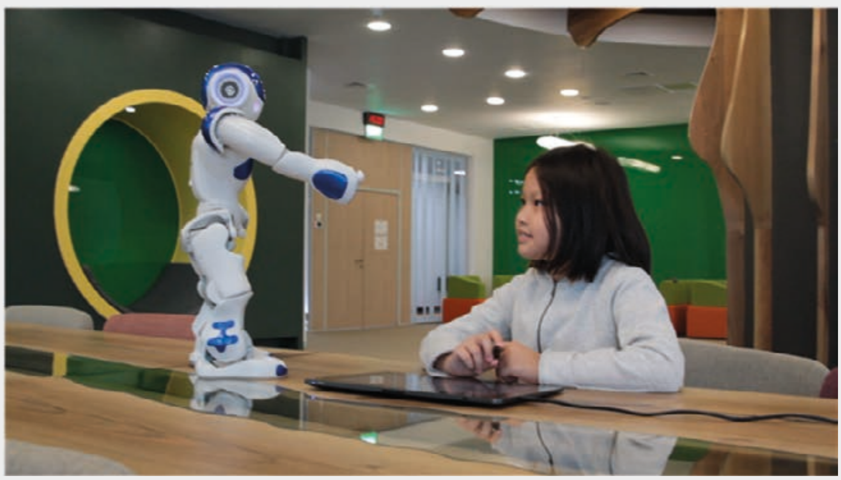
Robots have also shown promise in tutoring in the affective domain. This can be directly applied to education itself, for example when the robot encourages learners to complete more exercises or to practice at home (Kennedy et al., 2016[42]). But, robots can also be used to teach and promote soft skills and social skills. For example, robots are known to be effective at practicing social skills for children with Autism Spectrum Disorders (Robins et al., 2004[43]; Scassellati, 2007[44]). Here, the robot implements Applied Behaviour Analysis, a therapeutic approach in which social skills are internalised through repeated practice.

Robot tutoring comes into its own with the tutoring of physical skills, such as motor skills or physical rehabilitation. Here, the robot demonstrates and supports the learning process. A social robot has been used to teach handwriting skills to children aged between 6 and 8. For this peer learning was used: the robot was presented to children as having poor handwriting and the children were asked to demonstrate proper handwriting to the robot. As the robot improved, the poorest hand-writers were shown to improve as well (Hood, Lemaignan and Dillenbourg, 2015[20]; Zhexenova et al., 2020[45]) (Box 7.4).

Box 7.4 Robots supporting handwriting skills in Kazakhstan

The Kazakh government decided in 2017 on a transition from Cyrillic to Latin script for government communication and education. The transition will be phased over a seven-year period, and the role of formal education is essential in this process. A team of researchers studied how robots can be used to support children with learning the newly-adopted script and its handwriting system. They use a "learning by teaching" paradigm, in which the children taught the robot. While the robot approach achieved similar results as learning with a tablet computer or a teacher, the robot has significant benefits when it comes to lifting children's motivation and the robot was found to be the preferred method of learning (Zhexenova et al., 2020[45]).

Figure 7.6 Experimental setup



Source: Zhexenova et al. (2020[45])

The most challenging learning domain, but also the one where robot tutoring is likely to have the largest return on investment, is that where the robot relies on unconstrained social interaction for teaching. Language learning is a prime example, as it benefits from spoken interaction in the target language, something which could be practised with a robot. These robots would also be able to offer support going beyond formal teaching, and could for example offer psychosocial support. However, the technical challenges involved in unconstrained human-machine interaction mean that this application will probably take decades to mature. As mentioned above, in the meantime telepresence robots operated by humans can fill the gap.

Technical aspects of robots in education

Tutoring robots come in a variety of shapes and sizes, from 10cm-tall simple robots to tall, humanoid robots. The application and target audience often dictates which robots are more appropriate: for young children a small robot might be more appropriate, but a robot addressing a lecture theatre with adult students might need to have more authority and therefore be taller and human-like. The actual appearance of the robot seems to have little influence on the learning outcomes: research shows that more human-like robots do not necessarily achieve better learning outcomes, but rather that it is the presentation and social presence of the robot which matters for the learning outcomes. A meta-analysis (Belpaeme et al., 2018[3]) for example showed that small toy – like robots, such as the Keep On robot, which is a 15cm – tall robot, can achieve learning outcomes comparable to more expensive humanoid robots (such as the robot in Box 7.4).

In its simplest form, a social robot has very limited interactive abilities. It runs through simple scripts in response to minimal input, such as button presses or input on a tablet computer. Moving on from there, the robot can shape the interaction based on the learner's performance. It can monitor answers to quizzes and build a model of the learning trajectory of the student. This model is then used to shape the interaction, as already happens in Intelligent Tutoring Systems. In recent years, machine-learning techniques have been used to build more detailed models of the learner's performance and predict which responses by the robot will have the highest chance of increasing learning (Schodde, Bergmann and Kopp, 2017[46]; Spaulding, Gordon and Breazeal, 2016[47]). As such, robots can offer exercises and material that sufficiently challenges the student without being off-putting. Social robots can also provide encouragement or commiserate with wrong answers, for which they require a modest amount of information such as the answers given and the timing of the responses. While these approaches have for some time been available in educational software, they are only now becoming available in commercial social robots.

However, beyond adapting the teaching to the learner, the assumption is that social robots also complement their educational functionality with a range of social responses. One element of an increased social responsivity is the use of personalisation (Molenaar, 2021[48]). This not only means that the robot tailors its teaching or quizzing to the individual learner, but that it also stores and recalls personal information. This can range from using the learner's name, to storing facts about their family and hobbies and adapting its behaviour to the personality of the learner. Studies have shown that students bond with robots which appropriately use personal information (Belpaeme et al., 2012[9]), and that this has a positive impact on learning outcomes and motivation. Baxter et al. (2017[49]) embedded a social robot in two primary school classrooms in the United Kingdom for a period of two months. The robots acted as learning companions for familiar and novel subjects. One robot personalised its responses, by using the children's name when speaking and by adapting its personality to match that of the child, while the other robot did not. Results showed that there was increased learning of novel subjects when interacting with a robot that personalised its behaviours, with indications that this benefit extended to other class-based performance in which the robot did not take part. The study also showed increased acceptance of the personalised robot peer over a non-personalised version.

Finally, full social interaction requires the robot to be able to interpret verbal and non-verbal social signals and respond appropriately. Some elements are technically straightforward: a robot can use its camera to detect people and respond to their presence, it can make eye contact and read simple emotions. Speech recognition, the transcription of speech into written language, is fairly mature and, while it does not work well yet for younger users, the technology is sufficiently advanced to allow spoken responses to the robot. This can be used for simple, turn-based spoken interaction, for example the robot asks a multiple-choice question and listens for a spoken response.

However, unconstrained dialogue is still technically impossible. For this the robot not only needs to transcribe speech, but needs to understand what is being said and formulate an appropriate response. This is only possible in limited interaction settings, but the current state-of-the-art cannot yet handle open dialogue between a user and a robot. Still, this and other social skills, would make for a very persuasive and effective social robot in general, which would have large implications for the application of social robots, including in the educational landscape.

Attitudes of teachers

As teacher attitudes are a strong predictor of technology use in the classroom, a positive attitude towards robots will be a prime driver for adoption of social robots in education. Studies show that teaching professionals hold a range of attitudes towards the use of social robots in education, but that many have reservations about the introduction of robots in the classroom (Kennedy, Lemaignan and Belpaeme, 2016[50]; Kim and Lee, 2015[51]; Reich-Stiebert and Eyssel, 2016[52]; Serholt et al., 2014[53]).

Teachers with a more positive attitude towards technology in general, have perhaps unsurprisingly also a more positive outlook on robots in education. They also believe robots to be good at supporting teaching activities in science, technology, engineering, and mathematics (STEM), but do not consider other subjects, such as social skills or languages, to be appropriate for the use of robots. A survey of UK primary school teachers by Kennedy, Lemaignan and Belpaeme (2016[50]) showed that 57% of teachers believed robots could assist with STEM subjects, but only 13% believed the robot to be suitable for teaching humanities and arts. This may reflect the teachers' familiarity of robots as learning devices and tools rather than robots taking up an educational role.

Studies also surveyed teachers' concerns about the use of robots, the main ones are:

- Potential disruption to classroom activities, where the robot distracts from teaching rather than supports teaching.
- Fairness of access, especially if robots will be few and far between, access should be given to learners who will benefit most, not to privileged learners. Cost is indeed an issue, and if robots are expensive, they will be only available to learners and schools with sufficient financial means.
- The robot might add to the workload of the teaching staff. If a robot would act as a teaching aid in the classroom, it should be easy to use and require a minimum of effort to set up, operate or put away.
- Finally, an often heard concern was that social robots might have a negative impact on interpersonal relations and exacerbate social isolation. They could also introduce a "robotic" style of interacting to the classroom environment, as robots have simplistic interactions, exhibit a lack of empathy and a lack of flexibility, which would have negative consequences for young learners.

We need to stress however that these results are based on surveys completed by teachers who have never had access to social robots, or to robots used as teaching assistants. As such, responses will be coloured by preconceptions formed by popular media and science fiction, and might not reflect future attitudes. Anecdotal evidence suggests that teachers who have worked with social robots in the classroom are generally very positive about the potential applications of robot teachers.

One interesting element in all surveys is that job loss is less of a concern or no concern at all: teachers do not consider robots to be a replacement for human labour, but see robots as a technological aid for their profession.

Some of those concerns appear as legitimate, while others are less so. For example, while the concern about robot-like interactions, in which the interaction lacks depth and warmth, is currently valid, the idea that children (and adults) cannot engage in emotional interactions with robots is actually not true. Studies show that children and adults readily form emotional bonds with robots, and that these can be used to support people in a positive manner (Belpaeme et al., 2012[9]). As mentioned above, robots can help solve socio-emotional issues, and research has shown that, while understanding their limitations and the fact that they are just machines, children can "like" a well-designed robot (van Straten, Peter and Kühne, 2020[54]).

The questions around workload, disruption, technical maintenance and distraction are fair concerns that researchers and developers in education robotics have to factor in when designing their robots. The best and most effective uses of social and telepresence robots in education still have to be researched and experimented.

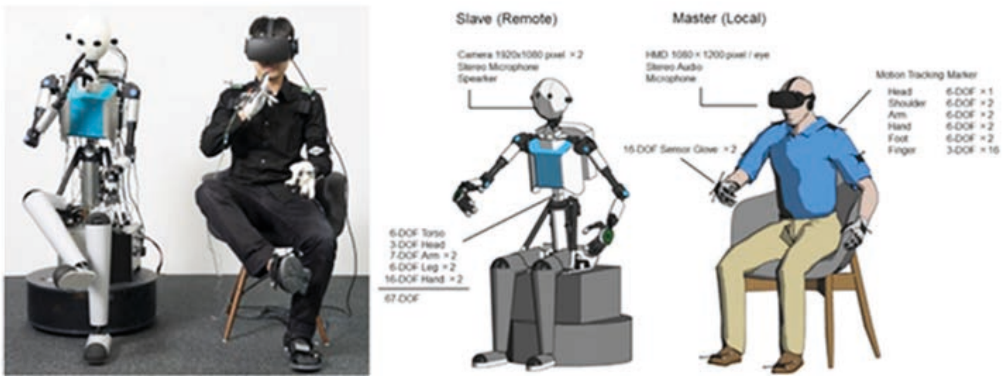
Commercial offerings

Cost and affordability are very legitimate concerns for the education community.

At the moment there is a very limited commercial availability of social robots for education. Some products are aimed at the edutainment market. Cody The Smart Cub by Vtech Playtime, an interactive bear toy, is advertised as a cuddly learning companion for children aged 6 to 36 months, but is limited to playing canned utterances and songs which can be adapted to the learner. The WEDRAW Educational Robot is a talking pen plotter aimed at children between 3 and 8 years old which can draw on paper and comes preloaded with content about shapes, figures and numbers. More advanced solutions exist as well, such as the Softbank Robotics Nao and Pepper robots, which are being presented as "educational robots" by the manufacturer. For example, an educational application for English learning was commercialised with Pepper robot (Tanaka et al., 2015[55]).

As for commercial telepresence robots, they are available at prices below USD 1 000. However, these affordable telepresence robots have limited capabilities in terms of sensors, actuators, and controlling interfaces. The latest remote avatar systems are equipped with higher-functioning sensors, actuators, and virtual reality interfaces (for robot operators), making them more expensive. Educational telepresence robots will thus probably be more expensive than those that are currently available. The TELESAR series is one of the most advanced "telexistence" system with a long history of research and development since the 1980s led by Susumu Tachi and his colleagues (Tachi, 2015[56]). The TELESAR VI (Figure 7.7) system comprises a remote avatar with 67 degrees of freedom to perform human-like movements. In principle, the system can also be equipped with rich sensing capabilities, including vision, audio, and haptics. Although these systems represent cutting-edge research and are highly expensive at present, the research lets us envision the future direction of remote education, with telepresence robots being able to behave with increasing complexity in a classroom.

Figure 7.7 TELESAR VI. (2019)



Source: Courtesy of Prof. Tachi.

A main limitation to the commercial success of these robots is twofold. On the one hand, there needs to be a substantial supply and demand for these robots. Often this forms a chicken and egg problem. When there is no demand for robot tutors, there is no commercial interest in developing these systems, and vice versa. Software and content need to be written (supply) that make these robots attractive to buyers, both families and educational institutions (demand). There currently is a lack of public and private investment here, as most content providers are reluctant to make content for educational robots. The other limitation is that social robots for education need appropriate processes and ecosystems before their introduction can be a success and they have a viable demand. At the moment there is a vacuum in teaching practice and teacher training programmes, which sits in the way of using social robots for education – even for those that are already available.

These limitations form a chicken and egg problem: content developers and robot manufacturers will withhold investment in these new technologies as long as the market is not developing, and market uptake is very low because of limited availability of affordable hardware and content. Still, there are early adopters. Some schools have been investing in the purchase of one or a few social robots and these are used for limited aspects of the curriculum, such as teaching English to Japanese school children.

Future prospects

There is limited commercial availability of robots used as teachers or tutors, and this is unlikely to change soon. The technology has to compete with other classroom educational technology tools and while research has shown that robots do offer a marked benefit in educational outcomes over screen-based technology, it is unclear if this is enough to convince technology companies and schools to invest in robots to assist teachers. As such, we need to take a long-horizon perspective and it is as yet unclear how the technology will mature and how it will find its way into mainstream education.

While research focuses on robots that have a very clear robotic appearance, with a head, eyes and a mouth, it is equally likely that aspects of the technology will be first introduced on other devices, such as digital voice assistants, which distinctly lack visual social features. Technical components, such as emotion perception or adaptive modelling of the learner, are likely to be found in other educational technology, and are not unique to robots.

A concern exists that robots in education will exacerbate the digital divide. There remain reservations about unequal access to educational technology, from access to high-speed Internet connections, access to computers and EdTech software, to differences in uptake of digital technologies across ethnicities, gender, socio-economic background, and geographic regions. As robots for education will in the foreseeable future be an exclusive technology, early adoption will likely be by affluent educational institutions and families. While this is typical for the innovation lifecycle and is typically followed by wider adoption, it is likely that the digital divide in education will persist and this will have an impact on the use of robots in education.

It is unclear if and when robots will support education, but it is very likely that insights from research into social robots for education will inspire future educational technology. The social aspects of these robots are important to motivation and learning, and the technology that powers social interaction – such as emotion recognition, the robot's ability to empathise, or the building of personalised interactions – is likely to find its way into future educational technology. It is also clear that robots will not be substitutes for human teachers; instead they will, in the best case, complement the teaching profession where human time and resources are scarce. It is unlikely that these robots will be seen in formal education in the next decade, instead, the first use of robots as tutors is expected to be in the home as educational toys. Still, the potential of robots for education is considerable, and it will only be a matter of time before we have robot assistants in the classroom.

References

Alemi, M., A. Meghdari and M. Ghazisaedy (2014), "Employing Humanoid Robots for Teaching English Language in Iranian Junior High-Schools", *International Journal of Humanoid Robotics*, Vol. 11/03, p. 1450022, http://dx.doi.org/10.1142/s0219843614500224. [25]

Aleven, V., E. Stahl, S. Schworm, F. Fischer and R. Wallace (2003), "Help Seeking and Help Design in Interactive Learning Environments", *Review of Educational Research*, Vol. 73/3, pp. 277-320, http://dx.doi.org/10.3102/00346543073003277. [40]

Alnajjar, F., C. Bartneck, P. Baxter and T. Belpaeme (2021), *Robots in education*, Routledge. [1]

Anthony, A. (2017), *The robot that staves off loneliness for chronically ill children*, The Guardian, https://www.theguardian.com/technology/2017/aug/13/robot-to-help-sick-children-norwegian-start-up (accessed on 18 May 2021). [29]

Bainbridge, W., J. Hart, E. Kim and B. Scassellati (2011), "The Benefits of Interactions with Physically Present Robots over Video-Displayed Agents", *International Journal of Social Robotics*, Vol. 3/1, pp. 41-52, http://dx.doi.org/10.1007/s12369-010-0082-7. [7]

Bartneck, C., T. Belpaeme, F. Eyssel, T. Kanda, M. Keijsers and S. Šabanović (2020), *Human-Robot Interaction*, Cambridge University Press, http://dx.doi.org/10.1017/9781108676649. [4]

Belpaeme, T., P. Baxter, R. Read, R. Wood, H. Cuayahuitl, B. Kiefer, S. Racioppa, I. Kruijff-Korbayova, G. Athasopoulos, V. Enescu, R. Looije, M. Neerincx, Y. Demiris, R. Ros-Espinoza, A. Beck, L. Cañamero, A. Hiolle, M. Lewis, I. Baroni, M. Nalin, P. Cosi, G. Paci, F. Tesser, G. Sommavilla and R. Humbert (2012), "Multimodal Child-Robot Interaction: Building Social Bonds", *Journal of Human-Robot Interaction*, Vol. 1/2, http://dx.doi.org/10.5898/jhri.1.2.belpaeme. [9]

Belpaeme, T., J. Kennedy, A. Ramachandran, B. Scassellati and F. Tanaka (2018), "Social robots for education: A review", *Science Robotics*, Vol. 3/21, p. eaat5954, http://dx.doi.org/10.1126/scirobotics.aat5954. [3]

Belton, P. (2018), *My robot makes me feel like I haven't been forgotten*, BBC, https://www.bbc.com/news/business-45332710 (accessed on 18 May 2021). [28]

Benitti, F. (2012), "Exploring the educational potential of robotics in schools: A systematic review", *Computers & Education*, Vol. 58/3, pp. 978-988, http://dx.doi.org/10.1016/j.compedu.2011.10.006. [2]

Beran, T., A. Ramirez-Serrano, R. Kuzyk, M. Fior and S. Nugent (2011), "Understanding how children understand robots: Perceived animism in child–robot interaction", *International Journal of Human-Computer Studies*, Vol. 69/7-8, pp. 539-550, http://dx.doi.org/10.1016/j.ijhcs.2011.04.003. [38]

Bethel, C., Z. Henkel, K. Stives, D. May, D. Eakin, M. Pilkinton, A. Jones and M. Stubbs-Richardson (2016), "Using robots to interview children about bullying: Lessons learned from an exploratory study", *2016 25th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN)*, http://dx.doi.org/10.1109/roman.2016.7745197. [24]

Bethel, C., M. Stevenson and B. Scassellati (2011), "Secret-sharing: Interactions between a child, robot, and adult", in *2011 IEEE International Conference on systems, man, and cybernetics (pp. 2489-2494)*, IEEE. [23]

Bhakta, R., M. Savin-Baden and G. Tombs (2014), "Sharing Secrets with Robots?", in Viteli, J. and M. Leikomaa (eds.), *Proceedings of EdMedia 2014--World Conference on Educational Media and Technology (pp. 2295-2301).* [13]

Breazeal, C. (2004), *Designing sociable robots*, MIT press. [5]

Catlin, D. (2014), "Using peer assessment with educational robots", in *International Conference on Web-Based Learning (pp. 57-65), Springer*, Cham. [14]

Chase, C., D. Chin, M. Oppezzo and D. Schwartz (2009), "Teachable Agents and the Protégé Effect: Increasing the Effort Towards Learning", *Journal of Science Education and Technology*, Vol. 18/4, pp. 334-352, http://dx.doi.org/10.1007/s10956-009-9180-4. [21]

Gordon, G., S. Spaulding, J. Kory Westlund, J. Joo Lee, L. Plummer, M. Martinez, M. Das and C. Breazeal (2016), "Affective personalization of a social robot tutor for children's second language skills", in *Proceedings of the Thirtieth AAAI Conference on Artificial Intelligence (pp. 3951-3957).* [36]

Hood, D., S. Lemaignan and P. Dillenbourg (2015), "When Children Teach a Robot to Write", *Proceedings of the Tenth Annual ACM/IEEE International Conference on Human-Robot Interaction*, http://dx.doi.org/10.1145/2696454.2696479. [20]

Kennedy, J., P. Baxter and T. Belpaeme (2015), "The Robot Who Tried Too Hard", *Proceedings of the Tenth Annual ACM/IEEE International Conference on Human-Robot Interaction*, http://dx.doi.org/10.1145/2696454.2696457. [10]

Kennedy, J., P. Baxter, E. Senft and T. Belpaeme (2016), "Social robot tutoring for child second language learning", *2016 11th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*, http://dx.doi.org/10.1109/hri.2016.7451757. [42]

Kennedy, J., S. Lemaignan and T. Belpaeme (2016), "The cautious attitude of teachers towards social robots in schools", in *Robots 4 Learning Workshop at IEEE RO-MAN 2016.* [50]

Kim, S. and Y. Lee (2015), "A Survey on Elementary School Teachers' Attitude toward Robot", in E-Learn: World Conference on *E-Learning in Corporate, Government, Healthcare, and Higher Education (pp. 1802-1807)*, iation for the Advancement of Computing in Education (AACE). [51]

Lemaignan, S., A. Jacq, D. Hood, F. Garcia, A. Paiva and P. Dillenbourg (2016), "Learning by Teaching a Robot: The Case of Handwriting", *IEEE Robotics & Automation Magazine*, Vol. 23/2, pp. 56-66, http://dx.doi.org/10.1109/mra.2016.2546700. [22]

Leyzberg, D., B. Scassellati, S. Spaulding and M. Toneva (2012), "The physical presence of a robot tutor increases cognitive learning gains", in *Proceedings of the annual meeting of the cognitive science society.* [11]

Li, J. (2015), "The benefit of being physically present: A survey of experimental works comparing copresent robots, telepresent robots and virtual agents", *International Journal of Human-Computer Studies*, Vol. 77, pp. 23-37, http://dx.doi.org/10.1016/j.ijhcs.2015.01.001. [8]

Molenaar, I. (2021), *Personalisation of learning: Towards hybrid human-AI learning technologies*, OECD Publishing. [48]

Mon cartable connecté (n.d.), http://www.moncartableconnecte.fr/index.php (accessed on 18 May 2021). [30]

Okamura, E. and F. Tanaka (2020), "Deployment of a Social Robot into a Classroom of Remote Teaching by Elderly People to School Children: a Case Report", *2020 29th IEEE International Conference on Robot and Human Interactive Communication (RO-MAN)*, http://dx.doi.org/10.1109/ro-man47096.2020.9223499. [26]

Okamura, E. and F. Tanaka (2016), "A pilot study about remote teaching by elderly people to children over a two-way telepresence robot system", *2016 11th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*, http://dx.doi.org/10.1109/hri.2016.7451820. [35]

Ramachandran, A., A. Litoiu and B. Scassellati (2016), "Shaping productive help-seeking behavior during robot-child tutoring interactions", *2016 11th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*, http://dx.doi.org/10.1109/hri.2016.7451759. [41]

Reich-Stiebert, N. (ed.) (2017), "Robot education peers in a situated primary school study: Personalisation promotes child learning", *PLOS ONE*, Vol. 12/5, p. e0178126, http://dx.doi.org/10.1371/journal.pone.0178126. [49]

Reich-Stiebert, N. and F. Eyssel (2016), "Robots in the Classroom: What Teachers Think About Teaching and Learning with Education Robots", in *Social Robotics, Lecture Notes in Computer Science*, Springer International Publishing, Cham, http://dx.doi.org/10.1007/978-3-319-47437-3_66. [52]

Robins, B., K. Dautenhahn, R. Boekhorst and A. Billard (2004), "Effects of Repeated Exposure to a Humanoid Robot on Children with Autism", in D*esigning a More Inclusive World, Springer London,* London, http://dx.doi.org/10.1007/978-0-85729-372-5_23. [43]

Saerbeck, M., T. Schut, C. Bartneck and M. Janse (2010), "Expressive robots in education: varying the degree of social supportive behavior of a robotic tutor", in *Proceedings of the SIGCHI conference on human factors in computing systems*, pp.1613-1622. [12]

Scassellati, B. (2007), "How Social Robots Will Help Us to Diagnose, Treat, and Understand Autism", in Springer Tracts in *Advanced Robotics, Robotics Research, Springer Berlin Heidelberg,* Berlin, Heidelberg, http://dx.doi.org/10.1007/978-3-540-48113-3_47. [44]

Schodde, T., K. Bergmann and S. Kopp (2017), "Adaptive Robot Language Tutoring Based on Bayesian Knowledge Tracing and Predictive Decision-Making", in *Proceedings of the 2017 ACM/IEEE International Conference on Human-Robot Interaction (pp. 128-136).* [46]

Serholt, S., W. Barendregt, I. Leite, H. Hastle, A. Jones, A. Paiva, A. Vasalou and G. Castellano (2014), "Teachers' views on the use of empathic robotic tutors in the classroom", *The 23rd IEEE International Symposium on Robot and Human Interactive Communication*, http://dx.doi.org/10.1109/roman.2014.6926376. [53]

Spaulding, S., G. Gordon and C. Breazeal (2016), "Affect-aware student models for robot tutors", in *Proceedings of the 2016 International Conference on Autonomous Agents & Multiagent Systems (pp. 864-872).* [47]

Tachi, S. (2015), *Telexistence* (2nd Edition), World Scientific. [56]

Tanaka, F., K. Isshiki, F. Takahashi, M. Uekusa, R. Sei and K. Hayashi (2015), "Pepper learns together with children: Development of an educational application", *2015 IEEE-RAS 15th International Conference on Humanoid Robots (Humanoids),* http://dx.doi.org/10.1109/humanoids.2015.7363546. [55]

Tanaka, F. and T. Kimura (2009), "The Use of Robots in Early Education: a Scenario based on Ethical Consideration", in *18th IEEE International Symposium on Robot and Human Interactive Communication (pp. 558-560),* IEEE. [18]

Tanaka, F. and S. Matsuzoe (2012), "Children Teach a Care-Receiving Robot to Promote Their Learning: Field Experiments in a Classroom for Vocabulary Learning", *Journal of Human-Robot Interaction*, pp. 78-95, http://dx.doi.org/10.5898/jhri.1.1.tanaka. [19]

Tanaka, F., T. Takahashi, S. Matsuzoe, N. Tazawa and M. Morita (2014), "Telepresence Robot Helps Children in Communicating with Teachers who Speak a Different Language", in *Proceedings of the 2014 ACM/IEEE international conference on Human-robot interaction (pp. 399-406).* [32]

Tanaka, F., T. Takahashi, S. Matsuzoe, N. Tazawa and M. Morita (2013), "Child-operated telepresence robot: A field trial connecting classrooms between Australia and Japan", *2013 IEEE/RSJ International Conference on Intelligent Robots and Systems*, http://dx.doi.org/10.1109/iros.2013.6697211. [27]

Tan, Q., M. Denojean-Mairet, H. Wang, X. Zhang, F. Pivot and R. Treu (2019), "Toward a telepresence robot empowered smart lab", *Smart Learning Environments,* Vol. 6/1, http://dx.doi.org/10.1186/s40561-019-0084-3. [34]

Telepresencerobots (n.d.), https://telepresencerobots.com/education-telepresence-robot (accessed on 15 May 2021). [31]

van den Berghe, R. et al. (2019), "Social Robots for Language Learning: A Review", *Review of Educational Research,* Vol. 89/2, pp. 259-295, http://dx.doi.org/10.3102/0034654318821286. [15]

van Straten, C., J. Peter and R. Kühne (2020), "Child–Robot Relationship Formation: A Narrative Review of Empirical Research", *International Journal of Social Robotics*, Vol. 12/2, pp. 325-344, http://dx.doi.org/10.1007/s12369-019-00569-0. [54]

VanLehn, K. (2011), "The Relative Effectiveness of Human Tutoring, Intelligent Tutoring Systems, and Other Tutoring Systems", *Educational Psychologist*, Vol. 46/4, pp. 197-221, http://dx.doi.org/10.1080/00461520.2011.611369. [6]

Vinciarelli, A., M. Pantic and H. Bourlard (2009), "Social signal processing: Survey of an emerging domain", *Image and Vision Computing*, Vol. 27/12, pp. 1743-1759, http://dx.doi.org/10.1016/j.imavis.2008.11.007. [39]

Vogt, P., R. van den Berghe, M. de Haas, L. Hoffman, J. Kanero, E. Mamus, J-M. Montanier, C. Oranç, O. Oudgeneg-Paz, D. Hernández Garcia, F. Papadoupolos, T. Schodde, J. Verhagen, C. Wallbridgell, B. Willemsen, J. de Wit, T. Belpaeme, T. Göksun, S. Kopp, E. Krahmer, A. Küntay, P. Leseman and A. Kumar Pandey (2019), "Second Language Tutoring Using Social Robots: A Large-Scale Study", *2019 14th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*, http://dx.doi.org/10.1109/hri.2019.8673077. [16]

Vogt, P., R. van den Berghe, M. de Haas, L. Hoffman, J. Kanero, E. Mamus, J-M. Montanier, C. Oranç, O. Oudgeneg-Paz, D. Hernández Garcia, F. Papadoupolos, T. Schodde, J. Verhagen, C. Wallbridgell, B. Willemsen, J. de Wit, T. Belpaeme, T. Göksun, S. Kopp, E. Krahmer, A. Küntay, P. Leseman and A. Kumar Pandey (2019), "Second Language Tutoring Using Social Robots: L2TOR - The Movie", *2019 14th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*, http://dx.doi.org/10.1109/hri.2019.8673016. [17]

Weber, K. and S. Zeaiter (2018), "Project H.E.A.R.T. (Humanoid Emotional Assistant Robots in Teaching)", in Buchner, J. et al. (eds.), I*nverted Classroom, Fachhochschule St. Pölten*, Fachhochschule St. Pölten GmbH. [37]

Yun, S., J. Shin, D. Kim, C. Gu Kim, M. Kim and M.T. Choi (2011), "Engkey: Tele-education Robot", in *Social Robotics, Lecture Notes in Computer Science*, Springer Berlin Heidelberg, Berlin, Heidelberg, http://dx.doi.org/10.1007/978-3-642-25504-5_15. [33]

Zhexenova, Z., A. Amirova, M. Abdikarimova, K. Kudaibergenov, N. Baimakhan, B. Tleubayev, W. Johal, P. Dillenbourg, A. Cohen Miller and A. Sandygulova (2020), "A Comparison of Social Robot to Tablet and Teacher in a New Script Learning Context", *Frontiers in Robotics and AI*, Vol. 7, http://dx.doi.org/10.3389/frobt.2020.00099. [45]

Part 2

8 Learning analytics for school and system management

Dirk Ifenthaler
University of Mannheim, Germany and Curtin University, Australia

Learning analytics for educational organisations has been a topic of discussion for the past decade. Yet, there are few examples of organisation-wide systematic and holistic adoptions of learning analytics. This chapter explores actionable frameworks and adopting models that could help successfully integrate learning analytics into educational organisations in an organisational change approach. While higher education organisations are aware of learning analytics and have started experimenting with dashboards for students and teachers this is far from organisational transformation. Research on the implementation and practice of learning analytics in K-12 schools is also scarce. A look at adoption models and policy recommendations in a broader international context may help push isolated experiments with learning analytics into the mainstream.

Introduction

With the increase in the amount of educational data, improved data storage and handling, and advances in computing and related analytics tools and algorithms, educational organisations are starting to embrace learning analytics. Learning analytics assess, elicit and analyse static and dynamic information about learners and learning environments for real-time modelling, prediction and optimisation of learning processes and environments, as well as educational decision making in an organisation (Ifenthaler, 2015[1]). To integrate learning analytics systems into educational organisations, actionable frameworks and adoption models are required (Buckingham Shum and McKay, 2018[2]; Dyckhoff et al., 2012[3]). However, these models may vary in different countries and organisations or within an individual organisation (Klasen and Ifenthaler, 2019[4]).

The potential benefits of learning analytics for educational organisations have been a topic of discussion for the past decade (Pistilli and Arnold, 2010[5]). However, there remains a lack of systematic and holistic adoption of learning analytics by organisations (Gašević et al., 2019[6]). The degree of adoption of learning analytics within an organisation is a measure of the number of stakeholders who use it or who have altered their practice because of it (Colvin et al., 2015[7]). For learning analytics to spread to more organisations, there needs to be communication between potential actors, who then follow a decision-making process as to whether to adopt or ignore it (Kotter, 2007[8]; Rogers, 1962[9]). Influencers who wish to promote the adoption of learning analytics must ensure and secure open communication channels and encourage people to start and stay on the journey of decision making from awareness to action (Colvin et al., 2015[7]; Drachsler and Greller, 2016[10]; Ifenthaler and Gibson, 2020[11]).

Higher education organisations have shown interest in adopting learning analytics but it is not yet a major priority (Ifenthaler and Yau, 2019[12]; Tsai and Gašević, 2017[13]; Lester et al., 2017[14]). Some have begun experimenting with dashboards for students and teachers but this is far from organisational transformation (Siemens, Dawson and Lynch, 2014[15]; Viberg et al., 2018[16]).

While studies on the adoption of learning analytics in higher education exist, the implementation of learning analytics in K-12 schools remains scarce (Andresen, 2017[17]; Gander, 2020[18]). There are two schools of thought on the usefulness of learning analytics for K-12 schools. Agasisti and Bowers (2017[19]) outline the importance of educational data and analytics for policy making, managing institutions, and classroom learning and teaching. However, Sergis and Sampson (2016[20]) argue that K-12 schools may not benefit as much from learning analytics as higher education organisations. K-12 schools require a holistic, multilevel analytics framework with several layers of data to produce sufficiently granular feedback to the school leadership and other stakeholders within the school while school analytics capture, analyse, and exploit organisation-wide educational data, allowing school leaders to monitor and (partially) influence their organisation's developments in order to better meet the needs of the students, teachers, parents, and external policy mandates. Therefore, Sergis and Sampson (2016[20]) propose a school analytics framework that includes elements on a micro layer (learning process monitoring, learning process evaluation, learner performance monitoring, learner performance evaluation), on a meso-layer (curriculum planning, teaching staff management, teaching staff professional development), and on a macro layer (district stakeholder accountability, infrastructural resource management, financial resource management, learner data management). However, research on how state and local policies can leverage data analytics for school improvement is scarce (Jimerson and Childs, 2017[21]).

This chapter focusses on organisational benefits from learning analytics and the challenges of adopting learning analytics in educational organisations. Three case examples provide insights into how organisations have been successful in adopting learning analytics and producing organisational benefits or overcoming organisational hurdles. The conclusion presents guidelines for policy makers, researchers and educational organisations adopting learning analytics and ends with a set of open questions to be addressed in future research and practice.

Organisational benefits from learning analytics

Current learning analytics research and practice in Australia, the United Kingdom and the United States is proving its value in addressing issues related to successful studying and identification of students at risk (Sclater and Mullan, 2017[22]), and monitoring and improving organisational capabilities (Ifenthaler, Yau and Mah, 2019[23]).

Governance (mega-level) facilitates cross-organisational analytics by incorporating data from all levels of learning analytics initiatives. Based on common data standards and open data schemas, rich datasets enable the identification and validation of patterns within and across organisations and therefore provide valuable insights for informing educational policy making. Examples on the government level include performance compressions across institutions including benchmarking and helping to inform policy development and resource allocations across school districts, states or countries. In crises such as the COVID-19 pandemic, learning analytics at the governance level may enable rapid response and coordination of expert support for institutions in need.

At the macro-level, organisation-wide analytics enable a better understanding of learner cohorts to optimise processes. These include allocating critical resources to do things like reduce dropout rates, and increase retention and success rates. It can be used to support school admission processes and predict school performance (based on individual student performance). Other applications support transitions between educational systems such as entry into higher education or job-seeking processes. The meso- and micro-level provide analytics insights within organisations and will not be further discussed in this contribution (Ifenthaler and Widanapathirana, 2014[24]).

An essential prerequisite of learning analytics benefits is knowing what the data and analytics is being used for. These can be broken down to: (1) summative and descriptive – detailed insights after completion of a learning phase (e.g., study period, semester, final degree) often compared against previously defined reference points or benchmarks; (2) formative and (near) real-time, which uses on-going information for improving processes through direct interventions on-the-fly; (3) predictive and prescriptive, which forecast the probability of outcomes in order to plan for future interventions, strategies, and actions. Table 8.1 provides examples at the governance and organisational level for all data and analytics perspectives (Ifenthaler, 2015[1]). These benefits can be mapped to different data profiles (e.g., student profile, learning profile, curriculum profile) including various analytics indicators (e.g., trace data, demographic background, course characteristics). Yau and Ifenthaler (2020[25]) provide an in-depth analysis of analytics indicators for specific learning analytics benefits.

Table 8.1 Example of benefits from learning analytics for the governance and organisation level

	Data and analytics perspective		
	Summative/Descriptive	Formative/Real-time	Predictive/Prescriptive
Governance	• Apply cross-institutional comparisons • Develop benchmarks • Inform policy making • Inform quality assurance processes	• Increase productivity • Apply rapid response to critical incidents • Analyse performance • Provide expert help	• Model impact of organisational decision making • Plan for change management
Organisation	• Analyse educational processes • Optimise resource allocation • Meet institutional standards • Compare units across programs and entities	• Monitor educational processes • Evaluate educational resources • Track enrolments • Analyse churn • Involve alumni	• Forecast educational processes • Project attrition • Model retention rates • Identify gaps

Source: Ifenthaler (2015[1]).

Three examples

Though there is rich research on the organisational benefits of learning analytics, the implementation of organisation-wide analytics systems is scarce (Buckingham Shum and McKay, 2018[2]). The following three case examples showcase how learning analytics can impact school and system management.

Case example 1: Adoption of learning analytics at the University of Wollongong, Australia

The University of Wollongong, Australia, faced the challenge of adopting learning analytics at the organisational level of the Deputy Vice-Chancellor (Academic) while finding ways to integrate different disciplinary learning and teaching cultures and demonstrating its value to students and teaching staff (Heath and Leinonen, 2016[26]). The university began by considering the existing tools and resources that could support learning analytics. Despite a mature data warehouse infrastructure, it was necessary to invest in additional support staff to focus on analytics, big data, and statistics. The University of Wollongong chose the 'blossoming' adoption of learning analytics – as opposed to a phased up-scaling adoption, which usually begins with a prototype, and is followed by up-scaling before reaching the stage of fully implemented learning analytics (Ferguson et al., 2014[27]). Faculties varied quite widely in the degree to which they could carry out a 'blossoming' form of adoption. This added additional complexity to the organisational change process. In addition to technical capabilities, a survey was used to collect students' views about learning analytics, including their preferences for functionalities, intervention strategies, and perceptions of privacy. Two governance committees were then formed: (a) the Learning Analytics Governance Committee, focussing on adopting learning analytics and (b) the Ethical Use of Data Advisory Group, focussing on student privacy and ethical issues regarding educational data (Heath and Leinonen, 2016[26]).

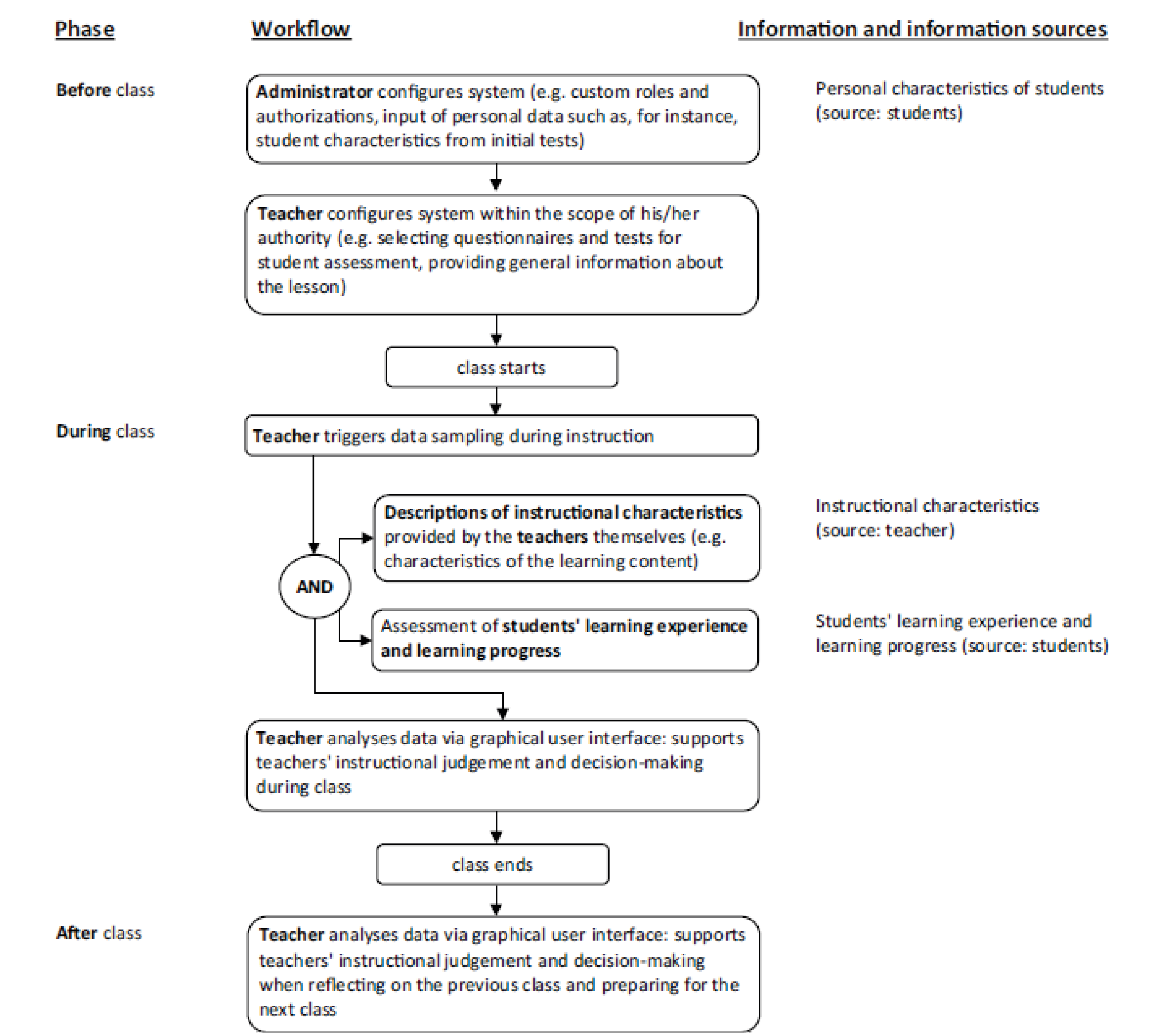
To conclude, four salient points shall be taken into consideration for a successful organisation-wide adoption (Heath and Leinonen, 2016[26]): (1) use common technological infrastructure such as a data warehouse; (2) involve students in all stages of the adoption process; (3) engage early adopters and establish communities of practice; (4) institute governance frameworks focussing on learning analytics strategy, data privacy, and ethics.

Case example 2: Teachers' Diagnostic Support System (TDSS) (Stuttgart, Germany)

Developed by researchers at the University Hohenheim in Stuttgart (Germany), the Teachers' Diagnostic Support System (TDSS) helps teachers adapt their teaching practices to the variety of student needs in the classroom. The stakeholders are teachers and students. The TDSS' collection and analysis of data are of particular interest. The TDSS allows data collection on (1) students' personal characteristics (e.g. domain-specific knowledge and competencies, emotional-motivational characteristics), (2) descriptions of instructional characteristics (e.g. characteristics of the learning content), and (3) students' learning experiences and learning progress (e.g. situational interest in the subject matter, actual knowledge about the topic) (Kärner, Warwas and Schumann, 2020[28]). Figure 8.1 provides an overview on the TDSS, which is a client–server-based software that is optimised for mobile devices.

TDSS allows the teacher to retrieve and analyse data during and after instruction to inform their teaching practice on-the-fly as well as prepare learning materials and future lessons. Micro-management through learning analytics may be expanded for cross-curricular teaching activities and school-wide diagnostic purposes.

Figure 8.1 TDSS workflow and information

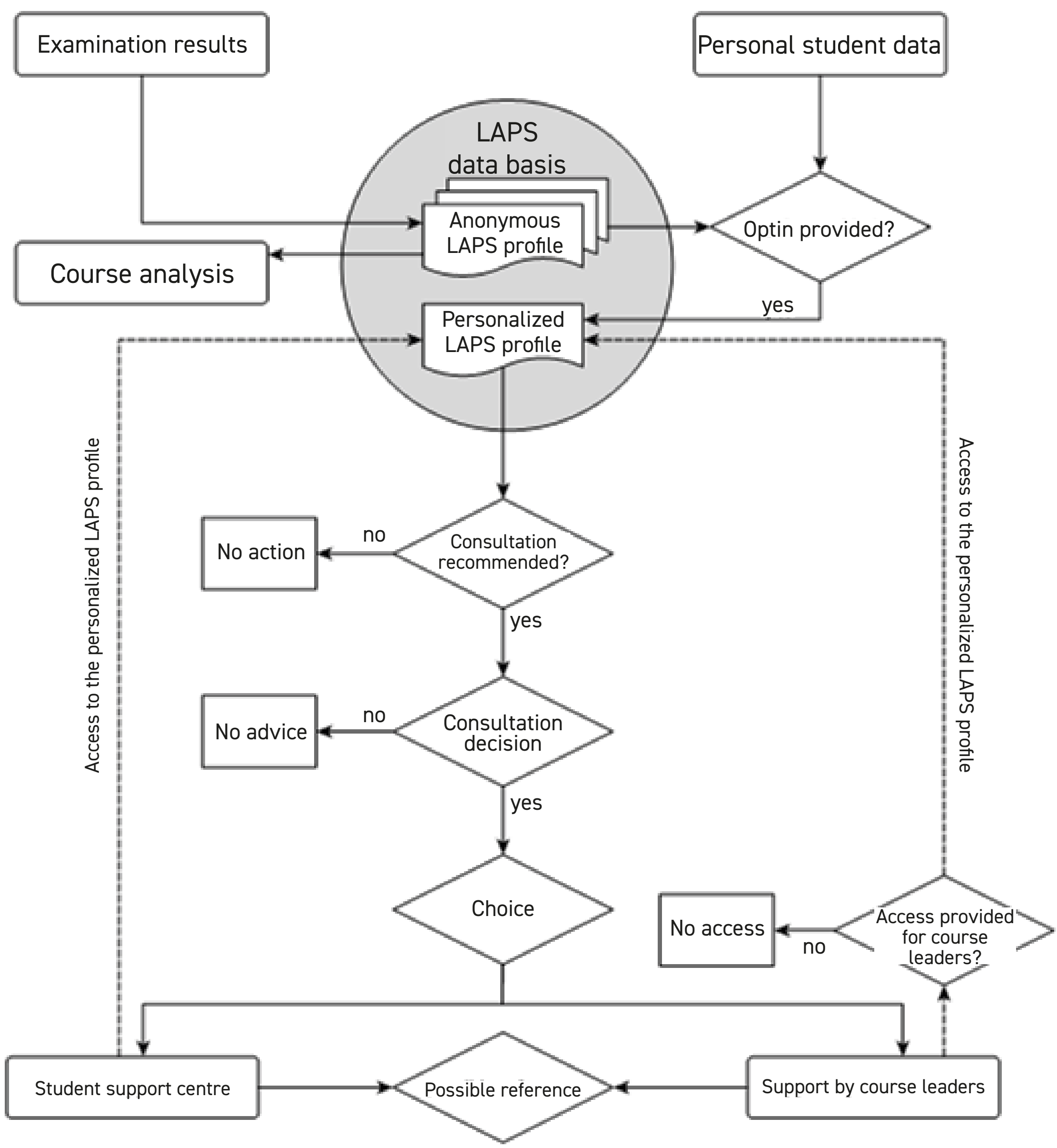


Source: Kärner, Warwas and Schumann (2020[28])

Case example 3: The LAPS project using machine-learning techniques for early student support at the Stuttgart Media University, Germany

The LAPS (Learning analytics for exams and study success) project was developed to identify students at risk of failing their studies. It was implemented at the Stuttgart Media University, Germany in 2014. The purpose of LAPS was to create an evidence-based discussion with academic staff and students at an early stage in their studies. Ethical issues of privacy, voluntariness, self-determination, self-responsibility, respect of individuality, confidentiality and anonymity are essential to the project. LAPS data, which is updated each semester, is used to produce a list of participating students' critical study progressions. It can identify more than 200 individual risk characteristics (see Figure 8.2) as well as study progressions with a high potential (Hinkelmann and Jordine, 2019[29]). In addition, LAPS is used for quality assurance, providing information about specific programmes, lectures and student cohorts. LAPS provides information on number of enrolled students, drop-outs, successful study progressions, average risk possibility, minimum/average/maximum student age, gender distribution, average grade of the university entrance qualification, and retreats from examinations for different programmes.

Figure 8.2 The LAPS process



Source: Hinkelmann and Jordine (2019[29])

The LAPS lectures insights allow detailed analysis for each semester of lectures and provides access to distribution of grades, number of successful examinations, average grade, number of retreats, number of registrations. The LAPS cohorts view allows comparison of the distribution of students' obtained ECTS (European Credit Transfer System) credits per semester. It also identifies possible structural problems when students do not achieve the required ECTS credits (Hinkelmann and Jordine, 2019[29]).

Learning analytics' challenges for organisations

The three case examples above show some possible benefits of learning analytics in triggering organisational change, and supporting teachers and students within the classroom or during their studies. There are many other possible benefits but also implementation challenges. How can educational organisations invest limited resources so as to achieve maximum benefits? Tsai and Gašević (2017[13]) point out several challenges for organisations implementing learning analytics initiatives:

- insufficient leadership in planning and monitoring learning analytics implementation and the organisational change process;
- uneven understanding of and commitment to the initiative by stakeholders, namely, administrative, technical, and teaching staff;
- a lack of pedagogical concepts and general awareness about the organisation's learning culture driving the expected benefits for learning and teaching;
- insufficient professional training for teaching staff, student service staff, technical staff and the like on the benefits and limitations of learning analytics or its technical infrastructure;
- insufficient rigorous empirical evidence on the effectiveness of learning analytics to support organisational decision making;
- not enough policies, regulations, and codes of practice on privacy and ethics in learning analytics.

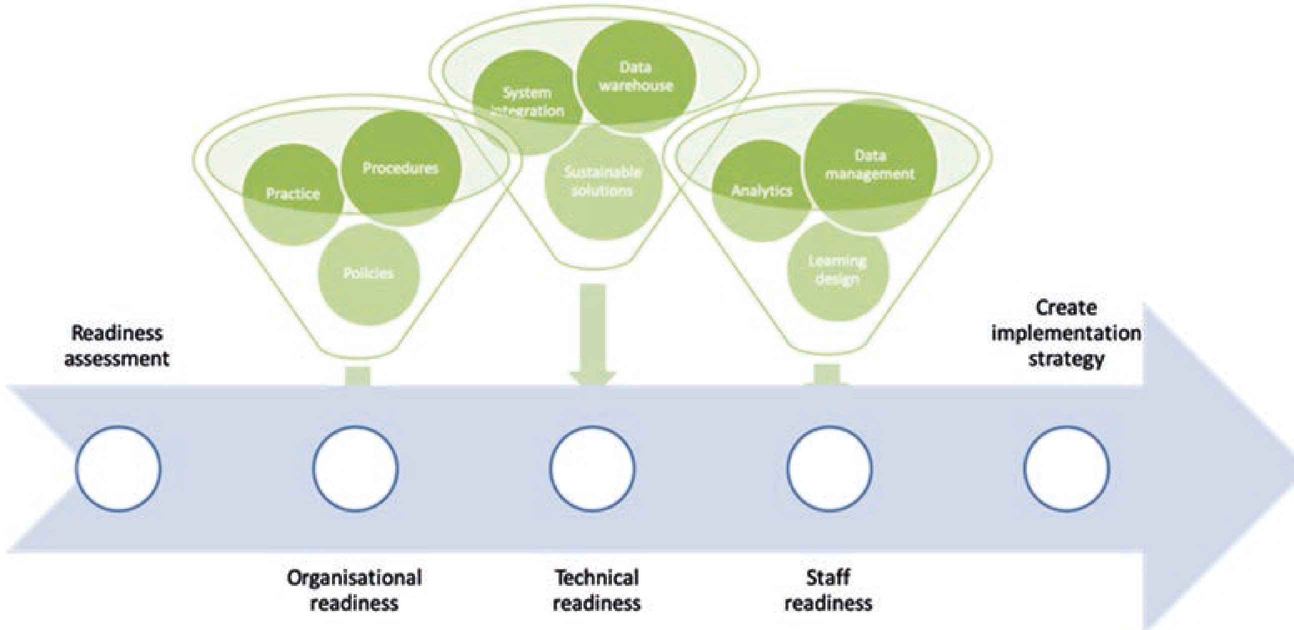
Leitner, Ebner and Ebner (2019[30]) recommend a framework of seven categories for driving organisation-wide learning analytics initiatives. These include: (1) defining how students, educators, researchers and administrators will gain from learning analytics; (2) stakeholders having access to actionable information on dashboards; (3) transparent communication with all stakeholders and ensuring of policies – particularly, privacy – that align with the organisation's core principles and others such as the EU's General Data Protection Regulation; (4) setting up and managing an information technology (IT) infrastructure that supports the requirements of the learning analytics initiative. Such advanced IT infrastructures may be provided through organisation-owned services or external service providers; (5) a scalable development and operation of learning analytics functions that the organisation can monitor and evaluate; (6) implementing a code of conduct and (7) procedures on an ethics of learning analytics that adapts to different cultural contexts.

Drachsler and Greller (2016[10]) have compiled an eight-point DELICATE checklist to facilitate trusted implementation of learning analytics which includes many of the points in Leitner, Ebner and Ebner's's framework (2019[30]). To these, they add the need to legitimise or provide reasons for the right to obtain data, secure consent through a contract with the data subjects, anonymise data and subjects, and ensure that external stakeholders adhere to national guidelines.

A change management strategy begins with a readiness assessment. The initial step consists of identifying the to-be-achieved benefits of learning analytics. The educational organisation then carries out an in-depth review of existing practices, procedures, and capabilities (see Case Example 1). The ensuing strategy thus lays out which benefits and specific learning analytics features are included as well as which infrastructure is required to successfully implement learning analytics.

A readiness assessment is also conducted using standardised instruments. Data on organisational readiness such as existing policy and data protection regulation is collected with a specific focus on the organisation's technical readiness (e.g. data warehouse, system integration) and the staff's level of educational data literacy (Schumacher, Klasen and Ifenthaler, 2019[32]).

Figure 8.3 Making it rain in ECHOES



Source: Ifenthaler (2020[31])

After the readiness assessment, any resulting implementation strategy should encompass monitoring and evaluating the learning analytics with regard to predefined and measurable KPIs (Key Performance Indicators). More importantly, the return on investment, defined as the expected gains (returns) per unit of cost (investment) needs to be monitored closely (Psacharopoulos, 2014[33]). These may be monetary returns, but also other gains, such as student retention, staff improvement (see Case Example 2) and student satisfaction (Gibson et al., 2018[34]).

In sum, a change management strategy could be guided by the following principles (Ifenthaler, 2020[31]; Leitner, Ebner and Ebner, 2019[30]; Tsai and Gašević, 2017[13]):

- definition of the learning analytics vision and objectives (e.g. using the abovementioned benefits matrix) and align them with the organisations mission and learning culture;
- identification of organisational, political, or technological factors that will affect the implementation;
- involvement and continuous information of all stakeholders including students, teachers, administrators, etc.;
- development of (and continuously update) a strategic plan focussing on short-term and long-term wins, including a needs and risk analysis as well as a clear timeline outlining responsibilities of involved stakeholders;
- allocation of resources and identification of expertise (inside and outside of the organisation) for achieving the learning analytics objectives;
- undertaking of a robust formative and summative evaluation of the learning analytics initiative to further refine the overall implementation and organisational change process.

Moving forward with learning analytics

Many learning analytics initiatives either follow an action research-based (Argyris and Schon, 1974[35]) or a design-based research approach (Huang, Spector and Yang, 2019[36]). For example, at the macro-level, learning analytics insights help staff target retention initiatives (see Case Example 3). At the meso-level, data reports enable improved teaching practices (see Case Example 2). At the micro-level, 'at-risk' student lists are provided to student support staff, enabling them to triage and prompt individual students to take corrective action (see Case Example 3) or to cope with student heterogeneity in classrooms (see Case Example 2).

Despite these benefits, implementing learning analytics in educational organisations is often a paradoxical exercise. A research organisation that is contemplating learning analytics may have world-class experts in data science, information systems, management, educational leadership and learning science. These experts may even be contributing to the development of robust learning analytics systems. But that does not necessarily mean they have clear insights into the "political" dimension of what is required to implement learning analytics within their organisation. Or, even if they have expertise in organisation development, perhaps their organisation's administration is not interested in it. Further, as bureaucracy takes over, these experts may not be interested in facilitating the change processes required.

In order to overcome such organisational barriers, Buckingham Shum and McKay (2018[2]) suggest the following: (a) The institution's IT services develops and implements the learning analytics system as it already oversees the learning management system, data warehouse, student information system, etc. In this approach, however, the academic faculty is unlikely to guide an evidence-centred development and implementation of the learning analytics system. (b) Faculty members conduct evidence-centred research and development, and they use their findings to drive the implementation of a learning analytics system. (c) An autonomous entity of the higher education institution – well connected to faculty and administration – drives implementation, requiring partnership among all stakeholders. This innovative approach combines the strengths of the other two mentioned above.

In a systematic review that included over 6 000 studies on learning analytics over the past six years, Ifenthaler and Yau (2020[37]) indicate that greater adoption of organisation-wide learning analytics systems is needed (or at least research about it). Establishing standards may speed organisational take-up. While standards for data models and data collection such as xAPI (Experience API) exist (Kevan and Ryan, 2016[38]), standards for indicators, visualisations, and design guidelines that would make learning analytics pedagogically effective are lacking (Seufert et al., 2019[39]; Yau and Ifenthaler, 2020[25]). This is something learning analytics research and development needs to address. This may be accomplished with the following guidelines (Ifenthaler, Mah and Yau, 2019[40]):

- develop flexible systems that can adapt to the needs of individual organisations, i.e., their learning culture, requirements of specific study programmes, student and teacher preferences, technical and administrative specifications;
- define requirements for data and algorithms;
- involve all stakeholders in developing a learning analytics strategy and implementation;
- establish organisational, technological and pedagogical structures and process for the application of learning analytics systems as well as providing support for all involved stakeholders for a sustainable operation;
- inform all stakeholders with regard to ethical issues and data privacy regulations including professional learning opportunities (e.g. educational data literacy);
- build a robust process for ensuring the validity and veracity of the system, data, algorithms and interventions;
- fund research on learning analytics;
- constitute local, regional and national learning analytics committees including stakeholders from science, economics and politics with a focus on adequate development and implementation (and accreditation) of learning analytics systems.

Conclusion

Learning analytics draw on an eclectic set of methodologies and data to provide summative, real-time and predictive insights for improving learning, teaching, organisational efficiency and decision making (Lockyer, Heathcote and Dawson, 2013[41]; Long and Siemens, 2011[42]). While there is a great amount of attention on the ability of learning analytics to predict possible student failure, this has been for individual isolated courses rather than educational organisations in general (Gašević, Dawson and Siemens, 2015[43]). Additionally, not all learning analytics seem to be effective for learning and teaching, as demonstrated in Dawson et al. (2017[44]). The adoption of learning analytics in educational organisations requires capabilities not yet fully developed (Ifenthaler, 2017[45]). There have not been any wide-scale organisational implementation of learning analytics and therefore no empirical evidence that learning analytics improves the performance of educational organisations. International perspectives on adoption models (Nouri et al., 2019[46]) as well as on policy recommendations (Ifenthaler and Yau, 2019[12]; Tsai et al., 2018[47]) may help push this forward.

There are also important questions: Who owns the data that is available to teachers and learners? What data should be made available, and what data should remain private? Who analyses it and what is the data analysed for? What can teachers do with the data? What feedback and monitoring of learning might students expect from learning analytics? How can techno-led or -enabled assessments be used fairly and what are the risks associated with data use for assessing students' achievements?

While considering the benefits of learning analytics for school and system management, the following developments may be implemented:

- Learning analytics can be used to develop specialised curricula, aligned with job market demand, to better prepare students for their future careers. Examples of job market intelligence for supporting learning at the workplace already exist (Berg, Branka and Kismihók, 2018[48]). This may be only a small step away.
- Learning analytics can facilitate course management and redesign of learning materials for flexible learning opportunities (Gosper and Ifenthaler, 2014[49]). This may also include adaptive and on-demand professional learning for educators.
- Learning analytics applications can support a plethora of administrative tasks (Bowers et al., 2019[50]). Examples include budgeting, purchasing and procurement activities as well as facilities management. In addition, human resources management may better support the demands of individual educators. These systems can support flexible teaching and learning in or out of schools in the most economical way. They can help manage a variety of different curricular specialisations, class sizes, teaching staff, technologies and general demands of the learning environment regarding aspects such as rooms or even furniture.
- Student applications and enrolment processes will best support student needs, and enable the school and teaching staff to micro-adapt the learning environment to students. Adaptive translation applications will support schools in communicating with parents and dealing with sick or absent students.

Policy makers are asked to develop and align policy that will encourage adoption of learning analytics for school and system management. This needs to be based on rigorous research findings, which are currently lacking. Research and development of trusted and effective learning analytics for school and system management should be encouraged.

In sum, policy makers, researchers and practitioners (Ifenthaler et al., In progress[51]) may consider the following strategies and actions:

- *Evidence-based practice led by analytics*: Researchers need to gather more evidence from the use of learning analytics in order to develop systems that positively impact learning. Policy makers can then develop learning analytics policies that focus on leadership, professional learning, enabling mechanisms, and data governance with added confidence.
- *Promote the adoption of learning analytics*: A readiness for cultural change sets the stage for acceptance and adoption, and helps guide the development of standards, principles and procedures by policy makers. These actions also address the challenge of updating principles and policies by engaging the impacted communities in the continual process of adapting and improving the organisational response to change.
- *Inform and guide data services providers and users*: Trustworthy, ethical learning analytics practices are supported by policy mechanisms such as standards, accreditation processes, audits and evidence-based recommendations informed by practice. Researchers play a critical role here in promoting sustainability and scalability of policy and practice, for example, by producing the knowledge needed to effectively embed analytics and provide just-in-time data services that support good decision making focussed on learning. This strategy of wisely balancing investment in data services as well as users supports both the supply and demand sides of the flow of information, which accelerates adoption and positive change.
- *Impact learning via analytics tools*: the priority for learning analytics should be optimising learning to achieve a more equitable and effective educational system and only, secondarily, accountability, testing, organisational change or financial efficiency. All stakeholders, including practitioners, researchers and policy makers, need new levels of data literacy to use this new tool and make use of the information learning analytics reveals.

- *The role of vendors in analytics solutions, adoption and implementation of analytics systems*: There is a growing supply of commercial vendors of these systems for school management. Vendors, such as BrightBytes (https://www.brightbytes.net), develop solutions and show evidence in self-funded studies concerning the benefits of implemented learning analytics systems. Researchers should critically consider these commercial solutions but seek rigorous independent evidence of benefits. Policy makers may take stock of existing commercial solutions and their growing adoption in educational organisations.

References

Agasisti, T. and A. Bowers (2017), "Data analytics and decision-making in education: towards the educational data scientist as a key actor in schools and higher education institutions", in Johnes, G. et al. (eds.), *Handbook of contemporary education economics*, Edward Elgar Publishing, Cheltenham, UK. [19]

Andresen, B. (2017), "Learning analytics for formative purposes", in Tatnall, A. and M. Webb (eds.), *Tomorrow's Learning: Involving Everyone. Learning with and about Technologies and Computing. WCCE 2017. IFIP Advances in Information and Communication Technology*, Springer, Cham. [17]

Argyris, C. and D. Schon (1974), *Theory in practice: Increasing professional effectiveness*, Jossey-Bass, San Francisco, CA. [35]

Berg, A., J. Branka and G. Kismihók (2018), "Combining learning analytics with job market intelligence to support learning at the workplace", in Ifenthaler, D. (ed.), *Digital workplace learning. Bridging formal and informal learning with digital technologies*, Springer, Cham. [48]

Bowers, A., A. Bang, Y. Pan and K. Edward (2019), *Education leadership data analytics (ELDA): A white paper report on the 2018 ELDA Summit*, Teachers College, Columbia University. [50]

Buckingham Shum, S. and T. McKay (2018), "Architecting for learning analytics. Innovating for sustainable impact", *EDUCAUSE Review*, Vol. 53/2, pp. 25-37. [2]

Colvin, C., T. Rogers, A. Wade, S. Dawson, D. Gašević, S. Buckingam Shum and K. Nelson (2015), *Student retention and learning analytics: A snapshot of Australian practices and a framework for advancement*, Australian Government Office for Learning and Teaching, Canberra, ACT. [7]

Dawson, S., J. Jovanovic, D. Gasevic and A. Pardo (2017), "From prediction to impact: Evaluation of a learning analytics retention program", in Molenaar, I., X. Ochoa and S. Dawson (eds.), *Proceedings of the seventh international learning analytics & knowledge conference*, ACM, New York, NY. [44]

Drachsler, H. and W. Greller (2016), *Privacy and analytics - it's a DELICATE issue. A checklist for trusted learning analytics.* [10]

Dyckhoff, A., D. Zielke, M. Bültmann, M. Chatti and U. Schroeder (2012), "Design and implementation of a learning analytics toolkit for teachers", *Educational Technology & Society*, Vol. 15/3, pp. 58-76. [3]

Ferguson, R., D. Clow, L. Macfadyen, A. Essa, S. Dawson and S. Alexander (2014), *Setting learning analytics in context: Overcoming the barriers to large-scale adoption*, http://dx.doi.org/10.1145/2567574.2567592. [27]

Gander, T. (2020), "Learning analytics in secondary schools", in Peters, M. and R. Heraud (eds.), *Encyclopedia of educational innovation*, Springer, Singapore. [18]

Gašević, D., S. Dawson and G. Siemens (2015), "Let's not forget: Learning analytics are about learning", *TechTrends*, Vol. 59/1, pp. 64-71, http://dx.doi.org/10.1007/s11528-014-0822-x. [43]

Gašević, D., Y.S. Tsai, S. Dawson and A. Pardo (2019), "How do we start? An approach to learning analytics adoption in higher education", *International Journal of Information and Educational Technology*, Vol. 36/4, pp. 342-353, http://dx.doi.org/10.1108/IJILT-02-2019-0024. [6]

Gibson, D., S. Huband, D. Ifenthaler and E. Parkin (2018), *Return on investment in higher education retention: Systematic focus on actionable information from data analytics.* [34]

Gosper, M. and D. Ifenthaler (2014), "Curriculum design for the twenty-first Century", in Gosper, M. and D. Ifenthaler (eds.), *Curriculum models for the 21st Century. Using learning technologies in higher education,* Springer, New York, NY. [49]

Heath, J. and E. Leinonen (2016), "An institution wide approach to learning analytics", in Anderson, M. and C. Gavan (eds.), Developing effective educational experiences through learning analytics, IGI Global, Hershey, PA. [26]

Hinkelmann, M. and T. Jordine (2019), "The LAPS project: using machine learning techniques for early student support", in Ifenthaler, D., J. Yau and D. Mah (eds.), *Utilizing learning analytics to support study success, Springer,* Cham. [29]

Huang, R., J. Spector and J. Yang (2019), "Design-based research", in Huang, R., J. Spector and J. Yang (eds.), *Educational technology. A primer for the 21st century,* Springer, Singapore. [36]

Ifenthaler, D. (2020), "Change management for learning analytics", in Pinkwart, N. and S. Liu (eds.), *Artificial intelligence supported educational technologies,* Springer, Cham. [31]

Ifenthaler, D. (2017), "Are Higher Education Institutions Prepared for Learning Analytics?", *TechTrends*, Vol. 61/4, pp. 366-371, http://dx.doi.org/10.1007/s11528-016-0154-0. [45]

Ifenthaler, D. (2015), "Learning analytics", in Spector, J. (ed.), *The SAGE encyclopedia of educational technology,* Sage, Thousand Oaks, CA. [1]

Ifenthaler, D. and D. Gibson (2020), *Adoption of data analytics in higher education learning and teaching*, Springer, Cham. [11]

Ifenthaler, D., D. Gibson, D. Prasse, A. Shimada and M. Yamada (In progress), "Putting learning back into learning analytics: actions for policy makers, researchers, and practitioners", *Educational Technology Research and Development.* [51]

Ifenthaler, D., D. Mah and J. Yau (2019), "Utilising learning analytics for study success. Reflections on current empirical findings", in Ifenthaler, D., J. Yau and D. Mah (eds.), *Utilizing learning analytics to support study success,* Springer, Cham. [40]

Ifenthaler, D. and C. Widanapathirana (2014), "Development and Validation of a Learning Analytics Framework: Two Case Studies Using Support Vector Machines", *Technology, Knowledge and Learning*, Vol. 19/1-2, pp. 221-240, http://dx.doi.org/10.1007/s10758-014-9226-4. [24]

Ifenthaler, D. and J. Yau (2020), "Utilising learning analytics to support study success in higher education: a systematic review", *Educational Technology Research and Development*, Vol. 68/4, pp. 1961-1990, http://dx.doi.org/10.1007/s11423-020-09788-z. [37]

Ifenthaler, D. and J. Yau (2019), "Higher Education Stakeholders' Views on Learning Analytics Policy Recommendations for Supporting Study Success", *International Journal of Learning Analytics and Artificial Intelligence for Education (iJAI),* Vol. 1/1, p. 28, http://dx.doi.org/10.3991/ijai.v1i1.10978. [12]

Ifenthaler, D., J. Yau and D. Mah (2019), *Utilizing learning analytics to support study success,* Springer, New York, NY. [23]

Jimerson, J. and J. Childs (2017), "Signal and symbol: how state and local policies address data-informed practice", *Educational Policy*, Vol. 31/5, pp. 584-614. [21]

Kärner, T., J. Warwas and S. Schumann (2020), "A Learning Analytics Approach to Address Heterogeneity in the Classroom: The Teachers' Diagnostic Support System", *Technology, Knowledge and Learning,* http://dx.doi.org/10.1007/s10758-020-09448-4. [28]

Kevan, J. and P. Ryan (2016), "Experience API: Flexible, Decentralized and Activity-Centric Data Collection", *Technology, Knowledge and Learning*, Vol. 21/1, pp. 143-149, http://dx.doi.org/10.1007/s10758-015-9260-x. [38]

Klasen, D. and D. Ifenthaler (2019), "Implementing learning analytics into existing higher education legacy systems", in Ifenthaler, D., J. Yau and D. Mah (eds.), *Utilizing learning analytics to support study success*, Springer, New York, NY. [4]

Kotter, J. (2007), "Leading change: Why transformation efforts fail", *Havard Business Review,* Vol. January, pp. 96-103. [8]

Leitner, P., M. Ebner and M. Ebner (2019), "Learning analytics challenges to overcome in higher education institutions", in Ifenthaler, D., J. Yau and D. Mah (eds.), *Utilizing learning analytics to support study success,* Springer, Cham. [30]

Lester, J., C. Klein, A. Johri and H. Rangwala (2017), *Learning analytics in higher education*, Wiley, Malden, MA. [14]

Lockyer, L., E. Heathcote and S. Dawson (2013), "Informing Pedagogical Action", *American Behavioral Scientist*, Vol. 57/10, pp. 1439-1459, http://dx.doi.org/10.1177/0002764213479367. [41]

Long, P. and G. Siemens (2011), "Penetrating the fog: Analytics in learning and education", *EDUCAUSE Review*, Vol. 46/5, pp. 31-40. [42]

Nouri, J., M. Ebner, D. Ifenthaler, M. Saqr, J. Malmberg, M. Khalil, J. Bruun, O. Viberg, M.A. Conde González, Z. Papamitsiou and U. Berthlsen (2019), "Efforts in Europe for Data-Driven Improvement of Education – A Review of Learning Analytics Research in Seven Countries", *International Journal of Learning Analytics and Artificial Intelligence for Education (iJAI)*, Vol. 1/1, p. 8, http://dx.doi.org/10.3991/ijai.v1i1.11053. [46]

Pistilli, M. and K. Arnold (2010), "Purdue Signals: Mining real-time academic data to enhance student success", *About campus: Enriching the student learning experience*, Vol. 15/3, pp. 22-24. [5]

Psacharopoulos, G. (2014), "The returns to investment in higher education", in Menon, M., D. Terkla and P. Gibbs (eds.), *Using data to improve higher education. Global perspectives on higher education*, Sense Publishers, Rotterdam. [33]

Rogers, E. (1962), *Diffusion of innovations*, Free Press of Glencoe, New York, NY. [9]

Schumacher, C., D. Klasen and D. Ifenthaler (2019), "Implementation of a learning analytics system in a productive higher education environment", in Khine, M. (ed.), *Emerging trends in learning analytics*, Brill, Leiden, NL. [32]

Sclater, N. and J. Mullan (2017), *Learning analytics and student success – assessing the evidence*, JISC, Bristol. [22]

Sergis, S. and D. Sampson (2016), "School analytics: a framework for supporting school complexity leadership", in Spector, J. et al. (eds.), *Competencies in teaching, learning and educational leadership in the digital age*, Springer, Cham. [20]

Seufert, S., C. Meier, M. Soellner and R. Rietsche (2019), "A pedagogical perspective on big data and learning analytics: a conceptual model for digital learning support", *Technology, Knowledge and Learning*, Vol. 24/4, pp. 599-619, http://dx.doi.org/10.1007/s10758-019-09399-5. [39]

Siemens, G., S. Dawson and G. Lynch (2014), *Improving the quality and productivity of the higher education sector - Policy and strategy for systems-level deployment of learning analytics. Canberra, Australia: Office of Learning and Teaching*, Australian Government, http://solaresearch.org/Policy_Strategy_Analytics.pdf. [15]

Tsai, Y. and D. Gašević (2017), *Learning analytics in higher education - challenges and policies: A review of eight learning analytics policies.* [13]

Tsai, Y., P. Moreno-Marcos, I. Jivet, M. Scheffel, K. Tammets, K. Kollom and D. Gaševic (2018), "The SHEILA Framework: Informing Institutional Strategies and Policy Processes of Learning Analytics", *Journal of Learning Analytics*, Vol. 5/3, http://dx.doi.org/10.18608/jla.2018.53.2. [47]

Viberg, O., M. Hatakka, O. Bälter and A. Mavroudi (2018), "The current landscape of learning analytics in higher education", *Computers in Human Behavior*, Vol. 89, pp. 98-110, http://dx.doi.org/10.1016/j.chb.2018.07.027. [16]

Yau, J. and D. Ifenthaler (2020), "Reflections on Different Learning Analytics Indicators for Supporting Study Success", *International Journal of Learning Analytics and Artificial Intelligence for Education (iJAI)*, Vol. 2/2, p. 4, http://dx.doi.org/10.3991/ijai.v2i2.15639. [25]

9 Early warning systems and indicators of dropping out of upper secondary school: the emerging role of digital technologies

Alex J. Bowers

Teachers College, Columbia University, United States

This chapter discusses the literature and practice on emerging technologies to predict and prevent dropping out of upper secondary school. First, it presents the current research on early warning systems and early warning indicators and discusses the accuracy of predictors of dropout. It shows the value of such research, including a typology of dropout profiles. Second, it provides an overview of current emerging digital methodologies from pattern analytics, data science, big data analytics, learning analytics, and machine learning as applied to identifying accurate predictors of dropping out. The conclusion looks to the future of early warning systems and indicators, both from a research and policy perspective, calling for the need for open access algorithms and code for early warning systems and indicators research and practice, and for the inclusion of the community in their design and application, proposing a framework of "Four A's of Early Warning Indicators" so that they are Accurate, Accessible, Actionable and Accountable.

Introduction

Students failing to graduate from upper secondary school (high school) is an issue globally. While the overall OECD average graduation rate from upper secondary school is 81%, there is large variation across national contexts, with lows at 60% graduation by age 25 in Mexico and highs around 90% in Greece, the Republic of Korea (South Korea), and Slovenia (OECD, 2019[1]). Failing to graduate upper secondary school (known as dropping out) is well-known to be linked with a wide array of student negative life outcomes, such as lower degree attainment, lower lifetime earnings, higher incarceration rates, as well as negative health effects (Belfield and Levin, 2007[2]; Lyche, 2010[3]; Rumberger, 2011[4]; Rumberger et al., 2017[5]). Thus, ensuring that students graduate upper secondary school is a high priority across education systems. This entails predicting early on which students are most likely to experience challenge in the schooling process that may lead to dropping out. With an accurate prediction and early warning, students may be provided additional resources that can help promote their persistence and success (Bowers and Zhou, 2019[6]). The field of education early warning systems and early warning indicators (EWS/EWI) is a recently emerging domain (Allensworth, 2013[7]; Balfanz and Byrnes, 2019[8]; Carl et al., 2013[9]; Davis, Herzog and Legters, 2013[10]; Frazelle and Nagel, 2015[11]; Kemple, Segeritz and Stephenson, 2013[12]; Mac Iver, 2013[13]; McMahon and Sembiante, 2020[14]) that is focused on providing actionable predictors of students failing to graduate from upper secondary school (Allensworth, Nagaoka and Johnson, 2018[15]). However, to date, while industries outside of education continually further leverage the technologies available through the emerging domains of big data, data science, data analytics, and machine learning (Piety, Hickey and Bishop, 2014[16]), the application of these technologies to drop out early warning systems has only recently come to the fore (Agasisti and Bowers, 2017[17]; Baker et al., 2020[18]; Bowers et al., 2019[19]). The purpose of this chapter is to discuss the present state of the field of dropout prediction and early warning systems and indicators. It focuses on the application of current innovations across the technology domains of machine learning, data analytics, and data science pattern analytics to discuss promising developments in this area. As the United States context has been the focus for much of the research in this field at the upper secondary level, I focus on primarily United States studies, informed by studies across other countries where possible. I conclude with a look forward at next-stage developments and an eye to the future.

Early Warning Systems and Indicators (EWS/EWI)

Across the research literature on dropping out, there is a focus on creating what have come to be known as Early Warning Systems (EWS) and Early Warning Indicators (EWI). An EWS is intended to provide actionable predictors of student challenge to help focus the efforts of a school system on specific interventions, systems, and student persistence and achievement (Allensworth, Nagaoka and Johnson, 2018[15]; Balfanz and Byrnes, 2019[8]; Mac Iver and Messel, 2013[20]; Davis, Herzog and Legters, 2013[10]; McMahon and Sembiante, 2020[14]). As a form of personalisation in education using data (Agasisti and Bowers, 2017[17]), an EWS collects EWIs under a single system which is designed to more efficiently allocate the limited resources of schools towards specific areas of challenge for students who are identified as highly likely of dropping out (also known as "at-risk" students) (Carl et al., 2013[9]; Dynarksi et al., 2008[21]; Dynarski and Gleason, 2002[22]; Mac Iver, 2013[13]; Rumberger et al., 2017[5]; Stuit et al., 2016[23]). Taking a positive orientation towards student persistence, much of the research that has come out of the Chicago city school system in the United States, which is a large and diverse urban education system, refers to these indicators as "on-track" indicators for graduation (Allensworth, 2013[7]; Allensworth and Easton, 2005[24]; Allensworth and Easton, 2007[25]; Hartman et al., 2011[26]; Kemple, Segeritz and Stephenson, 2013[12]). Importantly, rather than focus on student context, background, and demographic factors, such as family socio-economic status, which are highly related to student persistence (Rumberger, 2011[4]), these systems focus on predictors and indicators that are malleable so that schools can intervene and provide support to students (McMahon and Sembiante, 2020[14]).

However, much remains to be learned about exactly which indicators are the most accurate and predictive for which systems, how to provide that information for action and evidence-based practice in schools, and then what those schools can do with that information. For example, in a recent study that randomly assigned 73 schools in the United States to use an early warning intervention and monitoring system, after one year, treatment schools saw a decrease in at-risk student chronic absences and course failures but there was no effect on suspensions, low grade point averages, or student credit accrual (Faria et al., 2017[27]). Student credit accrual is seen as a central early outcome of these studies as a direct indicator of continued student persistence and a positive graduation trajectory. Similarly, in a recent randomised controlled experiment of 41 high schools in the United States in which treatment schools used a half-time staff member to monitor ninth grade early warning indicators and provide student supports, treatment schools decreased chronic absenteeism yet had no significant differences on student course failures or credits earned (Mac Iver et al., 2019[28]). One additional difficulty is that relevant early warning indicators may vary across cultures (Box 9.1 and Box 9.2). Thus, while much remains to be learned, the early warning systems and indicators domain is an intriguing and growing research and practice domain that attempts to identify students at risk of dropping out and positively intervene to support student persistence, using the most recent data mining techniques.

Accurate predictors of dropping out

A central concern across the early warning systems and indicators domain is the accuracy of indicators used to predict the probability of a student dropping out. While many studies examine a range of predictors using correlation, logistic regression, or similar types of statistics, and then report which variables are most significant in predicting student graduation or dropping out (Allensworth and Easton, 2007[25]; Balfanz, Herzog and Mac Iver, 2007[33]; Bowers, 2010[34]), recent research has begun to focus on comparing the accuracy of predictors using signal detection theory (Bowers, Sprott and Taff, 2013[35]; Bowers and Zhou, 2019[6]). In signal detection theory (Swets, 1988[36]; Swets, Dawes and Monahan, 2000[37]), all potential predictor variables are compared using a Receiver Operating Characteristic (ROC) plot which compares the sensitivity versus the specificity of a predictor, or the true-positives versus the false-positives, as one wishes for a predictor of an event to identify all of the cases who experience the event (here, dropping out), and not misidentify cases as likely to experience the event who did not (Bowers and Zhou, 2019[6]). For example, a dropout predictor may be very specific, in that all of the students with a combination of factors may drop out, but not sensitive as the ratio of students who have that combination of predictors who eventually drop out from the sample may be very small.

Box 9.1 Early warning system in India

In order to fight the high rate of school dropout, the government of Andhra Pradesh, a south-eastern state in India, deployed an application developed by Microsoft that uses machine learning to predict school dropout (Azure). Based on data such as enrolment, student performance, gender, socio-economic demographics, school infrastructure, and teachers' skills and experience, the software is designed to extract patterns and make predictions. Education officers use the results to guide their interventions and prioritise their investments. According to a government source, Azure identified more than 60 patterns for school dropout in the state. The use of outdated syllabuses for example holds back some students, and students with a lower performance in English or mathematics have a higher chance of dropping out, as they may feel less likely to get a good job or to get into a good college. Another pattern is for example that girls tend not to go to class if the school's toilets do not function. In 2019, it was reported that more than 10 000 schools in Andhra Pradesh were using Azure as an early warning system.

Sources: India AI, 2019[29]; The Wire, 2016[30]; MSV, 2016[31].

Box 9.2 Promoting graduation in vocational education and training (VET) in Helsinki (Finland)

In Finland, around half of the students completing their basic education continue on to vocational education (rather than high school, but can continue on to university after they graduated in VET if they want). Vocational education and training is competence-based and leaves much freedom to students: there are no courses, no classes, no traditional lessons, no school subjects, and no traditional weekly timetables. Some students can graduate in a couple of months, others, in three years. All students have a personal competence development plan to support them to acquire the qualification requirements in their domain. In Helsinki, an AI-based system (AI-HOKS) was developed to support VET students to graduate (and limit their risks of dropping out). Its main goal is to identify as early as possible the circumstances and the phase of learning when students will most likely need support, and to provide automated and semi-automated support (e.g. mobile coaching). The system scaffolds both students and teachers, focusing over time on four main dimensions: learning phase; learning activity; learning and social engagement; and study progress. The indicators for early warning and intervention are based on: 1) personal competence development plans (including implementation time lines, acquired competences and self-evaluation); 2) login and use of various tools and learning environments; 3) weekly mobile questionnaires sent to students' cell phones, and; 4) students' feedback provided through the system. With the aim to provide ethical learning analytics, traditional statistical methods that classify or profile students are not used. Instead, ground data is collected as a basis for building machine-learning models that can be used once the system has been used for a couple of years and larger data sets are available. As of 2021, the system is still being piloted so information about its effectiveness is not available yet.

Source: Interview between Pasi Silander and Stéphan Vincent-Lancrin, 2021[32]

Using these aspects of signal detection theory, Bowers, Sprott and Taff (2013) compared 110 dropout predictors identified from across the literature, and demonstrated that the vast majority of the predictors were no better than a random guess. However, the authors did identify two specific sets of early warning indicators that were more accurate than the others.

The first was the Chicago early warning indicator. Based on over a decade of research from the Chicago urban school system in the United States (Allensworth, 2013[7]; Allensworth and Easton, 2005[24]; Allensworth and Easton, 2007[25]; Allensworth et al., 2014[38]; Allensworth, Nagaoka and Johnson, 2018[15]), the "Chicago on-track" indicator was identified as the most accurate cross-sectional dropout predictor (cross-sectional meaning that the data for the predictor comes from a single year), focusing on the ninth grade and low or failing grades in core subjects such as mathematics or English, and accumulated credits. As a cross-sectional indicator that is very accessible for educators, the Chicago on-track indicator provides a strong means to use an EWI that has higher accuracy than comparable cross-sectional indicators that combine information that currently exists in the school's education data management system. Nevertheless, while the Chicago on-track indicator has higher accuracy than other cross-sectional indicators, the same study by Bowers et al. (2013) demonstrated much higher accuracy dropout predictors that rely on long-term longitudinal data.

In their comparison across 110 predictors of dropping out, the second set of most accurate early warning indicators was generated by one method that stood out as consistently providing the most accurate predictors of dropping out (Bowers, Sprott and Taff, 2013[35]), namely Growth Mixture Modelling (GMM). This technique has the ability to identify significantly different student data patterns over time (see Box 9.3). Three Growth Mixture Modelling analyses were identified for predicting dropping out with extremely high accuracy in comparison to all other predictors (Bowers, Sprott and Taff, 2013[35]). First, analysing a sample of more than 10 000 students from Québec (Canada) from ages 12 through 16, the authors examined student engagement with school over time, such as attendance, discipline, and subject enjoyment and interest, and identified specific low engagement trajectories predictive of dropping out (Janosz et al., 2008[39]). Second, a study using thousands of secondary student mathematics achievement in the United States identified specific growth or decline trajectories that proved predictive of dropout (Muthén, 2004[40]). And third, chief among these studies with high accuracy prediction was a study of trajectories of non-cumulative grade point averages (Bowers and Sprott, 2012[41]).

Box 9.3 Growth Mixture Modelling (GMM)

GMM is a pattern analysis framework that identifies statistically significant different growth or decline trajectories of respondents in a variable over time (Bowers and White, 2014[42]; Martin and von Oertzen, 2015[43]; Masyn, 2011[44]; Muthén, 2004[40]; Ram and Grimm, 2009[45]; Vermunt, Tran and Magidson, 2008[46]). In effect, GMM sits between data mining and inferential statistics as a method within the mixture model class of statistical models that is related to other pattern analytic techniques, such a cluster analysis, while providing a framework to inferentially test the likelihood of inclusion within clusters, including hypothesis testing and inclusion of control variables on the probability of group membership (Martin and von Oertzen, 2015[43]; Vermunt and Magidson, 2002[47]; Vermunt, Tran and Magidson, 2008[46]).

In their study, Bowers and Sprott (2012[41]) examined the non-cumulative grade point averages of over 5 000 students in the United States over the first three semesters of high school of ninth grade semester one, ninth grade semester two, and tenth grade semester one. A non-cumulative grade point average is the average grades for a student in any one grade level averaged across all subjects. The authors argued that this grade data is important to examine as school systems globally assign and collect teacher-assigned grades on a regular basis, yet historically this dataset has been rarely leveraged as important data within some education policy systems (Bowers, 2009[48]; Bowers, 2011[49]; Bowers, 2019[50]; Bowers and Sprott, 2012[41]; Brookhart, 2015[51]). A focus on non-cumulative GPAs is also an improvement over cumulative GPA as the variance in student data year-over-year can then be used to capture different trajectory types (Bowers, 2007[52]).

Bowers and Sprott (2012[41]) identified four significantly different trajectories of non-cumulative GPA: 1) students whose grades decline over this time period; 2) students whose grades start relatively low and rise slowly over time; 3) students who make up the majority and have flat grades through time near the average; and 4), students who have high grades throughout the time period. The main finding was that while the first two groups made up only 25% of the sample, they accounted for over 90% of all of the students who dropped out (Bowers and Sprott, 2012[41]). The majority of the dropouts experienced the low grades and slowly rising trajectory: their grades were going up, but apparently not fast enough. Students with the declining grades trajectory represented a much smaller fraction of dropouts. Together with previous studies using teacher-assigned grades (Allensworth and Luppescu, 2018[53]; Allensworth, Nagaoka and Johnson, 2018[15]; Battin-Pearson et al., 2000[54]; Bowers, 2010[34]; Bowers, 2010[55]; Finn, 1989[56]; Hargis, 1990[57]), these findings confirmed the strong predictive validity of grades on long-term student outcomes such as persistence in school and graduation. In spite of grades having a reputation in some of the psychometrics literature of being less reliable than standardised tests (Brookhart, 2015[51]), the research over the past 100 years has demonstrated that grades measure both academic achievement and a student's ability to negotiate the social processing of school, which is highly related to school engagement, persistence, and later life outcomes (Bowers, 2009[48]; 2011[49]; 2019[50]; Brookhart et al., 2016[58]; Kelly, 2008[59]; Willingham, Pollack and Lewis, 2002[60]).

In a follow-up study, using Latent Class Analysis (LCA) and a large United States nationally generalisable sample of students who drop out of school, Bowers and Sprott (2012[61]) further identified these two types of students who dropped out, and identified the remaining small group that accounts for less than 10% of the dropouts. LCA, like Growth Mixture Modelling, is a form of mixture modelling and allows for the identification of significantly different types of responders across a set of survey items (Collins and Lanza, 2010[62]; Masyn, 2011[44]; Muthén, 2004[40]; Vermunt and Magidson, 2002[47]), identifying a typology of responders. Here, the authors identified a three-group typology of dropouts, which corresponded to the previous dropout typology research and identified the proportions of each type across the United States (Bowers and Sprott, 2012[61]).

First, 38% of the students who dropped out represented the Jaded dropout type: these students correspond to the traditional conception of a dropout as "maladaptive" of the previous research, in that those students do not like school, do not think that teachers are there for them, and have a general jaded conception of the schooling process. These are students with low and declining grades. These students return to schooling the least and have low long-term outcomes.

Second, the Quiet dropout type amounted to about 53% of all dropouts. It corresponded to students with low and slowly rising grades. This was a significant finding as these students correspond to the majority of students who drop out, but are rarely identified by schooling systems as "at risk" because they generally like school, are connected to school, and have grades that are slowly rising – although their grades do not rise fast enough for them to eventually pass all of their classes and graduate.

The third type was the Involved dropout type, who accounted for about 9% of the dropouts. These students were highly involved in school, had generally high grades, and returned the most over time to complete their degrees and go on to post-secondary education. Hypothesised previously as students who are "lost at the last minute" (Menzer and Hampel, 2009[63]), these dropouts, while accounting for the smallest proportion of dropouts, would seem to be the first set of dropouts an intervention would be attempted on, as they persist the longest before dropping out, and often drop out due to either a significant life event (such as pregnancy or family mobility) or to the discovery of a mistake in their transcript and an unexpected need to take an additional course.

As noted across these studies (Bowers and Sprott, 2012[41]; 2012[61]; Bowers, Sprott and Taff, 2013[35]), this typology perspective for dropouts is a significant advance, as previously dropping out of school or being "at risk" was seen as a single monolithic category. The single category perspective for dropping out leads to a naïve view in designing interventions for students who drop out. This may be contributing to the difficulty in finding consistent intervention effects in randomised controlled trials (RCTs) focused on dropouts (Agodini and Dynarksi, 2004[64]; Freeman and Simonsen, 2015[65]). For example, given the findings above, RCTs that attempt to intervene and support student persistence through creating interventions that reconnect students with schooling assuming the traditional Jaded conception of dropping out, focus on only one third of the students who actually drop out, and ignore the vast majority of dropouts who are Quiet and Involved dropouts. Thus, it is not surprising that many experiments have struggled to demonstrate treatment effects, as to date dropout intervention RCTs have lacked a direct focus on creating very different interventions for the three different types of dropouts (McMahon and Sembiante, 2020[14]; Sansone, 2019[66]). For example, the Jaded students may need reconnection with schooling, while the Quiet students may benefit from additional academic instruction and tutoring, and the Involved students counselling on life events, enrolment, and early transcript requirement audits (Bowers and Sprott, 2012[61]).

Many of the studies on early warning systems and indicators estimate and evaluate a set of variables to predict dropping out by viewing the likelihood of students dropping out as a single category. As noted in the dropout predictor accuracy literature (Bowers, Sprott and Taff, 2013[35]) as well as the early warning system literature (McMahon and Sembiante, 2020[14]), it is the attention to the typology theory of dropping out that significantly contributes to the jump in accuracy for the group of Growth Mixture Modelling dropout predictors presented above, identifying more than 90% of the students who drop out. Additionally, perhaps just as important is the longitudinal nature of these accurate predictors that captures significantly different student trajectories over time. A central theory in the dropout literature that is often overlooked in some of the early warning systems and indicators literature (McMahon and Sembiante, 2020[14]) is what has been termed the "life-course perspective" (Alexander, Entwisle and Kabbani, 2001[67]; Dupéré et al., 2018[68]; Dupéré et al., 2015[69]; Finn, 1989[56]; Pallas, 2003[70]). Here, student failure to complete upper secondary school is seen as a long-term process in which, rather than a single event happening in time, multiple stressors accumulate over time and eventually reach a threshold in which the student no longer attends school. Capturing this long-term process then is important for accurately predicting early which students may drop out. This technology of Growth Mixture Models (GMM) provides a useful means to analyse exactly this type of data, relying on long-term over-time data rather than single individual time-point predictors, leading to the increased accuracy of growth mixture model-based predictors (Bowers, Sprott and Taff, 2013[35]).

In a study that used the full state-level dataset of all students over multiple years from the state of Wisconsin in the United States to examine a wide range of available individual variables to use for predicting dropping out, Knowles (2015[71]) used multiple data mining and machine-learning techniques. However, none proved more accurate than the growth mixture model predictors. Indeed, the analysis did neither consider the longitudinal nature of the data nor adopt a typology perspective, which is the core innovation and value added of Growth Mixture Models. Similarly, using the United States nationally generalisable High School Longitudinal Study of 2009 (HSLS:2009) of more than 21 000 high school students, Sansone (2019[66]) focused on single year grade 9 data and applied a similar set of machine learning and logistic regression models to a wide range of student achievement, behaviour, and attitudinal data. The study obtained similar accuracy findings as Knowles (2015[71]) with again none of the models as accurate as the growth mixture model predictors (Sansone, 2019[66]). Interestingly, a recent study that included more than 1 million students from the United States from 1998-2013 in the three states of Massachusetts, North Carolina, and the state of Washington, examined the accuracy of simply using the single time-point mathematics and English reading and writing standardised state test scores from grade 3 or grade 8 to predict dropping out in high school (Goldhaber, Wolff and Daly, 2020[72]). They find that their predictor accuracy was about the same as the Chicago on-track indicator, missing 25% or more of the students who eventually drop out. Looking beyond the United States context, a study that included single year data from hundreds of thousands of students in Guatemala and Honduras (Adelman et al., 2018[73]), as well as a range of variables such as student achievement, student, school and community demographics, obtained predictor accuracy with similar ranges as Knowles (2015[71]) and Sansone (2019[66]) with simple logistic regression methods. This highlights again the strength of the Growth Mixture Modelling method but also that the chosen data analysis method and indicators matter more for the accuracy of the dropout predictors than the size or comprehensiveness of the dataset.

Application of emerging digital technologies: pattern analytics and data science

Together, these findings from the early warning systems and indicators literature indicate a growing usefulness of emerging technologies to identify accurate predictors of dropping out of upper secondary school that focus on groups of patterns of individuals' data over time. While the research domain overall is still relatively new with much work yet to be done, the emerging and growing domains of pattern analytics, data analytics, data science, learning analytics, educational data mining, and machine learning (Koedinger et al., 2015[74]; Piety, 2019[75]) provide an attractive opportunity for researchers, education practitioners, and policymakers. They can take advantage of these new technologies to augment the ability of their current education systems to use the data already collected in schools in new ways that help inform decision making and instructional improvement (Agasisti and

Bowers, 2017[17]; Baker and Inventado, 2014[76]; Bienkowski, Feng and Means, 2012[77]; Bowers, 2017[78]). As noted in this literature, these techniques "make visible data that have heretofore gone unseen, unnoticed, and therefore unactionable" (Bienkowski, Feng and Means, 2012[77]), p.ix). In an effort to inform early warning systems and indicators research and practice, I argue here for increased attention to the opportunities that these techniques can provide to not only increase the accuracy of the indicators in early warning systems, but to increase the usefulness and actionable information provided to educators to take action in informed ways.

Dropout versus discharge

For example, a domain in the dropout research that could be informed through the application of data analytic technology is the issue of dropout versus school discharge (also known as pushout). This domain is one area that has received little attention in the EWS/EWI research, but is a well-known issue in the more general dropout literature. Historically, given the life-course typology perspective noted above, dropping out of school is theorised as a long-term continuous process of multiple accumulating stressors that may affect students in different ways (Alexander, Entwisle and Kabbani, 2001[67]; Dupéré et al., 2018[68]; Dupéré et al., 2015[69]). It is the ability of researchers to draw on this type of theory to match the realities of the dropout process through the analysis techniques noted above that may additionally contribute to the accuracy of predictors that incorporate this information, such as the growth mixture model predictors. However, a much smaller set of research has demonstrated that rather than a singular voluntary process, students who fail to graduate from secondary school may voluntarily leave while others may be involuntarily discharged (Riehl, 1999[79]; Rumberger and Palardy, 2005[80]). Thus, there is some evidence, especially under accountability pressures, that some schools may discharge low performing students through encouraging them to leave the system through a variety of techniques in an effort to increase the average test performance of the school (Rumberger and Palardy, 2005[80]). One process where this takes place is through transferring low performing students so as to avoid the students taking the mandatory tests on which school accountability are based. With enough successive transfers from school to school, the students end up leaving the system (Riehl, 1999[79]). This behaviour by school decision makers is highly problematic, unethical, and does not support student success. It should thus also be detected as early as possible with appropriate follow-on management and policy changes and professional development to refocus the schools on student persistence and supports. Similar types of data analytics as those described above could be applied to data from large schooling systems to identify if this discharge practice is taking place. Pattern analytic techniques such as the longitudinal typology analysis are designed to detect and determine exactly whether students are leaving the system through a process that differs substantively from that of voluntary dropouts. This is especially true when deep sets of longitudinal data across entire systems are available. Specifically, typology analytics such as latent class analysis and Growth Mixture Modelling may be quite useful in identifying these types of school behaviour. Thus, not only are these digital technologies helpful in informing early warning systems and indicators research and practice, but they can be used within and across systems to detect many different types of patterns.

To date, while the research across this domain continues to develop, there are a range of pattern analytic techniques that are being developed to further inform EWS/EWI. This is what the remainder of this section will focus on.

Machine-learning predictors: decision trees and random forests

Recent research has begun to expand the number of techniques available for early warning systems and indicators research and practice as researchers continue to explore the possibilities of using analytics from the data mining, data science, and longitudinal modelling fields (Agasisti and Bowers, 2017[17]; Piety, 2019[75]), including decision trees, longitudinal cluster analysis visualisations, and time-nested survival and hazard models.

First, Classification And Regression Decision Trees (CART) is a family of models that is well-known to be helpful for both researchers and decision makers, as decision trees provide empirically identified cut points, priorities, and weights on variables related to an outcome. As a form of data mining, the output of a decision tree is a figure that partitions and prioritises the most important variables identified by the algorithm in predicting the outcome, gives a cut point on that variable if it is continuous, and then proceeds to branches in the tree, showing the next set of highest priority variables, and so on (Breiman et al., 1993[81]; Quinlan, 1993[82]; Quinlan, 1990[83]). CART analysis has been used in education research and policy, to identify predictors of achievement and test score performance (Koon and Petscher, 2015[84]; Koon, Petscher and Foorman, 2014[85]; Martínez Abad and Chaparro Caso López, 2017[86]), and less often to predict on schooling completion and dropout. For example, Baker et al. (2020) used regression trees to analyse almost 5 000 students' data from Texas in the United States across 23 sets of variable features, confirming much of the literature on the predictors of dropping out as well as identifying additional interesting predictors, such as dress code violations and number of times a student absence was corrected to present (Baker et al., 2020[18]). In another example in dropout prediction, Soland (2013[87]; 2017[88]) used large generalisable datasets in the United States and over 40 variables to identify that college and career aspirations were important variables to include in dropout early warning prediction in addition to grade point average, standardised test scores, and teacher expectations (Soland, 2013[87]; 2017[88]).

Indeed, researchers using machine-learning decision trees have shown strong predictive accuracy for dropout across country contexts, including Mexico (Márquez-Vera, Morales and Soto, 2013[89]), Denmark (Sara et al., 2015[90]) and the Republic of Korea (South Korea) (Chung and Lee, 2019[91]). In Mexico, the study examined data from 670 middle school students from Zacatecas and included more than 70 predictors, finding that regression trees had the highest accuracy of predicting eventual school dropout (Márquez-Vera et al., 2013[92]; Márquez-Vera, Morales and Soto, 2013[89]).

Second, random "forests" is another powerful data analysis technique based on decision tree. As implied by the name, random forests have many regression "trees", estimating an ensemble of many CART analyses, while using a range of random starting parameters, and then testing the "forest" of tree models to find the best solution (Breiman, 2001[93]). In Denmark, using a large sample of over 72 000 students, multiple machine-learning procedures were examined for their accuracy in predicting dropout. The analysis was based on student data, including grades, absences, missing assignments, as well as school size and enrolment and community demographics (Sara et al., 2015[90]). Random forest techniques outperformed CART accuracy. Similarly, using random forest and analysing data from over 165 000 high school students in South Korea, Chung and Lee (2019) included a range of predictor variables such as absences, lateness to school or class, and amount of time on personal or extracurricular activities, clubs and volunteer work, and found strong accuracy in predicting dropout (Chung and Lee, 2019[91]). Using the same dataset, Lee and Chung (2019[91]) noted the extreme imbalance in South Korean dropout data, as the percentage of students who drop out is quite small, which can create problems for many standard regression tree-based machine-learning algorithms. To address this, the authors included a correction for this imbalance issue, demonstrating improvement in accuracy for both regression tree and random forest algorithms (Lee and Chung, 2019[94]).

The use of random forest machine-learning algorithms has also recently been examined with United States data. In a study analysing data from a mid-Atlantic regional school district of about 11 000 students (Aguiar et al., 2015[95]) the authors included a range of predictors such as student achievement, attendance, and mobility. Rather than predict the actual dropout event, the authors predicted students being in the top 10 percentile of their at-risk category, making it difficult to compare the results to other contexts or studies. More recently, a study using random forest machine learning and data from millions of students across the United States (Christie et al., 2019[96]) analysed a model that included over 70 student predictors including student achievement, attendance, behaviour, discipline, and course credits among many others. Unfortunately, while the authors demonstrate high accuracy of their model in predicting dropout, they do not list the 70 predictors, the algorithm, or how the predictors are weighted or used in the model, making it difficult to evaluate or use the results in further research.

Interestingly, across each of these studies using random forest machine learning, overall prediction accuracy was comparable to the growth mixture model predictors noted above, with accuracy over 80%-90%. Nevertheless, a consistent critique of such machine-learning models (Villagrá-Arnedo et al., 2017[97]), such as random forest, is that they are difficult to interpret and act upon, as exactly how the code and algorithm works is not apparent or simple to report, and across so many predictors, knowing how to use the predictor and intervene is not straightforward (Knowles, 2015[71]). While a model may be accurate in predicting dropout, if the algorithm is overly complex, hidden or unreported, it is more difficult to understand and implement models across contexts, and perhaps more importantly, design and test effective interventions.

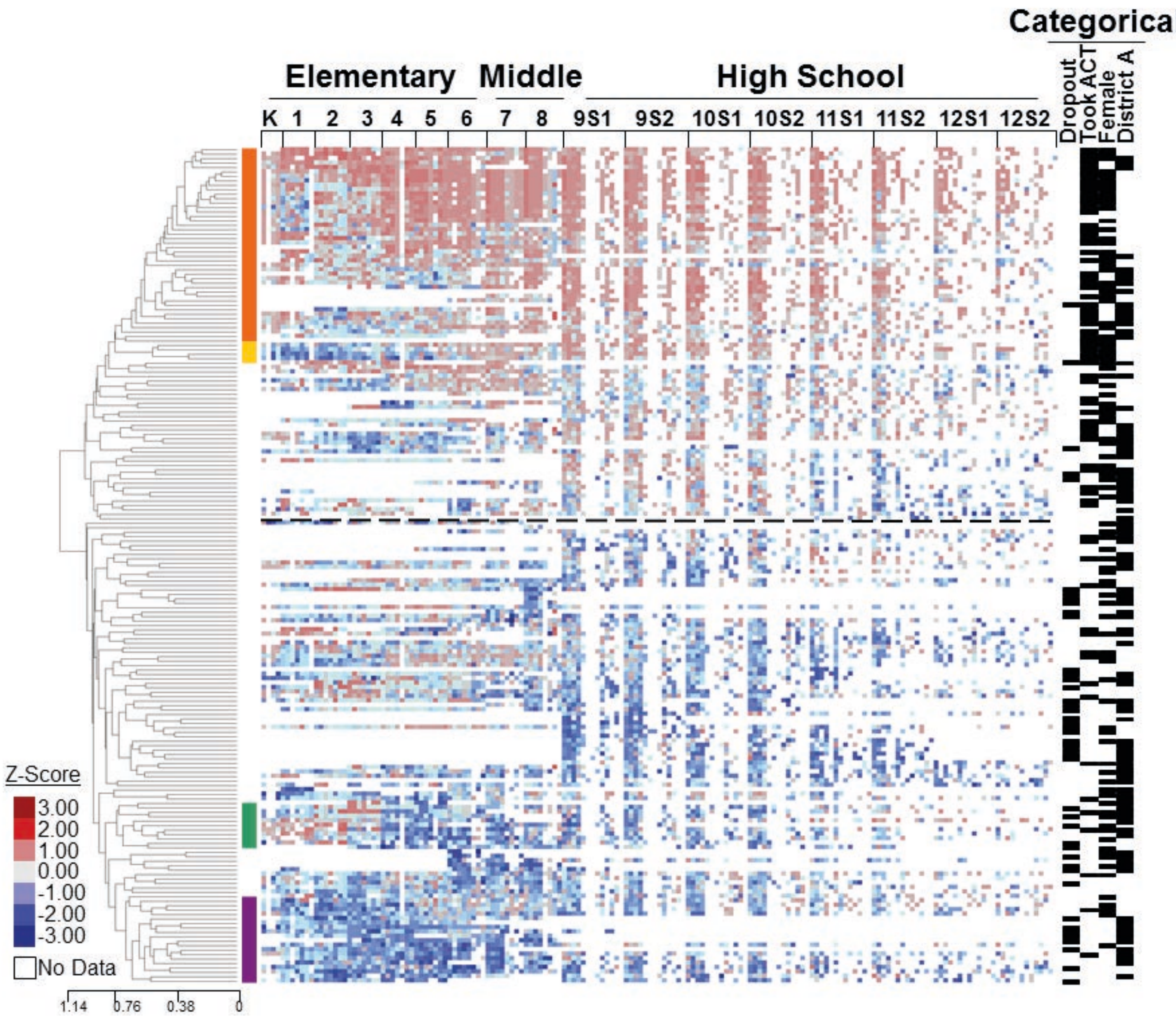
Hierarchical Cluster Analysis (HCA) heat maps

A persistent issue across the dropout EWS/EWI prediction domain is the need to unpack and open algorithms and predictions, examining not only the code but also the patterns identified through the method (Agasisti and Bowers, 2017[17]; Bowers et al., 2019[19]). This allows experts to not only see inside of the prediction algorithm and code to understand how it works, but also to work to not simply summarise students in overall averages, prediction scores, or at-risk status categories (Hawn Nelson et al., 2020[98]), but rather to view the entire life course of the student as an individual through the system that has been provided to them using the data available (Bowers, 2010[34]). One method proposed to address these issues is the application of the visual data analytic technology of hierarchical cluster analysis (HCA) heat maps. Adapted from the big data taxonomy and bioinformatics domains (Bowers, 2010[34]; Eisen et al., 1998[99]; Wilkinson and Friendly, 2009[100]), but used rarely in education research (Kinnebrew, Segedy and Biswas, 2014[101]; Lee et al., 2016[102]; Moyer-Packenham et al., 2015[103]), cluster analysis heat maps visualise clusters of patterns of individual student data and link that information to overall student outcomes, providing a means to examine not only clusters of patterns but the variance across clusters, students, and variables. For example, Figure 9.1. from Bowers (2010[34]) applies HCA heat maps to the entire longitudinal grading history of 188 students from two small school districts in the United States. It patterns the data across all of the students and visualises every individual course grade across all subjects from kindergarten through grade 12, and links them to overall outcomes such as graduation or drop out, as well as taking a college placement exam (in this case: ACT). Importantly for the visualisation, each student's data is represented in a "heat map" in which higher grades in each subject are a hotter red, and lower grades a colder blue. This replaces numbers with blocks of colour.

Hierarchical cluster analysis, relying on a long history of cluster analytic techniques from data mining (Bowers, 2007[52]; Romesburg, 1984[104]; Wilkinson and Friendly, 2009[100]), provides a useful means to provide organisation and pattern analysis to education data. Historically, education data is represented in data files with rows of students organised by either student name or identification number, and then columns of variables. When cluster analysis is applied to this type of data, rather than students being proximal to each other in the order of rows based on name or identification number, the similarity in the pattern of the variables included in the columns is used to reorganise the list, with indications of how similar or dissimilar every student is to their proximal cluster pattern in relation to all other clustered rows of students (Bowers, 2010[34]). The heat map then provides a visual means to display all of the data, and as similar rows are next to each other, blocks of colours form across the dataset, allowing researchers and decision makers to view the entire dataset and all students at the same time, linked to overall outcomes (for a review, please see Bowers, 2010a). In comparison, more traditional ways of viewing data in education, such as bar graphs or line plots of student data, requires a selection of variables to be displayed, and even for just 100 students can become uninterpretable with so many lines or bars. Rather than summarise student data to overall averages across only a small selection of variables, the HCA heat map analysis displays every student in the dataset along with every data point, patterned in a way that provides a means to examine each cluster linked to outcomes.

For the Bowers (2010[34]) study, each student's grades over their entire history in their K-12 education system were patterned and linked to their overall K-12 outcomes. Importantly, specific clusters of student longitudinal grading patterns could be identified that were previously unknown to the schooling system, such as a cluster of students who obtained high grades until grade 3 and 4, and then their grades fell, matching low graded students who more often drop out (see Figure 9.1. , bottom, green annotated cluster). Conversely, one of the smallest clusters of students were a group of students who had low grades for the first few years of schooling, but then rose quickly near the end of elementary school and all graduated (see Figure 9.1. , top, yellow annotated cluster). For both of these examples, all individual students in each cluster can be identified in the visualisation along with all of their actual data. This provides a unique opportunity for decision makers to see the overall scope of data that is important to upper secondary school graduation, patterned over time and across students. In addition to the overall patterns, HCA heat maps allow decision makers to focus on specific students to gain a much deeper understanding of the lived experience of that student through the system provided to them. This technology thus serves to help create more actionable interventions that can potentially be personalised to that student based on how the student has progressed through the system provided to them.

Figure 9.1 Hierarchical cluster analysis of K-12 student subject-specific grades identifies student dropout



Note: Each student in the dataset is represented as a row, columns are every subject in every grade K-12, ordered left to right from core subjects (mathematics, reading/English, science, etc.) to non-core subjects (languages, physical education, etc.). Each Z-scored (by column) grade in each subject for each student is represented as a heat map from cold blue (below the mean grade for that subject at that grade level) to grey (average), to hot red (above the mean). White represents no data. Left is the cluster tree with smaller horizontal lines representing higher similarity between rows. The annotation on the right represents dichotomous outcome or demographic variables. Note the dataset patterns into two large clusters (above and below the dotted horizontal line): the top generally high achieving students (red) do not dropout and do take the college entrance ACT exam, while the bottom lower achieving students (blue) drop out much more often and do not take the college entrance exam. Colour block annotations on the left indicate specific clusters of interest in which patterns emerge early in elementary school that are predictive of overall high school completion. See Bowers (2010[34]) for more information.

Source: Bowers (2010[34])

Issues of machine learning and data mining: generalisability, open access, and accuracy

Across these types of data mining and visual data analytic techniques, a central limitation is that the models are trained on specific school, district or state data but designed to provide actionable information about the broader education system. These data mining models historically are limited in their ability to generalise to a larger population beyond the training dataset. Indeed, recent early warning systems and indicators research across three school districts in the state of Ohio in the United States demonstrated different results across each district (Stuit et al., 2016[23]). This prompted the authors to recommend that researchers and practitioners validate findings from studies from outside of their system using the same variables but including the local system data, and then compare the findings. As the authors note:

Given the variability across school districts… in consistency of predicting failure to graduate on time, and in relative accuracy of indicators to predict failure to graduate on time, the findings suggest that it is important for school districts to examine and analyze their own student-level data in order to develop their own early warning system. (p.ii)

Similarly, at a larger scale, in a recent study that included student academic performance, attendance, and behaviour data from over 1 million students in the United States to predict dropout, when the data mining prediction algorithm was applied to 30 target districts outside of the dataset "none of the… models did as well as their average when applied to the target districts" (p.735) (Coleman, Baker and Stephenson, 2019[105]). Thus, a current central limitation of data mining techniques in dropout prediction is generalisability beyond the training dataset, yet this work is critically important for the school systems included in the training data. For school and system leaders, education data science techniques, such as data mining, pattern analysis, data visualisation, and prediction accuracy, are important recent innovations that can help support their decisions throughout the education system for actionable data use (Agasisti and Bowers, 2017[17]; Bowers, 2010[55]; Krumm, Means and Bienkowski, 2018[106]; Bowers, 2017[78]; Piety, 2019[75]). As noted in the recent literature on the intersection of data science, evidence-based improvement cycles, and school leadership, termed Education Leadership Data Analytics (ELDA) (Bowers et al., 2019[19]), this work includes:

ELDA practitioners working collaboratively with schooling system leaders and teachers to analyze, pattern, and visualize previously unknown patterns and information from the vast sets of data collected by schooling organisations, and then integrate findings in easy to understand language and digital tools into collaborative and community building evidence-based improvement cycles with stakeholders. (p.8).

Thus, this pattern analytic work within each school district and system is important to help inform decision making (Bowers, 2017[78]; Mandinach and Schildkamp, 2020[107]).

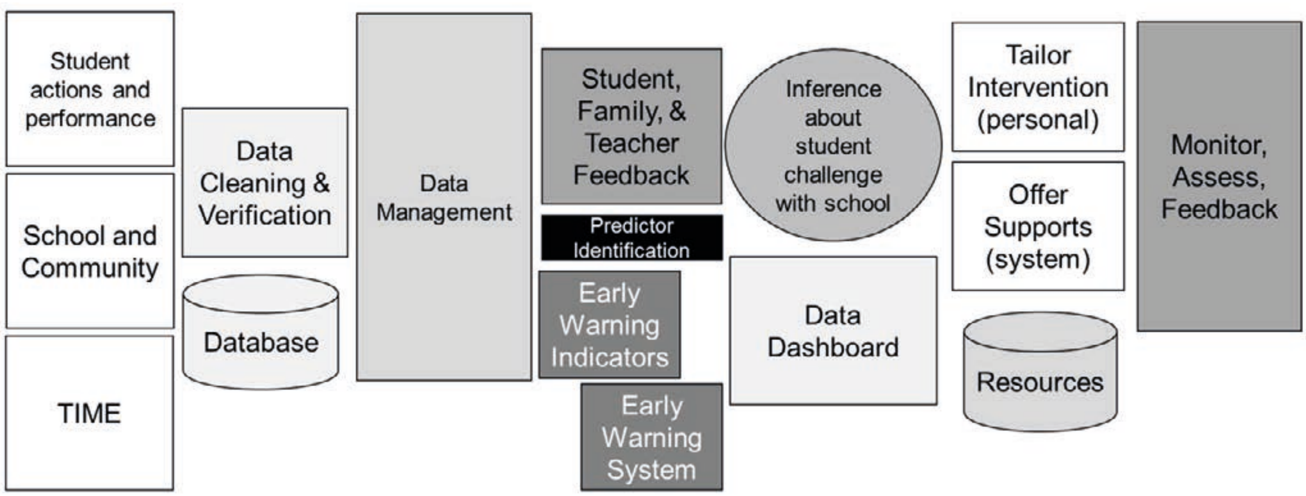
Nevertheless, this type of local replication and comparison of models can only take place if the methods and algorithms used are available publicly and open access. Following similar calls across multiple domains in research, health, industry, and government in which large datasets are analysed in similar ways to make recommendations (Stodden et al., 2016[108]; Wachter and Mittelstadt, 2019[109]), there are recent calls for transparency and open access publication of all algorithms in education. While education data must be private and confidential, if an algorithm makes a recommendation, prediction, or a decision for students, teachers, or schools, it is ethical that the code and algorithm be posted public and open access, free of hidden proprietary features (Agasisti and Bowers, 2017[17]; Bowers et al., 2019[19]). This open publication and access to algorithms in education decision making helps prevent issues noted in other sociological data science and data mining domains, such as bank loans and incarceration recommendations. Hidden algorithms and unintended consequences lead to bias and inequities in the algorithms and the outcomes they generate that go unidentified (Benjamin, 2019[110]; Hawn Nelson et al., 2020[98]; O'Neil, 2016[111]), whereas open algorithms with the appropriate data can be tested for bias and fairness (Corbett-Davies and Goel, 2018[112]; d'Alessandro, O'Neil and LaGatta, 2017[113]; Dudik et al., n.d.[114]; Loukina, Madnani and Zechner, 2019[115]; Zehlike et al., 2017[116]).

Recently, to help local school system practitioners assess the accuracy of their early warning systems and indicators, Bowers and Zhou (2019[6]) provided a guide for Receiver Operating Characteristic analysis. As noted above, Receiver Operating Characteristic analysis allows for the comparison of the accuracy of different predictors on an outcome, with the technique of receiver operating characteristic area under the curve (ROC AUC) providing a continuous measurement comparison of accuracy along with a statistical test of difference in the accuracy between two predictors (Bowers and Zhou, 2019[6]). However, as data practitioners in schooling systems come to their jobs through many different and idiosyncratic career paths (Bowers, 2017[78]; Bowers et al., 2019[19]), it is rare that they have received training in how to compare the accuracy of different predictors of educational outcomes. Additionally, in many instances, busy education data practitioners working on early warning systems and indicators may not have the time or training to generate the code for this type of EWI accuracy assessment and comparison. This issue is made more difficult in that few research studies in education currently publish their full algorithm and code to allow for replication and practitioner use (Bowers et al., 2019[19]; Knowles, 2015[71]). To address this issue, Bowers and Zhou (2019[6]) used large and comprehensive United States nationally generalisable datasets that are open and public to provide a guide and walkthrough for education data practitioners wishing to apply receiver operating characteristic area under the curve (ROC AUC) analysis of their EWIs within their EWS to assess the accuracy of each indicator and statistically compare significantly different levels of accuracy in prediction of education outcomes. Importantly, the study provides not only a guide demonstrating receiver operating characteristic area under the curve (ROC AUC) indicator accuracy analysis across a wide range of education outcomes, from dropout and upper secondary school completion, to post-secondary enrolment and completion, among others (Bowers and Zhou, 2019[6]), but also supplementary materials that include all of the code in the open source R statistical language for each table, equation, and figure in the study. This type of study helps to promote the sharing of algorithmic resources across contexts, testing of the code comparisons of results and application to local educational communities for decision making (Bowers et al., 2019[19]).

Conclusion and future directions

Across the early warning systems and indicators research and practice in reducing dropouts, emerging technologies from pattern analytics, data mining, learning analytics, and machine learning are providing an opportunity to expand what is known about how to accurately predict student outcomes and then provide interventions that help support student success. However, predictor identification and accuracy analysis is only one small part of a much larger system.

Figure 9.2 Only a small fraction of a schooling organisation's early warning systems are composed of predictor identification



Source: Alex J. Bowers.

While there is much recent research and interest in student outcome predictor identification and accuracy analysis in early warning indicators and systems, identifying predictors is only one small part of a much larger system. Adapting the central figure from a paper by engineers at Google Inc. in which they note that machine learning is only a very small piece in which «the required surrounding infrastructure is vast and complex» (p.4) (Sculley et al., 2015[117]), Figure 9.2. places the issues of "predictor identification" in the larger context and systems of preventing students' dropout. Multiple different pieces work together throughout the process, starting on the left of Figure 9.2. with collecting, cleaning, and managing a wide range of student and school data over time in collaboration with community members; providing that information to stakeholders through the early warning system and dashboard; and combining it with input from students, teachers, and family members. This then motivates inferences about student challenges and success in the system, which, when combined with appropriate and available resources, can be used to tailor interventions to student needs or modify current supports offered to all through the education system. Ultimately, as noted on the right of Figure 9.2. information can be gathered to feed back into the system and help support continuous improvement.

While my focus throughout this discussion has been on the specifics of early warning indicators and their accuracy, the point that this discussion is only one small part of the vast system in education organisations that can be leveraged to help support student success is exemplified in the work in the Chicago context in the United States (Allensworth, Nagaoka and Johnson, 2018[15]). As noted above, the Chicago on-track indicator is a well-known and fairly accurate cross-sectional early warning indicator of students dropping out of school. Over the last two decades, the city of Chicago has seen a dramatic increase in graduation rates (Allensworth et al., 2016[118]), from 52.4%

in 1998 to above 90% by 2019 (Issa, 2019[119]). However, as stated by researchers in Chicago, finding more accurate on-track indicators and providing these in an early warning system did not cause improvement, as this is a necessary but insufficient issue to address dropping out of school. Rather, the early warning system is a small piece of a much larger suite of systems combined with educator action that provides useful data to educators who then tailor interventions for students and create or modify current systems to help support student persistence. As noted by Allensworth (2013[7]):

> *Ten years ago, addressing high school dropout rates seemed like an intractable problem... Those were the days before wide access to student-level data systems. Now that educators can track students' progress through school, they have early warning indicators that are available, highly predictive of when students begin high school, and readily available to high school practitioners. Furthermore, the factors that are most directly tied to eventual graduation are also the factors that are most malleable through school practices—student attendance and effort in their courses. Not only can students be identified for intervention and support, but schools can use patterns in the indicators to address structural issues that make it more difficult for students to graduate (p.68-69) (Allensworth, 2013[7]).*

The four A's of early warning indicators: Accurate, Accessible, Actionable and Accountable.

In order to be a core component of an effective strategy supporting student persistence and graduation, early warning systems must have indicators that have four central tenets. Here I propose these tenets as "The Four A's of Early Warning Indicators", claiming that outcome predictors must be: Accurate, Accessible, Actionable and Accountable:

- *Accurate* in that the predictor actually identifies the outcome at some earlier time point, which is most easily assessed using accuracy metrics such as the receiver operating characteristic area under the curve (ROC AUC) as discussed above.
- *Accessible* in that the predictor is easy to understand and open to investigation. Accessible does not mean simple, but rather that the algorithm can be accessed, examined, and understood. Accessible is the opposite of proprietary, hidden, or machine learned algorithms that obfuscate how the prediction takes place, but instead is open, public, and understandable.
- *Actionable* in that the predictor can be used to take action to help tailor interventions, or modify the current system and organisation to address systemic issues. Actionable early warning indicators rely on predictors that are recent or real-time, malleable, and under the influence of stakeholders, in opposition to predictors that students, teachers, administrators, and family and community members have no control over.
- *Accountable* in that predictors are regularly checked for bias, are held up to and inspected by the communities for which they are predicting, and regularly audited and critiqued to examine the extent of algorithmic bias and promote fairness. Accountable early warning indicators include the community in the design and application of the predictors to issues of concern to the community, designing and using predictors in collaboration with the communities for which the system is designed to serve.

Of the four A's, accountability may be the most important aspect to consider for early warning systems and indicators. When Accountability is addressed well, the other three A's of Accurate, Accessible and Actionable become clearer. Indeed, accountability in algorithmic prediction is a growing area of concern globally. In discussing the legal issues around data privacy and algorithmic prediction in the European Union and the United States, Wachter and Mittelstadt (2019[109]) note:

> *Unfortunately, there is little reason to assume that organizations will voluntarily offer full explanations covering the process, justification for, and accuracy of algorithmic decision making unless obliged to do so. These systems are often highly complex, involve (sensitive) personal data, and use methods and models considered to be trade secrets... An explanation might inform the individual about the outcome or decision and about underlying assumptions, predictions, or inferences that led to it. It would not, however, ensure that the decision, assumption, prediction, or inference is justified. In short, explanations of a decision do not equal justification of an inference or decision. Therefore, if the justification of algorithmic decisions is at the heart of calls for algorithmic accountability and explainability... Individual-level rights are required that would grant data subjects the ability to manage how privacy-invasive inferences are drawn, and to seek redress against unreasonable inferences when they are created or used to make important decisions. (p.503-505) (Wachter and Mittelstadt, 2019[109]).*

This issue is made more problematic with the known racial, ethnic, and community bias of recent algorithmic prediction and recommendation systems in health care, finance, policing, and incarceration. As noted recently by Benjamin (2019[110]) writing in the journal Science:

> *Data used to train automated systems are typically historic and, in the context of health care, this history entails segregated hospital facilities, racist medical curricula, and unequal insurance structures, among other factors. Yet many industries and organizations well beyond health care are incorporating automated tools, from education and banking to policing and housing, with the promise that algorithmic decisions are less biased than their human counterpart. But human decisions comprise the data and shape the design of algorithms, now hidden by the promise of neutrality and with the power to unjustly discriminate at a much larger scale than biased individuals (p.422) (Benjamin, 2019[110]).*

Thus, accountability to the community for which it serves is a core issue for education early warning systems. This has resulted in recent calls in education data use, and predictive systems for an expanded and central role of the community in the planning, design, testing, and use of these data systems (Bowers et al., 2019[19]; Hawn Nelson et al., 2020[98]; Mandinach and Schildkamp, 2020[107]). These recommendations encourage the community to be an equal participant in the evidence-use cycle in collaboration with researchers, teachers, and administrators, helping to inform how early warning systems and indicators are designed, what the inferences about the outcomes are, and how those inferences will be used in positive and supportive ways, with specific and actionable steps (Hawn Nelson et al., 2020[98]).

Future directions for early warning indicator research and practice:

Given the current state of the research, three areas of interest to advance early warning indicator research and practice can be highlighted:

First, an issue across the early warning indicator research domain is replication of accuracy results, algorithms, and testing of new predictors across multiple contexts and datasets. Each study that identifies an early warning indicator typically analyses the data, notes the accuracy metrics (or not), and then encourages application of the identified indicator to practice. Yet, especially for machine-learning algorithms, but also for all indicators, there is a constant need to replicate each indicator across contexts, confirm accuracy and test for bias to provide a baseline comparison, and then innovate on what is already known. Thus, more studies of early warning indicators should replicate the previously known most accurate indicators for an outcome with the new dataset, following recent examples (Bowers, Sprott and Taff, 2013[35]; Coleman, Baker and Stephenson, 2019[105]; Knowles, 2015[71]), report the receiver operating characteristic area under the curve (ROC AUC) numbers, then compare those numbers to any new analytics, and then publish the code open access (Agasisti and Bowers, 2017[17]; Bowers et al., 2019[19]). To help encourage this type of code sharing and replication across datasets, implementing "FAIR" data standards in this domain would dramatically spur innovation by encouraging de-identified datasets and algorithms that are Findable, Accessible, Interoperable, and Reproducible (Austin et al., 2017[120]; Singh et al., 2019[121]). For detailed summary of FAIR data standards, see Austin et al (2017[120]).

Second, the vast majority of EWIs within most EWSs are cross-sectional single time-point variables, many of which are collected in lower secondary school. However, as shown above, the most accurate predictors use longitudinal data, examining the trajectories of students over time. Thus, a future area in this domain is to include ever more longitudinal and time-nested data that use long-term data starting much earlier than secondary school. However, time as a variable can create many problems for both traditional statistics and data mining, as a student's data over time is dependent on earlier and later time points, violating a central assumption of regression-based statistics. Additionally, dropout data has another time-dependent issue in that, as time continues forward, the sample dataset dynamically changes as students drop out and "leave" the dataset. This time-nested longitudinal data dependency at both the individual level and the sample level can present difficulties to researchers and practitioners looking to apply methods from the cross-sectional literature. While the dropout research to date has rarely confronted this issue, there are a few studies that have worked to apply models drawn from epidemiology, in which these longitudinal time-dependent conditional data issues are ever present (Bowers, 2010[55]; Lamote et al., 2013[122]). Interestingly, these issues can be modelled quite well using survival modelling techniques, specifically discrete-time hazard models (Singer and Willett, 2003[123]). As Bowers (2010[55]) shows, the hazard of dropping out is time dependent, and estimating that hazard depends on the sample of remaining students. When this is taken into account, multiple predictors of interest can then be tested for when they exert the largest effects on the hazard of dropping out at specific time points. This focus on hazard of dropping out at different time points can potentially inform interventions, as one intervention that might work early in secondary school may have no effect later, as well as the opposite.

Third, a well-known issue with dropping out in the United States specifically is what has been termed "dropout factories" (Balfanz et al., 2010[124]; Balfanz and Legters, 2006[125]; Balfanz and West, 2009[126]). This issue acknowledges that across a country there may be specific schools that account for a large proportion of students who drop out. In the United States, some schools graduate only 50% of their students. These "dropout factories" indicate that dropping out is to some extent located at the school and schooling system level of analysis. This implies a multilevel modelling framework of appropriately nesting students within schools (Hox, 2010[127]; Raudenbush and Bryk, 2002[128]) and then estimating the variance in dropping out that is attributable to the student level and school level (Lamote et al., 2013[122]), an issue implied in the dropout versus discharge example above (Rumberger and Palardy, 2005[80]). For early warning systems and indicators research, a productive future area of research is using these hierarchical modelling frameworks in conjunction with the pattern and predictive analyses described above to provide actionable information. For example, multilevel Growth Mixture Modelling or multilevel latent class analysis may show which schools have different proportions of students from different groups within a dropout typology. In an example of the importance of taking into account the dependent nature of the time-nested data, (Lamote et al., 2013[122]) used a multilevel discrete-time hazard model and demonstrated the importance of taking longitudinal student mobility across schools into account in the EWI literature. Thus, the school level is an important variable to include in EWS/EWI predictor research. If specific types of dropping out are co-located in a specific school building, then the school management and policy can focus on the specific organisation of this building to help improve student outcomes for their community.

In conclusion, there have been many recent advances in education early warning systems and indicators. Thanks to the use of data analytics techniques, some early warning systems predict dropout with an accuracy above 80-90% of cases. However, there are yet also specific domains in need of further research and application to practice. For example, most of those techniques relying on machine learning and data mining do not easily transfer from one educational context to another. More worrisome, many early warning indicators are still grossly inaccurate.

The four A's of early warning indicators of Accurate, Accessible, Actionable and Accountable provide a useful framework for advancing the field of predictive analytics and algorithms for helping support education outcomes. In the end, the technology of early warning systems and indicators is only a small piece of the much larger system of data use in schools, which includes a strong role and voice of the community, ethical use of data and algorithms throughout the process, and a continual focus on how to support student success through providing individual interventions and addressing system-level offerings and policies. Developing techniques and tools that make data actionable is just one of the steps towards effective actions supporting learning improvement and student success, but a promising one that digitalisation and innovations in data mining and data analytics should make sustainable in the near future.

References

Adelman, M., F. Haimovich, A. Ham and E. Vazquez (2018), "Predicting school dropout with administrative data: new evidence from Guatemala and Honduras", *Education Economics*, Vol. 26/4, pp. 356-372, http://dx.doi.org/10.1080/09645292.2018.1433127. [73]

Agasisti, T. and A. Bowers (2017), "Data Analytics and Decision-Making in Education: Towards the Educational Data Scientist as a Key Actor in Schools and Higher Education Institutions", in Johnes, G. et al. (eds.), *Handbook on the Economics of Education*, Edward Elgar Publishing, Cheltenham, UK, https://doi.org/10.7916/D8PR95T2. [17]

Agodini, R. and M. Dynarksi (2004), "Are experiments the only option? A look at dropout prevention programs", *The Review of Economics and Statistics*, Vol. 86/1, pp. 180-194. [64]

Aguiar, E., H. Lakkaraju, N. Bhanpuri, D. Miller, B. Yuhas and K. Addison (2015), *Who, when, and why: a machine learning approach to prioritizing students at risk of not graduating high school on time*, ACM, Poughkeepsie, New York, http://dx.doi.org/10.1145/2723576.2723619. [95]

Alexander, K., D. Entwisle and N. Kabbani (2001), "The dropout process in life course perspective: Early risk factors at home and school", *The Teachers College Record*, Vol. 103/5, pp. 760-822. [67]

Allensworth, E. (2013), "The Use of Ninth-Grade Early Warning Indicators to Improve Chicago Schools", *Journal of Education for Students Placed at Risk (JESPAR)*, Vol. 18/1, pp. 68-83, http://dx.doi.org/10.1080/10824669.2013.745181. [7]

Allensworth, E. and J. Easton (2007), *What matters for staying on-track and graduating in Chicago public high schools: A close look at course grades, failures, and attendance in the freshman year*, The University of Chicago, http://www.consortium-chicago.org. [25]

Allensworth, E. and J. Easton (2005), *The on-track indicator as a predictor of High School graduation*, http://www.consortium-chicago.org/publications/p78.html. [24]

Allensworth, E., J. Gwynne, M. de la Torre and P. Moore (2014), *Looking Forward to High School and College: Middle Grade Indicators of Readiness in Chicago Public Schools*, http://ccsr.uchicago.edu/sites/default/files/publications/Middle%20Grades%20Report.pdf. [38]

Allensworth, E., J. Gwynne, K. Healey and R. Crespin (2016), *High school graduation rates through two decades of district change: The influence of policies, data records, and demographic shifts.*, http://consortium.uchicago.edu/sites/default/files/publications/High%20School%20Graduation%20Rates-Jun2016-Consortium.pdf. [118]

Allensworth, E. and S. Luppescu (2018), *Why do students get good grades, or bad ones? The influence of the teacher, class, school, and student*, https://consortium.uchicago.edu/sites/default/files/publications/Why%20Do%20Students%20Get-Apr2018-Consortium.pdf. [53]

Allensworth, E., J. Nagaoka and D. Johnson (2018), *High School Graduation and College Readiness Indicator Systems: What We Know, What We Need to Know*, https://consortium.uchicago.edu/sites/default/files/publications/High%20School%20Graduation%20and%20College-April2018-Consortium.pdf. [15]

Austin, C., T. Bloom, S. Dallmeier-Tiessen, V. Khodiyar, F. Murphy, A. Nurnberger, L. Raymond, M. Stockhause, J. Tedds, M. Vardigan and A. Whyte (2017), "Key components of data publishing: using current best practices to develop a reference model for data publishing", *International Journal on Digital Libraries*, Vol. 18/2, pp. 77-92, http://dx.doi.org/10.1007/s00799-016-0178-2. [120]

Baker, R., A. Berning, S. Gowda, S. Zhang and A. Hawn (2020), "Predicting K-12 Dropout", *Journal of Education for Students Placed at Risk (JESPAR)*, Vol. 25/1, pp. 28-54, http://dx.doi.org/10.1080/10824669.2019.1670065. [18]

Balfanz, R., J. Bridgeland, L. Moore and J. Horning Fox (2010), *Building a grad nation: Progress and challenge in ending the high school dropout epidemic.*, https://www.americaspromise.org/resource/building-grad-nation-progress-challenge-ending-high-school-dropout-epidemic-november-2010 (accessed on 18 January 2021). [124]

Balfanz, R. and V. Byrnes (2019), "Early Warning Indicators and Intervention Systems: State of the Field", in Fredricks, J., A. Reschly and S. Christenson (eds.), *Handbook of Student Engagement Interventions, Elsevier*, http://dx.doi.org/10.1016/b978-0-12-813413-9.00004-8. [8]

Balfanz, R., L. Herzog and D. Mac Iver (2007), "Preventing Student Disengagement and Keeping Students on the Graduation Path in Urban Middle-Grades Schools: Early Identification and Effective Interventions", *Educational Psychologist*, Vol. 42/4, pp. 223-235, http://dx.doi.org/10.1080/00461520701621079. [33]

Balfanz, R. and N. Legters (2006), "Closing 'dropout factories': The graduation-rate crisis we know, and what can be done about it.", *Education Week*, Vol. 25, pp. 42-43. [125]

Balfanz, R. and T. West (2009), *Raising graduation rates: A series of data briefs: Progress toward increasing national and state graduation rates.*, https://new.every1graduates.org/raising-graduation-rates/ (accessed on 18 January 2021). [126]

Battin-Pearson, S., M. Newcomb, R. Abbott, K. Hill, R. Catalano and J. Hawkin (2000), "Predictors of early high school dropout: A test of five theories.", *Journal of Educational Psychology*, Vol. 92/3, pp. 568-582, http://dx.doi.org/10.1037/0022-0663.92.3.568. [54]

Belfield, C. and H. Levin (2007), "The education attainment gap: Who's affected, how much, and why it matters", in Belfield, C. and H. Levin (eds.), *The price we pay: Economic and social consequences of inadequate education*, Brookings Institution Press, Washington D.C. [2]

Benjamin, R. (2019), "Assessing risk, automating racism", *Science*, Vol. 366/6464, pp. 421-422, http://dx.doi.org/10.1126/science.aaz3873. [110]

Bienkowski, M., M. Feng and B. Means (2012), *Enhancing Teaching and Learning Through Educational Data Mining and Learning Analytics: An Issue Brief*, http://www.ed.gov/edblogs/technology/files/2012/03/edm-la-brief.pdf. [77]

Bowers, A. (2019), "Towards Measures of Different and Useful Aspects of Schooling: Why Schools Need Both Teacher Assigned Grades and Standardized Assessments", in Brookhart, S. and J. McMillan (eds.), *Classroom Assessment as Educational Measurement, National Council on Measurement in Education (NCME) Book Series*, Routledge, New York. [50]

Bowers, A. (2017), "Quantitative Research Methods Training in Education Leadership and Administration Preparation Programs as Disciplined Inquiry for Building School Improvement Capacity", *Journal of Research on Leadership Education*, Vol. 12/1, pp. 72-96, http://dx.doi.org/10.1177/1942775116659462. [78]

Bowers, A. (2011), "What's in a grade? The multidimensional nature of what teacher-assigned grades assess in high school", *Educational Research and Evaluation*, Vol. 17/3, pp. 141-159, http://dx.doi.org/10.1080/13803611.2011.597112. [49]

Bowers, A. (2010), "Analyzing the longitudinal K-12 grading histories of entire cohorts of students: Grades, data driven decision making, dropping out and hierarchical cluster analysis", *Practical Assessment Research and Evaluation*, Vol. 15/7, pp. 1-18, http://pareonline.net/pdf/v15n7.pdf. [34]

Bowers, A. (2010), "Grades and Graduation: A Longitudinal Risk Perspective to Identify Student Dropouts", *The Journal of Educational Research*, Vol. 103/3, pp. 191-207, http://dx.doi.org/10.1080/00220670903382970. [55]

Bowers, A. (2009), "Reconsidering grades as data for decision making: more than just academic knowledge", *Journal of Educational Administration*, Vol. 47/5, pp. 609-629, http://dx.doi.org/10.1108/09578230910981080. [48]

Bowers, A. (2007), *Grades and data driven decision making: Issues of variance and student patterns*, Michigan State University, East Lansing, http://files.eric.ed.gov/fulltext/ED538574.pdf. [52]

Bowers, A., A. Bang, Y. Pan and K. Grave (2019), *Education leadership data analytics (ELDA): A white paper report on the 2018 ELDA Summit*, Teachers College, Columbia University, https://doi.org/10.7916/d8-31a0-pt97. [19]

Bowers, A. and R. Sprott (2012), "Examining the Multiple Trajectories Associated with Dropping Out of High School: A Growth Mixture Model Analysis", *The Journal of Educational Research*, Vol. 105/3, pp. 176-195, http://dx.doi.org/10.1080/00220671.2011.552075. [41]

Bowers, A. and R. Sprott (2012), "Why Tenth Graders Fail to Finish High School: A Dropout Typology Latent Class Analysis", *Journal of Education for Students Placed at Risk (JESPAR)*, Vol. 17/3, pp. 129-148, http://dx.doi.org/10.1080/10824669.2012.692071. [61]

Bowers, A., R. Sprott and S. Taff (2013), "Do we know who will drop out? A review of the predictors of dropping out of high school: Precision, sensitivity and specificity", *The High School Journal*, Vol. 96/2, pp. 77-100, http://muse.jhu.edu/journals/high_school_journal/v096/96.2.bowers.html. [35]

Bowers, A. and B. White (2014), "Do Principal Preparation and Teacher Qualifications Influence Different Types of School Growth Trajectories in Illinois? A Growth Mixture Model Analysis.", *Journal of Educational Administration*, Vol. 52/5, pp. 705-736. [42]

Bowers, A. and X. Zhou (2019), "Receiver Operating Characteristic (ROC) Area Under the Curve (AUC): A Diagnostic Measure for Evaluating the Accuracy of Predictors of Education Outcomes", *Journal of Education for Students Placed at Risk (JESPAR)*, Vol. 24/1, pp. 20-46, http://dx.doi.org/10.1080/10824669.2018.1523734. [6]

Breiman, L. (2001), "Random forests", *Machine learning*, Vol. 45/1, pp. 5-32. [93]

Breiman, L., J. Friedman, C. Stone and R. Olshen (1993), *Classification and regression trees*, Routledge, New York. [81]

Brookhart, S. (2015), "Graded Achievement, Tested Achievement, and Validity", *Educational Assessment*, Vol. 20/4, pp. 268-296, http://dx.doi.org/10.1080/10627197.2015.1093928. [51]

Brookhart, S., T. Guskey, A. Bowers, J. McMillan, J. Smith, L. Smith, M. Stevens and M. Welsh (2016), "A Century of Grading Research", *Review of Educational Research*, Vol. 86/4, pp. 803-848, http://dx.doi.org/10.3102/0034654316672069. [58]

Carl, B., J. Richardson, E. Cheng, H. Kim and R. Meyer (2013), "Theory and Application of Early Warning Systems for High School and Beyond", *Journal of Education for Students Placed at Risk (JESPAR)*, Vol. 18/1, pp. 29-49, http://dx.doi.org/10.1080/10824669.2013.745374. [9]

Christie, S., D. Jarratt, L. Olson and T. Taijala (2019), *Machine-Learned School Dropout Early Warning at Scale*, Paper presented at the The 12th International Conference on Educational Data Mining, Montreal, Canada., https://www.infinitecampus.com/pdf/Machine-learned-School-Dropout-Early-Warning-at-Scale.pdf. [96]

Chung, J. and S. Lee (2019), "Dropout early warning systems for high school students using machine learning", *Children and Youth Services Review*, Vol. 96, pp. 346-353, https://doi.org/10.1016/j.childyouth.2018.11.030. [91]

Coleman, C., R. Baker and S. Stephenson (2019), "A Better Cold-Start for Early Prediction of Student At-Risk Status in New School Districts.", *Paper presented at the Proceedings of The 12th International Conference on Educational Data Mining (EDM 2019).* [105]

Collins, L. and S. Lanza (2010), *Latent Class and Latent Transition Analysis: With Applications in the Social, Behavioral, and Health Sciences*, Wiley, Hoboken, NJ. [62]

Corbett-Davies, S. and S. Goel (2018), *The Measure and Mismeasure of Fairness: A Critical Review of Fair Machine Learning.*, Stanford University. [112]

d'Alessandro, B., C. O'Neil and T. LaGatta (2017), "Conscientious Classification: A Data Scientist's Guide to Discrimination-Aware Classification", *Big Data*, Vol. 5/2, pp. 120-134, http://dx.doi.org/10.1089/big.2016.0048. [113]

Davis, M., L. Herzog and N. Legters (2013), "Organizing Schools to Address Early Warning Indicators (EWIs): Common Practices and Challenges", *Journal of Education for Students Placed at Risk (JESPAR)*, Vol. 18/1, pp. 84-100, http://dx.doi.org/10.1080/10824669.2013.745210. [10]

Dudik, M. et al. (n.d.), *fairlearn: A Python package to assess and improve fairness of machine learning models.*, Redmond, WA: Microsoft., https://fairlearn.github.io/. [114]

Dupéré, V. et al. (2018), "High School Dropout in Proximal Context: The Triggering Role of Stressful Life Events", *Child Development*, Vol. 89/2, pp. e107-e122, http://dx.doi.org/10.1111/cdev.12792. [68]

Dupéré, V., E. Dion, T. Leventhal, I. Archambault, R. Crosnoe and M. Janosz (2015), "Stressors and Turning Points in High School and Dropout", *Review of Educational Research*, Vol. 85/4, pp. 591-629, http://dx.doi.org/10.3102/0034654314559845. [69]

Dynarksi, M., L. Clarke, B. Cobb, J. Finn, R. Rumberger and J. Smink (2008), *Dropout prevention: A practice guide*, http://ies.ed.gov/ncee/wwc/pdf/practiceguides/dp_pg_090308.pdf. [21]

Dynarski, M. and P. Gleason (2002), "How can we help? What we have learned from recent federal dropout prevention evaluations", *Journal of Education for Students Placed at Risk*, Vol. 2002/1, pp. 43-69, http://dx.doi.org/10.1207/S15327671ESPR0701_4. [22]

Eisen, M., P. Spellman, P. Brown and D. Botstein (1998), "Cluster analysis and display of genome-wide expression patterns", *Proceedings of the National Academy of Sciences*, Vol. 95, pp. 14863-14868, http://rana.lbl.gov/papers/Eisen_PNAS_1998.pdf. [99]

Faria, A., N. Sorensen, J. Heppen, J. Bowdon, S. Taylor, R. Eisner and S. Foster (2017), *Getting students on track for graduation: Impacts of the Early Warning Intervention and Monitoring System after one year*, https://ies.ed.gov/ncee/edlabs/projects/project.asp?projectID=388. [27]

Finn, J. (1989), "Withdrawing from school", *Review of Educational Research*, Vol. 59/2, pp. 117-142. [56]

Frazelle, S. and A. Nagel (2015), *A practitioner's guide to implementing early warning systems*, http://ies.ed.gov/ncee/edlabs/regions/northwest/pdf/REL_2015056.pdf. [11]

Freeman, J. and B. Simonsen (2015), "Examining the Impact of Policy and Practice Interventions on High School Dropout and School Completion Rates", *Review of Educational Research*, Vol. 85/2, pp. 205-248, http://dx.doi.org/10.3102/0034654314554431. [65]

Goldhaber, D., M. Wolff and T. Daly (2020), *Assessing the Accuracy of Elementary School Test Scores as Predictors of Students' High School Outcomes: CALDER Working Paper No. 235-0520*, https://caldercenter.org/sites/default/files/CALDER%20WP%20235-0520_0.pdf. [72]

Hargis, C. (1990), *Grades and grading practices: Obstacles to improving education and helping at-risk students, Charles C. Thomas, Springfield.* [57]

Hartman, J., C. Wilkins, L. Gregory, L. Feagans Gould and S. D'Souza (2011), *Applying an on-track indicator for high school graduation: Adapting the Consortium on Chicago School Research indicator for five Texas districts*, http://ies.ed.gov/ncee/edlabs/regions/southwest/pdf/REL_2011100.pdf. [26]

Hawn Nelson, A., D. Jenkins, S. Zanti, M. Katz, E. Berkowitz, TC. Burnett and D. Culhane (2020), *A Toolkit for Centering Racial Equity Throughout Data Integration*, https://www.aisp.upenn.edu/wp-content/uploads/2020/05/AISP-Toolkit_5.27.20.pdf. [98]

Hox, J. (2010), Multilevel Analysis, Routledge, http://dx.doi.org/10.4324/9780203852279. [127]

India AI (2019), *AI is being used to identify potential school dropout rate in Andhra Pradesh*, https://indiaai.gov.in/case-study/ai-is-being-used-to-identify-potential-school-dropout-rate-in-andhra-pradesh (accessed on 29 April 2021). [29]

Interview between Pasi Silander and Stéphan Vincent-Lancrin (2021), *Private communication between Pasi Silander, City of Helsinki, and Stéphan Vincent-Lancrin, OECD.* [32]

Issa, N. (2019, August 22), *Chicago high school dropout rate hits all-time low, CPS says.*, https://chicago.suntimes.com/2019/8/22/20828653/cps-chicago-public-schools-dropout-rate (accessed on 18 January 2021). [119]

Janosz, M., I. Archambault, J. Morizot and L. Pagani (2008), "School engagement trajectories and their differential predictive relations", *Journal of Social Issues*, Vol. 64/1, pp. 21-40. [39]

Kelly, S. (2008), "What Types of Students' Effort Are Rewarded with High Marks?", *Sociology of Education*, Vol. 81/1, pp. 32-52, http://dx.doi.org/10.1177/003804070808100102. [59]

Kemple, J., M. Segeritz and N. Stephenson (2013), "Building On-Track Indicators for High School Graduation and College Readiness: Evidence from New York City", *Journal of Education for Students Placed at Risk (JESPAR)*, Vol. 18/1, pp. 7-28, http://dx.doi.org/10.1080/10824669.2013.747945. [12]

Kinnebrew, J., J. Segedy and G. Biswas (2014), "Analyzing the temporal evolution of students' behaviors in open-ended learning environments", *Metacognition and Learning*, Vol. 9/2, pp. 187-215, http://dx.doi.org/10.1007/s11409-014-9112-4. [101]

Knowles, J. (2015), "Of Needles and Haystacks: Building an Accurate Statewide Dropout Early Warning System in Wisconsin", *Journal of Educational Data Mining*, Vol. 7/3, pp. 18-67, http://www.educationaldatamining.org/JEDM/index.php/JEDM/article/view/JEDM082. [71]

Koedinger, K., S. D'Mello, E. McLaughlin, Z. Pardos and C. Rosé (2015), "Data mining and education", *Wiley Interdisciplinary Reviews: Cognitive Science*, Vol. 6/4, pp. 333-353, http://dx.doi.org/10.1002/wcs.1350. [74]

Koon, S. and Y. Petscher (2015), *Comparing methodologies for developing an early warning system*, http://ies.ed.gov/ncee/edlabs/regions/southeast/pdf/REL_2015077.pdf. [84]

Koon, S., Y. Petscher and B. Foorman (2014), *Using evidence-based decision trees instead of formulas to identify at-risk readers*, http://dx.doi.org/REL 2014–036. [85]

Krumm, A., B. Means and M. Bienkowski (2018), *Learning Analytics Goes to School*, Routledge, New York, NY : Routledge, 2018., http://dx.doi.org/10.4324/9781315650722. [106]

Lamote, C., J. Van Damme, W. Van Den Noortgate, S. Sperybroeck, T. Boonen and J. Bilde (2013), "Dropout in secondary education: an application of a multilevel discrete-time hazard model accounting for school changes.", *Quality and Quantity*, Vol. 47/5, pp. 2425-2446. [122]

Larusson, J. and B. White (eds.) (2014), *Educational data mining and learning analytics*, Springer, New York, http://link.springer.com/chapter/10.1007%2F978-1-4614-3305-7_4. [76]

Lee, J., M. Recker, A. Bowers and M. Yuan (2016), *Hierarchical Cluster Analysis Heatmaps and Pattern Analysis: An Approach for Visualizing Learning Management System Interaction Data*, Paper presented at the International Conference of Educational Data Mining (EDM), Raleigh, NC, http://www.educationaldatamining.org/EDM2016/proceedings/paper_34.pdf. [102]

Lee, S. and J. Chung (2019), "The Machine Learning-Based Dropout Early Warning System for Improving the Performance of Dropout Prediction", *Applied Sciences*, Vol. 9/15, p. 3093, https://doi.org/10.3390/app9153093. [94]

Loukina, A., N. Madnani and K. Zechner (2019), "The many dimensions of algorithmic fairness in educational applications", *Proceedings of the Fourteenth Workshop on Innovative Use of NLP for Building Educational Applications*, http://dx.doi.org/10.18653/v1/w19-4401. [115]

Lyche, C. (2010), *"Taking on the Completion Challenge: A Literature Review on Policies to Prevent Dropout and Early School Leaving"*, OECD Education Working Papers, No. 53, OECD Publishing, Paris, https://dx.doi.org/10.1787/5km4m2t59cmr-en. [3]

Mac Iver, M. (2013), "Early Warning Indicators of High School Outcomes", *Journal of Education for Students Placed at Risk (JESPAR)*, Vol. 18/1, pp. 1-6, http://dx.doi.org/10.1080/10824669.2013.745375. [13]

Mac Iver, M. and M. Messel (2013), "The ABCs of Keeping On Track to Graduation: Research Findings from Baltimore.", *Journal of Education for Students Placed at Risk (JESPAR)*, Vol. 18/1, pp. 50-67. [20]

Mac Iver, M. et al. (2019), "An Efficacy Study of a Ninth-Grade Early Warning Indicator Intervention", *Journal of Research on Educational Effectiveness*, Vol. 12/3, pp. 363-390, http://dx.doi.org/10.1080/19345747.2019.1615156. [28]

Mandinach, E. and K. Schildkamp (2020), "Misconceptions about data-based decision making in education: An exploration of the literature", *Studies in Educational Evaluation*, http://dx.doi.org/10.1016/j.stueduc.2020.100842. [107]

Márquez-Vera, C., A. Cano, C. Romero and S. Ventura (2013), "Predicting student failure at school using genetic programming and different data mining approaches with high dimensional and imbalanced data", *Applied Intelligence*, Vol. 38/3, pp. 315-330, http://dx.doi.org/10.1007/s10489-012-0374-8. [92]

Marquez-Vera, C., C. Morales and S. Soto (2013), "Predicting School Failure and Dropout by Using Data Mining Techniques", *IEEE Revista Iberoamericana de Tecnologias del Aprendizaje*, Vol. 8/1, pp. 7-14, http://dx.doi.org/10.1109/rita.2013.2244695. [89]

Martin, D. and T. von Oertzen (2015), "Growth Mixture Models Outperform Simpler Clustering Algorithms When Detecting Longitudinal Heterogeneity, Even With Small Sample Sizes", *Structural Equation Modeling: A Multidisciplinary Journal*, Vol. 22/2, pp. 264-275, http://dx.doi.org/10.1080/10705511.2014.936340. [43]

Martínez Abad, F. and A. Chaparro Caso López (2017), "Data-mining techniques in detecting factors linked to academic achievement", *School Effectiveness and School Improvement*, Vol. 28/1, pp. 39-55, http://dx.doi.org/10.1080/09243453.2016.1235591. [86]

Masyn, K. (2011), *Latent class analysis and finite mixture modeling*, Oxford University Press, Oxford. [44]

McMahon, B. and S. Sembiante (2020), "Re-envisioning the purpose of early warning systems: Shifting the mindset from student identification to meaningful prediction and intervention", *Review of Education*, Vol. 8/1, pp. 266-301, http://dx.doi.org/10.1002/rev3.3183. [14]

Menzer, J. and R. Hampel (2009), "Lost at the last minute", *Phi Delta Kappan*, Vol. 90/9, pp. 660-664. [63]

Moyer-Packenham, P., S. Tucker, A. Westenkow and J. Symanzik (2015), "Examining Patterns in Second Graders' Use of Virtual Manipulative Mathematics Apps through Heatmap Analysis", *International Journal of Educational Studies in Mathematics*, Vol. 2/2, pp. 1-16, http://dx.doi.org/10.17278/ijesim.2015.02.004. [103]

MSV, J. (2016), *Forbes*, https://www.forbes.com/sites/janakirammsv/2016/07/30/how-microsoft-is-making-big-impact-with-machine-learning/?sh=784705a02f16 (accessed on 29 April 2021). [31]

Muthén, B. (2004), "Latent variable analysis: Growth mixture modeling and related techniques for longitudinal data", in Kaplan, D. (ed.), *The Sage handbook of quantitative methodology for the social sciences*, Sage Publications, Thousand Oaks, CA, http://www.statmodel.com/papers.shtml. [40]

OECD (2019), *Education at a Glance 2019: OECD Indicators*, OECD Publishing, Paris, https://dx.doi.org/10.1787/f8d7880d-en. [1]

O'Neil, C. (2016), *Weapons of math destruction: How big data increases inequality and threatens democracy.*, Broadway Books. [111]

Pallas, A. (2003), "Educational transitions, trajectories, and pathways", in Mortimer, J. and M. Shanahan (eds.), *Handbook of the life course*, Kluwer Academic/Plenum Publishers, New York. [70]

Piety, P. (2019), "Components, Infrastructures, and Capacity: The Quest for the Impact of Actionable Data Use on P–20 Educator Practice", *Review of Research in Education*, Vol. 43/1, pp. 394-421, http://dx.doi.org/10.3102/0091732x18821116. [75]

Piety, P., D. Hickey and M. Bishop (2014), *Educational data sciences: Framing emergent practices for analytics of learning, organizations, and systems*, ACM. [16]

Quinlan, J. (1993), *C4.5: Programs for Machine Learning, Morgan Kaufmann*, San Mateo, CA. [82]

Quinlan, J. (1990), "Probablistic decision trees", in Kodratoff, Y. and R. Michalski (eds.), *Machine learning*, Morgan Kaufmann, San Francisco (CA), https://doi.org/10.1016/B978-0-08-051055-2.50011-0. [83]

Ram, N. and K. Grimm (2009), "Methods and Measures: Growth mixture modeling: A method for identifying differences in longitudinal change among unobserved groups", *International Journal of Behavioral Development*, Vol. 33/6, pp. 565-576, http://dx.doi.org/10.1177/0165025409343765. [45]

Raudenbush, S. and A. Bryk (2002), *Hierarchical linear models: Applications and data analysis methods (2nd ed.)*, Thousand Oaks: Sage. [128]

Riehl, C. (1999), "Labeling and letting go: An organizational analysis of how high school students are discharged as dropouts", in Pallas, A. (ed.), *Research in sociology of education and socialization*, JAI Press, New York. [79]

Romesburg, H. (1984), *Cluster analysis for researchers*, Lifetime Learning Publications, Belmont, CA. [104]

Rumberger, R. (2011), *Dropping Out: Why Students Drop Out of High School and What Can Be Done About It*, Harvard University Press, Cambridge, Mass. [4]

Rumberger, R., H. Addis, E. Allwensworth, R. Blafanza, J. Bruch, E. Dillon, D. Duardo, M. Dynarski, J. Fugerson, M. Jayanthi, R. Newman-Gonchar, K. Place and C. Tuttle (2017), *Preventing dropout in secondary schools (NCEE 2017-4028)*, https://ies.ed.gov/ncee/wwc/Docs/PracticeGuide/wwc_dropout_092617.pdf. [5]

Rumberger, R. and G. Palardy (2005), "Test Scores, Dropout Rates, and Transfer Rates as Alternative Indicators of High School Performance", *American Educational Research Journal*, Vol. 42/1, pp. 3-42, http://dx.doi.org/10.3102/00028312042001003. [80]

Sansone, D. (2019), "Beyond Early Warning Indicators: High School Dropout and Machine Learning", *Oxford Bulletin of Economics and Statistics*, Vol. 81/2, pp. 456-485, http://dx.doi.org/10.1111/obes.12277. [66]

Sara, N. et al. (2015), *High-School Dropout Prediction Using Machine Learning: A Danish Large-scale Study*, Paper presented at the European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning, Bruges, Belgium. [90]

Sculley, D., G. Holt, D. Golovin, E. Davydov and T. Phillips (2015), "Hidden technical debt in machine learning systems.", *Paper presented at the Advances in neural information processing systems.* [117]

Singer, J. and J. Willett (2003), *Applied Longitudinal Data Analysis*, Oxford University Press, http://dx.doi.org/10.1093/acprof:oso/9780195152968.001.0001. [123]

Singh, L., A. Deshpande, W. Zhou, A. Banerjee, A. Bowers, S. Friedler, H. Jagadish, G. Karypis, Z. Obradovic, A. Vullikanti and W. Zuo (2019), "NSF BIGDATA PI Meeting - Domain-Specific Research Directions and Data Sets", *ACM SIGMOD Record*, Vol. 47/3, pp. 32-35, *http://dx.doi.org/10.1145/3316416.3316425*. [121]

Soland, J. (2017), "Combining Academic, Noncognitive, and College Knowledge Measures to Identify Students Not on Track For College: A Data-Driven Approach", *Research & Practice in Assessment*, Vol. 12/Summer 2017, pp. 5-19, http://www.rpajournal.com/dev/wp-content/uploads/2017/07/Summer_2017.pdf#page=5. [88]

Soland, J. (2013), "Predicting High School Graduation and College Enrollment: Comparing Early Warning Indicator Data and Teacher Intuition", *Journal of Education for Students Placed at Risk (JESPAR)*, Vol. 18/3-4, pp. 233-262, http://dx.doi.org/10.1080/10824669.2013.833047. [87]

Stodden, V., M. McNutt, D. Bailey, E. Deelman, Y. Gil, B. Hanson, M. Heroux, J. Ioannidis and M. Taufer (2016), "Enhancing reproducibility for computational methods", *Science*, Vol. 354/6317, pp. 1240-1241, http://dx.doi.org/10.1126/science.aah6168. [108]

Stuit, D., M. O'Cummings, H. Norbury, J. Heppen, S. Dhillon, J. Lindsay and B. Zhu (2016), *Identifying early warning indicators in three Ohio school districts*, http://ies.ed.gov/ncee/edlabs/regions/midwest/pdf/REL_2016118.pdf. [23]

Swets, J. (1988), "Measuring the accuracy of diagnostic systems", *Science*, Vol. 240/4857, pp. 1285-1293, http://dx.doi.org/10.1126/science.3287615. [36]

Swets, J., R. Dawes and J. Monahan (2000), "Psychological Science Can Improve Diagnostic Decisions", *Psychological Science in the Public Interest*, Vol. 1/1, pp. 1-26, http://dx.doi.org/10.1111/1529-1006.001. [37]

The Wire (2016), *Aadhaar in Andhra: Chandrababu Naidu, Microsoft Have a Plan For Curbing School Dropouts*, https://thewire.in/politics/aadhaar-in-andhra-chandrababu-naidu-microsoft-have-a-plan-for-curbing-school-dropouts (accessed on 29 April 2021). [30]

Vermunt, J. and J. Magidson (2002), "Latent class cluster analysis", in Hagenaars, J. and A. McCutcheon (eds.), *Applied latent class analysis*, Cambridge University Press. [47]

Vermunt, J., B. Tran and J. Magidson (2008), "Latent class models in longitudinal research", in Menard, S. (ed.), *Handbook of longitudinal research: Design, measurement and analysis*, Elsevier, Burlington, MA, http://www.statisticalinnovations.com/articles/VermuntTranMagidson2006.pdf. [46]

Villagrá-Arnedo, C., F. Gallego-Durán, F. Llorens-Largo, P. Compañ-Rosique, R. Satorre-Cuerda and R. Molina-Carmona (2017), "Improving the expressiveness of black-box models for predicting student performance", *Computers in Human Behavior*, Vol. 72, pp. 621-631, http://dx.doi.org/10.1016/j.chb.2016.09.001. [97]

Wachter, S. and B. Mittelstadt (2019), "A right to reasonable inferences: re-thinking data protection law in the age of big data and AI.", *Columia Business Law Review*, Vol. 494. [109]

Wilkinson, L. and M. Friendly (2009), "The History of the Cluster Heat Map", The American Statistician, Vol. 63/2, pp. 179-184, http://dx.doi.org/10.1198/tas.2009.0033. [100]

Willingham, W., J. Pollack and C. Lewis (2002), "Grades and test scores: Accounting for observed differences", *Journal of Educational Measurement*, Vol. 39/1, pp. 1-37, http://www.jstor.org/stable/1435104. [60]

Zehlike, M., F. Bonchi, C. Castillo, S. Haijan, M. Megahed and R. Baeza-Yates (2017), *FA*IR: A Fair Top-k Ranking Algorithm*, ACM, Singapore, Singapore, http://dx.doi.org/10.1145/3132847.3132938. [116]

10 Game-based assessment for education

Jack Buckley, Laura Colosimo, Rebecca Kantar, Marty McCall, and Erica Snow

Imbellus, Inc., United States

This chapter discusses how recent advancements in digital technology could lead to a new generation of game-based standardised assessments in education, providing education systems with assessments that can test more complex skills than traditional standardised tests can. After highlighting some of the advantages of game-based standardised assessment compared to traditional ones, this chapter discusses how these tests are built, how they work, but also some of their limitations. While games have strong potential to improve the quality of testing and expand assessment to complex skills in the future, they will likely supplement traditional tests, which also have their advantages. Three examples of game-based assessments integrating a range of advanced technologies illustrate this perspective.

Introduction

Rapid technological developments such as virtual/augmented reality, digital user interface and experience design, machine learning/artificial intelligence, and educational data mining have led to the improvement of simulated digital environments, and accelerated progress in the quality and design of digital simulations and video games. While this has led to the development of a range of "e-learning" applications to be used both inside and outside of the classroom (from virtual labs to medical e-learning tools with simulations), this technological advancement has also opened avenues for a new generation of standardised assessments. Such game-based assessments allow for the assessment of a broader range of skills (e.g. creativity, collaboration or socioemotional skills), as well as better measurement of some aspects of the "thinking" of respondents, including in traditional domains like science and mathematics. Moreover, the use of simulated environments enables assessing knowledge and skills in settings that are more authentic to "real life" applications of those skills.

While promising, this new generation of assessments brings its own challenges. For examples, they are more costly and difficult to develop than traditional standardised tests based on a simple succession of discrete questions or small tasks. Nevertheless, some game-based standardised assessments have already been successfully developed and will likely be one part of how the learners of tomorrow will be assessed. The chapter is organised as follows: we first argue that game-based assessments address many of the critiques of traditional assessment and have the potential of being aligned more closely to teaching and learning in the classroom; we then explain how these assessments work, what kind of technology they use, what kind of data they draw on, and highlight the challenges in building them; we provide some examples of game-based standardised assessments, before reflecting on the role they could have in the future, and what national infrastructure may be required to deliver them at scale.

Why game- or simulation-based assessment in education?

The use of standardised assessment in education – increasingly coupled with well-defined standards for academic content – is far from a new idea, dating back some four decades in some high income countries and at least 20 years internationally (Braun and Kanjee, 2006[1]). More recently, leaders in education policy, teaching and learning, and cognitive theory have come together to call for greater coherence among instruction, curriculum, and assessment and for a comprehensive assessment system that informs decisions "from the statehouse to the schoolhouse" (Gong, 2010[2]). As part of this improvement initiative there has been a growing movement around and interest in new assessment technologies and approaches, including immersive, game- or simulation-based assessments (GBAs) (DiCerbo, 2014[3]; Shaffer et al., 2009[4]; Shute, 2011[5]). As we discuss below, these new approaches take advantage of the increasing prevalence of educational technology – primarily personal computers and high-speed connectivity – in schools, as well as advances in psychometrics, computerised assessment design, educational data mining, and machine learning/artificial intelligence available to test makers.

In education, traditional standardised assessments have long been dominated by a model centred on collections of discrete questions (or "items") designed to cover content in an assessment framework by addressing parts of the domain to be measured (Mislevy et al., 2012[6]). GBAs, on the other hand, aim to blur the line between traditional assessment and more engaging learning activities through the use of games and simulations designed to measure constructs in an environment that maximises "flow" and rewards students for demonstrating their cognitive processes in more engaging and authentic situations, not just their ability to memorise key facts (Shute et al., 2009[7]).

While traditional educational assessments are designed to generally meet standards of technical quality in areas like *validity* (does the assessment measure what it is supposed to measure?), *reliability* (does it do this consistently and with minimal error?), and *fairness* (is the assessment culturally sensitive, accessible, and free of bias against any groups of test-takers?), they have nevertheless been criticised from a range of perspectives. We briefly review each of the following specific criticisms of traditional standardised assessment (e.g. Sanders and Horn, 1995[8]) and how game-based assessment may ameliorate them:

- the need to apply modern psychological theory to assessment;
- insufficient alignment of assessment with curriculum and instruction (Duncan and Hmelo-Silver, 2009[9]);
- lack of integration of assessments for different purposes, including formative, interim, and summative (Perie, Marion and Gong, 2009[10]);
- inability of traditional assessment to measure some important and increasingly policy-relevant constructs (Darling-Hammond, 2006[11]), and;
- declines in student engagement and motivation (Nichols and Dawson, 2012[12])

Application of modern psychological theory to assessment

The seminal volume *Knowing What Students Know* (National Research Council, 2001[13]) brought cognitive theory into the assessment realm using a framework accessible to teachers and policy makers. It called for an examination of mental functions involved in deep understanding, concepts that are difficult to assess with the sort of short, disconnected questions typical of standardised tests (Darling-Hammond et al., 2013[14]). New task types (or assessment items frequently used in classrooms but not on standardised tests) requiring complex performance on more realistic tasks were called for, including essays, projects, portfolios, and observation of classroom performance. Games and simulations have become more central due to the potential for eliciting evidence of deeper understanding and cognitive processes. Interpretation of streaming data from gameplay or interaction with a carefully-designed digital user interface allows researchers to evaluate how people go about solving problems and can lead to more targeted feedback (Chung, 2014[15]). For example, modern academic content standards in science increasingly require students to learn and demonstrate scientific practices (i.e. that they can think and reason like a scientist) as well as science facts. Accordingly, game-based-assessment allows the test maker to build scenarios and simulations where the students' reasoning and process can be observed through their complex interactions with elements in the game or simulation.

The need for assessment to be better aligned with curriculum and instruction

In keeping with the push to improve education, curriculum has seen a shift, incorporating learning theory and evidence-centred design approaches that provide students with grounding phenomena and hands on examples (Mislevy et al., 2012[6]; Arieli-Attali et al., 2019[16]). However, traditional standardised assessments have remained relatively stagnant, providing only limited information for teachers and learners, and furthering the divide between what is learned (content of curriculum) and what is tested (content of assessments) (Martone and Sireci, 2009[17]). As researchers and policy makers continue to call for new assessment frameworks that incorporate theories of learning and foundational transferrable skills consistent with classroom activity (National Research Council, 2012[18]; Darling-Hammond et al., 2013[14]; Conley, 2018[19]), this has led to increased interest in the development of games, simulations, and intelligent tutoring systems designed around learning progressions or specific instructional units.

Increasing assessment coherence

Assessments are broadly categorised by their purpose: how are scores used and interpreted? *Summative* tests are given at the end of instruction to evaluate what has been learned. Examples of summative assessment applications in education include annual large-scale accountability tests and college entrance exams, but also "drop from the sky" monitoring tests like PISA, TIMSS, and various national assessments (Oranje et al., 2019[20]). Summative assessments may be high-stakes for students (college entrance exams or graduation tests) but often are low-stakes for students but higher-stakes for other actors in the education system. Interim tests are given during the instructional period to evaluate progress toward summative goals and suggest instructional changes. Formative tests are also given during instruction but are closely linked to specific instruction and individual performance. Unlike interim assessments which can be aggregated at various education levels and are related to broad summative goals, formative assessments are adjusted to individual needs and to immediate teaching strategy (Shepard, Penuel and Pellegrino, 2018[21]). Each of these levels of educational assessment has different purposes and often requires appropriate measurement models and validation methods (Ferrara et al., 2017[22]).

The amalgamation of all these various types of assessment can create confusion for educators and parents and often comes at the expense of students' instructional time. Accordingly, there is growing interest in rationalising this confusing and fractured system. In the United States, for example, as theoretically-based, instructionally-relevant assessments have become more prominent, there have been widespread calls for "coherence" in the assessment enterprise across all levels (Gong, 2010[2]; Marion et al., 2019[23]). That is, policy makers and educators increasingly want all assessments that students encounter throughout the school year to function together as a single, coherent system.

While games and simulations in assessment have been most often targeted at the formative level, recent advances in development and scoring have made their use in large-scale summative tests in national accountability systems and international comparisons more feasible (Verger, Parcerisa and Fontdevila, 2019[24]; Klieme, 2020[25]). For example, in a coherent system, an immersive GBA could be used in a variety of ways. Formatively, the GBA could provide continuous feedback and personalised suggestions in the course of instruction. As an interim measure, the student can be assessed under more standardised simulation conditions to gauge progress toward summative goals. In summative mode, the student could be presented with a novel but related GBA scenario to solve without formative supports, allowing for a better understanding of what students have learned and are able to do. Box 10.1 highlights an example for assessing vocational skills in Germany.

Measuring different constructs - "hard-to-measure" skills

Another critique of traditional standardised assessments is that they are ineffective in measuring knowledge, skills, and abilities beyond very simple content in very circumscribed domains (Madaus and Russell, 2010[27]). For example, while a traditional standardised test may be a valid, reliable, fair, and efficient way to measure algebra, it may not be a modality suitable for measuring constructs like creative thinking or collaborative problem solving. This is an especially relevant critique in education for two reasons. First, modern curricular frameworks around the world increasingly are multidimensional, including cross-cutting skills as well as more traditional academic content. For example, the Next Generation Science Standards in the United States (www.nextgenscience.org/) include not only disciplinary core concepts, but also cross-cutting ideas in science, and scientific and engineering practices. Second, there is growing realisation among international policy makers of the importance of so-called

"21st Century Skills" or skills associated with "deeper learning," such as critical thinking, communication, collaboration, and creativity (Trilling and Fadel, 2009[28]; Vincent-Lancrin et al., 2019[29]; Fadel, Bialik and Trilling, 2015[30]). The use of games or simulations is a very promising way to assess these complex constructs either as part of a revised curricular framework or as a novel addition to the content covered by the usual standardised tests (Stecher and Hamilton, 2014[31]; Seelow, 2019[32]).

Box 10.1 Using simulations to train and assess vocational skills in Germany

With the ASCOT+ projects, Germany's Federal Ministry for Education and Research is supporting the development of digital training and assessment for vocational skills in different domains (car mechatronics, problem solving for technical systems, commercial problem solving, inter-professional and socio-emotional skills in nursing). In addition to digital training units using videos and simulations, the project is developing assessments that will be used as exams to certify apprentices' skills. For example, in the domain of commercial professions, a competency-oriented assessment task creator is being developed to allow assessors to design exams that certify students' and workers' competences, leading to a shift from knowledge-based to competence-based examination. An assessment bank of digital examination tasks that can be slightly modified or combined is proposed for assessors' customisation. It will be launched (and legally recognised for exams in Germany) in 2022. In the domain of car mechatronics, examination tasks are also developed to test trainees' competences in a simulated environment – and also to develop their skills.

Source: Bundesministerium für Bildung und Forschung (n.d.[26])

Measuring constructs differently - interaction and engagement

It goes without saying that most test-takers do not enjoy the traditional assessment experience (Nichols and Dawson, 2012[12]; Madaus and Russell, 2010[27]). One of the attractions of game-based assessment – beyond those noted above – is the promise of delivering valid and reliable measurement of complex constructs while bringing some of the engagement and immersive properties of modern video games. While there is growing evidence supporting this benefit across a broad range of operational game-based assessments (Hamari et al., 2016[33]), it is important to remember the inherent difference in purpose between games played for enjoyment versus those used for measurement (particularly, but not limited to, those used in high-stakes contexts). In addition, the need for GBA to meet assessment's more stringent scientific criteria of validity, reliability, and fairness, means that the transferability of engagement and immersion may be somewhat limited or at least different in nature (Oranje et al., 2019[20]). Simply put, game-based assessments might not be as fun as "real" games.

We now turn to a closer examination of the features of GBA and a brief discussion of how to design game-based tests.

How do we build game-based tests?

Designing from the ground up

Using games and game-based features as a means to increase engagement and capture hard-to-measure constructs is not a new idea (Cordova and Lepper, 1996[34]). However, the assessment field's knowledge and understanding of how best to implement this type of assessment and how to best use the data it provides continues to grow and mature. There are many ways to incorporate games and game-based features into a system or assessment that have varying impact on the learner. Thus, building a GBA requires forethought about the exact types of features and their potential impact on the learner and data collection (Shute and Ventura, 2013[35]).

The assessment designer must determine, a priori, exactly what the game is attempting to measure and how each game-based element provides evidence that allows such measurement. This includes storyboarding out the measures of interest, determining the evidence needed to capture them, and the exact quantification of that evidence. As Mislevy (2018[36]) notes, *"the worst way to design a game- or simulation-based assessment is to design what seems to be a great game or simulation, collect some observations to obtain whatever information the performances provide, then hand the data off to a psychometrician or a data scientist to 'figure out how to score it.'"* While there are benefits to post hoc exploratory analyses, they should not be the driving mechanism for how one scores the assessment. Before the assessment design team starts to develop the game specifications, they must first outline what they intend to measure and how this will be accomplished. This includes the quantification of evidence and scales that will be used.

This important foundational work cannot be done post hoc as this will often result in poor psychometric performance or lack of interpretability. For example, while adapting an existing game for use as an assessment may, at first glance, appear to generate a large amount of data for each test-taker, it is often the case that such data may yield items or measurement opportunities that are poorly-aligned to the desired content domain, exhibit high intercorrelation (rendering many of them useless), or are at the wrong level of difficulty (i.e. too easy or too hard for the target population). Therefore, the item design process should take place nearer to the beginning of the entire project, as designing a GBA takes a significant amount of forethought and discipline – and mistakes can be very costly.

However, this is not to say that analysis of actual test-taker data in the GBA development process is not important. Not only should the designers conduct traditional empirical psychometric analyses necessary to create valid and reliable assessments, they should also take advantage of the wealth of additional data generated by GBA to apply novel methods from domains like machine learning to extract more useable information about test-takers' ability or other constructs where possible (e.g. Gobert, Baker and Wixon, 2015[37]).

Games for assessment vs. "gamification"

We draw an important distinction here between designing games or simulations explicitly for measurement purposes and "gamification" or the addition of game-like elements to existing tasks or activities to increase engagement, flow, or motivation (Deterding et al., 2011[38]). An example of gamification would be adding a leader board, badges, personalised avatars, or progress bars to classroom activity and, while this may be useful in improving student engagement or motivation, it is not the sort of designed game-based assessment we discuss here. It is important to note, however, that the distinction between GBA and gamification is somewhat more grey in the assessment of social-emotional learning in education and "non-cognitive" skills in education and workplace selection. Nevertheless, it is still possible to assess these other types of skills and dispositions via designed games (Yang et al., 2019[39]) and not merely the addition of game-like elements to traditional tests.

Telemetry and the question of "stealth"

Game-based or simulation-centred assessments collect a wealth of data that is often missed or unable to be captured by traditional tests – sometimes "stealthily" or unbeknownst to the test-taker (Shute and Ventura, 2013[35]). This includes patterns of choices, search behaviours, time-on-task behaviours, and, in some cases, eye movement or other biometric information. These rich data sources can be used to help illustrate the cognitive processes that a student engages in as they complete a task (Sabourin et al., 2011[40]; Snow et al., 2015[41]), rather than just focusing on the end product of their performance. However, in order to collect and quantify this information, GBA developers need to carefully prescribe the data that the system collects, often referred to as "telemetry." This process involves mapping out every action a user can take during the design phase and assigning that action a value or name in the data infrastructure. This can most often be accomplished using data collection or measurement frameworks such as Evidence-Centred Design (ECD) (Mislevy et al., 2012[6]). Successful mapping of telemetry to measurement objectives requires a concentrated effort between designers, software engineers, and measurement scientists. As with any assessment, stakeholders should feel confident in what is being measured and how. To use telemetry in educational assessment – particularly in high-stakes and summative applications – we need to be very clear about the actions we are capturing, their interpretation, and how they should be stored and quantified.

How hard is this to do? How do costs compare to more traditional approaches?

Our experience suggests that building valid, reliable, and fair game-based assessments is considerably more complex and challenging than traditional test development. Success requires an interdisciplinary team with a broad range of skills, including game designers, software engineers ideally with a background in game, and cognitive scientists, as well as the test designers, content experts, educational researchers, and psychometricians usually needed to develop an assessment. For this reason, building GBAs is relatively expensive and is thus not always an efficient way to measure simple constructs. For example, while the benefits of GBA have led several operational assessment programs, such as PISA and the U.S. National Assessment of Educational Progress (NAEP), to add game or simulation components, due to cost they have done so in a limited fashion as part of a hybrid approach combined with more traditional item types and assessment strategies (Bergner and von Davier, 2018[42]).

An additional challenge to consider with GBAs is the need to make them accessible for students with disabilities. While the field of educational measurement has made significant progress in this area in recent decades, extending frameworks like Universal Design for Learning (Rose, 2000[43]) to game-based assessment requires careful design, extensive testing, and, in some cases, the invention of new approaches and novel technologies such as haptic feedback technologies that enable the implementation of touch-based user interfaces allowing the assessment of visually-impaired students (Darrah, 2013[44]).

New psychometric methods – and challenges

In addition to requiring a broader range of technical expertise, GBAs can also require innovation in technologies or statistical approaches to measurement. For example, psychometricians have suggested new measurement models reflecting task complexity (Mislevy et al., 2000[45]; Bradshaw, 2016[46]; de la Torre and Douglas, 2004[47]). These new models and others in development are intended to better address theories of cognition and learning and to capture complex latent ability structures. They provide measurement models appropriate to the new data streams generated by games and simulations.

Game-based assessment in education also brings new fairness and equity concerns. For example, differential access to computers in the home or school environment as well as (possibly gendered) differences in familiarity with video game mechanics or user interface components could exacerbate existing achievement gaps or create new ones. Part of the responsible development of GBA is to monitor these gaps and also to minimise differential item functioning (DIF – defined as when items don't behave as expected for test-takers of the same ability but different backgrounds) for both the usual subgroups (gender, ethnicity, language status) but potentially new ones like gaming experience (see Box 10.2). One key design element that reduces risk of differential item functioning in game-based assessment is the design of effective tutorials within each game or simulation that quickly teach the necessary game mechanics to those test-takers possibly less familiar with the user interface.

The promise of AI and machine learning

Beyond psychometric innovation, game- and simulation-based assessment also poses new opportunities for technical innovation based on recent developments in machine learning and artificial intelligence (Ciolacu et al., 2018[49]). For example, in high-stakes applications requiring multiple forms of an assessment to ensure security, computationally-intensive artificial intelligence algorithms (AI) enable the calibration of difficulty of variants of a game as part of the equating process required to ensure fairness for all test-takers. In other words, the AI can be used to "play" all of the proposed variations of the GBA as means of increasing the likelihood that they are all comparable in difficulty before moving to expensive and time-consuming pilot testing with human test-takers.

More broadly, similar AI play, as well as the application of machine learning techniques to large pilot datasets of game performance, can be used as part of the process of deriving meaningful information from telemetry data logs (i.e. the data collected during the assessment game/simulation process). That is, a key part of the development of game-based assessment should include a stage where item scores are refined and improved as via exploratory data analysis and educational data mining as larger amounts of test-taker data become available. Although this data mining should not replace the design process described above, experience suggests that computer-aided iteration here can improve the reliability and efficiency of game-based assessment by increasing the amount of useful information on test-taker performance available (Mislevy et al., 2014[50]).

Box 10.2 Choice of technology matters for gender equity: Insights from an experiment in Chile

Evidence from an experiment in a public school in Santiago suggests that gender differences in learning in educational games may depend on the technological platform used (Echeverría et al., 2012[48]). This may also apply to performance in game-based assessment. In the experiment, 11th grade students played First Colony, an educational game that requires students to apply concepts from electrostatics. Players assume the role of astronauts sent on a mission to bring back a precious crystal. Since the crystal is fragile, the astronauts can only move it using electrical force. In the version of the game implemented on a platform with multiple computer mice, students play in groups of three with each student controlling one mouse. With the mouse, students can move their astronaut, change their charge value and polarity and activate their charge to interact with the crystal. In the augmented reality version, students can perform the same actions using a tablet. Here, the classroom blends with the game world: each desk is covered with a set of markers that allow the augmented reality system to place virtual objects over the desks. Using the webcam at the top of the screen, the system determines the location of each student's astronaut by detecting the relative position of each student to the paper markers. While no gender differences in performance were observed when students played using the multiple-mice platform, boys outperformed girls when playing the same game using an augmented reality platform, with a statistically significant difference.

The results from the experiment revealed statistically significant differences in performance between boys and girls after they used the augmented reality platform. Given that there was no gender difference following the use of the multiple-mice version of the game, this suggests that the choice of platform can create a gender gap in learning that is unrelated to the game. Since girls seem to struggle more to use the augmented reality platform, it is possible that using the technology for GBA would put them at a disadvantage. Educators should select the technology to be used in GBA with care, as their choices may have unintended effects.

Some examples of game-based assessment in education

SimCityEDU: Pollution Challenge (GlassLab)

SimCityEDU: Pollution Challenge was a GBA released in 2014 by GlassLab, a collaborative development initiative funded by the John D. and Catherine T. MacArthur and Bill and Melinda Gates Foundations. It built on the design of the popular SimCity game series and put the test-taker in the role of a virtual city's mayor, tasked with balancing economic growth and environmental protection over a series of four levels of increasing complexity. SimCityEDU was designed as a formative assessment of problem solving, systems thinking, and causality in systems for students at approximately the ISCED-2 level.

The assessment content was explicitly aligned to the Framework for 21st Century Learning and Council for Economic Education standards as well as to aspects of the United States' Next Generation Science Standards and Common Core State Standards in Mathematics. Much as in the source game, problem-solving tasks were very engaging and largely spatial and economic in nature. GlassLab also devoted considerable resources to solving issues such as "tutorialisation" and telemetry processing to create useful assessment items as well as pioneering new psychometric models to support inference and reporting (Mislevy et al., 2014[50]; Mislevy, 2018[36]).

Crisis in Space (ACTNext)

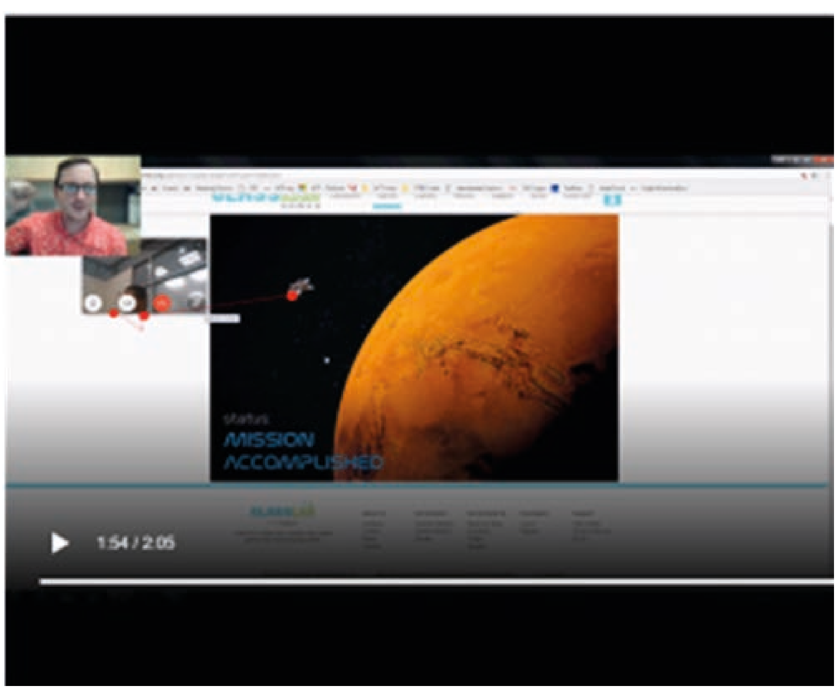
In Crisis in Space, ACTNext developed a pilot version of a GBA designed to assess the collaborative problem solving and related socioemotional skills of middle school (ISCED-2) students. In this game, a pair of two test-takers is tasked with working together to troubleshoot a series of problems on a space station, with one of them in the role of an astronaut on the station and the other in mission control on the ground. By having the students actually collaborate in a cooperative game, Crisis in Space delivers an authentic and engaging experience and improves upon earlier attempts to measure collaboration via student-"agent" (chatbot) interaction.

Figure 10.1 SImCityEDU: Pollution Challenge (GlassLab)

Note: Introduced in 2013 by the now-defunct GlassLab, SimCityEDU: Pollution Challenge was a transformation of the popular SimCity videogame franchise into an assessment of middle school (primarily ISCED 2). The test-taker engages with a modified version of the SimCity interface to solve various urban issues. Telemetry data are processed via sophisticated psychometric models.
Source: Games for change (n.d.[50]); Glasslab (n.d.[51])

Crisis in Space, which won the innovation prize at the 2020 e-Assessment Awards, is particularly notable for its use of a wide range of data types, including user interface-generated telemetry, audio recordings of student conversation, and test-taker eye-tracking data. ACTNext also has implemented advanced machine learning technology such as natural language processing (NLP) to process these data and score instances of collaboration as successful or unsuccessful (Chopade et al., 2019[53]).

Figure 10.2 Crisis in Space (ACTNext)



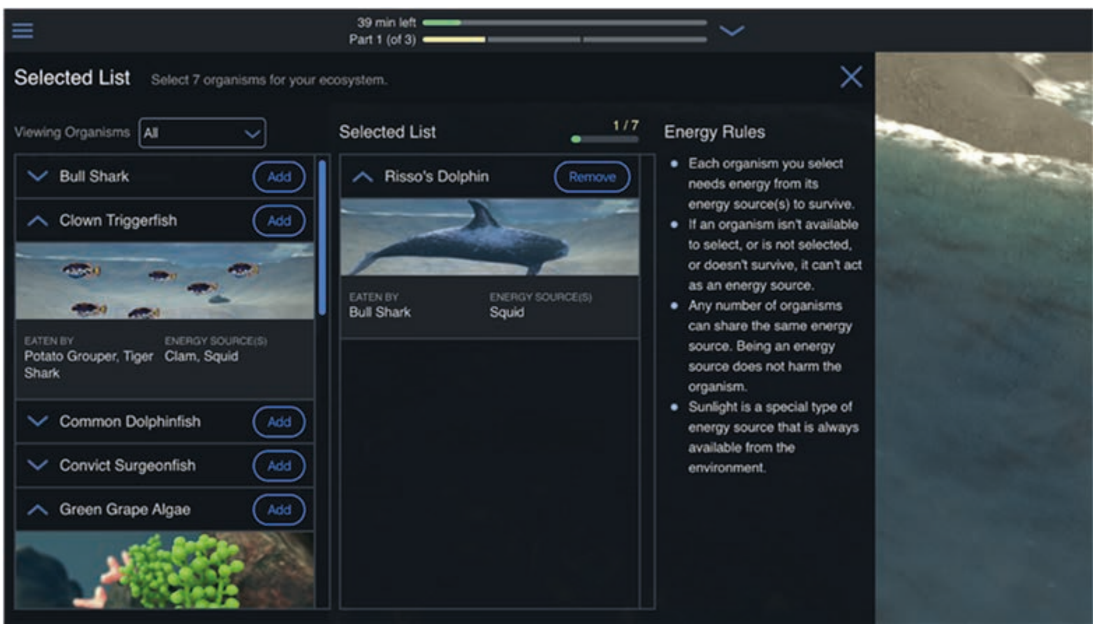
Note: Crisis in Space is a pilot game-based-assessment under development by ACT, Inc. as part of an ongoing program of research and development in collaborative problem-solving assessment by their research arm, ACTNext. In the scenario, two players work together to troubleshoot a space station. Technologies used for measurement include eye-tracking and natural language processing.
Source: https://actnext.org/collaboration-assessment-online-games/ ; https://actnext.org/research-and-projects/cps-x-crisis-in-space/ (reproduced with permission)

PEEP – Project Education Ecosystem Placement (Imbellus)

The third example game-based assessment, Project Education Ecosystem Placement (PEEP) is also in the pilot phase and is intended to measure problem solving in the ISCED-2 and 3 population via a game where test-takers are required to construct a viable food web or ecosystem and place it in the natural environment where it can thrive. PEEP, funded by the Walton Family Foundation, is an adaptation of a game-based assessment originally designed for employment selection that is currently used by the global consultancy, McKinsey and Company, to select new business analysts.

The education version has been adapted to reflect more accurate life sciences content as well as made developmentally appropriate for students. PEEP is designed to be eventually used in high-stakes, summative assessment and supports the creation of many parallel forms or versions to improve test security. To develop these, PEEP uses an algorithm to create viable ecosystem solutions of approximately equivalent difficulty based on a large library of organisms. PEEP can also be delivered as a "stage adaptive" assessment task where test-takers are presented with a series of problems to solve whose difficulty varies algorithmically depending on prior performance.

Figure 10.3 PEEP - Project Education Ecosystem Placement (Imbellus)



Note: Designed to be used as part of a longer assessment of problem solving, the PEEP task challenges students to build a viable ecosystem and place it in a natural environment where it can thrive. The assessment task was adapted from an assessment created by Imbellus for use in employment selection with support from the Walton Family Foundation and in partnership with Summit Public Schools and several other school systems. It uses student telemetry to create a range of both process and product measurement opportunities suitable for scoring via item-response theory.
Source: Courtesy of Imbellus.

What is the long-term promise of this approach and what is necessary to get us there?

Educators, administrators, and policymakers should consider integrating GBA in their educational assessment systems, as it offers unique advantages over more traditional approaches. Game-based assessments are special because they can mirror the dynamic interaction, structural complexity, and feedback loops of real-world situations. In the long term, integrated assessment systems should rely on game- and simulation-based scenarios to evaluate how students integrate and apply knowledge, skills, and abilities. Robust scenarios can involve a subset of content from an academic domain of interest, but perhaps their greatest advantage lies in facilitating the measurement of 21st Century skills like problem solving and collaboration.

The advantages of GBA, including the ability to assess historically hard-to-measure cognitive processes, better alignment with modern curricula, and increased student engagement in the measurement process, make it an important part of the future of all educational assessment systems. However, game-based approaches often do not produce as many useable item scores as we might hope given their relatively high development cost when compared to more traditional, discrete items. Thus a cost-effective and robust system of assessments might use game-based scenarios in combination with traditional and less costly discrete technology-enhanced assessment items. A good design principle for such an assessment system would be to use relatively inexpensive, traditional assessments where feasible (e.g. measuring proficiency with academic content) and reserve the more expensive GBA scenarios and simulations for the measurement of more complex cognitive constructs. Moreover the use of GBA should not be limited to summative assessment alone but should instead be part of a coherent system of assessment throughout the academic year. Such an efficient, hybrid system of assessment could theoretically be designed for many uses, including accountability reporting, driving local instruction, and individual student growth modelling.

In order to realise the promise of game- and simulation-based assessment at the national level, education ministries need to invest in the infrastructure needed to design, implement, and operationally deliver such tests. While some of this capacity can be contracted out to private-sector vendors, successful implementation will require public capabilities as well. These include sufficient computer hardware in schools (although there is a growing trend to consider "bring your own device" policies) and a networking backbone capable of acceptable data transfer speeds.

References

Arieli-Attali, M., S. Ward, J. Thomas, B. Deonovic and A. von Davier (2019), "The Expanded Evidence-Centered Design (e-ECD) for Learning and Assessment Systems: A Framework for Incorporating Learning Goals and Processes Within Assessment Design", *Frontiers in Psychology*, Vol. 10, http://dx.doi.org/10.3389/fpsyg.2019.00853. [16]

Bergner, Y. and A. von Davier (2018), "Process Data in NAEP: Past, Present, and Future", *Journal of Educational and Behavioral Statistics*, Vol. 44/6, pp. 706-732, http://dx.doi.org/10.3102/1076998618784700. [42]

Bradshaw, L. (2016), "Diagnostic Classification Models", in *The Handbook of Cognition and Assessment*, John Wiley & Sons, Inc., Hoboken, NJ, USA, http://dx.doi.org/10.1002/9781118956588.ch13. [46]

Braun, H. and A. Kanjee (2006), "Using assessment to improve education in developing nations.", in Braun, H. et al. (eds.), *Improving Education Through Assessment, Innovation, and Evaluation.*, Cambridge, Mass.: American Academy of Arts and Sciences. [1]

Bundesministerium für Bildung und Forschung (n.d.), *ASCOT+*, https://www.ascot-vet.net (accessed on 30 May 2021). [26]

Chopade, P., D. Edwards, S. Khan, A. Andrade and S. Pu (2019), "CPSX: Using AI-Machine Learning for Mapping Human-Human Interaction and Measurement of CPS Teamwork Skills", *2019 IEEE International Symposium on Technologies for Homeland Security (HST)*, http://dx.doi.org/10.1109/hst47167.2019.9032906. [53]

Chung, G. (2014), "Toward the Relational Management of Educational Measurement Data.", *Teachers College Record*, Vol. 116/11. [15]

Ciolacu, M., A. Fallah Tehrani, L. Binder and P. Mugur Svasta (2018), "Education 4.0 - Artificial Intelligence Assisted Higher Education: Early recognition System with Machine Learning to support Students' Success", *2018 IEEE 24th International Symposium for Design and Technology in Electronic Packaging (SIITME)*, http://dx.doi.org/10.1109/siitme.2018.8599203. [49]

Conley, D. (2018), *The Promise and Practice of Next Generation Assessment.*, Harvard University Press, Cambridge, MA. [19]

Cordova, D. and M. Lepper (1996), "Intrinsic motivation and the process of learning: Beneficial effects of contextualization, personalization, and choice.", *Journal of Educational Psychology*, Vol. 88/4, pp. 715-730, http://dx.doi.org/10.1037/0022-0663.88.4.715. [34]

Darling-Hammond, L. (2006), "Constructing 21st-Century Teacher Education", *Journal of Teacher Education*, Vol. 57/3, pp. 300-314, http://dx.doi.org/10.1177/0022487105285962. [11]

Darling-Hammond, L., J. Herman, J. Pellegrino, J. Abedi, J. Lawrence Aber, E. Baker, R. Bennett, E. Gordon, E. Haertel, K. Hakuta, A. Ho, R. Lee Linn, P.D. Pearson, J. Popham, L. Resnik, A. Schoenfeld, R. Shalveson, L. Shepard, L. Shulman and C. Steele (2013), "Criteria for High-quality Assessment.", *Stanford Center for Opportunity Policy in Education.* [14]

Darrah, M. (2013), "Computer Haptics: A New Way of Increasing Access and Understanding of Math and Science for Students Who are Blind and Visually Impaired", *Journal of Blindness Innovation and Research*, Vol. 3/2, http://dx.doi.org/10.5241/3-47. [44]

de la Torre, J. and J. Douglas (2004), "Higher-order latent trait models for cognitive diagnosis", *Psychometrika*, Vol. 69/3, pp. 333-353, http://dx.doi.org/10.1007/bf02295640. [47]

Deterding, S., D. Dixon, R. Khaled and L. Nacke (2011), "From game design elements to gamefulness", *Proceedings of the 15th International Academic MindTrek Conference on Envisioning Future Media Environments - MindTrek '11*, http://dx.doi.org/10.1145/2181037.2181040. [38]

DiCerbo, K. (2014), "Game-Based Assessment of Persistence.", *Educational Technology & Society*, Vol. 17/1, pp. 17-28. [3]

Duncan, R. and C. Hmelo-Silver (2009), "Learning progressions: Aligning curriculum, instruction, and assessment", *Journal of Research in Science Teaching*, Vol. 46/6, pp. 606-609, http://dx.doi.org/10.1002/tea.20316. [9]

Echeverría, A., M. Améstica, F. Gil, M. Nussbaum, E. Barrios and S. Leclerc (2012), "Exploring different technological platforms for supporting co-located collaborative games in the classroom", *Computers in Human Behavior*, Vol. 28/4, pp. 1170-1177. [48]

Fadel, C., M. Bialik and B. Trilling (2015), *Four-dimensional Education: the Competencies Learners Need to Succeed.*, Center for Curriculum Redesign, Cambridge, MA. [30]

Ferrara, S., E. Lai, A. Reilly and P. Nichols (2017), "Principled Approaches to Assessment Design, Development, and Implementation", in The *Handbook of Cognition and Assessment*, John Wiley & Sons, Inc., Hoboken, NJ, USA, http://dx.doi.org/10.1002/9781118956588.ch3. [22]

Games for change (n.d.), Games for change, http://www.gamesforchange.org/game/simcityedu-pollution-challenge/ (accessed on 30 April 2021). [51]

Glasslab (n.d.), *SIMCITY edu: pollution challenge*, https://s3-us-west-1.amazonaws.com/playfully-games/SC/brochures/SIMCITYbrochure_v3small.pdf (accessed on 30 April 2021). [52]

Gobert, J., R. Baker and M. Wixon (2015), "Operationalizing and Detecting Disengagement Within Online Science Microworlds", *Educational Psychologist*, Vol. 50/1, pp. 43-57, http://dx.doi.org/10.1080/00461520.2014.999919. [37]

Gong, B. (2010), *Using balanced assessment systems to improve student learning and school capacity: An introduction.*, Council of Chief State School Officers, Washington, DC. [2]

Hamari, J., D. Shernoff, E. Rowe, B. Coller, J. Asbell-Clarke and T. Edwards (2016), "Challenging games help students learn: An empirical study on engagement, flow and immersion in game-based learning", *Computers in Human Behavior*, Vol. 54, pp. 170-179, http://dx.doi.org/10.1016/j.chb.2015.07.045. [33]

Klieme, E. (2020), "Policies and Practices of Assessment: A Showcase for the Use (and Misuse) of International Large Scale Assessments in Educational Effectiveness Research", in *International Perspectives in Educational Effectiveness Research*, Springer International Publishing, Cham, http://dx.doi.org/10.1007/978-3-030-44810-3_7. [25]

Madaus, G. and M. Russell (2010), "Paradoxes of High-Stakes Testing", *Journal of Education*, Vol. 190/1-2, pp. 21-30, http://dx.doi.org/10.1177/0022057410190001-205. [27]

Marion, S., J. Thompson, C. Evans, J. Martineau and N. Dadey (2019), "A Tricky Balance: The Challenges and Opportunities of Balanced Systems of Assessment.", in *Paper Presented at the Annual Meeting of the National Council on Measurement in Education Toronto, Ontario April 6, 2019.*, National Center for the Improvement of Educational Assessment, https://www.nciea.org/sites/default/files/inline-files/Marion%20et%20al_A%20Tricky%20Balance_031319.pdf (accessed on 2 January 2020). [23]

Martone, A. and S. Sireci (2009), "Evaluating Alignment Between Curriculum, Assessment, and Instruction", *0*Vol. 79/4, pp. 1332-1361, http://dx.doi.org/10.3102/0034654309341375. [17]

Mislevy, R. (2018), *Sociocognitive Foundations of Educational Measurement.*, Routledge, New York. [36]

Mislevy, R., R. Almond, D. Yan and L. Steinberg (2000), "Bayes nets in educational assessment: Where do the numbers come from?", *(CSE Report 518). Los Angeles: University of California, National Center for Research on Evaluation, Standards, and Student Testing (CRESST).* [45]

Mislevy, R., J. Behrens, K. Dicerbo and R. Levy (2012), "Design and discovery in educational assessment: Evidence-centered design, psychometrics, and educational data mining.", *Journal of educational data mining*, Vol. 4/1, pp. 11-48. [6]

Mislevy, R., A. Oranje, M. Bauer, A. von Davier, J. Hao, S. Corrigan, E. Hoffman, K. DiCerbo and M. John (2014), *Psychometric Considerations in Game-Based Assessment.*, Glasslab Games, Redwood City, CA. [50]

National Research Council (2001), *What Students Know: The Science and Design of Educational Assessment.*, National Academies Press, Washington, D.C., http://dx.doi.org/10.17226/10019. [13]

Nichols, S. and H. Dawson (2012), "Assessment as a Context for Student Engagement", in *Handbook of Research on Student Engagement*, Springer US, Boston, MA, http://dx.doi.org/10.1007/978-1-4614-2018-7_22. [12]

Oranje, A., B. Mislevy, M. Bauer and G. Tanner Jackson (2019), "Summative Game-based Assessment.", in Ifenthaler, D. and Y. Kim (eds.), Ga*me-based Assessment Revisited*, Springer. [20]

Pellegrino, J. and M. Hilton (eds.) (2012), *Education for Life and Work: Developing Transferable Knowledge and Skills in the 21st Century. Committee on Defining Deeper Learning and 21st Century Skills.*, National Academies Press, Washington, D.C., http://dx.doi.org/10.17226/13398. [18]

Perie, M., S. Marion and B. Gong (2009), "Moving Toward a Comprehensive Assessment System: A Framework for Considering Interim Assessments", *Educational Measurement: Issues and Practice*, Vol. 28/3, pp. 5-13, http://dx.doi.org/10.1111/j.1745-3992.2009.00149.x. [10]

Rose, D. (2000), "Universal Design for Learning", *Journal of Special Education Technology*, Vol. 15/3, pp. 45-49, http://dx.doi.org/10.1177/016264340001500307. [43]

Sabourin, J., J. Rowe, B. Mott and J. Lester (2011), "When Off-Task is On-Task: The Affective Role of Off-Task Behavior in Narrative-Centered Learning Environments", in *Lecture Notes in Computer Science, Artificial Intelligence in Education*, Springer Berlin Heidelberg, Berlin, Heidelberg, http://dx.doi.org/10.1007/978-3-642-21869-9_93. [40]

Sanders, W. and S. Horn (1995), "Educational Assessment Reassessed", *education policy analysis archives*, Vol. 3, p. 6, http://dx.doi.org/10.14507/epaa.v3n6.1995. [8]

Seelow, D. (2019), "The Art of Assessment: Using Game Based Assessments to Disrupt, Innovate, Reform and Transform Testing.", *Journal of Applied Testing Technology*, Vol. 20/S1, pp. 1-16. [32]

Shaffer, D., D. Hatfield, G. Navoa Svarovsky, P. Nash, A. Nulty, E. Bagley, K. Frank, A. Rupp and R. Mislevy (2009), "Epistemic Network Analysis: A Prototype for 21st-Century Assessment of Learning", *International Journal of Learning and Media*, Vol. 1/2, pp. 33-53, http://dx.doi.org/10.1162/ijlm.2009.0013. [4]

Shepard, L., W. Penuel and J. Pellegrino (2018), "Using Learning and Motivation Theories to Coherently Link Formative Assessment, Grading Practices, and Large-Scale Assessment", *Educational Measurement: Issues and Practice*, Vol. 37/1, pp. 21-34, http://dx.doi.org/10.1111/emip.12189. [21]

Shute, V. (2011), *Stealth assessment in computer-based games to support learning. Computer games and instruction.*, Information Age Publishers, Charlotte, NC, http://myweb.fsu.edu/vshute/pdf/shute%20pres_h.pdf. [5]

Shute, V. and M. Ventura (2013), "Stealth Assessment: Measuring and Supporting Learning in Video Games.", in John, D. and C. MacArthur (eds.), *Foundation Reports on Digital Media and Learning.*, The MIT Press, Cambride, MA, http://dx.doi.org/10.7551/mitpress/9589.001.0001. [35]

Shute, V., M. Ventura, M. Bauer and D. Zapata-Riviera (2009), "Melding the power of serious games and embedded assessment to monitor and foster learning.", *Serious games: Mechanisms and effects*, Vol. 2, pp. 295-321. [7]

Snow, E., L. Allen, M. Jacovina and D. McNamara (2015), "Does agency matter?: Exploring the impact of controlled behaviors within a game-based environment", *Computers & Education*, Vol. 82, pp. 378-392, http://dx.doi.org/10.1016/j.compedu.2014.12.011. [41]

Stecher, M. and L. Hamilton (2014), Measuring hard-to-measure student competencies: A research and development plan., RAND Corporation, Santa Monica, CA, https://www.rand.org/pubs/research_reports/RR863.html. [31]

Trilling, B. and C. Fadel (2009), *21st century skills: Learning for Life in Our Times.*, Jossey-Bass. [28]

Verger, A., L. Parcerisa and C. Fontdevila (2019), "The growth and spread of large-scale assessments and test-based accountabilities: a political sociology of global education reforms", *Educational Review*, Vol. 71/1, pp. 5-30, http://dx.doi.org/10.1080/00131911.2019.1522045. [24]

Vincent-Lancrin, S., C. Gonzalez-Sancho, M. Bouckaert, F. de Luca, M. Fernandez-Barrerra, G. Jacotin, J. Urgel and Q. Vidal (2019), *Fostering Students' Creativity and Critical Thinking: What it Means in School*, Educational Research and Innovation, OECD Publishing, Paris, https://dx.doi.org/10.1787/62212c37-en. [29]

Yang, F., L. Leqi, Y. Wu, Z. Lipton, P. Ravilkimar, W. Cohen and T. Mitchel (2019), "Game Design for Eliciting Distinguishable Behavior.", *Paper prepared for the 33rd Conference on Neural Information Processing Systems.*, https://papers.nips.cc/paper/8716-game-design-for-eliciting-distinguishable-behavior.pdf (accessed on 2 January 2020). [39]

11 Blockchain for Education: A New Credentialing Ecosystem

Natalie Smolenski
Hyland Credentials, United States

Blockchain technology is revolutionising the world of financial services by providing distributed networks for transacting digital currencies. This same digital infrastructure can be used to verify important claims and credentials, including educational and academic records. Within education, significant momentum exists worldwide to use blockchain technology for issuing, sharing, and verifying educational experiences and qualifications. This chapter provides an overview of blockchain technology and spotlights its use in education to create portable, interoperable, user-controlled digital credentials. These verifiable claims constitute a form of social currency that empowers students and workers with the ability to transfer their competencies and skills anywhere in the world they choose to live, study, and work.

Introduction

This chapter demonstrates the value of blockchain technology – a new digital verification infrastructure – for education. In short, blockchain technology enables anyone to validate claims about an individual or institution, including their characteristics and qualifications, and to do this instantly and with a very high level of certainty. This helps eliminate records fraud; facilitates the movement of learners and workers between institutions and geographies; and empowers individuals by giving them increased control over their own data.

The chapter is divided into several sections: 1) a general introduction to blockchain technology; 2) an explanation of how it benefits education; 3) a global survey of blockchain implementations for education led by both public and private sector entities; 4) and finally, a series of recommendations for policymakers and educational institutions regarding how to best approach the new technology.

The purpose of the first section of this chapter, "Understanding Blockchain Technology", is to help readers unfamiliar with blockchain to understand how it works before applying it to use cases in education. It provides an overview of the history and main concepts of blockchain as well as its main functions and use cases. In short, the main value of blockchain is to provide a mechanism to achieve consensus between multiple untrustworthy parties. The focus of this section is on digital currencies as the first concrete application of blockchain technology, but it also highlights other possible use cases.

Readers who understand blockchain technology well may therefore skip this first section and move directly to section 2, "Benefits of Blockchain for Education", which demonstrates how blockchain can be used to reliably verify and transfer degrees, academic records, and other types of credentials. While blockchain shows potential promise for other educational use cases, such as streamlining administrative processes, credentialing is by far the most mature application of blockchain technology in education today and therefore the focus of this report.

The third section, "Real-World implementations," provides the reader with a survey of the global market for verifiable digital credentialing solutions employing blockchain technology. As this market is rapidly evolving, this section can be considered current as of the second quarter of 2021 but will likely change significantly in the following months and years. One of the most exciting testimonies to the promise of blockchain technology has been its immediate global uptake; this is true for education as well as for other industries.

As a truly global network technology, therefore, blockchain mirrors the World Wide Web. The fourth section, "Driving Change", therefore recommends that governments, solution providers, and educational institutions privilege international portability and platform interoperability for blockchain credentialing solutions they adopt or build. The COVID-19 crisis has only underscored the need for secure, instant, verifiable digital data transfer between institutions as well as between individuals and institutions. As with any digitisation project, moving to a verifiable digital credentialing model requires both budgetary investment and change management. However, by choosing solutions and providers building on open standards, organisations can future-proof their projects and ensure their long-term viability.

Ultimately, a shift to blockchain-enabled, verifiable credentials creates a global ecosystem of interoperable records, a new scaffolding of trust to underpin the global mobility of labour and services. It also supports a "lifelong learning" model of education, in which traditional educational institutions are stops along a pathway of ongoing personal development that occurs both within and beyond educational institutions. The 21st century economy, and the digital natives who have been born and come of age in this economy, have internalised this model of personal development and expect their credentials to reflect the same convenience, security, and portability as other digital data formats. Educational institutions are therefore in the exciting and perhaps unprecedented position of serving as a global vanguard for new technology adoption: academic records, secured by transnational digital ledgers that become the first non-currency use cases for a next-generation social currency of trust.

Understanding blockchain technology

A breakthrough in computer science

Until 2009, digital money never saw wide adoption. Why? Because it was too easy to "double spend" it: too easy for someone to pretend they have digital money they do not have and spend it over and over. The ability for anyone to create money at will would render any form of such money worthless.

Computer scientists spent decades attempting to solve this problem, which finally saw an elegant resolution with the Bitcoin protocol for peer-to-peer electronic cash transactions (Nakamoto, 2008[1]). Bitcoin was the first "blockchain": the first chain of uneditable transactions validated by adversarial nodes in a network. But we will return to what blockchain and Bitcoin are in a moment. Before we can truly understand the significance of the blockchain breakthrough, we first need to understand how it solves an even more foundational computer science problem: the Byzantine Generals Problem (BGP).

The BGP involves coordinating action among multiple parties, some of whom are untrustworthy and unreliable. In the canonical BGP example, multiple Byzantine Generals are planning to attack a city, but the attack will only be successful if they *all* attack at the same time. If *all* of these parties cannot coordinate on a single action, everyone loses, but if they all coordinate on the same action, everyone wins (Moskov, 2018[2]). It is a binary situation: complete failure or total victory. A total victory requires something very hard to come by in any social group: consensus.

Computer systems that reliably produce consensus among multiple untrustworthy or adversarial parties are known as "byzantine fault-tolerant" (BFT). Examples of systems that require byzantine fault tolerance are aircraft flight control systems, spacecraft flight control systems, nuclear power systems, and digital currency systems (Wikipedia, n.d.[3]; Binance Academy, n.d[4]). These systems utilise different solutions to the Byzantine Generals Problem, or different *consensus mechanisms.* A consensus mechanism is a way of arriving at consensus among parties who do not trust one another. Different consensus mechanisms emerged to address the different needs and purposes of various byzantine fault-tolerant systems.

For example, if you are trying to track a stock-keeping unit (SKU) throughout multiple stops in a supply chain, and you do not trust all the links in that supply chain to always tell you the truth, you may want a BFT system with certain characteristics. On the other hand, if you are trying to track the spending of digital currency, and everyone in the network is incentivised to pretend they have more currency than they actually do, an even more robust consensus mechanism is necessary. Not all BFT systems are blockchains. But Bitcoin, the first implementation of blockchain technology, is a BFT system that emerged to track the ownership and spending of digital currency. In addition to solving the BGP, it also solved the double-spend problem through a consensus mechanism known as "proof of work" (Wikipedia, n.d.[5]).

Think of proof of work as CAPTCHA, but for computers. With CAPTCHA, a website requires you to prove that you are a human by solving a puzzle that is easy for a human to solve, but hard for a computer to solve. If you can solve this puzzle, the receiving computer will allow you to send data or subscribe to a service. Similarly, computers that are asked by other computers to do something for them – like send money from one party to another – may also require that requesting computers demonstrate that they have solved a complex puzzle before their request is granted. Verifying that this puzzle has been solved by requesting computers is "proof of work."

Proof of work is intended to make communication between computers more costly in order to eliminate frivolous and fraudulent use of the network. For example, proof of work prevents Distributed Denial of Service (DDoS) attacks and SPAM – where adversarial users throw sand in the gears of a network by flooding it with malicious or nonsensical data. As computers get faster and smarter, proof of work becomes harder and harder – more expensive in terms of processing power. This is necessary in order to keep BFT networks secure, stable, and usable. But the tendency towards more processing power also means increased energy usage, which is why proof-of-work systems like Bitcoin have been criticised for their environmental impacts (Temple, 2019[6]). In response, clusters of computers running the Bitcoin protocol now frequently use renewable energy sources and are located in cold climates in order to save on energy required to cool them (Baydakova, 2019[7]; Morris, 2018[8]). Despite Bitcoin's relatively high energy usage, however, several influential studies that quantified that energy usage in the past have been demonstrated to be profoundly inaccurate (DiChristopher, 2017[9]). Claims that the Bitcoin network would use as much electricity as the rest of the world by 2020 have manifestly not come to pass (Cuthbertson, 2017[10]). And as computer processing power becomes more energy efficient, the Bitcoin network does too – without sacrificing the security and usability needed to serve as a long-term store of value for digital currency (American Chemical Society, 2019[11]). In fact, a recent white paper by investment management firm Ark Invest suggests that Bitcoin mining could play a key role in incentivising the shift towards renewable energy by helping to manage supply and demand in energy markets (Ark Invest, 2021[12]).

So, Bitcoin is a byzantine fault-tolerant network designed to securely maintain and even grow the value of its digital currency by preventing the double-spend problem through proof of work. But why is it called a blockchain?

Blockchains and distributed ledgers

As mentioned above, not all byzantine fault-tolerant systems are blockchains. While there is debate about what exactly makes a blockchain different from other BFT systems, there is broad agreement that the word "blockchain" refers to a type of distributed ledger that records an append-only, immutable database of transactions. The word "blockchain" was first used in connection with Bitcoin to refer to a ledger of digital currency transactions. Bitcoin's inventor, Satoshi Nakamoto, initially called it a "proof-of-work chain" or a "timechain" because transactions are appended to the ledger in a sequential, timestamped manner that preserves the record permanently in the order it was created (Messari, 2019[13]). In addition, these transactions are grouped together in "blocks", which are deposited to the chain after a set period of time. (In the case of Bitcoin, new blocks of transactions are added to the chain every 10 minutes.) This block-based architecture easily lent itself to the term "blockchain," which was "coined" by the early Bitcoin community.

What about BFT systems that do not rely on blocks of transactions appended to an immutable database? Those are sometimes called "distributed ledgers" (DLTs). Distributed ledgers may or may not track the allocation of currency in a network and usually use consensus mechanisms different from proof of work. They may even be editable by parties with the right permissions. Some people separate DLTs from blockchains and treat them as different things, while others consider blockchains to be a specific type of DLT.

Box11.1 Public/Private, open/permissioned - What's the difference?

Blockchains have rules about who can write transactions to their network and who can run nodes of the chain and validate those transactions. The following guidelines provide a general rubric for understanding the differences between different types of blockchains.

Writing to chain:

Public: Anyone can write transactions to the chain and view those transactions once they have been written.

Private: Only permissioned parties may write transactions to the chain and view transactions once they have been written.

Running nodes and validating transactions:

Open: Anyone can run nodes and validate transactions on the chain. ("Open" in this context does not mean "open source". Many private blockchains are based on an open source codebase but are private because use of the network is gated.)

Permissioned: Only permissioned parties may run nodes and validate transactions on the chain.

Blockchains can be any of the following combinations: Public/Open, Public/Permissioned, or Private/Permissioned. (Although Private/Open blockchains are possible in theory, in practice all private chains are also permissioned.)

Note that just because a transaction has occurred on a public blockchain and anyone can view it does not necessarily mean anyone can understand what it means. This is because public blockchains usually record transactions pseudonymously or anonymously and hash the transaction content so that it is not readable by people. This is a necessary step to preserve privacy on a public platform. (This is also why the criticism of public blockchains as not privacy preserving is often misleading.)

In order to verify transactions on public blockchains, a human user relies on technical standards that provide a series of technical steps for verifying those transactions. This verification is generally facilitated by an application, like a web browser or mobile wallet. Applications that verify past transactions may do so in an open and free manner or gate verification behind a paywall or some other access barrier.

For example, some credentialing companies that use public blockchains to anchor transactions charge for credential verification as part of their business model. With permissioned or private blockchains, credential verification may also be gated either behind a paywall or behind some type of authentication service.

We do not have to choose a side in this debate to understand that the distinction between blockchains and DLTs arose because different BFT networks are created for different purposes. The major differentiating factor between different BFT networks is their relative level of centralisation or decentralisation. In other words, to what extent is a network's infrastructure supported by, and consensus determined by, an authority or authorities? For some use cases, a more centralised blockchain architecture makes sense; for others, it does not. Below we examine different implementations of more or less decentralised blockchains in light of the stakeholders making use of them for particular purposes.

Decentralisation vs. centralisation

It is useful to see Bitcoin as an archetypal example of a decentralised blockchain. It is decentralised because of the main problem it arose to solve: the problem of an infinitely growing money supply that eventually devalues the money itself. Governments throughout history have been incentivised to create more and more money in order to fund wars, prop up markets, pay back debts, and for many other reasons (Bhatia, 2021[14]). The problem with this is that over time, it results in devaluation of the currency, which destroys the savings and wealth of ordinary people (Alden, 2021[15]). To prevent currency devaluation, Bitcoin was created to be digitally scarce, with a fixed supply of

21 million coins. No individual or government can change this due to the rigidity of the underlying bitcoin protocol. This means that the bitcoin currency was designed to be scarce and therefore to increase in value over time. (The Bitcoin protocol is usually referred to in uppercase, while the bitcoin cryptocurrency is not capitalised. The bitcoin cryptocurrency is sometimes abbreviated as BTC.)

Some in the Bitcoin community expect a dramatic and permanent spike in the value of bitcoin due to a process known as "hyperbitcoinisation" – in effect, the opposite of hyperinflation (Krawisz, 2014[16]; Kenobit, 2018[17]). Although Nakamoto never said so explicitly, the way he architected Bitcoin's monetary policy suggests that he expected the value of one bitcoin to rise so high that most people would never own a full coin. For this reason, Nakamoto subdivided each bitcoin into 100 million satoshi. Thanks to satoshi, people do not need to own even one bitcoin in order to transact with the currency and use it as a long-term store of value.

Nakamoto's intent to protect the value of bitcoin from the interventions of activist monetary policy makes it clear that bitcoin is designed to be a stateless currency. That means it is architected to avoid the possibility of being captured by any centralised authority. Because the Bitcoin blockchain is open and public, anyone is able to run full or partial nodes of the protocol and commit transactions to the network using proof of work. Transactions on the network are pseudonymous and moving towards anonymity (Nuzzi, 2019[18]). All of the network code is open source, meaning anyone can contribute to it, and it is maintained on a volunteer basis by a group of core developers and cryptographers. People who are unhappy with how Bitcoin is governed can also fork the code to create competing blockchains, and the open marketplace lets users decide which chain they want to use to store their value and transact their money. So far, bitcoin has by far the largest market capitalisation of any blockchain-based cryptocurrency. As of April 2021, the cryptocurrency is worth over USD 1 trillion, the fastest asset to reach this milestone in recent history (Ali, 2021[19]). The next most valuable cryptocurrency, ether (ETH), is currently worth about one-third of the Bitcoin network, at over USD 300 billion (CoinMarketCap, n.d.[20]). We will return to Ethereum, the blockchain network underpinning ether, below.

Understandably, this level of decentralisation has not endeared Bitcoin to some governments (McWhinney, 2019[21]). Banks – particularly central banks – have been especially reticent about the stateless cryptocurrency (Steenis, 2017[22]), partly because it offers an alternative store of value and medium of exchange that remains unaffected by their monetary policy interventions. In addition, many companies and industry consortia are reluctant to use open, public blockchains, preferring private or permissioned-public models that allow them to exert more control over who gets to use the network and in what capacity. In addition, decentralisation does come with technical trade-offs: greater near-term efficiencies in network speed and transaction processing are often easier to achieve in more centrally-controlled networks. This is because networks that gate access only to trusted parties eliminate some of the need for byzantine fault tolerance: the more you trust the participants in a network to behave honestly, the less need you have for robust BFT, which tends to slow down a network's decision-making.

In part due to dissatisfaction and discomfort with Bitcoin – but also because Bitcoin inspired a generation of developers with the potential to generate enormous personal wealth on the basis of token seigniorage by minting their own cryptocurrencies – other cryptocurrencies and DLT architectures quickly arose which may preserve some aspects of the decentralisation of the Bitcoin network, but with more centralised characteristics that presume a higher concentration of trusted actors. Some are quite open about being fully centralised systems. As noted above, many of these centralised blockchains have no associated cryptocurrency and instead act as infrastructure for the delivery of services, for which their architects may charge. Accordingly, we can think of decentralisation as a spectrum; it can manifest to a greater or lesser degree in different parts of a blockchain's architecture.

Here it is also important to note that there is considerable controversy over what constitutes decentralisation. Some architectural decisions that read as decentralised to some may appear highly centralised to others. Part of the difficulty is that the blockchain space is laden with value judgments and strong ideological commitments. This means that for some, decentralisation is simply and straightforwardly good, while centralisation is bad. The opposite also holds true; some have castigated decentralised architectures as immoral, criminal, or worse. This is an area of ongoing and often heated debate. In this paper, I do not adopt "good" or "bad" value judgments with regards to decentralisation or centralisation. Rather, I stress that these characteristics of a system will necessarily be adapted to what that system is made for.

Ethereum, the second most valuable blockchain today in terms of market capitalisation (CoinMarketCap, n.d.[20]) is an example of a blockchain modelled on the highly decentralised architecture of Bitcoin. Like Bitcoin, Ethereum is an open, public blockchain; anyone can transact on the network, run nodes, and contribute to the network's open source code. In 2018, the US Securities and Exchange Commission determined that both bitcoin and ether cannot be considered securities because no third party carries out "entrepreneurial and managerial functions" with regards to their issuance and value (Sharma, 2018[23]; Pisani, 2018[24]).

However, there are two major differences between Bitcoin and Ethereum which have contributed to their very different developmental trajectories. One is that, whereas Bitcoin was built to be the most secure form of digital money available, Ethereum was built to be a distributed "virtual machine"—an infrastructure for running decentralised applications. Ethereum's associated cryptocurrency, ether, is called "gas" because its primary (though not only) function is to power applications on the network rather than to serve as a long-term store of value or universal medium of exchange. As a virtual computing platform, Ethereum allows the inclusion of executable code known as "smart contracts" within its blockchain transactions. Smart contracts allow organisations to automate and coordinate action at scale. Using *if* -> then logic, smart contracts execute certain functions (generally the payment of money) when contractually-defined conditions are met (CoinTelegraph, n.d.[25]). This feature of Ethereum has given rise to a class of applications called "decentralised applications," or DApps, whose bylaws and procedures are embedded in smart contracts (Wikipedia, n.d.[26]) Although smart contracts can also be executed on Bitcoin, Ethereum's more extensive and developer-friendly smart contracts functionality has made it the most popular blockchain choice for DApps and the most actively used blockchain today (Leising and Kharif, 2020[27])

Another major difference between Bitcoin and Ethereum is that Bitcoin's founder chose to remain anonymous and withdrew from visible participation in the network's management two years after the network was launched (Bernard, 2018[28]). Although many individuals have claimed to be Satoshi Nakamoto in the ensuing years, none have been able to prove this conclusively, and none have been able to exercise significant influence over the network's technical roadmap. By contrast, Ethereum's main inventor, Vitalik Buterin, is well-known and maintains an active role in the governance of the Ethereum network. For example, Buterin has been one of the primary drivers behind the decision to transition Ethereum from a proof-of-work (PoW) to a proof-of-stake (PoS) consensus mechanism, a change referred to as Ethereum 2.0 (Foxley, 2020[29]).

Buterin's prominent role in Ethereum network governance was perhaps most visible after a famous 2016 incident known as the "DAO hack". An attacker exploited a bug in an Ethereum smart contract run by a Decentralised Autonomous Organisation (DAO) to appropriate at least USD 89 million worth of ether (Kar, 2016[34]). Buterin led a group that hacked the hacker and re-appropriated the funds, but an anonymous letter purportedly from the attacker claimed that they had taken them legally, as they were only performing actions allowed by the smart contract (and "code is law") (Siegel, 2016[35]). Buterin, however, feared that acquiescing to the hacker's logic would destroy investor trust in Ethereum and any future DApps and DAOs built on it. So he proposed a "hard fork" of the network, which would roll it back to a time before the attack and return the stolen money to DAO investors.

The decentralised nature of the Ethereum network did mean that the majority of miners running the Ethereum protocol still had to choose to adopt Buterin's proposal, but his stature in the community, and the support he enjoyed from other Ethereum founders, made it likely that they would. Indeed, most Ethereum developers did choose to run the forked code, making the new, rolled-back Ethereum (ETH) the most widely-accepted version of the blockchain. Some, however, refused; they continued running the un-forked blockchain, which is now known as Ethereum Classic (ETC). By contrast, none of the people claiming to be Satoshi Nakamoto were able to influence the technical roadmap of Bitcoin during the "Bitcoin Civil War" of 2017 (Dinkins, 2017[36]). Instead, the Civil War resulted in a number of minor Bitcoin forks – Bitcoin Cash (BCH), Bitcoin Gold (BTG), Bitcoin Satoshi's Vision (BSV) – which to this day remain much smaller chains from a value and adoption standpoint compared to the original Bitcoin (BTC) (CoinMarketCap, n.d.[20]).

The ongoing involvement of Buterin and the other Ethereum founders in the network's governance has enabled different possibilities for the network – in particular its adaptation for use by enterprises and governments looking to streamline core business applications and services. Ethereum's smart contracts functionality enables organisations to execute tasks that require a considerable amount of coordination between parties who may not trust or even know one another: settle payments, provide escrow services, pay out investor dividends, enable secure electronic voting, track SKUs through a supply chain, and many other functions. The company ConsenSys, founded by Ethereum co-founder Joseph Lubin, focuses entirely on delivering enterprise-ready DApps using Ethereum (ConsenSys, n.d.[37]).

Box 11.2 Proof of stake

Proponents of proof of stake (PoS) argue that the ever-increasing energy cost of proof-of-work (PoW) mining has made running blockchain nodes prohibitive for ordinary people because it has resulted in a centralisation of mining power among powerful groups called "mining pools" (Muzzy, 2020[30]). To remedy this, PoS requires the miner (or "forger," as they are sometimes called to differentiate them from users of PoW) to put their own cryptocurrency at risk to validate a transaction rather than expending energy to solve complex mathematical problems. The penalty for dishonest dealing in a PoS system is having one's cryptocurrency confiscated to various degrees. In the case of Ethereum, moving to a PoS system has also been explained as the necessary precondition for securely scaling the network's transactional capacity.

However, PoS has its critics as well. Research firm Messari has pointed out that PoS has the potential to create a system of class-based blockchain governance, where the wealthiest participants in a blockchain network have the most say over what is true, or validated on the network (Watkins, 2020[31]). This risk is particularly acute for those blockchains that implement "on-chain governance" – a way of placing control over a network's technical roadmap in the hands of token holders rather than core developers, who are sometimes perceived to be unaccountable to network users (Messari.io., 2019[32]). Simply put, there is a risk that wealthy token holders could skew decisions about the governance of PoS systems to maximise profit for themselves.

Wealthy token holders may be the initial investors in a network who receive large cryptocurrency allocations in return for early support. They may also be later investors who purchase large amounts of the network's cryptocurrency and use it to frequently validate transactions, thus accruing more and more rewards for validation – a "rich get richer" dynamic that eventually squeezes out smaller crypto holders from the network (Watkins, 2020[31]). Other critics of PoS argue that scaling the transactional capacity of a blockchain network doesn't necessarily need to occur on-chain. For example, the need to scale "slow" PoW systems like Bitcoin has given rise to a thriving industry of "Layer 2" applications that allow users to conduct most transactions off-chain while using the blockchain only as a final settlement layer (Coin Telegraph, n.d.[33]).

It is important to note here that no blockchain network – Bitcoin, Ethereum, or otherwise – has arisen to solve the problem of wealth inequality or redistribution. Today, owners of every cryptocurrency are a very small number of people compared to the world's population, which necessarily skews the distribution of cryptocurrency wealth. While blockchain networks are critical interventions for helping us reimagine governance in the 21st century, their social effects are highly targeted and should not be mistaken as replacements for all other policy and social platforms.

But many established institutional actors are not interested in using a public, open digital infrastructure to implement smart contract-based applications. The Enterprise Ethereum Alliance (n.d.[38]) was formed in order to build versions of Ethereum that are public/permissioned or private and which have purpose-built functionality for the specific use cases targeted by client enterprises (Sharma, 2019[39]). The EEA has grown to over 200 companies implementing centralised versions of Ethereum that are funded and managed by specific companies or governments. One prominent example of an enterprise Ethereum implementation is the European Blockchain Services Infrastructure (EBSI), a public-permissioned blockchain created by the European Commission to deliver cross-border public services (CEF Digital, n.d.[40]). In the private sector, EEA member J.P. Morgan spearheaded the development of Quorum, a permissioned implementation of Ethereum built specifically for the financial industry (Quorum, n.d.[41]).

In 2015, the Linux Foundation announced the Hyperledger Project, an initiative by a consortium of technology, finance, and supply chain companies to build private blockchains that allegedly improve upon the scalability and reliability of existing public chains (Wikipedia, n.d.[42]). The Hyperledger Project now includes multiple blockchain "toolkits" that enable organisations or consortia to create private blockchains that meet their specific business requirements. None of these toolkits create blockchains with cryptocurrencies, leading many to refer to them as distributed ledgers instead. Popular toolkits today include Hyperledger Fabric and Hyperledger Sawtooth. Hyperledger Besu is an Enterprise Ethereum codebase that was contributed in 2019 (Castillo, 2019[43]). In the credentialing world, Hyperledger toolkits have been used by companies including Salesforce, Workday, and Greenlight to build blockchain networks that they manage (Lemoie and Souares, 2020[44]). Companies including Amazon, IBM, Oracle and SAP also use Hyperledger toolkits to provide "Blockchain-as-a-Service" (BaaS) offerings for clients using their cloud infrastructures (Amazon, n.d.[45]; IBM, n.d.[46]; Oracle, n.d.[47]; SAP, n.d.[48]). Some non-profit associations also use Hyperledger toolkits (among others) to build blockchain infrastructures for the delivery of public and commercial services. A few well-known examples include the Alastria consortium in Spain (Alastria, n.d.[49]) and the LACChain initiative from the IDB Lab (ConsenSys, 2020[50]).

The blockchain platform R3 Corda is another example of a private, enterprise-ready blockchain implementation. Founded in 2014, R3 is a private corporation that built a blockchain to manage and synchronise financial agreements (Wikipedia, n.d.[51]). The Corda platform, open sourced in 2016, was built to conform to the technical requirements and standards of the financial industry. Since that time, Corda has evolved to accommodate the needs of multiple industries (its main competitor is arguably Hyperledger Fabric, which was developed a few years later) (Shin, 2020[52]). Corda's conservative approach to programming languages and cryptographic standards appeared to be rewarded recently by an IEEE study which identified it as the only blockchain platform compatible with NIST standards and therefore qualified for use by the US Federal Government (Jessel, 2020[53]). In an interesting twist, however, the first US Federal Government implementation of blockchain is actually a procurement platform for the Department of Health and Human Services built on Hyperledger Fabric (Malone, 2020[54]).

Finally, a growing crop of cryptocurrencies pegged to the value of national fiat currencies – so-called "stablecoins" – have arisen to facilitate exchange of fiat currency into cryptocurrencies and vice versa. These stablecoins are typically "run and entirely controlled by the private companies that issue them and receive traditional bank deposits in return for issuing their respective currencies to their customers" (Dagger, 2020[55]). However, decentralised stablecoins are emerging as well, largely to circumvent a number of cumbersome and contradictory legal injunctions against the use of centralised stablecoins (Fitzpatrick, 2020[56]). As state central banks begin issuing digital fiat currencies, however, the transition to government-run centralised stablecoins will likely accelerate.

Costs and limitations

Blockchains are no small undertaking. They emerged to solve the Byzantine Generals Problem: how do you create consensus among a large number of adversarial actors who cannot be trusted? The answer – a decentralised network of nodes validating transactions according to a shared consensus mechanism – only works if it is deployed at scale. If a blockchain has too few validators, it is easy for malicious actors to attack it by gaining enough network power to override the consensus in their favour or to someone else's detriment. This is known as a 51% attack. Conversely, if a small network consists only of trusted actors, then it doesn't really need byzantine fault tolerance, making it generally much less expensive and more efficient to use traditional databases and applications.

Creating, deploying, and operating a blockchain network is a massive undertaking. Public blockchains finance their operations through stakeholder incentives tied to their native cryptocurrencies: the reward for mining bitcoin and ether by validating transactions on these networks (and thereby keeping them running and secure) is the bitcoin or ether cryptocurrency itself. In addition, transaction fees are assessed and paid to block validators in the network's cryptocurrency.

Satoshi Nakamoto designed Bitcoin with a limited supply of bitcoin (21 million) and the assumption of hyperbitcoinisation: that, if the network became successful, the value of bitcoin would skyrocket and much of the global economy would eventually run on the basis of bitcoin transactions. This means that as revenue from mining decreases and use of the network increases, the cost of running and maintaining Bitcoin could be subsidised by transaction fees alone. (Think of VISA or MasterCard as contemporary analogues.)

Ethereum founders, by contrast, did not cap the absolute amount of ether. Ether are awarded for every block mined, with an amount that has decreased over time (EthHub, n.d.[57]; ConsenSys, 2019[58]) and an annual cap of 18 million new ether (Sharma, 2019[59]). Vitalik Buterin made a bet that ether would not reach 100 million in circulation for at least a century (Buterin, 2016[60]), but due to the network's popularity, this milestone was instead achieved in 2018 (Varshney, 2018[61]). As a result, Buterin has proposed capping the total amount of ether at

120 million to avoid inflation (or 144 million if that cap is missed), but this proposal has not yet been accepted (Varshney, 2018[61]). Like Bitcoin, users of Ethereum also pay transaction fees, but more importantly, smart contracts require ether (gas) to run. In other words, Ethereum staked its future dominance not on becoming the world's reserve currency, like Bitcoin, but on becoming the world's computer. This combination of transaction fees and gas could finance the Ethereum network in the event that mining rewards disappear.

Permissioned and private blockchains, by contrast, usually do not have corresponding cryptocurrencies. This means they must be funded at a large cost that easily adds up to hundreds of thousands to millions of dollars each year. In some cases, like managing a supply chain for an entire industry, such a cost may be justified and can be distributed among member organisations to generate return on investment. But in cases where only one organisation is seeking to streamline internal processes, the cost usually far outweighs the value. Using blockchains where a private database will do then becomes an expensive, over-engineered experiment likely to end in failure.

The sheer scale needed to make operating blockchains cost-effective is beginning to swing the pendulum of the market back towards public blockchains. After all, blockchains only provide value through network effects, but each private blockchain is, by design, its own gated network. If each company develops its own private blockchain, the world may end up with a sea of private networks built on different standards that are not interoperable. The problem of data silos repeats itself all over again – this time, after considerable R&D expenditure.

Ernst & Young has been the most vocal industry player, arguing that the era of private blockchains is effectively over. Paul Brody, EY's Global Blockchain Leader, writes:

> *"what has typically happened thus far is that for every organization willing to join a consortium or someone else's network, two companies started their own. Not much hope for a network effect here. Blockchains, like most network technologies, are natural monopolies. The more users they have, the more useful they become, and once they achieve a position of dominance, [private blockchains] tend to start aggressively extracting "rents" from their members. Having seen this pattern in several consumer applications, at EY we believe most enterprises will prefer to stick to their legacy tools rather than let themselves experience a similar fate facing an all-powerful private blockchain monopolist in the B2B space. Public blockchains like Ethereum offer a better choice for enterprise users. Even if they do achieve monopoly-like dominance, there is no controlling entity to extract excess profits — there is only an ecosystem of competing service providers." (Brody, 2019[62])*

Brody compares public blockchains to the internet, which is a shared infrastructure that is not controlled by a single enterprise but underpins value for all enterprises. This is an apt analogy. During the early days of the internet, many companies and governments insisted on having their own "intranets" – private networks that could only be accessed internally. Over time, however, only the most mission-critical intranets remained, while most digital business processes transitioned to the World Wide Web. A similar trend is playing out now in the blockchain world, with many enterprises insisting on private blockchains in the belief that they are more secure and efficient, only to realise after considerable investment that they have undercut the network's value by limiting access to it.

Brody's contention that public blockchains have matured considerably with regards to privacy may benefit from some context. From the beginning, public blockchains have restricted the kind of data that can be placed on-chain to public keys, hashes and, in some cases, smart contracts. However, it is possible for hackers to "game" public smart contracts – as happened with the 2016 DAO hack – by exploiting vulnerabilities. Similarly, public keys tied to known identifiers can be used to perform "network analysis," tracking a user's transactions on a blockchain (Chainalysis, n.d.[63]). This capability is used by law enforcement to track criminal financial activity on blockchains (Weinstein, 2015[64]). Although hashes are virtually impossible to reconstruct to reveal transaction data, there is ongoing debate about whether or not they could constitute personally identifiable information if used incorrectly (Mearian, 2018[65]; Longstaff, 2018[66]).

Zero-knowledge proofs (ZKPs) address these concerns by concealing all transaction data on public blockchains. While this is attractive to many users of blockchains, some governments and law enforcement agencies are concerned about the use of ZKPs due to the limits it would place on their ability to investigate financial crimes. Until recently, Zcash and Sovrin (a codebase also donated to the Linux Foundation under the name Hyperledger Indy) have been the main public blockchains employing zero-knowledge proofs (Sharma, 2019[67]; The Sovrin Foundation, 2018[68]). Enabling this functionality for Bitcoin and Ethereum could be particularly valuable due to the large network effects of those networks. Today, however, there are a variety of approaches to data minimisation (restricting the type and amount of data that can be verified) using blockchain networks in addition to ZKPs.

Is there a role for some private blockchain implementations going forward? Undoubtedly, as there is a role for some intranets. But the purpose of blockchain technology – to facilitate transactions, agreements, and exchange of data among untrusted actors at a very large scale – can only be realised in the long term through interoperability. This does not necessarily mean that there will be one blockchain to rule them all, but rather that blockchain architectures that facilitate exchange and data validation at the largest scale will likely gain the greatest adoption.

Open standards

Blockchains alone do not guarantee data portability or system interoperability. As indicated in Box 11.1, data anchored to a blockchain is only useful if it can be reliably understood and verified. Since public blockchains do not encode plaintext data on-chain, and permissioned and private blockchains have widely varying standards as to what is and is not stored on-chain, validating identity, financial transactions, agreements, and many other things requires a standardised way of sharing, verifying, and reading blockchain-anchored data.

This is where open standards come into play. Open standards are not applications; they are standardised ways of formatting, writing, sharing, retrieving, and verifying data. The advent of blockchain technology has spurred considerable innovation in open standards as companies, governments, and individual contributors realise the need for a shared method for verifying what is true. Shared methods reduce vendor-dependency, helping organisations and individuals avoid the kind of lock-in that comes from platform-dependent ways of managing data.

Perhaps the most prominent organisation developing open standards for blockchain-anchored data today is the World Wide Web Consortium (W3C). The W3C was formed during the early days of the internet (1994) by Tim Berners-Lee, MIT, CERN, DARPA, and the European Commission to codify open standards that are used to transact information across the global digital infrastructure of the web (W3C, n.d.[69]). The W3C is the leading independent internet standards body in the world, sourcing volunteer contributions from many individuals and organisations.

Among the first standards the W3C is developing for blockchain-anchored data are credentialing standards: how to verify whether claims being made about one party by another are valid. The most notable of these standards are Verifiable Credentials (VCs) (W3C, n.d.[70]) and Decentralised Identifiers (DIDs) (W3C, n.d.[71]). The verifiable credentials open standard was made a full W3C recommendation in 2019. It is designed to work in tandem with a new standard for blockchain-anchored digital identifiers, known as Decentralised Identifiers (DIDs), which was submitted for recommendation status in April 2021.

The intent of the W3C DID standard is to add privacy and flexibility functionality to the public key cryptography which is currently used to identify individuals and organisations making use of blockchain technology. Taken together, verifiable credentials and DIDs enable the evolution of a blockchain-agnostic, privacy-protecting, feature-rich ecosystem for secure, tamper-evident digital credentials. For these reasons, the US Department of Homeland Security has been issuing grants to companies working in this space to speed along the development of these new standards (US Department of Homeland Security, 2019[72]; US Department of Homeland Security, 2017[73]; US Department of Homeland Security, 2019[74]). VCs and DIDs are also seeing uptake by governments including Canada (VON, n.d.[75]; Benay, 2019[76]), Singapore (OpenCerts, n.d.[77]), South Korea (Insights, 2020[78]), and the European Commission (Du Seuil and Pastor, 2019[79]; EU Blockchain Observatory and Forum, 2019[80]).

It is important to note that the VC and DID standards are agnostic with regards to a digital record's form factor or content. They function similarly whether records are educational, financial, professional, or vital. To address the need for standardisation on record content and form factors within industries, additional standards bodies have been formed. As this paper focuses on the applications of blockchain to the education industry, our focus here will be on those implementations. At the W3C, this includes the Verifiable Credentials for Education Task Force (W3C, n.d.[81]) and the Educational and Occupational Credentials in Schema.org Community Group (W3C, n.d.[82]). Other standards for educational credential descriptions include the Common Microcredential Framework (CMF) launched by the European MOOC Consortium (Konings, 2019[83]), the Credential Transparency Description Language developed by the Credential Engine (Credential Engine, n.d.[84]) and the Badge Alliance (Alliance, n.d.[85]), as well as the Comprehensive Learner Record (CLR) standards under the custodianship of IMS Global (IMS Global, n.d.[86]). The W3C works closely with some of these institutions – like the Credential Engine and IMS Global – to complement their efforts and create globally interoperable records definitions.

The W3C is not, however, the only organisation working on interoperability for decentralised credentialing standards. Credentialing is only one application of verifiable claims under the broader umbrella of digital identity. Other groups working on the next generation of web-based identity infrastructure include Rebooting Web of Trust (RWoT) (Web of Trust, n.d.[87]), the Internet Identity Workshop (IIW) (Internet Identity Workshop, n.d.[88]), the T3 Innovation Network at the US Chamber of Commerce Foundation (US Chamber of Commerce Foundation, n.d.[89]), the Digital Credentials Consortium (MIT Open Learning, 2020[90]), the Trust Over IP Foundation (hosted at the Linux Foundation) (Trust Over IP Foundation, n.d.[91]) (The Linux Foundation, 2020[92]), and the Decentralised Identity Foundation (DIF) (Decentralized Identity Foundation, n.d.[93]). Most recently, the Learning Technology Standards Committee (LTSC) at the IEEE (IEEE, n.d.[94]) and the Learning Economy Foundation (Learning Economy Foundation, n.d.[95]) have proposed an "Internet of Education (IoE)" (Internet of Education, n.d.[96]) to unify these and other consortia working on technical standards for education and credentialing (Education, n.d.[97]).

There is considerable overlap among participants in these groups, although they have different target memberships. The Trust Over IP Foundation and DIF are primarily corporate membership organisations; the T3 Innovation Network is a public-private membership organisation; the Learning Economy Foundation attracts members with a background in education technology and policy; the Digital Credentials Consortium is an academic membership organisation; and RWoT and the IIW attract technologist volunteers passionate about internet privacy and autonomy regardless of affiliation. The W3C functions as a kind of node for technical consensus formation between these groups as well as others.

Self-sovereign identity

Efforts to build a standards-based architecture for decentralised identity over the web have strong overlaps with the Self-Sovereign Identity (SSI) movement. There is wide disagreement about what exactly "self-sovereignty" is or whether it is desirable, but those who identify themselves as part of this movement typically are motivated by giving individual users of digital technologies greater control over their personal data (Smolenski, 2016[98]). This means, among other things, removing their dependence on platforms and governments to manage this data. Indeed, data portability and platform independence may be the two most defining characteristics of self-sovereign digital identity, with some arguing that these descriptions are more useful in practice than the term "self-sovereign" (Renieris, 2020[99]).

Arguably the most well-known articulation of the principles of self-sovereign identity was made by computer scientist Christopher Allen in a widely-circulated 2016 blog post (Allen, 2016[100]). Since that time, there has been considerable debate about what each of these principles mean and what technological implementations best address them. Adherents of various SSI or SSI-adjacent philosophies actively participate in developing standards for digital identity through the aforementioned standards bodies, and their points of view often differ widely. In addition, SSI standards development must contend with existing legal regulations around the use of cryptography and data protection. This sometimes results in the proliferation of technical implementations within standards definitions rather than their consolidation. Nevertheless, progress continues to be made towards interoperable, user-centric standards for verifiable credentials and digital identity (T3 Innovation Network, 2020[101]).

Benefits of blockchain for educational credentialing

Blockchains arose as infrastructures to verify transactions across an adversarial network – that is, a network with untrusted parties. Although blockchain functionality can be extended to execute automated business processes via smart contracts, its primary value remains in serving as a shared, single source of truth. Blockchain's conjoined verification and smart contract functionality has the potential to automate educational business processes like credential transfers, academic record transfers, credential equivalency establishment, and even the administrative apparatus of the university.

Blockchains are not a prerequisite for automating business processes, however. Many existing software applications do this, and those employing workflows, intelligent process automation (IPA), and robotic process automation (RPA) have significantly advanced much of this functionality. So when is using blockchains mission-critical for achieving automation objectives? It may be the case if said automation must occur across a network of adversarial, uncoordinated actors. If it does, a follow-up question is whether sufficient network effects can be achieved to produce a return on the high ongoing cost of building and maintaining a blockchain network.

To make this more concrete, let's take an example that is often used to illustrate the benefits of blockchain in education: automating the establishment of credit equivalencies as students transfer from one school to another. While some argue that blockchains will solve this problem (and some vendors have built blockchain networks specifically to do so), credit equivalencies could also be automatically determined by existing applicant tracking systems and databases using widely agreed-upon definitions of what credits mean and widely-accepted standards for educational data transfer. In other words, the primary obstacle to automated credit transfers to date has not been technological, but social: every jurisdiction (down to even the university or school) often has its own credit definitions and may be reluctant to make those transferable by using definitions established elsewhere. This example illustrates that, at times, "blockchain" may be presented as a solution to problems that actually require off-chain forms of social consensus building. In some regions, this may be easier at the higher education level, as common definitions and standards for credits have been established and endorsed. This is, for example, the case in the European Union, with its European Credit Transfer System.

Perhaps blockchain could be the missing piece that enables an application to coordinate action across many different parties who would otherwise require a trusted third party to act as a mediator. But the most prominent example of an education provider who took such an approach supports the case made above: that blockchain may not yet provide a level of automation value that warrants its use. Woolf University, initially hailed as a "Blockchain University" in global media reports (Young, 2018[102]; Vander Ark, 2018[103]), was originally built on a distributed ledger platform (Parisi, 2018[104]). As Woolf Founder Joshua Broggi explained, "We use a blockchain to create efficiencies by managing custodianship of student tuition, enforcing regulatory compliance for accreditation, and automating a number of processes." (Vander Ark, 2018[103]). Woolf experimented with different blockchain models in its early years of operation, but later moved away from blockchain technology. Instead, Woolf now highlights its offering as a cloud-based SaaS platform for organisations and students, delivering the social mission of access to quality education from anywhere in the world. Woolf now refers to itself as "The Borderless University," and is indeed used by faculty and students from all over the world (Woolf University, n.d.[105]). Web-based applications enabled streamlining the administrative apparatus of the university without blockchain.

By contrast, other companies creating blockchain-based education platforms prominently feature their use of blockchain. These include ODEM (Maaghul, 2019[106]) and BitDegree (BitDegree, 2017[107]), both of which use the public Ethereum blockchain and have their own ERC20 tokens. In the United States, the Learning Economy Foundation (Learning Economy Foundation, n.d.[95]) has proposed a vision for a blockchain-based "ecosystem" that unites credentialing authorities, students, employers, education technology firms, and others in a shared marketplace where educational outcomes are incentivised through monetisation ("rewards"). Learning Economy Foundation co-founders, Chris Purifoy and Jacksón Smith, describe their vision as follows in a cover story for *G20 Summit Magazine*: "By quantifying the true value of education, a whole economy can be built around it to pay students to learn, educators to create substantive courses, and stewards to help the Learning Economy grow. Blockchain provides a decentralised way for everyone adding value to global education to coordinate around the commonwealth without the friction of individual partnerships." (Purifoy and Smith, 2018[108])

Learning Economy founders anticipate that building their blockchain platform will create a "Learning Gold Rush" (the title of their Implementation Roadmap) (Learning Economy Foundation, n.d.[109]) where everyone is incentivised to teach, learn, and employ through financial rewards administered by the network.

The Learning Economy model has stretched the promise of blockchain the furthest, conjecturing that blockchain networks will be able to not only coordinate social action, but do so in a way that creates sufficiently motivating financial rewards that are automatically administered to all stakeholders in a marketplace, realising the hazily-defined goal of "improving learning outcomes" from the bottom up. While the nobility of this vision is laudable, it borders on a utopianism that is out of step with the actual capabilities of any blockchain network. In the long term, it is the standards-based credentialing projects in which Learning Economy is engaged that will likely see the most success (see "Real-World Implementations", below).

In the near term, therefore, the primary value of blockchain for education is likely more basic: it makes credential verification an order of magnitude faster, cheaper, and more secure. When used in combination with open standards, blockchains remove ongoing dependencies on issuing institutions, software providers, and third parties to verify official records. Moreover, blockchains enable direct ownership of digital credentials by both issuers and recipients.

It cannot be emphasised enough that use of *open technology standards* for digital credential issuance, storage, sharing, and verification is the precondition for realising these benefits of blockchain technology (Jagers, 2018[110]). Just like any other type of software, a blockchain can become a walled garden if there is no way to make the data it references communicable to others. Accordingly, the development of shared open standards for verifiable credentials is one of the most rapidly evolving areas of the blockchain technology stack.

The W3C Verifiable Credentials and Decentralised Identifier standards described in the previous section are seeing some of their first real-world applications in the education industry. This is largely due to the digital credentialing work educational institutions, non-profits, and education technology companies have engaged in for years and which has laid the groundwork for the formulation of the W3C standards. An early version of verifiable credentials, Blockcerts (Blockcerts, n.d.[111]), was developed by MIT and the company Learning Machine specifically for educational certificates like diplomas and transcripts. Blockcerts, in turn, grew out of the Open Badges standard for portable digital credentials, which was spearheaded by the Mozilla Foundation in 2011 and came into the custodianship of IMS Global in 2017 (Wikipedia, n.d.[112]).

Blockcerts made notable security and portability improvements to Open Badges that enabled its use for a wider range of high-stakes credentialing use cases, like diplomas and transcripts (Federation of State Medical Boards, 2019[113]). The Blockcerts reference libraries were published in 2016 under an MIT Free and Open Source Software (FOSS) license (Learning Machine Newsroom, 2016[114]). In 2017, Learning Machine was the first company to launch a commercial Blockcerts issuing platform, which it sold to educational institutions to issue digital diplomas and other educational certificates (Hyland Credentials, n.d.[115]). Other companies soon followed suit, and some schools developed their own Blockcerts issuing platforms as well (McMaster University Office of the Registrar, n.d.[116]; Unviersidad Carlos III de Madrid, 2018[117]). In a related project, the Government of Singapore forked the Blockcerts codebase to create their own open standard for digital credentials, OpenCerts (OpenCerts, n.d.[118]). In 2020, Learning Machine was acquired by enterprise content services firm Hyland and rebranded Hyland Credentials.

As the W3C Verifiable Credentials specification has matured, custodians of both Blockcerts and Open Badges (as well as OpenCerts) have committed to updating the standards to be compliant with the verifiable credentials specification. This is possible because the VC standard is flexible enough to accommodate many different credential types. As the education industry standardises around using VCs, credential exchange and interoperability across software platforms will be significantly enhanced (T3 Innovation Network, 2020[101]). In addition, the work being done to standardise the definitions of educational credentials will significantly facilitate the processing and exchange of educational records between schools and across borders.

Benefits of using blockchain technology in combination with these emerging open standards for educational credentials include the following: eliminate records fraud; streamline and reduce the cost of records sharing and verification, and; return control of personal data to individuals and reduce institutional risk.

Eliminate records fraud

Current state

Academic records fraud is pervasive and widespread. Studies estimate that over 100 000 degrees are simply purchased each year in the United States (Accredited Online Colleges, n.d.[119]); this would possibly include more than half of all PhDs (Ezell and Bear, 2012[120]). Moreover, validating the authenticity of a record is a separate process from validating the authenticity of a school: a "real" diploma may be purchased from a fake school, and fake diplomas that look like they are from real schools may also be purchased. In 2012, there were more than 3 300 unrecognised universities worldwide, many of them simply diploma mills (Ezell and Bear, 2012[120]); today the number is likely much higher. Moreover, real schools may have their academic records tampered by recipients looking to increase their chances of admission to particular jobs or programmes of study. The problem of records modification is so acute in some areas that, according to anecdotal reports, some secondary schools have simply stopped issuing academic transcripts (Smolenski, 2018[121]).

Professional licenses are forged with a frequency similar to academic records. In the United States alone, an Ohio State University Study estimated that up to 2 million medical practitioners may be practicing with fraudulent diplomas or licenses (Gibson, 2017[122]). Major news stories routinely surface in countries around the world about professionals practicing without licenses or with forged licenses (Gibson, 2017[122]) (CNN, 2020[123]). For the above reasons, an ecosystem of Credential Verification Organisations (CVO's) has arisen to validate the authenticity of schools and records; however, their efficacy has been limited.

The problem of fraud and of non-accredited higher education institutions delivering worthless degrees has been identified as a problem limiting the benefits of the internationalisation of higher education (OECD, 2004[124]). In 2005, UNESCO and the OECD issued "Guidelines for quality provision in cross-border higher education", which was followed by the establishment of a database of accredited institutions that is maintained by UNESCO (OECD, 2005[125]). This database is, however, difficult to make comprehensive and keep updated.

Blockchain with open standards

Blockchain provides a decentralised, transnational verification infrastructure to avoid fraud, facilitate international student mobility, and safeguard the public from professionals with illegitimate credentials anywhere in the world. Blockchain technologies, used in combination with leading open standards like Blockcerts and Verifiable Credentials, make use of advanced cryptography in combination with digital signatures to validate both the provenance of a credential (which institution issued it) and the authorised recipient (to whom it was issued). Digital signatures and hashed data, combined with an immutable blockchain ledger, ensure that a credential has not been tampered. Decentralised verification allows any third party (employer, government, school, or individual) to validate whether a credential was really issued by the claimed organisation to the claimed individual and whether any changes to the credential have been made since issuance. Credential revocation and expiration can also be instantly validated with the highest level of confidence. And just as these technical standards can be used to validate credentials issued to individuals, they can be used to validate institutional credentials as well, including elements like accreditation status.

Streamline and reduce the cost of records sharing and verification

Current state

Today, an individual applying for employment or further study must request their official academic records from their school, often for a fee, depending on the country or institution. They must also request that those records be sent by the school or by a software provider with whom the school contracts to the institution to which she is applying. The receiving institution then validates the credential by checking its integrity using the software vendor's solution, by contacting the school, or by checking with a third-party Credential Verification Organisation. The process is time consuming and often costly.

That is the best-case scenario, in which the issuing institution still exists and is able to locate, validate, or send the records to the receiving institution (relying party). In many cases, however, particularly in countries where political disturbances or natural disasters have struck, records have been destroyed, and issuing institutions may no longer be functional. In such cases, many record owners (the subjects of their own records) will have opportunities for employment, residency, citizenship, or further study closed to them because their qualifications or even identities cannot be validated.

When the records that must be verified were issued in a foreign country, an even more complex process called the Apostille is triggered. In 1961, a group of countries signed a treaty abolishing the requirement to legalise foreign public documents (Hague Convention, 1961[126]). Countries that are signatories to the Apostille Convention make use of an Apostille process to authenticate documents issued by designated authorities. Countries that are not signatories of the Convention require an Authentication Certificate from federal authorities of the issuing country to validate the authenticity of a foreign public document (US Department of State, Bureau of Consular Affairs, n.d.[127]). The authentication process involves conveying the original physical document, often with a supplementary request form and a fee, to a government institution authorised to issue the Apostille or Authentication Certificate. This institution will then issue a stamp on the document or a supplementary Certificate and mail the document(s) back to the recipient, who then must convey it to the authority of the foreign country in which they intend to live, work, or study. This process takes several months on average and causes the recipient to incur a financial cost. If documents cannot be authenticated, the recipient loses opportunities for travel, employment, and further study.

Blockchain with open standards

The combination of blockchain with open standards allows issuing institutions to issue records to recipients only once. After that, recipients can prove cryptographically that the record was issued to them and by which institution.

Recipients can share their records at their own discretion, and even choose which data within a record to disclose. Records may be verified by anyone to whom the recipient grants record access, instantly and for free. This facilitates not only the validation of academic credentials, like diplomas and transcripts, but also any document that must be Apostilled (as well as the Apostille Certificate itself). In 2020, companies Hyland and Hedera Hashgraph announced a proof of concept to anchor electronic Apostilles to a blockchain in partnership with the Texas Secretary of State (Texas Blockchain Council, 2021[128]). This proof of concept leveraged both verifiable credentials and PDF hashing, along with application-based workflow management to streamline and automate much of the process of issuing and verifying Apostilles.

With blockchain credentialing based on open standards, even if an issuing institution ceases to function, records are lost, or the software vendor whose product was used to issue such digital credentials no longer exists, the recipient still owns verifiable versions of their records and can share and verify them at will. As verifiable credentials are also fully machine-readable, schools verifying incoming records can pre-screen them for qualifications for particular programs of study. This dramatically cuts the receiving institution's overhead associated with validating records while removing friction for recipients as they seek opportunities for further learning.

When verifiable credentials standards are used, it also doesn't matter whether the organisation needing to verify a credential is located in the same country or in a different country from the one in which it was issued. Any authority in a foreign country can validate an individual's identity documents and qualifications by checking the document integrity, the cryptographic identifiers of issuer and recipient, and the status of the record. In short, use of the blockchain combined with open standards could, over time, significantly reduce or even eliminate the need for an international Apostille process or certificates of authentication.

As a result, open standards for blockchain-based credentialing stand to increase trust in the immigration process and international travel in general. This is why government departments like the US Department of Homeland Security have been investing in the development of Verifiable Credentials and Decentralised Identifiers: they see them as anti-counterfeiting technology standards for use by federal agencies like Customs and Border Protection (CBP) and the Transportation Security Administration (TSA).

Ultimately, the use of verifiable digital records, enabled by blockchain technology and open standards, will significantly reduce both the human and technological overhead currently associated with validating records issued both on paper and using legacy digital formats.

Return control of personal data to individuals and reduce institutional risk

Current state

Today, individuals may have in their possession copies or even originals of their official records, but they generally cannot have them validated by a third party without first requesting that the institutions that issued those records re-authenticate or re-send authenticated versions of those records. This often renders the official records actually in the possession of the individual functionally worthless. It disempowers the very person the records are about by removing their ability to share and verify them, re-inscribing their dependency on issuing institutions and third-party verification organisations. In effect, individuals cannot be said to own their records, even if those records are in their possession.

The current state also increases liability for institutions, who are tasked with being the lifelong curators and validators of official records on behalf of their constituents. Not only must institutions retain and safeguard growing troves of personal data, but since they are the ones sharing documents on behalf of individuals, they must comply with onerous regulations to protect the privacy of record subjects. Fines for mishandling data can be punitively high, particularly under new legislation such as the EU General Data Protection Regulation (GDPR).

Blockchain with open standards

Once a blockchain-anchored credential is issued to an individual using open standards, the individual is in possession of a digital file that attests to the provenance and integrity of the credential, as well as to their custodianship of it. Now individuals have a usable record that they can share instantly, without charge, with anyone they choose. Digital blockchain records can be re-used indefinitely without any need for re-issuance or additional validation by the issuing authority or any other party.

Since the locus of records ownership and sharing shifts to the individual, while verification is decentralised, issuing institutions no longer need to share records on behalf of their constituents and can dramatically reduce their internal records maintenance burden. This eliminates the need for storing large amounts of personal data for long periods of time and de-risks institutional data management by dramatically reducing the circumstances under which an institution must share a constituent's personal data.

Summary

The benefits of using blockchain technology in combination with open standards for educational credentialing – like W3C Verifiable Credentials and Decentralised Identifiers, among others – amount to an impressive list which can be grouped into three major categories:

1. *Security*. Open standards for blockchain credentialing prevent fraud, mitigate risk for institutions that both issue and validate official documents, and protect the brands and reputations of education providers as well as qualification holders. As verifiable credentials adoption increases, their use in credentialing, screening, and hiring will likely become an industry best practice or even a legal requirement.

2. *Public good*. Blockchain records that use open standards remove barriers to accessing opportunity and facilitate economic development by increasing trust in institutions and workforce qualifications while mitigating the devastating effects of political conflicts and natural disasters. When people control their own records, they are free to continue using them anywhere in the world, while relying parties enjoy a high level of confidence in credentials issued anywhere.

3. *Efficiency*. Blockchain records that use open standards dramatically increase the efficiency and convenience of records management: the need to re-issue or manually verify documents drops by an order of magnitude; the need to use other record formats is removed over time; and records sharing and verification is instant and free. Processes like credit transfers, which are notoriously complex and time consuming for educational institutions, can be automated with the advent of machine-readable, verified course content.

A consensus is growing behind these benefits. A recent report by the US Chamber of Commerce Foundation's T3 Innovation Network on the benefits of adopting self-sovereign approaches to learner records summarised them as follows:

- "Learners have control over who can access their record(s), including aspects of their records, and when they are allowed to access them.
- Authentication is cryptographically secured, most typically on distributed ledgers, making the credentials verifiable, and accessible regardless of the state of the issuing organisation at the time of verification.
- Online verification of learners and issuers can be secured and streamlined.
- Verifiable credentials can support verification of non-traditional achievements providing evidence of learning in various contexts." (T3 Innovation Network, 2020[101]).

The benefits of an approach to credentialing that leverages both blockchain technology and open standards make it a valuable component of digitisation projects for any institution that issues or validates official documents. Such an approach is the prerequisite for creating a truly interoperable global ecosystem for digital credentials: where anyone can exchange academic records with anyone else and instantly screen and verify them without fear of fraud. Due to their deep investment in credentialing economies, education providers have been among the first to implement blockchain credentialing programs. The following section reviews several examples of educational institutions that have embarked on this path.

Real-World implementations

Any situation in which fraud must be prevented is a potential blockchain use case. This is because, as the previous sections have described, the blockchain is primarily an infrastructure for verifying claims. For this reason, blockchain is set to revolutionise every «trust industry,» from banking to insurance to law enforcement, healthcare, and supply chain management (McCauley, 2019[129]). It is education, however, that has been at the forefront of blockchain technology adoption, as the verification of academic credentials continues to be an ongoing and pressing need (Grech and Camilleri, 2017[130]).

But verifying qualifications is a requirement far beyond the education industry; employers worldwide have a clear interest in receiving credentials they can rely upon with the highest degree of confidence. Multilateral institutions like the Inter-American Development Bank have noted that credentialing human capital in a verifiable way is one of the easiest ways to create trust in a global economy (Cabrol, 2018[131]). For this reason, the American Council on Education observes, "blockchain, in particular, holds promise to create more efficient, durable connections between education and work. It can provide the technological fabric to help displaced workers translate their skills for new education opportunities and employers, and may hold particular value for those currently underserved by the existing education-to-employment paradigm." (Lemoie and Souares, 2020[44]).In particular, the shift towards a lifelong learning model for education and the ever-increasing mobility of learners and workers make verifiable credentials a prerequisite for economic momentum in the twenty-first century. The term "ecosystem" is frequently used by blockchain technologists, employers, and educational institutions to describe the ideal state of credential exchange and verification.

This section features several examples of real-world implementations of blockchain-based credentialing in education from countries around the world. Note that not all of the projects that have been built to date have real users yet (or customers, if they are commercial). Some of these projects have been announced, but little public information is available to track their implementation or uptake. Accordingly, this list is to be taken as a snapshot of what the blockchain credentialing market looks like in the first few years of its development, with the understanding that this landscape will continue to rapidly evolve.

United States

It can be said that the birth of open technical standards for verifiable digital credentials occurred at MIT. In 2015, the Learning Initiative division of the MIT Media Lab began working on a project to anchor academic credentials to the blockchain (MIT Media Lab, n.d.[132]). Soon thereafter, the company Learning Machine (since 2020, Hyland Credentials) joined forces with this team to bring the project to fruition (MIT Media Lab, n.d.[133]). In 2016, Blockcerts, the open standard for digital credentials, was published under an MIT Free and Open Source Software (FOSS) license at blockcerts.org (Blockcerts, n.d.[111]). In 2017, the Blockcerts references libraries allowed anyone to create their own software applications for issuing, storing, sharing, and verifying secure digital credentials. The verifiable credentials technology was intentionally made open source in order to avoid vendor lock-in and create an open ecosystem for credentialing applications. Blockcerts is also blockchain-agnostic, meaning that it supports anchoring to most blockchain network types. Current implementations include Bitcoin, Ethereum (Learning Machine, 2018[134]), and Hyperledger Fabric (Castro-Iragorri, 2018[135]). In 2017, Learning Machine launched a commercial credential issuing system. In 2019, it announced that it would update Blockcerts to become a W3C Verifiable Credential (Jagers, n.d.[136]).

A number of US universities and K-12 institutions issued credentials using Blockcerts to their students and graduates (Hargrave and Karnoupakis, 2020[137]). These include MIT (Durant, 2017[138]), SNHU (Kelly, 2018[139]), Union Public Schools in Tulsa, Oklahoma (Friedman, 2019[140]), ECPI University (Southside Daily Staff, 2018[141]), Maryville University (Learning Machine Company Newsroom, 2019[142]), and Central New Mexico Community College (Salas, 2018[143]), among others. In 2020, the first statewide implementation of blockchain credentialing in the United States was launched by the New Mexico Higher Education Department (Hyland Newsroom, 2020[144]).

In February 2020, a consortium of research institutions around the world announced bloxberg, a fork of Blockcerts designed to anchor proof of scientific research to a private blockchain network (Bloxberg, 2020[145]).

A host of new projects are also emerging. In February 2020, the Digital Credentials Consortium (DCC) issued a white paper on the digital credentials infrastructure for the future, aligned with the W3C Verifiable Credential model and with the EU General Data Protection Regulation (Digital Credentials Consortium, 2020[146]). In 2018, MIT and eleven other international universities from Canada, Germany, Italy, Mexico the Netherlands, the United States founded the Digital Credentials Consortium to develop an infrastructure for issuing, sharing, and verifying digital credentials of academic achievement in higher education. A report released by the American Council on Education and the US Department of Education's Office of Educational Technology provides a useful survey of blockchain credentialing initiatives in the United States (Lemoie and Souares, 2020[44]):

- Workday Credentials is a platform for issuing blockchain-anchored records using the W3C Verifiable Credentials standard (Workday Credentials, n.d.[147]). It uses a private implementation of Hyperledger Fabric to anchor records (Ledger Insights, 2020[148]; Meetup, 2020[149]).

- Pistis.io provides web and mobile wallets for recipients to upload documents and hash them to a blockchain built on Hyperledger Fabric (Pistis.io, n.d.[150]).
- Greenlight provides a blockchain network currently used by five independent school districts in North Texas to exchange academic records and help lower-income students apply to multiple colleges at once. The network is a private implementation of Hyperledger Fabric (Lemoie and Souares, 2020[44]).
- Salesforce has launched the Trusted Learner Network, a private blockchain network powered by Hyperledger Sawtooth.
- Arizona State University is currently using Salesforce's Trusted Learning Network to send course completion records to the two-year schools from which their students have transferred (Lemoie and Souares, 2020[44]).
- ODEM and BitDegree also include credentialing as one of the functionalities of their Ethereum-based education marketplaces.
- The San Jose State University School of Information is applying for a grant to use an SSI universal resolver to "enable individuals with verifiable digital credentials to gain access to resources at all participating libraries." (Lemoie and Souares, 2020[44]).
- The Learning Economy Foundation is involved with the Arizona State University Trusted Learner Network implementation (Arizona State University, n.d.[151]). It is also leading blockchain credentialing projects in Colorado (the C-Lab) (Learning Economy Foundation, n.d.[152]), North Dakota (the ND ILR Co-Lab) ((n.a.), n.d.[153]), Florida (the Broward County OpenCLR Lab) (Learning Economy Foundation, n.d.[154]), and the Asia-Pacific Region (Asia-Pacific AP Lab) (Learning Economy Foundation, n.d.[155]).

Canada

The Government of Canada may have taken the lead internationally in terms of self-sovereign credentialing applications. In 2018, the Digital Identity Office at the Treasury Board of Canada Secretariat released "Canada's trusted digital identity vision," a short film demonstrating a future in which citizens can apply for and manage a variety of government services and benefits online using private, secure forms of authentication (Treasury Board of Canada Secretariat, 2018[156]). The Province of British Columbia was among the first to implement a live digital credentialing project using SSI: the OrgBook, which uses Hyperledger Indy to manage verifiable credentials for over 1 million active businesses in the province (Lemoie and Souares, 2020[44]).

In 2019, the Government of Canada Talent Cloud initiative started to certify "Free Agent" public sector workers using Blockcerts (Benay, 2019[76]; Talent Cloud | Nuage de talents, 2018[157]). The Talent Cloud is a public sector talent marketplace that reimagines a skill and credential recognition ecosystem for Canada's Public Service (Government of Canada Talent Cloud, n.d.[158]; Greenspoon, 2018[[159]; World Government Summit et al., 2018[160]). Public sector workers can maintain a profile on the Talent Cloud that includes validated qualifications and experience, which serve as evidence of skill when applying for any public sector job.

McMaster University in Hamilton, Ontario, also built its own Blockcerts issuing system, which is used to issue digital diplomas to graduating students today (McMaster University Office of the Registrar, n.d.[116]).

Most recently, the Association of Registrars of the Universities and Colleges of Canada (ARUCC) has announced a partnership with Digitary to build a national credentialing network for Canadian higher education institutions (Hamdani, 2020[161]). Digitary has, in turn, partnered with Evernym to implement a blockchain-based credentialing solution (Crace, 2019[162]). Evernym's solutions implement W3C Verifiable Credentials leveraging the Sovrin and Hyperledger Indy blockchain networks.

European Union and United Kingdom

Within the European Union, the country that has made the farthest advances in applying blockchain for educational credentialing is Malta. Since 2017, Malta has branded itself the "Blockchain Island" as a result of Prime Minister Joseph Muscat's decision to embark on the "calculated risk" of investing in blockchain technology to fight corruption, cut bureaucracy, and diversify the country's booming tech sector (Al Ali and van der Walt, 2018[163]).

In January 2017, Malta's Ministry for Education and Employment (MEDE) started to implement the world's first nationwide pilot project for issuing academic credentials to a blockchain (Sixtin, 2017[164]; Cocks, 2017[165]). Since that time, the project has expanded in scope to include all educational institutions in Malta (Sansone, 2019[166]).

In Spain, the Alastria blockchain network was designed and built by a consortium of over 500 businesses, government institutions, and universities with the aim of facilitating digital services, including academic credentialing (Alastria, n.d.[49]). SmartDegrees also emerged, using the Ethereum-based Quorum blockchain to anchor digital credentials (SmartDegrees, n.d.[167]). Vottun is another Spanish blockchain credentialing firm, this one using the public Ethereum blockchain (Vottun, n.d.[168]).

In Central Europe, SAP launched its TrueRec credentialing platform using Ethereum, which it piloted with KU Leuven University (Jonkers, 2018[169]). The Slovenian firm 0xcert created their own open standard for tokenising and transferring ownership of digital credentials as well as other digital assets (0xcert, n.d.[170]).

Within the United Kingdom, GradBase offers a blockchain credentialing system using the Bitcoin blockchain (GradBase, n.d.[171]). PwC UK has introduced "SmartCredentials," a credentialing platform built on a permissioned version of Ethereum (PwC UK, n.d.[172]). The Open University's Open Blockchain initiative (The Open University, n.d.[173]) explores a number of applications of blockchain credentialing, primarily using Ethereum: blockchain-anchored Open Badges and Blockcerts. These include QualiChain (The Open University, n.d.[174]), a platform for matching employers with job seekers, and PeerMiles (The Open University, n.d.[174]), an initiative to recognise the reviewing records of researchers. These are both research projects rather than commercial software applications.

All of these initiatives will need to reckon with the standardisation being introduced not only by the W3C, IEEE, and ISO, but also by the European Blockchain Services Infrastructure (EBSI), a collection of blockchain networks under development by the European Commission and European Blockchain Partnership, a coalition of EU member states dedicated to researching public sector applications of blockchain technology. The EBSI is funded by the Connecting Europe Facility, which is responsible for supporting the mandate for a Digital Single Market throughout the EU (CEF Digital, n.d.[40]). The purpose of the EBSI is to serve as an infrastructure for cross-border digital services. In this sense it is similar in scope to Spain's Alastria network, but with a pan-European mandate. A number of participants in the Alastria project are now coordinating build-out of the EBSI.

The EBSI currently makes use of Hyperledger Besu (a permissioned implementation of the Ethereum codebase) and Hyperledger Fabric, but its intention is to become blockchain-agnostic. Technical documentation of this open source project is available to the public (CEF Digital, n.d.[175]). The four initial use cases targeted by the EBSI are:

1. Notarisation of documents for auditing purposes;
2. Certification of diplomas;
3. The European Self-Sovereign Identity Framework (ESSIF); and
4. Trusted data-sharing (Allen, 2016[100]; CEF Digital, n.d.[176]).

Version 1.0 of the EBSI launched in February 2020, and future releases are anticipated on an annual basis (Smolenski, n.d.[177]). Open market consultations have begun to help shape future development grants for private firms intending to build services on the network.

The EBSI is separate from the Europass project, another EU-wide digital credentialing initiative. The Europass is overseen by the Commission's Directorate-General for Education and Culture and aims to make skills portable and recognisable across Europe. Since 2012, the Europass web portal has allowed individuals to create an electronic portfolio of their academic credentials and other qualifications, their "European Skills Passport". These five standard documents are not anchored to a blockchain, but are stored in XML format and may be digitally signed by the issuing institution using signing keys provided by authorities trusted under the eIDAS regulation for electronic identification and signatures. Through the European Skills/Competencies, Qualifications and Occupations (ESCO) initiative, Europass profile holders can also be matched with job opportunities via the ESCO semantic classification (European Commission, n.d.[178]).

In 2018, a set of requirements for digitally signed Europass documents were elaborated, drawing heavily on the established eIDAS trust framework. This set of requirements is referred to as the European Digital Credentials Infrastructure (EDCI) (everis, n.d.[179]). Distinct from the EBSI, the EDCI does not employ blockchains for verification, with the exception of the accredited status of issuing authorities. These are to be anchored to a blockchain—presumably, eventually this will be the EBSI. The EDCI is in effect an open standard for issuing, receiving, storing, sharing, and verifying digitally signed documents, much like Blockcerts and Verifiable Credentials. As the EDCI was not elaborated in partnership with the W3C, however, it diverges from the verifiable credentials specification in several respects, most notably by requiring the use of XML rather than JSON. These discrepancies are currently being reconciled within the W3C Credentials Community Group.

Middle East and North Africa

The Middle East and North Africa (MENA) region has been the site of considerable enthusiasm for blockchain technology, particularly since the announcement of the United Arab Emirates' "Blockchain Strategy 2021" in 2018 (The United Arab Emirates' Government portal, n.d.[180]). The initiative aims to place 50% of government transactions on a blockchain by 2021 and has been closely involved with the Smart Dubai 2021 (Smart Dubai 2021, n.d.[181]) initiative. As part of its strategic push towards blockchain, the UAE government has sponsored numerous start-up competitions to identify local vendors who may be able to supply public and private sector services.

One UAE-based provider in the academic credentialing space is Educhain. Having won the Techstars start-up competition, they went on to implement several blockchain credentialing pilots with UAE educational institutions in both higher education and K-12 segments. These include United Arab Emirates University (UAEU, 2019[182]), the University of Dubai (CNN, n.d.[183]), and AMSI (AMSI, n.d.[184]).

In 2020, the firm Shahada was launched (Smartworld, 2020[185]). It sells a SaaS platform, also called Shahada, to create and issue blockchain-anchored digital credentials. Its website states that "Shahada builds on the open standards developed by MIT" (Shahada, n.d[186]). Shahada is a joint venture between Smartworld, a leading UAE systems integrator, and Grape Technology, a UAE-based start-up specialising in blockchain technology (Shahada, n.d.[186]). The firm has successfully issued blockchain-anchored credentials for the University of Dubai (Zawya, 2020[187]). It has also integrated its platform with UAE Pass, which ensures recipient identity verification for government and commercial services throughout the country (Smartworld, 2020[185]).

In Egypt, Zewail City of Science and Technology signed an MOU with start-up Intelli Coders to build "BlockCred," (Abdou, 2019[188]) a blockchain credentialing system for educational and professional training programmes in the city (BlockCred, n.d.[189]). BlockCred is a DApp built on the Blockstack blockchain platform.

First in the MENA region to issue blockchain-anchored academic credentials was King Abdullah University of Science and Technology (KAUST) in Saudi Arabia (Company Newsroom of Learning Machine, 2018[190]). In 2018, it used Learning Machine (now Hyland Credentials) to issue its first Blockcerts-compliant digital diplomas. The University of Bahrain soon followed suit (Global Blockchain Business Council, n.d.[191]). As of November 2020, both universities continue with their digital credentialing initiatives.

Latin America and the Caribbean

Latin America has been the site of a growing number of blockchain credentialing initiatives in recent years. The projects that have announced implementation have generally employed the Blockcerts open standard for digital credentialing. These include the issuance of digital diplomas by Tecnológico de Monterrey, Mexico's premier technical research university (Longino Torres, 2019[192]). Universidad Autónoma de Nuevo León implemented a Blockcerts credentialing project in partnership with Learning Machine via a local implementation partner, SYSARTEC (SYSARTEC, 2020[193]). The Ministry of Labor in the Bahamas and the Caribbean Examinations Council also delivered Blockcerts pilots for workforce training, graduation certificates, and examination results (Munro, 2018[194]; Jamaica Observer, 2019[195]).

Another solution provider offering a Blockcerts credentialing system in the region is Xertify, a Colombia-based digital credentialing company (Xertify, n.d.[196]). The Dirección Estatal de Profesiones in the State of Querétaro in Mexico has been using Xertify to issue a digital Cédula Profesional (professional license) for many types of practitioners (La Fuente Querétaro, 2020[197]; @profesionesqro, 2020[198]). Xertify also works with universities in the region, including Universidad ECCI (@xertifyco, 18 June 2020[199]) and Universidad Quindío (@xertifyco, 12 June 2020[200]) in Colombia, to issue digital diplomas.

Prince Consulting, a software services firm based in Argentina, has also implemented Blockcerts credentialing projects for educational institutions through its subsidiary, OSCity (Prince Consulting, n.d.[201]).

A major regional project, sponsored by the Inter-American Development Bank, is modelled on Alastria: LACChain. Led by a coalition of public and private organisations, the project aims to build a blockchain for the delivery of commercial and public services in the region (ConsenSys, 2020[50]).

Asia Pacific

In Singapore, blockchain credentialing saw quick uptake. A couple of notable start-ups created platforms for issuing credentials anchored to the Ethereum network: Attores, a blockchain development firm offering "smart contracts as a service," and Indorse, a reputation platform designed to match candidates with job opportunities. In 2017, Attores piloted blockchain diplomas with Ngee Ann Technical University (McSpadden, 2017[202]). Since their initial launch, Attores and Indorse have merged, preserving the Indorse brand and pivoting to become a professional development tool for software engineers (Indorse, n.d.[203]). In the same timeframe, the Government of Singapore developed OpenCerts, an open standard for issuing blockchain-anchored digital credentials built by forking and modifying the Blockcerts codebase (OpenCerts, n.d.[118]). It has since strongly encouraged universities to use OpenCerts by partnering with local software development firms who can build issuing applications using the OpenCerts open source reference libraries (OpenCerts, n.d.[77]).

In 2018, the Ministry of Education in Malaysia announced a project to issue academic credentials to the NEM blockchain, built by the Council of ICT Deans at Malaysian Universities (Asia Blockchain Review, 2018[204]).

In 2019, Pallavan Learning Systems (PLS) became the first K-12 institution in India – and potentially the world – to issue Blockcerts to their students. The Pallavan School in Jhalawar, Rajhastan and Vasant Valley School in New Delhi issued multiple types of credentials: School Leaving Certificates, Language Certificates, Character Certificates, Letters of Recommendation, and Five Areas of Development Mark Sheets (Company Newsroom of Learning Machine, 2020[205]).

In 2019, the Hong Kong University of Science and Technology (HKUST) began issuing award certification letters as Blockcerts using a system it built in-house (HKUST Academic Registry, n.d.[206]). HKUST continues to evolve the system to issue degree diplomas and transcripts.

Driving change

The adoption of new technologies presents opportunities and challenges for institutions looking to adapt existing practices and workflows. These challenges broadly fall into two categories: ideas and logistics.

Ideas challenges include:

1. Needing to think in new ways.
2. Emotionally accepting new trade-off decisions: new ideas have different advantages and disadvantages than do old ones.
3. Difficulty imagining what change will look like.
4. Changes to roles and responsibilities that people may feel comfortable with.
5. Difficulty imagining what success will look like.
6. Separating means from ends: the same ends (i.e. security, empowerment, confidence, risk mitigation) may be better achieved using different means from what is employed today.

Logistical challenges include:

1. How do we implement a new technology and associated processes?
2. What happens to our legacy technology investments?
3. How expensive will it be?
4. Who will be there to guide and support us through the change?

5. How can we utilise our internal resources most effectively?
6. Can we go back to the old way of doing things if the new way does not work?
7. Are there hidden traps or tricks we are unaware of?

These challenges can be navigated by skilful leadership within organisations that helps teams implementing new technologies understand both the why (ideas) and the how (logistics) of change. This means helping people think about change in a way that is both enthusiastic about the real benefits it promises and realistic about the work required to successfully bring it about.

In the transition to verifiable credentialing for education, there are two broad stakeholder groups who have an outsized say in how that change comes about: policymakers and educational institutions. This section presents a few suggestions for how these groups may approach this aspect of technological transformation.

What can policymakers do?

Many look to policymakers to understand if what they are doing is in line with the values and laws of their community; policymakers are also often sought out for normative guidance: what should be done? As such, policymakers should consider what kinds of goals they have in mind for their communities and whether or not a new technology is instrumental in achieving those goals.

In the case of verifiable credentials, the advantages are quite clear: they increase trust across the board, speed up economic transactions, streamline the process of applying to schools and jobs, facilitate and add security to cross-border mobility, and empower holders with verifiable records of lifelong learning. Within education, it seems clear that implementing blockchain-based credentialing solutions benefits educational institutions, students, employers, and as a result, entire economies.

When examining existing laws and policies, it is good to ask if a regulation intended to achieve a particular objective (for example, validating that a credential was issued by the correct authority) is tied to a technical implementation that is too specific. Legislation that mandates particular technical implementations can end up stifling innovation, as technology evolves at a rapid clip that often can't be anticipated. As a general rule, effective policy provides frameworks for a solution to a problem rather than being overly prescriptive about the solution itself. This enables the ingenuity of a community to flourish while setting clear guidelines about the boundaries of what is legally acceptable.

As people in positions to set priorities for a community, policymakers can also determine how funds are allocated to new technology projects. This may include spending on R&D initiatives, a public investment fund providing capital to young start-ups, investing in education and training in underserved communities, and providing support for new technology pilots or more mature implementations.

Some governments may choose to fund the build-out of their own blockchain networks or technology standards. However, these projects should bear in mind that the value of verifiable digital credentials lies in two things: 1) international portability and 2) platform independence. Any technology infrastructure, whether built by the public or private sector, that creates lock-in or limits utility across borders will create challenges to adoption and scalability. It will also create additional bureaucratic hassles, which is at odds with the promise of verifiable credentials. Accordingly, government-led technology projects should keep closely in sync with the work of global standards bodies like the W3C, IEEE, ISO, and others, to ensure that what is developed for their populations has usability internationally and across software platforms.

Resources can also be allocated in the form of time and attention. Policymakers taking the time to understand the motivations and concerns of technologists working on new solutions to problems can go a long way towards speeding along the rate of innovation in a community. In taking this initiative, policymakers also open the door for technologists to understand their concerns and ideas as well. Where policymakers and technologists see one another as hostile, cooperation stalls and mistrust can polarise a society. But time spent building something together across differences of opinion and background is time well spent.

What can educational institutions do?

When approaching the idea of using blockchain-anchored verifiable credentials, one of the first things educational institutions can do is recognise that the technology landscape is changing. The 21st-century economy requires portable, verifiable digital credentials, and students will increasingly expect them from any institution they attend. The COVID-19 pandemic, in particular, has highlighted the need for schools to issue portable, secure documents that can be received, shared, and verified safely at a distance. While using blockchain technology may not be in an institution's immediate plan, nevertheless being informed about the landscape of offerings and thinking through implementation best practices can begin today.

As for the question of how to implement the technology, there are many ways for educational institutions to begin issuing and verifying blockchain credentials. Issuing institutions who opt for solutions based on open standards, like W3C Verifiable Credentials and Blockcerts, have two main paths forward: either by building their own applications for credential issuing using the open source reference libraries or by licensing a vendor-supported product or service to issue standards-compliant records. Those who choose solutions that are not based on open standards limit their options to either vendor-supported services or idiosyncratic solutions built and maintained in-house. These may run into data portability and interoperability issues over time.

Whatever path they choose, any institution contemplating the issuance of its official records using the blockchain may use the following brief recommendations as a guide.

1. *Identify use cases.* The blockchain is the most secure verification structure available for credentials, so it is best suited for records that must be verified with a high level of certainty. In education, these records include, but are not limited to, diplomas and degrees, transcripts, school leaving certificates, diploma supplements, comprehensive learner records, student IDs, and examination results. For education professionals, teaching licenses or warrants or certification of continuing professional education are promising places to start. Blockchain is less needed for more ephemeral credentials that only need temporary validation. However, the latter may "stack up" to higher-level credentials that are anchored to a blockchain.

2. *Future-proof your initiative by committing to open standards.* Many vendors in the blockchain credentialing space have created custom solutions in which credentials only display and verify within their vendor-controlled software system or blockchain network. This ties ongoing access, sharing, and verification of credentials to that particular vendor. Rather than making your organisation's use of official records dependent upon a software provider and risking having to "re-do" a digital credentialing implementation if you switch services, choose open architectures to make your credentialing solution a lasting one. Insist upon the use of open standards (Blockcerts, verifiable credentials) and blockchains that are either public (like Bitcoin or Ethereum) or permissioned and private blockchains maintained by institutions with a public mandate, like governments or consortia of well-established organisations (for example, the EBSI). This increases the likelihood that those organisations will have the resources and longevity to maintain an expensive permissioned or private blockchain network over time.

If, as an institution, you choose to contract with a vendor to provide a blockchain credentialing solution rather than building one in-house, below are a series of recommendations that will help you make the best vendor selection for your institution.

1. *Choose a software provider experienced in the delivery of credentialing solutions based on open standards.* The market for digital credentials is growing, and with it a growing number of vendors are offering solutions. A subset of those vendors are committed to using open standards, and a smaller subset of those vendors have experience successfully implementing digital credentialing projects. An experienced vendor should be able to point to case studies in which their software was used to successfully deliver and build on a high-stakes digital credentialing project, then describe how their process could be adapted to your institution's needs.

2. *Select a vendor who will help you create and execute on an implementation plan.* Blockchain credentialing is a relatively new field. As such, a change management component is a critical part of any successful implementation. Ask vendor candidates about their typical onboarding and implementation process and how

they will help you introduce this new credentialing initiative alongside your current credentialing practices. How will you define objectives? How will you measure success? A strong vendor partner is available to assist with these questions.

3. *Be prepared to budget for your blockchain credentialing initiative.* Blockchain credentialing is an exciting field that leverages cutting-edge new technology to deliver unparalleled convenience and certainty with regards to digital claims. This is an advanced technology, however; not a commodity. Your institution should be prepared to set aside an annual budget item to cover the costs of the initiative, much in the same way you budget for issuing paper or PDF digital records. Over time, you can sunset older ways of credentialing and potentially create savings by switching fully to a verifiable credentials model. Financial return on investment will vary based on your institution's current credentialing practices, blockchain credentialing vendor selection, and implementation model. You may want to request a return on investment (ROI) analysis from vendors prior to selecting your provider(s), but be aware that this will require you reveal the current costs (in time, personnel, and money) of your existing credentialing practices.

While there is always a switching cost associated with the adoption of new technologies, it is time for educational institutions to begin the transition to blockchain-anchored verifiable credentials. Not only must the proliferation of academic and professional fraud be firmly rooted out, but the continuing growth in the global mobility of learners and workers alongside mass displacement due to conflict, natural disasters (including climate change), and most recently the pandemic make verifiable credentials a critical precondition not only for streamlining the movement of learners between educational institutions and from education to employment, but of preserving public safety as people live and work across highly varied jurisdictions and geographies (Jagers, n.d.[207]). As educational institutions have been forced to move operations online and internationalise in unprecedented ways due to the COVID-19 pandemic, the shift to verifiable digital credentials forms an integral part of any institution's overall digitisation strategy.

Conclusion

The global education landscape is rapidly changing. By 2030, over 7 million students are expected to travel internationally for higher education (Holon IQ, 2018[208]). The rapid growth of developing economies in Asia and Africa is expected to fuel a massive expansion of the education sector, adding over 350 million post-secondary graduates and 800 million secondary graduates to the global marketplace (Holon IQ, 2018[209]). The world's teaching capacity is already strained, but over 100 million new teachers will be required to meet the projected need (Holon IQ, 2018[209]).

This represents a significant opportunity for new technologies, including artificial intelligence, process automation, digital education marketplaces, and blockchain to help scale global education infrastructure. The vital role of technology in meeting the needs of this generation of students is expected to propel the global education technology market to USD 10 trillion by 2030 (Holon IQ, 2018[209]). In all economies, the need for technology-driven upskilling and reskilling is keenly felt.

COVID-19 has slowed enrolment in higher education somewhat, but not uniformly, and not permanently (Hess, 2020[210]; Miller, 2020[211]). It is mostly a response to the still heavily brick-and-mortar status quo for the delivery of education. As valuable as such a model is, it is seeing difficulty both with responding to a communicable disease pandemic and scaling its operations to accommodate virtual education delivery. In addition, many brick-and-mortar higher education institutions are loss-making, relying on donations and grants to make ends meet. Many were already financially precarious before COVID-19 struck, and the pandemic has accelerated insolvency for some (Thys, 2020[212]).

As institutions adapt, however, we will likely see enrolments continue to grow as a result of projected population growth. Educational institutions will increasingly rely on time-saving and cost-saving technologies to streamline administrative operations, deliver educational content, conduct assessments, credential students, and connect them with employment opportunities.

What does this mean for credentialing? The most easily credentialed skills – and the most valuable credentials to have in the labour marketplace – will likely continue to be associated with hard skills, or skills that can be most easily quantified and tested (Trilling and Fadel, 2009[213]). Jobs and professions requiring credentials are most likely to remain in technical and analytical fields. This does not mean, of course, that soft skills are not critical for professional success (Beheshti, 2020[214]); only that they are less easily quantified, tested, and credentialed. Despite flourishing movements to credential a broader range of learning experiences (Parrish, Fryer and Parks, 2017[215]), competencies, and skills (Reed, 2016[216]), employers are likely to continue relying on personal recommendations, evidence of past achievements, and reputation to assess soft skills in prospective applicants.

Nevertheless, where credentials are required or desired, being able to share, receive, and verify them instantly in a portable, interoperable digital format will quickly become a table-stakes expectation. The blockchain is an ideal technology to enable this, since it acts as a distributed, digital verification infrastructure for records and claims. With growing momentum among technologists and policymakers to return control of personal data back to users, an international standards movement has arisen to ensure that blockchain-anchored credentials are issued, stored, shared, and verified in a manner that protects user privacy and is independent of any particular vendor infrastructure.

This paper has described how these international technology standards for verifiable credentials are being applied in the sphere of education. As this field is evolving rapidly, the market landscape described here will likely look very different in even a few months. However, the standards described here will likely remain critical connective tissue for a rapidly accelerating global ecosystem of credential exchange. Institutions looking to implement a verifiable credentialing program should look to these standards to ensure that their projects protect privacy, empower users, and ensure credential mobility for all.

The early adoption of verifiable credentials for academic records by some educational institutions has placed the education industry at the forefront of one of the most significant technological innovations since the advent of the internet: a new "trust layer" for digital claims that has broad applicability across industries and use cases. Much as blockchain technology represents a distributed, tamper-proof version of an accounting ledger, verifiable credentials standards are mechanisms to detect tampering and validate a credential's provenance and recipient in a decentralised manner. This technology will likely see parabolic adoption in coming years as institutions move to prevent fraud, streamline the processing and verification of claims, and return control of personal data to end users. The end result is an ecosystem that enables individuals to maintain a lifelong record of achievement and seamlessly transfer between institutions and geographies to live, work, and study.

As the world of learning and work continues to become more mobile and interconnected, people require higher levels of digital trust in order to live, learn, and do business together. Facilitating that trust is both the value and the promise of blockchain technology.

Annex 11.A
Appendix: Technical glossary

Blockchain: A type of distributed ledger that records an append-only, immutable database of transactions. Blockchains were originally used to maintain a record of who owns digital currency, thereby preventing the reproduction and tampering of digital assets. This same technology can be employed to verify the integrity of and track ownership of any digital asset, including academic credentials.

Cryptography: "Secret writing." A way of protecting information by using codes so that only intended recipients may read or use it.

Decentralised Identifier (DID): A globally unique identifier that does not require a centralised registration authority because it is registered with distributed ledger technology or other form of decentralised network.

Distributed Ledger Technology (DLT): A database that is consensually shared and synchronised across multiple sites. Each site of the DLT network runs a part of its infrastructure and can write or access the entries shared across that network based on permission controls. Any changes or additions made to the ledger are reproduced across all sites. A blockchain is a type of distributed ledger, usually with an attendant cryptocurrency.

Self-Sovereign Identity (SSI): "A set of technical standards and a set of community-promulgated principles seeking to enable a shift towards more individual control over digital identities and personal data."[1]

Verifiable Credential (VC): A digital credential that is tamper evident and whose provenance (authorship) can be cryptographically verified.

Note

1. T3 Innovation Network. "Applying Self-Sovereign Identity Principles to Interoperable Learning Records: Principles, Challenges, and Community Guidance." US Chamber of Commerce Foundation. June 2020. p. 2; https://www.uschamberfoundation.org/sites/default/files/media-uploads/Applying%20SSI%20Principles%20to%20ILRs%20Report.pdf

References

(n.a.) (n.d.), *ND Integrated Learner Record: Phase II Project Charter*, https://docs.google.com/document/d/1hiKCNepcPQkCDHy7yFT9x613RWjK7Mjk/edit (accessed on 5 February 2021). [153]

@profesionesqro (2020), *Twitter*, https://twitter.com/profesionesqro/status/1256225566631567360. [198]

@xertifyco (12 June 2020), *Twitter*, https://twitter.com/xertifyco/status/1271571738044792832. [200]

@xertifyco (18 June 2020), *Twitter*, https://twitter.com/xertifyco/status/1273813449311760390. [199]

0xcert (n.d.), *Blockchain API gateway*, https://0xcert.org/ (accessed on 16 November 2020). [170]

Abdou, M. (2019), *MOU between Zewail City and Intelli Coders, LLC to use blockcred.io for academic credentials issuing*, Medium, https://medium.com/blockcred/mou-between-zewail-city-and-intelli-coders-llc-to-use-blockcred-io-a3124bafe143 (accessed on 26 June 2020). [188]

Accredited Online Colleges (n.d.), *Fake Schools, Fake Degrees: Avoiding Diploma Mills*, https://www.accredited-online-college.org/avoiding-diploma-mills/ (accessed on 16 November 2020). [119]

Al Ali, N. and E. van der Walt (2018), *Blockchain Island Dream Is 'Calculated Risk' Says Malta Leader*, Bloomberg, https://www.bloomberg.com/news/articles/2018-09-10/blockchain-island-dream-is-calculated-risk-says-malta-leader (accessed on 16 November 2020). [163]

Alastria (n.d.), *Alastria*, https://alastria.io/en/la-red/ (accessed on 13 November 2020). [49]

Alden, L. (2021), *A Century of Fiscal and Monetary Policy: Inflation vs Deflation, LynAlden.com.*, https://www.lynalden.com/fiscal-and-monetary-policy/. [15]

Ali, A. (2021), *Bitcoin is the Fastest Asset to Reach a $1 Trillion Market Cap*, Visual Capitalist, https://www.visualcapitalist.com/bitcoin-is-the-fastest-asset-to-reach-a-1-trillion-market-cap/. [19]

Allen, C. (2016), *The Path to Self-Sovereign Identity, Life with Alacrity*, http://www.lifewithalacrity.com/2016/04/the-path-to-self-soverereign-identity.html (accessed on 17 November 2020). [100]

Alliance, B. (n.d.), *Badge Alliance*, https://www.badgealliance.org (accessed on 28 June 2020). [85]

Amazon (n.d.), *Blockchain on AWS*, https://aws.amazon.com/blockchain/ (accessed on 13 November 2020). [45]

American Chemical Society (2019), *Estimating the environmental impact of Bitcoin mining*, Science Daily, https://www.sciencedaily.com/releases/2019/11/191120080246.htm. [11]

AMSI (n.d.), *Our Schools | First in the World to Implement Blockchain-based Digital Academic Credentials for Students*, https://amsi.ae/amsi-news/183-our-schools-first-in-the-world-to-implement-blockchain-based-digital-academic-credentials-for-students (accessed on 17 November 2020). [184]

Arizona State University (n.d.), *Trusted Learner Network*, https://trust.asu.edu/ (accessed on 16 November 2020). [151]

Ark Invest (2021), *SolarBatteryBitcoin.*, Github, https://github.com/ARKInvest/SolarBatteryBitcoin (accessed on 4 May 2021). [12]

Asia Blockchain Review (2018), *Malaysian Ministry of Education Applies Blockchain Technology as a Certificate Verification against Degree Fraud*, https://www.asiablockchainreview.com/malaysian-ministry-of-education-apply-blockchain-technology-as-a-certificate-verification-against-degree-fraud/ (accessed on 17 November 2020). [204]

Baydakova, A. (2019), *Bitcoin Mining Farms Are Flourishing on the Ruins of Soviet Industry in Siberia*, CoinDesk, https://www.coindesk.com/bitcoin-mining-farms-are-flourishing-on-the-ruins-of-soviet-industry-in-siberia (accessed on 17 November 2020). [7]

Beheshti, N. (2020), *5 Of The Most In-Demand Soft Skills Companies Are Looking For This Year*, Forbes, https://www.forbes.com/sites/nazbeheshti/2020/01/28/5-of-the-most-in-demand-soft-skills-companies-are-looking-for-this-year/?sh=5393bb2e56c7 (accessed on 17 November 2020). [214]

Benay, A. (2019), *Securing the future of talent mobility in the Government of Canada*, https://tbs-blog.canada.ca/en/securing-future-talent-mobility-government-canada (accessed on 13 November 2020). [76]

Bernard, Z. (2018), *Everything you need to know about Bitcoin, its mysterious origins, and the many alleged identities of its creator*, Business Insider, https://www.businessinsider.com/bitcoin-history-cryptocurrency-satoshi-nakamoto-2017-12?IR=T (accessed on 17 November 2020). [28]

Bhatia, N. (2021), *Layered Money: From Gold and Dollars to Bitcoin and Central Bank Digital Currencies.*, Nikhil Bhatia. [14]

Binance Academy (n.d.), *Byzantine Fault Tolerance Explained*, Binance Academy, https://academy.binance.com/blockchain/byzantine-fault-tolerance-explained (accessed on 25 June 2020). [4]

BitDegree (2017), *Revolutionizing Global Education with Blockchain - White Paper*, https://www.bitdegree.org/bitdegree-vision.pdf (accessed on 17 November 2020). [107]

Blockcerts (n.d.), *Blockcerts*, https://www.blockcerts.org (accessed on 16 November 2020). [111]

BlockCred (n.d.), *BlockCred*, https://www.blockcred.io/ (accessed on 17 November 2020). [189]

Bloxberg (2020), *"bloxberg: The Trusted Research Infrastructure"*, https://bloxberg.org/wp-content/uploads/2020/02/bloxberg_whitepaper_1.1.pdf (accessed on 5 February 2021). [145]

Brody, P. (2019), *How public blockchains are making private blockchains obsolete*, Ernst & Young, https://www.ey.com/en_us/innovation/how-public-blockchains-are-making-private-blockchains-obsolete (accessed on 17 November 2020). [62]

Buterin, V. (2016), *Let's talk about the projected coin supply over the coming years*, Reddit, https://www.reddit.com/r/ethereum/comments/5izcf5/lets_talk_about_the_projected_coin_supply_over/dbc66rd/ (accessed on 17 November 2020). [60]

Cabrol, M. (2018), *Technology is Creating a Human-Centric Economy, and It's Based on Trust, Inter-American Development Bank*, https://blogs.iadb.org/caribbean-dev-trends/en/technology-is-creating-a-human-centric-economy-and-its-based-on-trust/ (accessed on 16 November 2020). [131]

Castillo, M. (2019), *Hyperledger Unanimously Approves First Ethereum Codebase For Enterprises*, Forbes, https://www.forbes.com/sites/michaeldelcastillo/2019/08/29/hyperledger-unanimously-approves-first-ethereum-codebase-for-enterprises/#838c037794cc (accessed on 17 November 2020). [43]

Castro-Iragorri, C. (2018), *Hyperledger Supporting Blockcerts Compliant Digital Diplomas across Colombian Universities*, Hyperledger Global Forum, https://hgf18.sched.com/event/G8rr/hyperledger-supporting-blockcerts-compliant-digital-diplomas-across-colombian-universities-carlos-castro-iragorri-universidad-del-rosario (accessed on 16 November 2020). [135]

CEF Digital (n.d.), *EBSI Documentation home*, https://ec.europa.eu/cefdigital/wiki/display/CEFDIGITALEBSI/EBSI+Documentation+home (accessed on 17 November 2020). [175]

CEF Digital (n.d.), *European Blockchain Services Infrastructure*, European Commission, https://ec.europa.eu/cefdigital/wiki/display/CEFDIGITAL/EBSI (accessed on 13 November 2020). [40]

CEF Digital (n.d.), *Use Cases and Functional Documentation*, https://ec.europa.eu/cefdigital/wiki/display/CEFDIGITALEBSI/Use+Cases+and+Functional+Documentation (accessed on 17 November 2020). [176]

Chainalysis (n.d.), *Chainalysis*, https://www.chainalysis.com/ (accessed on 13 November 2020). [63]

CNN (2020), *Almost 1 in 3 pilots in Pakistan have fake licenses, aviation minister says*, https://edition.cnn.com/2020/06/25/business/pakistan-fake-pilot-intl-hnk/index.html (accessed on 16 November 2020). [123]

CNN (n.d.), *University blockchain experiment aims for top marks*, https://edition.cnn.com/videos/tv/2018/06/28/blockhain-university-dubai-global-gateway.cnn (accessed on 17 November 2020). [183]

Cocks, P. (2017), *Malta first to launch education blockchain certification*, Malta Today, https://www.maltatoday.com.mt/news/national/80704/government_launches_first_blockhain_initiative_in_education (accessed on 16 November 2020). [165]

Coin Telegraph (n.d.), *What Is Lightning Network And How It Works*, Cointelegraph, https://cointelegraph.com/lightning-network-101/what-is-lightning-network-and-how-it-works (accessed on 27 June 2020). [33]

CoinMarketCap (n.d.), *CoinMarketCap*, https://coinmarketcap.com/ (accessed on 25 June 2020). [20]

CoinTelegraph (n.d.), *What Are Smart Contracts? Guide For Beginners*, https://cointelegraph.com/ethereum-for-beginners/what-are-smart-contracts-guide-for-beginners (accessed on 27 July 2020). [25]

Company Newsroom of Learning Machine (2020), *First Blockcerts Issued in India: Pallavan School and Vasant Valley School, Facilitated by Pallavan Learning Systems and Learning Machine, Issue Blockchain Credentials*, https://learningmachine.newswire.com/news/first-blockcerts-issued-in-india-pallavan-school-and-vasant-valley-21072917 (accessed on 17 November 2020). [205]

Company Newsroom of Learning Machine (2018), *Press Release: KAUST Set to Be First University in Middle East to Issue Blockchain Credentials Using Blockcerts Open Standard*, https://learningmachine.newswire.com/news/kaust-set-to-be-first-university-in-middle-east-to-issue-blockchain-20703496 (accessed on 17 November 2020). [190]

ConsenSys (2020), *Latin American and Carribean Citizens Are Coordinating COVID-19 Deconfinement With This New Ethereum Application*, ConsenSys, https://consensys.net/blog/news/latin-american-and-carribean-citizens-are-coordinating-covid-19-deconfinement-with-this-new-ethereum-application/ (accessed on 17 November 2020). [50]

ConsenSys (2019), *The Thirdening: What You Need To Know*, ConsenSys, https://media.consensys.net/the-thirdening-what-you-need-to-know-df96599ad857 (accessed on 17 November 2020). [58]

ConsenSys (n.d.), *About*, https://consensys.net/about/ (accessed on 26 June 2020). [37]

Crace, A. (2019), *Digitary & Evernym Collaborate*, The PIE News, https://thepienews.com/news/digitary-evernym-collaborate-boost-credentialing/ (accessed on 16 November 2020). [162]

Credential Engine (n.d.), *Credential Transparency Description Language (CTDL) Handbook*, https://credreg.net/ctdl/handbook (accessed on 13 November 2020). [84]

Cuthbertson, A. (2017), *Bitcoin Mining on Track to Consume All of the World's Energy by 2020, Newsweek*, https://www.newsweek.com/bitcoin-mining-track-consume-worlds-energy-2020-744036 (accessed on 17 November 2020). [10]

Dagger, M. (2020), *A Comprehensive Guide to Decentralized Stablecoins*, Hackernoon, https://hackernoon.com/a-comprehensive-guide-to-decentralized-stablecoins-22f66553c807 (accessed on 17 November 2020). [55]

Decentralized Identity Foundation (n.d.), *Decentralized Identity Foundation*, https://identity.foundation/ (accessed on 13 November 2020). [93]

DiChristopher, T. (2017), *No, bitcoin isn't likely to consume all the world's electricity in 2020.*, CNBC, https://www.cnbc.com/2017/12/21/no-bitcoin-is-likely-not-going-to-consume-all-the-worlds-energy-in-2020.html (accessed on 17 November 2020). [9]

Digital Credentials Consortium (2020), *Building the digital credential infrastructure for the future*, https://digitalcredentials.mit.edu/wp-content/uploads/2020/02/white-paper-building-digital-credential-infrastructure-future.pdf. [146]

Dinkins, D. (2017), *SegWit Locked-in, But Questions Remain About Bitcoin's Level of Decentralization*, CoinTelegraph, https://cointelegraph.com/news/Bitcoin-is-decentralized-but-not-distributed-and-that-fact-likely-contributed-to-Bitcoins-civil-war (accessed on 17 November 2020). [36]

Du Seuil, D. and C. Pastor (2019), *Understanding the European Self-Sovereign Identity Framework (ESSIF)*, https://ssimeetup.org/understanding-european-self-sovereign-identity-framework-essif-daniel-du-seuil-carlos-pastor-webinar-32/ (accessed on 13 November 2020). [79]

Durant, E. (2017), *Digital Diploma debuts at MIT*, MIT News, https://news.mit.edu/2017/mit-debuts-secure-digital-diploma-using-bitcoin-blockchain-technology-1017 (accessed on 16 November 2020). [138]

Education, T. (n.d.), *The Internet of Education*, https://docs.google.com/document/d/1tWiaiU95OvS_ktVDbSjQlwl6CZVUzgikWuTunTde8U0/edit (accessed on 13 November 2020). [97]

Enterprise Ethereum Alliance (n.d.), *Entethalliance*, https://entethalliance.org/ (accessed on 13 November 2020). [38]

EthHub (n.d.), *Monetary Policy*, https://docs.ethhub.io/ethereum-basics/monetary-policy/ (accessed on 13 November 2020). [57]

EU Blockchain Observatory and Forum (2019), *EU Blockchain Observatory and Forum: Workshop Report, e-Identity, Brussels, November 7 2019*, https://www.eublockchainforum.eu/sites/default/files/reports/workshop_5_report_-_e-identity.pdf. [80]

European Commission (n.d.), *ESCO and the new Europass: occupations and skills for lifelong career management*, https://ec.europa.eu/esco/portal/news/fd542061-178a-4a04-81ae-b3b0eccb3edd (accessed on 17 November 2020). [178]

Everis (European Commission) (n.d.), *Europass framework for digitally-signed credentials*, https://ec.europa.eu/futurium/en/system/files/ged/europass_background-info_framework-digitally-signed-credentials.pdf (accessed on 17 November 2020). [179]

Ezell, A. and J. Bear (2012), *Degree Mills: The Billion-dollar Industry That Has Sold Over A Million Fake Diplomas*, Prometheus Books. [120]

Federation of State Medical Boards (2019), *Healthcare and Digital Credentials: Technical, Legal, and Regulatory Considerations*, https://www.fsmb.org/digitalcredentials/ (accessed on 16 November 2020). [113]

Fitzpatrick, L. (2020), *Expect A Huge Backlash If Stablecoins Are Banned*, Forbes, https://www.forbes.com/sites/lukefitzpatrick/2020/04/27/expect-a-huge-backlash-if-stablecoins-are-banned/#634415c261d2. [56]

Foxley, W. (2020), *Vitalik Buterin Clarifies Remarks on Expected Launch Date of Eth 2.0.*, CoinDesk, https://www.coindesk.com/vitalik-buterin-clarifies-remarks-on-expected-launch-date-of-eth-2-0. [29]

Friedman, S. (2019), *Union Public Schools Launches Blockchain-Based Credentials*, https://thejournal.com/articles/2019/04/30/union-public-schools-launches-blockchain-based-credentials.aspx. [140]

Gibson, K. (2017), *Your MD may have a phony degree*, CBS News, https://www.cbsnews.com/news/your-md-may-have-a-phony-degree/ (accessed on 16 November 2020). [122]

Global Blockchain Business Council (n.d.), *Chain Reaction: Blockchain Enters the Mainstream - 2020 annual report*, https://www.lw.com/thoughtLeadership/gbbc-report-blockchain-enters-mainstream (accessed on 17 November 2020). [191]

Government of Canada Talent Cloud (n.d.), *Talent Cloud*, https://talent.canada.ca/en (accessed on 16 November 2020). [158]

GradBase (n.d.), *GradBase*, https://www.gradba.se/en/ (accessed on 16 November 2020). [171]

Grech, A. and A. Camilleri (2017), *Blockchain in Education*, European Commission Joint Research Centre | Edited by Andreia Inamorato dos Santos, https://publications.jrc.ec.europa.eu/repository/bitstream/JRC108255/jrc108255_blockchain_in_education(1).pdf. [130]

Greenspoon, A. (2018), *The value of 'free agents' inside the public service*, Policy Options, https://policyoptions.irpp.org/magazines/november-2018/the-value-of-free-agents-inside-the-public-service/ (accessed on 16 November 2020). [159]

Hague Convention (1961), *Convention Abolishing the Requirement of Legalisation for Foreign Public Documents*, https://www.hcch.net/en/instruments/conventions/full-text/?cid=41. [126]

Hamdani, Z. (2020), *ARUCC teams up with Digitary to build Made for Canada network*, Global Education Times, https://www.globaleducationtimes.org/news/north-america/arucc-teams-up-with-digitary-to-build-made-for-canada-network/1690/?fbclid=IwAR271A6g9UK3yjtND-Uclm3ymmD3h8hBToEsPa_MG4hUdjAjzvODMD81I0E (accessed on 16 November 2020). [161]

Hargrave, S. and E. Karnoupakis (2020), *Blockchain Success Stories: Case Studies from the Leading Edge of Business*, O'Reilly Media. [137]

Hess, A. (2020), *College enrollment is down because of the pandemic—and community colleges have been hit the hardest*, CNBC, https://www.cnbc.com/2020/10/01/how-the-coronavirus-pandemic-has-impacted-college-enrollment.html (accessed on 17 November 2020). [210]

HKUST Academic Registry (n.d.), *Electronic Documents Verification Program*, https://registry.ust.hk/verify/info (accessed on 17 November 2020). [206]

Holon IQ (2018), *Education in 2030: The $10 Trillion dollar question*, https://www.holoniq.com/2030/ (accessed on 17 November 2020). [209]

Holon IQ (2018), *Global Student Flows*, https://www.holoniq.com/global-student-flows (accessed on 17 November 2020). [208]

Hyland Credentials (n.d.), *Customer Stories*, https://www.hylandcredentials.com/customer-stories (accessed on 16 November 2020). [115]

Hyland Newsroom (2020), New Mexico Higher Education Department selects Hyland Credentials for Higher Education, https://news.hyland.com/new-mexico-higher-education-department-selects-hyland-credentials-for-higher-education/ (accessed on 16 November 2020). [144]

IBM (n.d.), *IBM Blockchain Services. Blockchain consulting that turns strategy into business outcomes.*, https://www.ibm.com/blockchain/services/ (accessed on 13 November 2020). [46]

IEEE (n.d.), *Learning Technology Standards Committee (LTSC)*, https://sagroups.ieee.org/ltsc/ (accessed on 13 November 2020). [94]

IMS Global (n.d.), *Comprehensive Learner Record*, https://www.imsglobal.org/activity/comprehensive-learner-record (accessed on 13 November 2020). [86]

Indorse (n.d.), *Indorse*, https://indorse.io/metamorph (accessed on 17 November 2020). [203]

Insights, L. (2020), *South Korea has global ambitions for decentralized identity systems*, https://www.ledgerinsights.com/south-korea-decentralized-identity-ambitions/ (accessed on 13 November 2020). [78]

Internet Identity Workshop (n.d.), *Internet Identity Workshop*, https://internetidentityworkshop.com/ (accessed on 28 June 2020). [88]

Internet of Education (n.d.), *Internet of Education*, https://www.internetofeducation.org/ (accessed on 13 November 2020). [96]

Jagers, C. (2018), *The Blockchain in Education*, Learning Machine Blog, https://medium.com/learning-machine-blog/the-blockchain-in-education-5a322fe9fe86. [110]

Jagers, C. (n.d.), *A Passport to Greater Opportunity*, Learning Machine Blog, https://www.learningmachine.com/passport-greater-opportunity/ (accessed on 17 November 2020). [207]

Jagers, C. (n.d.), *Future Proof*, Learning Machine Blog, https://www.learningmachine.com/future-proof (accessed on 16 November 2020). [136]

Jamaica Observer (2019), *CXC to provide students with e-certificates after successful trial*, http://www.jamaicaobserver.com/latestnews/CXC_to_provide_students_with_e-certificates_after_successful_trial?profile=1228 (accessed on 17 November 2020). [195]

Jessel, B. (2020), *Ethereum, Fabric, Corda, And Multichain. Only One Is Government Ready - New Report*, Forbes, https://www.forbes.com/sites/benjessel/2020/04/21/ethereum-fabric-corda-and-multichain-only-one-is-government-readynew-report/#4f0443ae263b. [53]

Jonkers, R. (2018), *SAP Blockchain for Higher Education*, https://www.slideshare.net/RobJonkers/sap-blockchain-for-higher-education (accessed on 16 November 2020). [169]

Kar, I. (2016), *The developers behind Ethereum are hacking the hacker that hacked it*, QZ, https://qz.com/713078/the-developers-behind-ethereum-are-hacking-the-hacker-that-hacked-it/ (accessed on 24 November 2020). [34]

Kelly, R. (2018), *Southern New Hampshire U Issues Blockchain Credentials to College for America Grads*, https://campustechnology.com/articles/2018/06/11/southern-new-hampshire-u-issues-blockchain-credentials-to-college-for-america-grads.aspx (accessed on 16 November 2020). [139]

Kenobit, O. (2018), *Hyperbitcoinization: Winner Takes All*, Medium, https://medium.com/coinmonks/hyperbitcoinization-winner-takes-all-69ab59f9695f (accessed on 24 November 2020). [17]

Konings, L. (2019), *Electronic Platform for Adult Learning in Europe (EPALE): EMC launches Common Microcredential Framework*, https://epale.ec.europa.eu/en/content/emc-launches-common-microcredential-framework (accessed on 24 November 2020). [83]

Krawisz, D. (2014), *Hyperbitcoinization, Satoshi Nakamoto Institute*, Satoshi Nakamoto Institute, https://nakamotoinstitute.org/mempool/hyperbitcoinization/ (accessed on 24 November 2020). [16]

La Fuente Querétaro (2020), *Profesionistas podrán tramitar cédula en línea*, https://periodicolafuente.com/profesionistas-tramitar-cedula-en-linea/ (accessed on 17 November 2020). [197]

Learning Economy Foundation (n.d.), *Broward OpenCLR Lab*, https://www.learningeconomy.io/browardlab (accessed on 16 November 2020). [154]

Learning Economy Foundation (n.d.), *C-Lab: A Learning Economy Co-Lab*, https://www.learningeconomy.io/clab (accessed on 16 November 2020). [152]

Learning Economy Foundation (n.d.), *Learning Economy Foundation*, https://www.learningeconomy.io/ (accessed on 13 November 2020). [95]

Learning Economy Foundation (n.d.), *Memo: AP-Lab Co-Lab Learning Ecosystems for the Asia-Pacific Region*, https://docs.google.com/document/d/1pzdcZdus734OF2olk1ZZjFY6vdjZY3G8-IVPZlqf4cA/edit#heading=h.9tijzbwkqbxf (accessed on 24 November 2020). [155]

Learning Economy Foundation (n.d.), *Roadmap*, https://www.learningeconomy.io/#Roadmap (accessed on 16 November 2020). [109]

Learning Machine (2018), *Blockcerts Enables Multi-Chain Issuing and Verification of Official Documents*, https://learningmachine.newswire.com/news/blockcerts-enables-multi-chain-issuing-and-verification-of-official-20680828 (accessed on 16 November 2020). [134]

Learning Machine Company Newsroom (2019), *Press Release: Maryville University Partners With Learning Machine to Issue Digital Diplomas Using Blockcerts*, https://learningmachine.newswire.com/news/maryville-university-partners-with-learning-machine-to-issue-digital-21036653 (accessed on 24 November 2020). [142]

Learning Machine Newsroom (2016), *Learning Machine and MIT Media Lab Release Blockchain Technology for Educational Credentials*, https://learningmachine.newswire.com/news/learning-machine-and-mit-media-lab-release-blockchain-technology-for-11633933 (accessed on 16 November 2020). [114]

Ledger Insights (2020), *Workday sees blockchain as key to employee credentialing*, https://www.ledgerinsights.com/workday-blockchain-employee-credentialing/ (accessed on 16 November 2020). [148]

Leising, M. and O. Kharif (2020), *Ethereum Races Clock to Collect Enough Coins for Big Upgrade*, Bloomberg, https://www.bloomberg.com/news/articles/2020-11-23/ethereum-races-clock-to-collect-enough-coins-for-huge-upgrade. [27]

Lemoie, K. and L. Souares (2020), *Connected Impact: Unlocking Education and Workforce Opportunity through Blockchain*, American Council on Education, https://www.acenet.edu/Documents/ACE-Education-Blockchain-Initiative-Connected-Impact-June2020.pdf (accessed on 24 November 2020). [44]

Longino Torres, J. (2019), *¡Hola, futuro! Inicia Tec era de títulos universitarios en blockchain*, Tecnologico de Monterrey, https://tec.mx/es/noticias/nacional/educacion/hola-futuro-inicia-tec-era-de-titulos-universitarios-en-blockchain?&utm_source=twitter&utm_medium=social-media&utm_campaign=addtoany (accessed on 17 November 2020). [192]

Longstaff, B. (2018), *Why hashed Personally identifiable information (PII) on the blockchain can be safe*, Medium, https://medium.com/meeco/why-hashed-personally-identifiable-information-pii-on-the-blockchain-can-be-safe-b842357b9663 (accessed on 24 November 2020). [66]

Maaghul, R. (2019), *Why ODEM is building on the Ethereum blockchain*, ODEM Blog, https://blogs.odem.io/why-odem-is-building-on-the-ethereum-blockchain (accessed on 24 November 2020). [106]

Malone, K. (2020), *Blockchain Saving HHS $30M on First Accelerate Contract*, MeriTalk, https://www.meritalk.com/articles/blockchain-saving-hhs-30m-on-first-accelerate-contract/ (accessed on 24 November 2020). [54]

McCauley, A. (2019), *Unblocked: How Blockchain will Change Your Business (and What to Do About It)*, O'Reilly. [129]

McMaster University Office of the Registrar (n.d.), *Digital diplomas*, https://registrar.mcmaster.ca/digitaldiplomas/#:~:text=McMaster%20University%20issues%20graduates%20a,protect%20and%20verify%20the%20diploma (accessed on 16 November 2020). [116]

McSpadden, K. (2017), *Ngee Ann Polytechnic to pilot blockchain diplomas*, partners Attores, e27, https://e27.co/ngee-ann-polytechnic-to-pilot-blockchain-diplomas-partners-attores-20170526/ (accessed on 17 November 2020). [202]

McWhinney, J. (2019), *Why Governments Are Afraid of Bitcoin.*, Investopedia, https://www.investopedia.com/articles/forex/042015/why-governments-are-afraid-bitcoin.asp (accessed on 24 November 2020). [21]

Mearian, L. (2018), *Will Blockchain Run Afoul of GDPR? (Yes and no)*, Computerworld, https://www.computerworld.com/article/3269750/will-blockchain-run-afoul-of-gdpr-yes-and-no.html (accessed on 24 November 2020). [65]

Meetup (2020), *[Hyperledger SF] Workday and AWS discuss Hyperledger Fabric [Mon, Feb 24]*, https://www.meetup.com/Hyperledger-SF/events/267518767/ (accessed on 16 November 2020). [149]

Messari (2019), *Timechains*, https://messari.io/article/timechains (accessed on 24 November 2020). [13]

Messari.io. (2019), *On-chain Governance*, https://messari.io/article/on-chain-governance (accessed on 24 November 2020). [32]

Miller, L. (2020), *Four Bright Spots in a Year That's Irreversibly Changed Higher Ed Admissions*, Liaison International for Inside Higher Ed, https://narratives.insidehighered.com/bright-spots-year-changed-admissions/index.html (accessed on 17 November 2020). [211]

MIT Media Lab (n.d.), *Digital Certificates Project*, https://certificates.media.mit.edu/ (accessed on 16 November 2020). [133]

MIT Media Lab (n.d.), *Project: Digital Academic Credentials*, https://www.media.mit.edu/projects/media-lab-digital-certificates/overview/ (accessed on 16 November 2020). [132]

MIT Open Learning (2020), *University-led Digital Credentials Consortium explores technology for digital academic credentials*, https://openlearning.mit.edu/news/university-led-digital-credentials-consortium-explores-technology-digital-academic-credentials (accessed on 24 November 2020). [90]

Morris, C. (2018), *Bitcoin Mining Uses More Energy Than Homes Do in Iceland*, Fortune, https://fortune.com/2018/02/13/iceland-bitcoin-mining-electricity/ (accessed on 24 November 2020). [8]

Moskov, A. (2018), *What is the Byzantine Generals Problem?*, CoinCentral, https://coincentral.com/byzantine-generals-problem/ (accessed on 9 November 2020). [2]

Munro, A. (2018), *Good clean use case: Bahamian students graduate on the blockchain*, finder, https://www.finder.com.au/good-clean-use-case-bahamian-students-graduate-on-the-blockchain (accessed on 17 November 2020). [194]

Muzzy, E. (2020), *What Is Proof of Stake?*, Consensys Blog, https://consensys.net/blog/blockchain-explained/what-is-proof-of-stake/ (accessed on 24 November 2020). [30]

Nakamoto, S. (2008), *Bitcoin: A Peer-to-Peer Electronic Cash System*, http://www.bitcoin.org (accessed on 24 November 2020). [1]

Nuzzi, L. (2019), *Schnorr Signatures & The Inevitability of Privacy in Bitcoin*, Medium, https://medium.com/digitalassetresearch/schnorr-signatures-the-inevitability-of-privacy-in-bitcoin-b2f45a1f7287 (accessed on 24 November 2020). [18]

OECD (2005), *Guidelines for quality provision in cross-border higher education*, http://www.oecd.org/education/innovation-education/35779480.pdf. [125]

OECD (2004), *Internationalisation and Trade in Higher Education*, OECD, http://dx.doi.org/10.1787/9789264015067-en. [124]

OpenCerts (n.d.), *Collaborate*, https://opencerts.io/collaborate (accessed on 13 November 2020). [77]

OpenCerts (n.d.), *OpenCerts*, https://opencerts.io/ (accessed on 16 November 2020). [118]

Oracle (n.d.), *Oracle Blockchain Enterprise Edition*, https://www.oracle.com/database/technologies/blockchain-platform-enterprise-edition.html (accessed on 13 November 2020). [47]

Parisi, B. (2018), *Woolf University: Will Blockchain Make College Affordable?*, In the Mesh, https://inthemesh.com/archive/woolf-university-will-the-blockchain-make-college-more-affordable/ (accessed on 24 November 2020). [104]

Parrish, J., J. Fryer and R. Parks (2017), *Expanding the Academic Record: Revolutionizing Credentials*, NACE, https://www.naceweb.org/job-market/trends-and-predictions/expanding-the-academic-record-revolutionizing-credentials/ (accessed on 17 November 2020). [215]

Pisani, B. (2018), *Bitcoin and ether are not securities, but some initial coin offerings may be*, SEC official says, CNBC, https://www.cnbc.com/2018/06/14/bitcoin-and-ethereum-are-not-securities-but-some-cryptocurrencies-may-be-sec-official-says.html#close (accessed on 24 November 2020). [24]

Pistis.io (n.d.), *Pistis.io*, https://pistis.io/welcome/ (accessed on 16 November 2020). [150]

Prince Consulting (n.d.), *Prince Consulting*, http://princeconsulting.biz/ (accessed on 17 November 2020). [201]

Purifoy, C. and J. Smith (2018), *G20: Of Currency, New Gold Standards & Rocket Fuel to Coordinate Global Impact*, G20 Leaders Summit Magazine, https://www.learningeconomy.io/post/g20-cover-story-of-currency-new-gold-standards-and-rocket-fuel-to-coordinate-global-impact (accessed on 24 November 2020). [108]

PwC UK (n.d.), *Smart Credentials*, https://www.pwc.co.uk/blockchain/smart-credentials.html (accessed on 16 November 2020). [172]

Quorum (n.d.), *Quorum*, https://www.goquorum.com/ (accessed on 13 November 2020). [41]

Reed, M. (2016), *Co-Curricular Transcripts? Questions about a potentially good idea*, Inside Higher Ed., https://www.insidehighered.com/blogs/confessions-community-college-dean/co-curricular-transcripts (accessed on 17 November 2020). [216]

Renieris, E. (2020), *SSI? What we really need is full data portability*, Women in Identity Blog, https://womeninidentity.org/2020/03/31/data-portability/ (accessed on 24 November 2020). [99]

Salas, R. (2018), *CNM named 'Innovator of the Year' for their use of blockchain tech*, KOB4, https://www.kob.com/albuquerque-news/cnm-named-innovator-of-the-year-for-their-use-of-blockchain-tech-/5189370/ (accessed on 16 November 2020). [143]

Sansone, K. (2019), *Malta is first country to put education certificates on blockchain*, Malta Today, https://www.maltatoday.com.mt/news/national/93148/malta_is_first_country_to_put_education_certificates_on_blockchain#.XG67xpNKg0o (accessed on 16 November 2020). [166]

SAP (n.d.), *Intelligent Technologies*, https://www.sap.com/products/intelligent-technologies/blockchain.html (accessed on 13 November 2020). [48]

Shahada (n.d.), *Shahada*, https://www.shahada.ae/ (accessed on 2 November 2020). [186]

Sharma, R. (2019), *What is the Enterprise Ethereum Alliance?*, Investopedia, https://www.investopedia.com/tech/what-enterprise-ethereum-alliance/ (accessed on 24 November 2020). [39]

Sharma, R. (2019), *Why Is Ethereum Co-founder Proposing a Hard Cap?*, Investopedia, https://www.investopedia.com/news/why-ethereum-cofounder-proposing-hard-cap/ (accessed on 24 November 2020). [59]

Sharma, R. (2019), *Zero Knowledge Proofs' Could Boost Blockchain Adoption on Wall Street*, Investopedia, https://www.investopedia.com/news/zero-knowledge-proofs-could-boost-blockchain-adoption-wall-street/ (accessed on 24 November 2020). [67]

Sharma, R. (2018), *SEC Official Declares Ether Is Not a Security*, Investopedia, https://www.investopedia.com/news/sec-officially-declares-ether-not-security/#:~:text=Ether%20is%20not%20a%20security%2C%20according%20to%20Bill%20Hinman%2C%20U.S.,Commission%20Director%20of%20Corporate%20Finance.&text=After%20Bitcoin%2C%20ether%20is%20th (accessed on 24 November 2020). [23]

Shin, A. (2020), *Hyperledger Fabric vs R3 Corda: A Business Perspective*, Medium, https://medium.com/@kaishinaw/hyperledger-fabric-vs-r3-corda-a-business-perspective-cf935824de9 (accessed on 24 November 2020). [52]

Siegel, D. (2016), *Understanding the DAO Attack*, CoinDesk, https://www.coindesk.com/understanding-dao-hack-journalists (accessed on 24 November 2020). [35]

Sixtin, E. (2017), *Malta Introduces First Blockcerts Pilot in Europe*, BTCManager, https://btcmanager.com/malta-introduces-first-blockcerts-pilot-in-europe/ (accessed on 16 November 2020). [164]

Smart Dubai 2021 (n.d.), *Smart Dubai 2021*, https://2021.smartdubai.ae/ (accessed on 24 November 2020). [181]

SmartDegrees (n.d.), *SmartDegrees*, https://www.smartdegrees.es/en/home-en/ (accessed on 16 November 2020). [167]

Smartworld (2020), *Press Release: Smartworld and Grape Technology announced the launch of Shahada, a digital certificate protocol built on blockchain technology*, https://www.smartworld.com/smartworld-and-grape-technology-announced-the-launch-of-shahada-a-digital-certificate-protocol-built-on-blockchain-technology/ (accessed on 17 November 2020). [185]

Smolenski, N. (ed.) (n.d.), *Interview with Lluís Alfons Ariño Martín, Convenor of the Digital Diplomas Use Case for the EBSI.* [177]

Smolenski, N. (ed.) (2018), *Conversation with a private tutor living and working in China.* [121]

Smolenski, N. (2016), *Identity and Digital Self-Sovereignty*, Medium, https://medium.com/learning-machine-blog/identity-and-digital-self-sovereignty-1f3faab7d9e3 (accessed on 24 November 2020). [98]

Southside Daily Staff (2018), *Here's a new way ECPI University is issuing its degrees*, Wydaily, https://wydaily.com/local-news/2018/11/05/heres-a-new-way-ecpi-university-is-issuing-its-degrees/ (accessed on 16 November 2020). [141]

Steenis, H. (2017), *Bitcoin's bite: why central banks should clamp down on cryptocurrencies*, Schroders, https://www.schroders.com/en/uk/private-investor/insights/markets/bitcoins-bite-why-central-banks-should-clamp-down-on-cryptocurrencies/ (accessed on 24 November 2020). [22]

SYSARTEC (2020), *LinkedIn*, https://www.linkedin.com/feed/update/urn:li:activity:6745799560867700736/ (accessed on 4 May 2021). [193]

T3 Innovation Network (2020), *Applying Self-Sovereign Identity Principles to Interoperable Learning Records: Principles, Challenges, and Community Guidance*, US Chamber of Commerce Foundation, https://www.uschamberfoundation.org/sites/default/files/media-uploads/Applying%20SSI%20Principles%20to%20ILRs%20Report.pdf (accessed on 24 November 2020). [101]

Talent Cloud | Nuage de talents (2018), *Twitter*, https://twitter.com/GC_Talent/status/1066820323490897921 (accessed on 24 November 2020). [157]

Temple, J. (2019), *Bitcoin mining may be pumping out as much CO2 per year as Kansas City.*, MIT Technology Review, https://www.technologyreview.com/2019/06/12/873/bitcoin-mining-may-be-pumping-out-as-much-cosub2-sub-per-year-as-kansas-city/ (accessed on 24 November 2020). [6]

Texas Blockchain Council (2021), *Hyland and Hedera Hashgraph Present Blockchain Proof of Concept for Records Verification to Texas Secretary of State*, https://www.prnewswire.com/news-releases/hyland-and-hedera-hashgraph-present-blockchain-proof-of-concept-for-records-verification-to-texas-secretary-of-state-301204696.html?tc=eml_cleartime. [128]

The Linux Foundation (2020), *Cross-Industry Coalition Advances Digital Trust Standards*, https://www.linuxfoundation.org/press-release/2020/05/cross-industry-coalition-advances-digital-trust-standards/ (accessed on 13 November 2020). [92]

The Open University (n.d.), *Open Blockchain*, https://blockchain.open.ac.uk/ (accessed on 16 November 2020). [173]

The Open University (n.d.), *QualiChain*, https://qualichain-project.eu/ (accessed on 16 November 2020). [174]

The Sovrin Foundation (2018), *The Sovrin Network and Zero Knowledge Proofs*, The Sovrin Foundation, https://sovrin.org/the-sovrin-network-and-zero-knowledge-proofs/ (accessed on 13 November 2020). [68]

The United Arab Emirates' Government portal (n.d.), *Emirates Blockchain Strategy 2021*, https://u.ae/en/about-the-uae/strategies-initiatives-and-awards/federal-governments-strategies-and-plans/emirates-blockchain-strategy-2021 (accessed on 17 November 2020). [180]

Thys, F. (2020), *One Third Of Private 4-Year Colleges Are At High Risk Financially, Model Predicts*, WBUR, https://www.wbur.org/edify/2020/05/08/higher-education-financial-crisis (accessed on 17 November 2020). [212]

Treasury Board of Canada Secretariat (2018), *Canada's trusted digital identity vision*, https://www.canada.ca/en/treasury-board-secretariat/corporate/news/canada-trusted-digital-identity-vision.html (accessed on 16 November 2020). [156]

Trilling, B. and C. Fadel (2009), *21st Century Skills: Learning for Life in Our Times*, John Wiley And Sons, Inc. [213]

Trust Over IP Foundation (n.d.), *Trust Over IP Foundation*, https://trustoverip.org/ (accessed on 13 November 2020). [91]

UAEU (2019), *United Arab Emirates University Launches UAEU Passport, University-Wide Blockchain Applications to Streamline Student Services*, https://www.uaeu.ac.ae/en/news/2019/feb/uaeu_passport.shtml (accessed on 17 November 2020). [182]

Unviersidad Carlos III de Madrid (2018), A*creditaciones de Competencias Utilizando la Tecnología Blockchain en Cursos Spocs*, https://www.uc3m.es/sdic/articulos/2018/acreditaciones-utilizando-blockchain (accessed on 16 November 2020). [117]

US Chamber of Commerce Foundation (n.d.), *The T3 Innovation Network*, https://www.uschamberfoundation.org/t3-innovation (accessed on 13 November 2020). [89]

US Department of Homeland Security (2019), *News Release: DHS Awards $199K for Blockchain Credential Lifecycle Management*, https://www.dhs.gov/science-and-technology/news/2019/11/14/news-release-dhs-awards-199k-blockchain-tech (accessed on 13 November 2020). [74]

US Department of Homeland Security (2019), *News Release: DHS Awards 159K for Infrastructure to Prevent Credential Fraud*, https://www.dhs.gov/science-and-technology/news/2019/11/12/news-release-dhs-awards-159k-prevent-credential-fraud (accessed on 13 November 2020). [72]

US Department of Homeland Security (2017), *News Release: DHS S&T Awards $750K to Virginia Tech Company for Blockchain Identity Management Research and Development*, https://www.dhs.gov/science-and-technology/news/2017/09/25/news-release-dhs-st-awards-750k-virginia-tech-company (accessed on 13 November 2020). [73]

US Department of State, Bureau of Consular Affairs (n.d.), *Apostille Requirements*, https://travel.state.gov/content/travel/en/records-and-authentications/authenticate-your-document/apostille-requirements.html (accessed on 16 November 2020). [127]

Vander Ark, T. (2018), *Imagining a Blockchain University*, Forbes, https://www.forbes.com/sites/tomvanderark/2018/06/13/woolf-building-the-first-blockchain-university/?sh=41e45cff5ae5. [103]

Varshney, N. (2018), *Ethereum's supply has crossed 100M, here's what that means*, TheNextWeb, https://thenextweb.com/hardfork/2018/06/11/ethereums-total-supply/ (accessed on 24 November 2020). [61]

VON (n.d.), *Verifiable Organizations Network: Global Digital Trust for Organizations*, https://vonx.io/ (accessed on 13 November 2020). [75]

Vottun (n.d.), *Vottun*, http://vottun.com (accessed on 16 November 2020). [168]

W3C (n.d.), *Decentralized Identifiers (DIDs) v1.0*, https://w3c.github.io/did-core/ (accessed on 28 June 28). [71]

W3C (n.d.), *Educational and Occupational Credentials in Schema.org Community Group*, https://www.w3.org/community/eocred-schema/ (accessed on 28 June 2020). [82]

W3C (n.d.), *Facts about W3C*, https://www.w3.org/Consortium/facts.html (accessed on 13 November 2020). [69]

W3C (n.d.), *Verifiable Credentials Data Model 1.0*, https://www.w3.org/TR/vc-data-model/ (accessed on 24 November 2020). [70]

W3C (n.d.), *Verifiable Credentials for Education Task Force*, https://w3c-ccg.github.io/vc-ed/ (accessed on 13 November 2020). [81]

Watkins, R. (2020), *Power and Seigniorage in Proof of Stake*, Messari.io, https://messari.io/article/power-and-seigniorage-in-proof-of-stake (accessed on 24 November 2020). [31]

Web of Trust (n.d.), *Rebooting Web of Trust*, https://www.weboftrust.info/ (accessed on 13 November 2020). [87]

Weinstein, J. (2015), *How can law enforcement leverage the blockchain in investigations?*, Coin Center, https://coincenter.org/entry/how-can-law-enforcement-leverage-the-blockchain-in-investigations (accessed on 24 November 2020). [64]

Wikipedia (n.d.), *Byzantine fault*, https://en.wikipedia.org/wiki/Byzantine_fault (accessed on 25 June 2020). [3]

Wikipedia (n.d.), *Decentralized applications*, https://en.wikipedia.org/wiki/Decentralized_application#:~:text=A%20decentralized%20application%20(DApp%2C%20dApp,referred%20to%20as%20smart%20contracts (accessed on 27 July 2020). [26]

Wikipedia (n.d.), *Hyperledger*, https://en.wikipedia.org/wiki/Hyperledger (accessed on 28 June 2020). [42]

Wikipedia (n.d.), *Mozilla Open Badges*, https://en.wikipedia.org/wiki/Mozilla_Open_Badges (accessed on 26 June 2020). [112]

Wikipedia (n.d.), *Proof of work*, https://en.wikipedia.org/wiki/Proof_of_work (accessed on 25 June 2020). [5]

Wikipedia (n.d.), *R3 (company)*, https://en.wikipedia.org/wiki/R3_(company) (accessed on 28 June 2020). [51]

Woolf University (n.d.), *Woolf University*, https://woolf.university/ (accessed on 26 June 2020). [105]

Workday Credentials (n.d.), *Workday Credentials & WayTo™ By Workday*, https://credentials.workday.com/docs/overview/ (accessed on 16 November 2020). [147]

World Government Summit et al. (2018), *Case Study: Free Agents and GC Talent Cloud – Canada*, *Embracing Innovation in Government*, https://www.oecd.org/gov/innovative-government/Canada-case-study-UAE-report-2018.pdf (accessed on 24 November 2020). [160]

Xertify (n.d.), *Xertify*, https://xertify.co/ (accessed on 17 November 2020). [196]

Young, J. (2018), *Academics Propose a 'Blockchain University,' Where Faculty (and Algorithms) Rule*, EdSurge, https://www.edsurge.com/news/2018-10-25-academics-propose-a-blockchain-university-where-faculty-and-algorithms-rule. [102]

Zawya (2020), *Press Release: University of Dubai successfully publish e-Credentials of graduates on blockchain platform*, https://www.zawya.com/mena/en/press-releases/story/University_of_Dubai_successfully_publish_eCredentials_of_graduates_on_blockchain_platform-ZAWYA20200901071849/ (accessed on 17 November 2020). [187]

Authors

Ryan Baker

Ryan Baker is Associate Professor at the University of Pennsylvania, and Director of the Penn Center for Learning Analytics. His lab conducts research on engagement and robust learning within online and blended learning, seeking to find actionable indicators that can be used today but which predict future student outcomes. Baker has developed models that can automatically detect student engagement in over a dozen online learning environments, and has led the development of an observational protocol and app for field observation of student engagement that has been used by over 160 researchers in 7 countries. Predictive analytics models he helped develop have been used to benefit over a million students, over a hundred thousand people have taken MOOCs he ran, and he has coordinated longitudinal studies that spanned over a decade. He was the founding president of the International Educational Data Mining Society, is currently serving as Editor of the journal Computer-Based Learning in Context, is Associate Editor of two journals, was the first technical director of the Pittsburgh Science of Learning Center DataShop, and currently serves as Co-Director of the MOOC Replication Framework (MORF). Baker has co-authored published papers with over 300 colleagues.

Tony Belpaeme

Tony Belpaeme is Professor at Ghent University and Professor in Robotics and Cognitive Systems at the University of Plymouth, UK. He received his PhD in Computer Science from the Vrije Universiteit Brussel (VUB) and currently leads a team studying cognitive robotics and human-robot interaction. He coordinated the H2020 L2TOR project, studying how robots can be used to support children with learning a second language, and coordinated the FP7 ALIZ-E project, which studied long-term human-robot interaction and its use in paediatric applications. He worked on the FP7 DREAM project, studying how robot therapy for Autism Spectrum Disorder. Starting from the premise that intelligence is rooted in social interaction, Belpaeme and his research team try to further the science and technology behind artificial intelligence and social human-robot interaction. This results in a spectrum of results, from theoretical insights to practical applications.

Alex J. Bowers

Alex J. Bowers is an Associate Professor of Education Leadership at Teachers College, Columbia University, where he works to help school leaders use the data that they already collect in schools in more effective ways to help direct the limited resources of schools and districts to specific student needs. His research focuses on the intersection of effective school and district leadership, organization and HR, data driven decision making, student grades and test scores, student persistence and dropouts. His work also considers the influence of school finance, facilities, and technology on student achievement. Dr. Bowers studies these domains through the application of data science, and big data analytics, such as data visualization analytics, multilevel and growth mixture modeling, and cluster analysis heatmap data dashboards. He earned his Ph.D. in K12 Educational Administration from Michigan State University, and previous to teaching and education research, spent a decade as a cancer researcher in the biotechnology industry, with a M.S. in Biochemistry, Microbiology and Molecular Biology, and a B.S. in Biochemistry.

Jack Buckley

Jack Buckley is Head of Assessment and Learning Sciences at Roblox, where our mission is to bring the world together through play. Roblox enables anyone to imagine, create, and have fun with friends as they explore millions of immersive 3D experiences, all built by a global community of developers. Prior to this role, he was president and chief scientist at Imbellus, a game-based assessment technology startup. He was previously senior vice president at the American Institutes for Research (AIR), where he led their research and evaluation area, and he still serves as Institute Fellow on several projects. Before joining AIR, he helped lead the redesign of the SAT at the College Board, where he served as senior vice president of research. Before that, he served as Commissioner of the US Department of Education's National Center for Education Statistics (NCES), where he was responsible for the measurement of all aspects of US education, including conducting the National Assessment of Educational Progress and coordinating participation in international assessments. While at NCES, he also acted as a senior technical adviser to Department of Education leadership and co-chair of its data strategy team. In addition, Jack has been a tenured associate professor of applied statistics at New York University and an an assistant professor of educational measurement at Boston College, an analytic methodologist in the intelligence community, and a US Navy surface warfare officer and nuclear reactor engineer.

Laura Colosimo

Pierre Dillenbourg

A former teacher in elementary school, Pierre Dillenbourg graduated in educational science (University of Mons, Belgium). He started his research on learning technologies in 1984. In 1986, he has been on of the first in the world to apply machine learning to develop a self-improving teaching system. He obtained a PhD in computer science from the University of Lancaster (UK), in the domain of artificial intelligence applications for education. He has been assistant professor at the University of Geneva. He joined EPFL in 2002. He has been the director of Center for Research and Support on Learning and its Technologies, then academic director of Center for Digital Education, which implements the MOOC strategy of EPFL (over 2 million registrations). He is full professor in learning technologies in the School of Computer & Communication Sciences, where he is the head of the CHILI Lab: «Computer-Human Interaction for Learning & Instruction ». He is the director of the leading house DUAL-T, which develops technologies for dual vocational education systems (carpenters, florists,...). With EPFL colleagues, he launched in 2017 the Swiss EdTech Collider, an incubator with 80 start-ups in learning technologies. He (co-)founded 4 start-ups, does consulting missions in the corporate world and joined the board of several companies or institutions. In 2018, he co-founded LEARN, the EPFL Center of Learning Sciences that brings together the local initiatives in educational innovation. He is a fellow of the International Society for Learning Sciences. He currently is the Associate Vice-President for Education at EPFL.

Sidney D'Mello

Sidney D'Mello (PhD in Computer Science) is an Associate Professor in the Institute of Cognitive Science and Department of Computer Science at the University of Colorado Boulder. He is interested in the dynamic interplay between cognition and emotion while individuals and groups engage in complex real-world tasks. He applies insights gleaned from this basic research program to develop intelligent technologies that help people achieve to their fullest potential by coordinating what they think and feel with what they know and do. D'Mello has co-edited seven books and published almost 300 journal papers, book chapters, and conference proceedings. His work has been funded by numerous grants and he currently serves(d) as associate editor for Discourse Processes and PloS ONE. D'Mello is the Principal Investigator for the NSF National Institute for Student-Agent Teaming.

Judith Good

Judith Good is a Professor of Human Factors and the Internet of Things at the University of Amsterdam. She wrote her chapter while Professor of Interaction Design and Inclusion at the University of Sussex. She has a background in psychology and artificial intelligence, and has worked at universities in both Europe and the United States. Her broad research focus has been on better understanding how people learn, and on how the development of innovative technologies can better support their learning. Within this area, she is particularly interested in looking at how best to support people with disabilities, particularly autism. In so doing, she aims to develop effective methods for involving the end users of such technologies in their design from the outset.

Dirk Ifenthaler

Dirk Ifenthaler is Professor and Chair of Learning, Design and Technology at University of Mannheim, Germany and UNESCO Deputy Chair of Data Science in Higher Education Learning and Teaching at Curtin University, Australia. Dirk's research focuses on the intersection of cognitive psychology, educational technology, data analytics, and organisational learning. He is the Editor-in-Chief of the Technology, Knowledge and Learning, Senior Editor of Journal of Applied Research in Higher Education, and Deputy Editor-in-Chief of International Journal of Learning Analytics and Artificial Intelligence for Education.

Contact: dirk@ifenthaler.info

Rebecca Kantar

Rebecca serves as the Entrepreneur in Residence at Roblox Corporation, a technology platform bringing the world together through play. Roblox acquired, Imbellus, a simulation-based assessment technology company that Rebecca founded and led. Rebecca started Imbellus with conviction that better educational assessments would drive better curriculum and instruction. Imbellus developed high-stakes tests that measured deep thinking skills like problem solving, systems thinking, and decision-making. Rebecca led Imbellus in raising over $23M, signing multimillion-dollar development contracts with Fortune 500 corporations and delivering operational, fair, reliable, and valid assessments. Prior to launching Imbellus, Ms. Kantar founded an expert network that Gerson Lehrman Group (GLG) acquired in 2012. Rebecca attended public school in Massachusetts and then Harvard College and currently lives in Massachusetts.

Marty McCall

Marty McCall is an Expert for the Game-based Innovation Laboratory at McKinsey & Company. The Laboratory combines scientific analysis with modern game design to create immersive digital experiences that engage users and provide meaningful insights for clients.

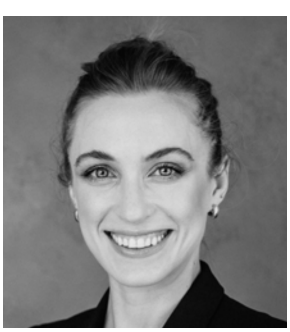
Before working at McKinsey, she was the Lead Psychometrician at Imbellus, a start-up producing game-based software. Prior to that she was the Director of Psychometrics for the Smarter Balanced Assessment Consortium, which produced assessments for public schools in a coalition of states. She was director of operational assessments at Northwest Evaluation Association, an organization that produces assessments for measuring academic growth throughout the school year. Her work at NWEA was preceded by work as a psychometrician for the states of Oregon and Washington.

Inge Molenaar

Inge Molenaar is associated professor Educational Sciences at the Behavioural Science Institute at Radboud University in the Netherlands. She has over 20 years of experience in the field of technology enhanced learning taking multiple roles from entrepreneur to academic.

Her research in the Adaptive Learning Lab focusses on technology empowered innovation to optimize students' learning. The application of data, learning analytics and Artificial Intelligence to understand how learning unfolds over time is central in her work. Artificial Intelligence offers a powerful way to make new steps towards measuring, understanding and designing innovative learning scenarios. Dr Molenaar envisions Hybrid Human-Systems that augment human intelligence with artificial intelligence to empower learners and teachers in their quest to make education more efficient, effective and responsive. In this endeavor collaboration between governments, schools, research and industry is essential to develop the next generation educational systems. Dr Molenaar has just received an ERC Starting Grant to develop the first Hybrid Human-AI Regulation system to train young learners' Self-regulated learning skills with the help of AI and she also recently became a Jacobs Foundation Fellow.

Dr Molenaar holds Master's degrees in Cognitive Psychology and International Business studies (Maastricht University) and a PhD in Educational Sciences (University of Amsterdam).



Natalie Smolenski

Head of Business Development, Hyland Credentials

Natalie Smolenski leads business development for Hyland Credentials, a solution for verifiable digital credentials anchored in blockchain technology. Hyland Credentials was previously Learning Machine, a software firm she helped found, grow, and successfully exit.

As an author and public speaker, Natalie focuses on the intersections of identity, technology, and government. Her doctoral work focused on the economic dimensions of human development and mental health. By bringing a scientific perspective to distributed digital technologies and social transformation, she helps audiences from all backgrounds understand how individuals connect to form communities and build the infrastructures of the future.

You can find a list of Natalie's publications and public speaking at
www.nataliesmolenski.com.

Her latest project, VALUED., can be found at valued.nataliesmolenski.com.

Erica Snow

Erica Snow is the Director of Learning and Data Science at Roblox, where our mission is to bring the world together through play. Roblox enables anyone to imagine, create, and have fun with friends as they explore millions of immersive 3D experiences, all built by a global community of developers.

Prior to this role, she was the Director of Learning and Data science at Imbellus, a game-based assessment technology startup. She was previously the Learning Analytics Lead Scientist at SRI international, where she led work focused on the evaluation and implementation of educational technologies within the classroom. Erica has over 60 peer-reviewed publications in the fields of Data Science, Cognitive Science, Educational Technology and Learning Science. In addition Erica is an adjunct professor at American University where she teaches predictive analytics for business.

Fumihide Tanaka

Fumihide Tanaka is an associate professor in the Faculty of Engineering, Information and Systems of the University of Tsukuba, Japan. After he obtained a Ph.D. from the Tokyo Institute of Technology in 2003, he joined Sony Corporation and worked for the research and development of entertainment robots. He then started the research of human-robot interaction (HRI) during the time he stayed in the University of California in San Diego between 2004 and 2007. A long-term field study conducted in a nursery school in that period has been regarded as a classic study of child-robot interaction, having been reported in major news media such as CNN, Nature, and Science. He moved to academia in 2008, and since then he has been working for robotics for education in the University of Tokyo and the University of Tsukuba. He supervised the development of an educational application for Pepper robot during the time he worked for SoftBank Corp. as a consultant. Currently he serves as a director for foreign affairs in the Robotics Society of Japan.

Stéphan Vincent-Lancrin

Stéphan Vincent-Lancrin is a Senior Analyst and Deputy Head of Division at the Organisation for Economic Co-operation and Development (Directorate for Education and Skills). He currently leads work on education during the covid-19 crisis, but also leads the OECD work on the digitalization in education, notably the project on "Smart data and digital technology in education: AI, learning analytics and beyond" -- including a blockchain component. He also leads work around disciplined innovation and change management, showing with the work on "Fostering and Assessing Creativity and Critical Thinking in Education" what kind of support, environment and tools school teachers and university professors could be given to improve their teaching and their students' learning. An example of capacity development through international professional learning communities. More generally, speaking he work on innovation, research and how new trends influence the futures of learning and education policy at the schooling and higher education levels.

OECD Digital Education Outlook 2021
PUSHING THE FRONTIERS WITH AI, BLOCKCHAIN, AND ROBOTS

How might digital technology and notably smart technologies based on artificial intelligence (AI), learning analytics, robotics, and others transform education? This book explores such question. It focuses on how smart technologies currently change education in the classroom and the management of educational organisations and systems. The book delves into beneficial uses of smart technologies such as learning personalisation, supporting students with special learning needs, and blockchain diploma credentialing. It also considers challenges and areas for further research. The findings offer pathways for teachers, policy makers, and educational institutions to digitalise education while optimising equity and inclusivity.

PAPERBACK ISBN 978-92-64-64199-0
PDF ISBN 978-92-64-90464-4

9 789264 641990