

Improving learning

Interest in learning has grown well beyond class-based interactions between teachers and students. The value of higher learning keeps growing, spurring the need for education innovation and reform (Coates, 2017). Global population growth keeps driving an increasing need for higher education to reach more people than ever before.

Economic development is shifting more people in more countries into professional roles. Longer lifespans are leading people to work for longer, requiring re-skilling, re-directing, and re-certification. The global race for top-talent has intensified, creating a need to understand not just learning at scale but also in extremely novel contexts. To understand and advance the growth of learning, an increasing suite of research and development is advancing economic models, platforms, professions and networks to shape the future (Coates, 2020). Such work confers important impetus for this paper, as it establishes the broader value of learning and major foundations and rationales for assessment reform work.

Assessment plays a huge role in learning. Done well, assessment forms a core part in articulating what learners already know, in helping people learn, and in spotlighting what learners need to learn. Done poorly or without reflection, assessment can waste time and money, spur anxiety and distaste for learning, provide misleading information and generate adverse outcomes. Assuming most education practice lies between these extremes, the goal of assessment reform is to shift education practice, therefore learning, towards the positive end of the spectrum. Reformed assessment gives people the feedback and encouragement they need to learn in smarter ways.

Here lies a problem, for despite the crucial role it plays in higher education, assessment has yet to have its transformational moment. For instance, online learning

and workforce changes have revolutionised curriculum and teaching, while major platforms and business transformations have disrupted and reconfigured admissions and broader forms of student management. Yet much assessment is still being done today as it was a century ago (Coates, 2014, 2020) and there remain pressing reasons to reform. Of course, reform rationales have only proliferated and diversified since the early 2020 disruptions to international and campus-based education. Coates, Xie and Hong (2020) have detailed the extent of accelerated reconfiguration of education fundamentals.

This paper introduces the 'smarter learning' initiative, spotlights reforms required to achieve underpinning 'next-generation' forms of assessment and articulates feasible steps ahead. The argument driving this paper is that learning is becoming more valuable, that improving learning hinges on assessment reform, and that such innovation will yield major productivity advances for higher education and broader communities.

This paper provides education leaders with an overview of the field, and a guide to thinking about current practices as well as institution and policy innovations. It unpacks these ideas and sketches one perspective for innovation. It discusses the characteristics of smarter forms of learning and argues that smarter learning hinges on particular kinds of assessment reform, specifically deft integration of a constellation of technological, practical and educational factors. It presents the results of a multi-university feasibility study and concludes by articulating what appear to be the best steps ahead.

Smarter Learning

'Smarter learning' is an initiative being developed for contemporary higher education. It is grounded in the normative perspective that 'learning is smarter' when assessment is reformed in ways that enhance integrity and productivity. Such reform rests on deft education

design, robust platforms and careful reconfiguration of management and business processes. Smarter learning blends world-leading assessment capability and learning software with education expertise to advance learning experiences and outcomes. It is better for learners, teachers, institutions and stakeholders.

What, then, does it mean to enhance the integrity of assessment? The robustness of assessment can be aligned with a range of technical criteria which are far from specific to higher education. Gleaned from a century of measurement science research, such criteria ultimately go to spotlighting and emphasising different forms of validity (Coates & Richardson, 2012; Richardson & Coates, 2014; AERA, APA & NCME, 2015). These very

general principles have been clarified into standards which are far more relevant to higher education.

For instance, Table 1 lists indicative standards which are broadly relevant to all assessment tasks. Such standards are necessarily normative ideals and it would be difficult, if not impossible, for any task to fully meet these expectations. It is important, and distinctive, that critical consideration be given to each standard in task design, development, implementation and review. The quest is not that a task be flawless, but that it has known and optimised properties.

What does it mean to make assessment more productive, to make quality-informed advances in efficiency?

Table 1: Indicative assessment standards with prompt questions

Standards	Prompt questions
Coverage	Does the task cover sufficient range and depth of content and all relevant material?
Authenticity	Does the task seem relevant and real? Does it appear useful and meaningful?
Criterion	Does the task correlate with other indicators of similar topics?
Discrimination	Does the task distinguish varying performance levels?
Practicality	Is it easy for students to engage with the task? Is the task 'user friendly'?
Efficiency	Is the task efficient for staff to implement and use standard equipment and procedures?
Responsiveness	Does the task yield timely feedback for students? Does it support lively learning?
Interpretability	Are task requirements understood by all students? Is task language easy to read?
Transparency	Are task requirements and expectations clear to students?
Educational	Does the task prompt students to learn and contribute seamlessly to the experience?
Consequential	Does feedback have expected consequences and promote improvement?
Production	Are task materials produced to a high standard? Have they been designed and proofed?
Clearance	Have relevant legal and cultural approvals have been secured for the task?
Consistency	Does the task perform consistently across people, time and contexts?
Alignment	Does the task align with students, curriculum, teaching and outcomes?
Distinctiveness	Is the task sufficiently distinctive and does it add unique value and insights?
Scoring	Do rubrics enable sound and generalisable scoring?
Validation	Have task materials been validated and improved by students?

Source: Coates (2020)

Enhancing the productivity of assessment goes in general terms to increasing the ratio of outcomes to inputs by delivering the same or more outcomes for the same or fewer inputs, while holding quality constant. Outcomes, in this respect, include the number and variety of assessments. Inputs include staff and student time as well as direct and indirect costs. With assessment, boosting productivity seems most likely to be achieved through input reduction rather than a greater number of assessments. This necessitates a different production function.

Such change is required given the growth of higher education alone. For instance, recent conservative reckoning (Coates, 2018) found that the number of pieces of assessment each year in Australian higher education had risen from around 7 million to nearly 30 million in the last twenty years. With no change in production function this implies a quadrupling in recurrent annual costs to more than \$400 million for marking alone, leaving aside the other large costs for assessment development, administrative and support staff, capital and intermediaries. All up, including indirect costs, it is easy to see how largely unreformed assessment practices might be costing close to a billion dollars annually. Given that Australia has less than half a percent of the world's higher education students, these figures balloon into a much broader need for assessment productivity reform.

What is required to spur such progress? Analysis of recent advances in higher education, spanning many operational levels and functions and crossing many countries, indicates that substantial gain has been derived from careful integration of educational, technological and practical resources (Coates, 2020). Education expertise, of all substantive and technical varieties, can furnish theories and ideas but not the resources or infrastructure for change. Despite compelling rhetoric, technology alone has proved insufficient to advance education, though it serves well in facilitating and catalysing roles. Practical efforts, including financial, operational or political rationales, have a track record in higher education of sparking inadequate, unstable and precarious change.

The success of many productive contemporary reforms is that they put education first and position technology and practical matters as enablers. Figure 1 depicts this 'value-creating constellation' in which practical nous is the glue that binds these ingredients in ways which yield additional quality and productivity. As Wong (2019) and Liu, Wong and Coates (2019) have clarified through

case study and conceptual analysis, contemporary academic reforms have flourished not because they involve technology, but because they engage technology in practical solutions which advance the productivity of higher learning.

Figure 1: Value-creating constellation

It is important to clarify that 'smarter learning' goes beyond the typical shift from instructivist to constructivist forms of education, or from teacher- to

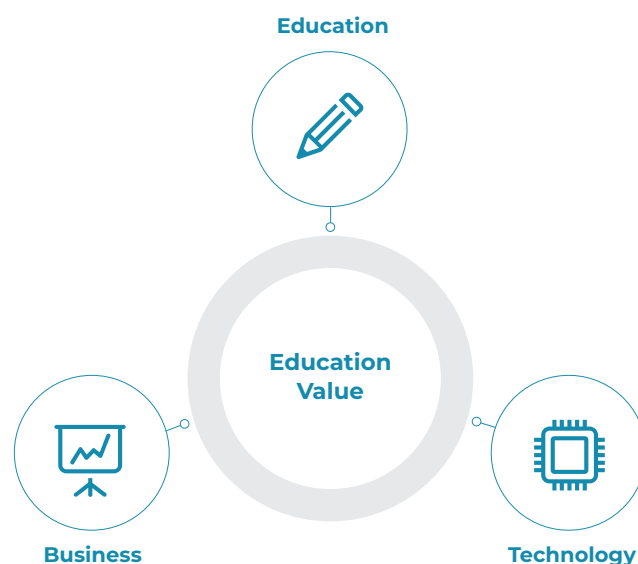


Figure 1: Value-creating constellation

student-centred forms of exchange. The big difference is that, as the initial definition conveys, 'smarter learning' is about the quality and productivity of learning, not activities and conditions relating to people and institutions. This is important, for it immediately shifts focus onto the value, outcomes and impact of learning, and it immediately dislocates learning from particular people and places. As Coates (2020) puts forward, contemporary shocks and transformations to higher education have clarified the necessity and significance of such shifts.

Assessment reform

Assuming such progress is merited, then how should it be directed? To begin, it's helpful to position the progress of assessment reform to date. This highlights the value of parameterising assessment in ways which can be used to probe weaknesses and, more constructively, to explore and synthesise options for development.

The evolution of assessment reform can be framed in three eras. Coates (2014, 2018) unpacks the historical, educational and technical analyses which underpin this articulation. Table 2 highlights shifting practices in the transition from traditional, through stretched, to next-generation practices. Key characteristics of traditional approaches to assessment, which are highly individual and collegial in nature, are shown towards the left. These traditional practices have been stretched in recent decades as higher education has expanded. Much of this growth has been expansionary rather than transformative in nature. Legacy practices have been

stretched and patched for bigger delivery as higher education has expanded. Technical, institutional and educational analysis reveals, however, that traditional assessment is costly when scaled, and quality may suffer. Shifting to next-generation assessment, by definition and design, represents the kind of reformed assessment which carries potential to under-gird 'smarter' forms of learning. The depiction in Table 2 is of course an oversimplification, but it does help to highlight general characteristics, developments and trajectories.

Table 2: Three eras of assessment

	Traditional	Stretched	Next-generation
Timeframe	1990s and before...	1990s to 2020	...2020s and after
Authority	University	University or regulator	Shared
Production	Solo academics	Academic teams	Co-creation
Format	Paper	Paper and online	Online
Location	Campus	Campus and online	Online
Implementation	Universities	Universities	Engineers
Scoring	Solo academics	Moderated practice	Automated
Reporting	Generic	Contextualised	Customised

Source: Coates (2018)

How then can assessment be understood in ways that illuminate weaknesses and options for development? Adopting a value chain perspective and deconstructing assessment into key phases and activities is helpful. Table 3 models this, presenting an evaluative model which draws from measurement science research (e.g. Mislevy, Almond & Lukas, 2003; Bennett, 2015; Richardson & Coates, 2016). Borrowing these mechanisms situates higher education in the broader world and helps garner the most scientifically relevant insights. The basic model provides a compelling lens for reviewing existing bottlenecks and options for reform.

Each specific assessment context does of course have its own situation, hence evaluative particularities. For current purposes, the framework is helpful to highlight the overarching innovation need and opportunity. Meta-analysis of thousands of contexts (Coates, 2014, 2016, 2018, 2020) has affirmed that such need and opportunity is indeed 'across-the-board'. Circumstances vary, but from a broad perspective every assessment phase and activity could benefit from 'next-generation' reform.

Table 3: Assessment phases and activities

Phase	Indicative activities
Planning	Governance, leadership, management
Development	Mapping resources, specifying outcomes, selecting formats, drafting materials, qualitative review, quantitative review, material production
Implementation	Designing administration, organising facilities, managing students, administering assessments, resolving problems
Analysis	Collating results, marking, producing data, cross-validating results
Reporting	Producing grades, analysing and commenting, reporting, reviewing and improving



It is important to emphasise that next-generation assessment involves enhancing the quality and productivity of assessment. It enables, enriches and augments traditional collegial practices. It does not impose top-down reform, which rarely resonates with everyday academic practice. Conceptually, it occupies or builds out the middle space by strengthening academic practice. This differentiates it strategically from existing or prior initiatives such as OECD AHELO (Coates & Richardson, 2012), the European Commission's CALOHEE (2021), the Australian Learning and Teaching Council's AMAC (Edwards, Wilkinson, Canny, Pearce & Coates, 2014) or a plethora of commercial assessment initiatives.

From an academic, institutional and technological perspective, this broad need can be distilled into four main modules. These are presented here, and go to making tasks, managing administrations, implementing and proctoring, and marking and reporting. Each module is already being deployed at scale using the Examina+ software (Genix Ventures, 2021), as the subsequent case study conveys.

From a functional perspective, making assessment tasks involves creating, validating, aligning and optimising online or paper-based tasks. Academics have been creating assessment tasks forever, though in traditional ways and not necessarily in ways which optimise educational, practical or technological resources. Shifting away from individual or ad hoc task development and adopting task authoring tools boosts the integrity, efficiency and quality of assessment. Such tools enable the adoption of richer and dynamic tasks, collaborative task development, task sharing, and more sophisticated mapping to education and vocation competencies. Deploying such tools within courses or across institutions helps to create and edit tasks, integrate scattered resources, tag and map tasks, align assessment with learning outcomes, proof materials, and generate hybrid assignments and exams.

Managing assessment administrations is about scheduling people, infrastructure and resources. While all higher education institutions have existing arrangements for handling this, these are typically configured in ways which suit traditional rather than next-generation forms of assessment. For instance,

they may process programs, courses or people in batches to undertake static tasks, whereas much more dynamic and nuanced forms of sampling can be deployed to enable more astute and authentic matching of students to tasks. Dedicated scheduling software can also help institutions and learners reduce assessment risks and costs by managing schedules, rostering staff and absences, coordinating paper delivery and third-party logistics, logging and investigating incidents, and handling special needs and situations.

Next-generation assessment entails implementation reform, and to the extent required, changes to proctoring arrangements. Particularly when linked with more intentional assessment design, step-change improvements in delivery and security become feasible. In general, this involves deploying technologies which assure quality implementation. Integral tools here involve planning assessments, registering and verifying students, delivering automated and coordinated communications, allocating and aligning people and venues, delivering practice, scheduled and on-demand assessments, enabling real-time monitoring and authentication, and reporting performance metrics. The adoption of contemporary implementation platforms helps shift beyond batch processing of students and assessments, enabling much finer-grained and even task/competence-level alignment of students with assessments.

The final cluster of assessment phases and activities to be distilled into a software module does to marking and reporting. Typically, these activities are handled by academics working alone or in small teams, with reporting taking place as a somewhat private activity between teachers and students. Confidentialities around reporting must be protected, of course, but there are quality and efficiency limitations with keeping all aspects of this process secret. As most large-scale assessments signpost, huge quality dividends can be derived from collaborative marking and from sufficiently anonymised benchmarking. Well-managed collaboration can increase the reliability and efficiency of marking and reduce reporting errors and delay. Platforms have ample means of supporting secure marking, automating marking, managing markers, dealing with response validation, and generating benchmark and competence-level reports.

Using contemporary platforms to activate assessment either in modular ways or through end-to-end solutions is giving real-life to next-generation assessment,

bringing about the qualities flagged in Table 3. Without impacting important governance, leadership or management arrangements, the deployment of such software enables various forms of collaboration and co-creation, activation of dynamic and rich task formats, hybrid paper and online implementation, nuanced matching of tasks to learners, and quality assured and customised reporting. Leaving aside obstacles arising from entrenched institutional and faculty practices, advancing such reform seems like an obvious way ahead.

Feasibility study

A feasibility study was conducted in 2020 to evaluate the quality and productivity of the next-generation assessment which fuels smarter learning. This work did affirm feasibility, and it offers useful case study insights into the growth of this field.

In June 2020, dozens of faculty and over 1,000 students from five Chinese university business schools participated in a pilot of next-generation assessments of macro- and micro-economics, accounting and statistics. This assessment project was led by a small research team and driven by the Examina+ platform (Genix Ventures, 2020).



The assessment design was extensive. The project focused on engaging universities, developing frameworks, mapping assessment tasks, assessment implementation, and reporting. The project was co-designed operationally with participating institutions.

A sample of university business schools was engaged, deliberately diverse in terms of location, context, and concentration. Each institutional team consisted of a senior leader, key faculty in each field to be tested, administrative staff, and students. Collaborative communication groups were formed for each field and for all universities, and several meetings were held. Establishing open communication channels which help faculty share non-confidential insights on assessment is a major value of next-generation assessment.

As these five universities bootstrapped the collaboration, the research team drafted learning outcomes frameworks. Initial frameworks took account of existing frameworks, curriculum materials, and assessment materials. These initial frameworks were modified in an iterative fashion by the research team and university teachers and specified broad and narrow curriculum areas and learning outcomes, as well as the likely difficulty of each task. After the frameworks had stabilised, they were configured into the assessment system. Framework development took two months.

Participating universities provided a large array of assessment tasks, which were then used to build the library of shared assessments. The tasks focused on the introductory level, targeted accounting, statistics or economics, and had a multiple-choice or short-text response format. The research team tagged all tasks with meta-data, edited and formatted them, then loaded them onto the assessment system. The tasks were mapped to the frameworks. The university experts reviewed the shared task library, and selected item samples which took account of focus, difficulty, format and quality.



Assessment administration was coordinated by participating universities using standard protocols and supported by researchers in China and internationally. Specifications for each assessment (e.g. student numbers, time, locations, security, staffing) were finalised with each university. The assessment platform provided end-to-end support for each assessment in terms of schedule optimisation, registration and enrolment, resource management, secure assessment delivery, and automated marking and analysis. Constructed-response tasks were distributed across universities for marking and cross-validation.

Reports were prepared for each university and data was analysed psychometrically. Marker and task performance were analysed, proving novel information to universities. University reports contained diagnostic and benchmarking insights, and basic descriptive statistical information. The average scores provided a benchmark for comparing performance of tasks, students and scorers.

Overall, the universities identified specific benefits through participating in the feasibility study, including that next-generation assessment:

- Helped teachers improve courses and enhance student learning
- Assured the quality of learning and academic programs for leaders
- Enhanced the efficiency of assessment, saving time, resources and costs
- Improved the quality and value of assessment resources and approaches
- Provided independent third-party assessment data to continuously improve programs
- Provided scale necessary to support micro-credentials
- Supported business accreditation requirements
- Produced reports which enables universities to benchmark against peers and standards
- Built innovative communities around platforms, methods and research; and
- Engaged faculty in teaching professional development.

Overall, the technically designed, technologically supported and collaborative approach strengthened and streamlined university assessment. The universities further saw that foundations had been laid which would mature in following replications.

Steps ahead

If such aspirations and advance have merit, what changes are required to spur these forms of assessment reform and development of smarter learning? Three stagewise developments seem required to trigger the overarching reform, which could play out in different ways in different contexts.

Figure 2: Stagewise developments to guide reform



Figure 2: Stagewise developments to guide reform

First, there is a need to evaluate, diagnose and re-design prevailing assessment arrangements. To propel this an 'academic risk ranking' has been developed, which invites faculty, academic leaders or whole institutions to evaluate and compare existing assessment arrangements. This furnishes assessors of all types with insight into the health or strength of current arrangements and how they compare with other operations. Specifically, users can respond to a selection of quick questions which together yield insight into salient assessment activities and conditions, and a report is produced of existing assessment arrangements. This carries intrinsic value, not least as it is a first of its kind, but also serves as a guide for reform. Thinking differently about assessment and giving consideration to methodological, technological and practical factors,

sparks consideration of the redesign that underpins next-generation assessment.

Second, there is a need for teachers, institutions and students to engage with sufficiently sophisticated assessment platform. This of course involves all the issues associated with software adoption, as well as being made more complex given academic and workforce sensitivities around assessment and entrenched institutional practices. A fertile way forward, beyond system-level deployment, is to enable faculty to experiment with modular software and literally 'play

around' with the systems and experience the potential. Figure 3 presents facets of the Examina+ platform (Genix Ventures, 2021), which together give life to these modules, and the interplay between these facets.

Figure 3: The Examina+ modular platform

Third, the activation of localised modular assessment software is a precondition for realising perhaps the broadest value of assessment reform, being the formation of academic collaboratives. The modular nature of the system means that such individual experimentation can be joined-up into a synchronised network. These, in many respects, mirror the collaborative networks which sustain and advance academic research.

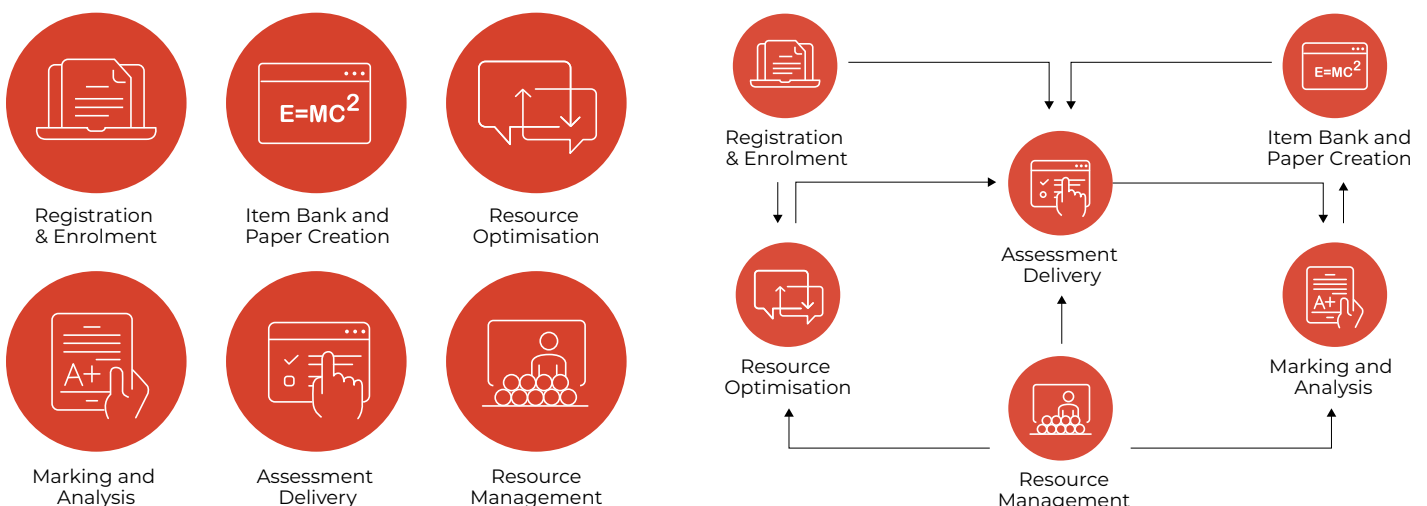


Figure 3: The Examina+ modular platform

Figure 4 gives life to this concept, illustrating various kinds of collaboration. Such collaborations could be woven together across institutional and disciplinary boundaries.

Figure 4: Smarter Learning collaboratives

As this paper has conveyed, various forms of structured collaboration are embedded in the next-generation assessment which underpins smarter learning. Next-generation assessment has been designed over two decades with hundreds of experts, universities, governments, teachers and students (Coates, 2012, 2014, 2016, 2018, 2020). It involves collaboration among

been tested in hundreds of higher education contexts and this work has affirmed the value of technology-based assessment reform in terms of improvements to quality and productivity. Work is well underway to take them to scale. Rough estimation conveys that about nine billion pieces of assessment are conducted in higher education each year, with about one billion being done online. Given that not all online assessment will be optimised, there seems ample scope for growth.

What, then, could help take these reforms to scale? As an example, it is helpful to sketch one option. Structured collaboration (Figure 4), managed by a collaborative consisting of academics from various institutions and experts from industries, is an interesting model to explore further. In this model, a group of participating

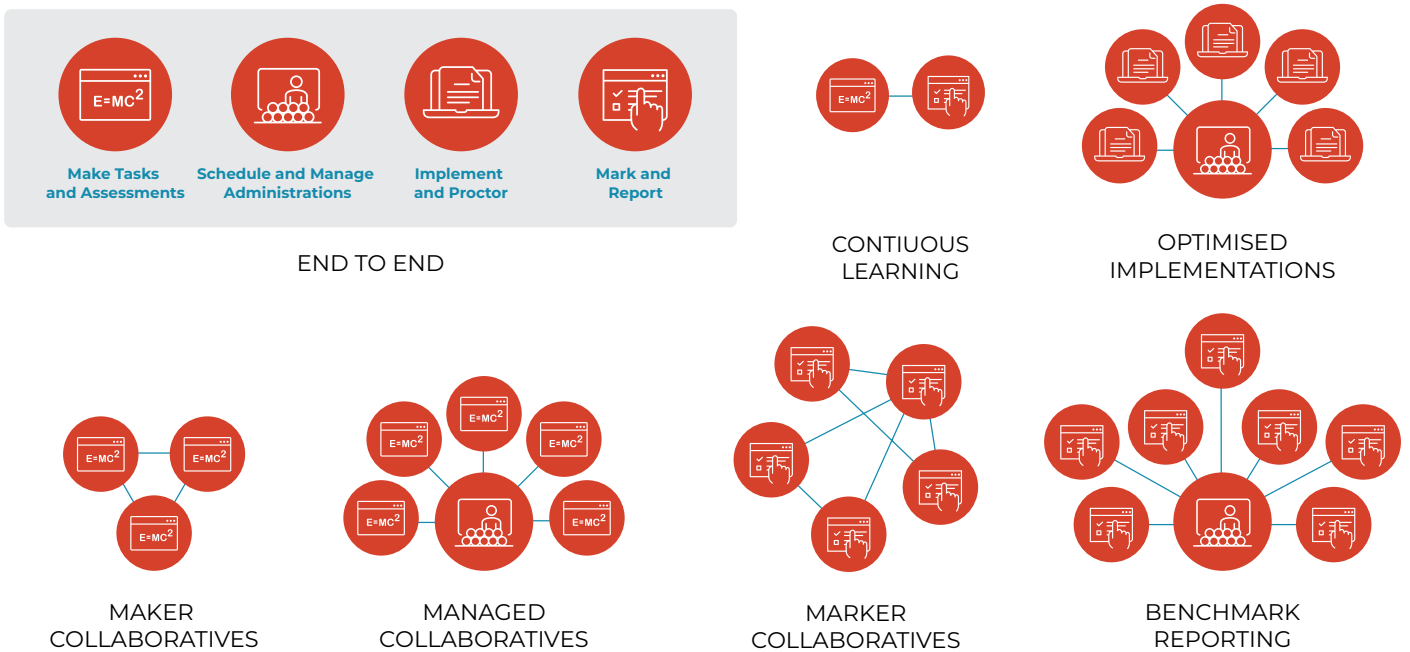


Figure 4: Smarter Learning collaboratives

universities, assessment experts and technology specialists, as well as building shared assessment frameworks and materials, validating materials, online delivery, and the production of sophisticated benchmark reports. Faculty are engaged in structured and formative ways, which also boosts professional competence. Authenticity is critical. Next-generation assessment builds on the expertise of teachers and universities which, distinguishes it from off-the-shelf products. Participating universities play a key role, contributing expertise, leadership and resources.

These are interesting ideas, but also ideas which have been developed. These reform options have already

universities work closely with each other who can share best practices in terms of learning design, contribute to the development of database, standardise assessment content and compare results to derive valuable insights. In addition, experts from professional bodies can be invited to facilitate the mapping between desired skillset in the industry and what is being assessed at university. Chartered Accountants, for instance, might work with universities to help standardise the curriculum and exams in first-year accounting courses and provide the necessary assurance on the relevance of the content as well as accounting accreditation for students who complete the 'audited' courses.

Figure 5: Potential structured collaboration model

This model has several distinctive advantages. The first, is improved efficiency through reduction of duplicated efforts. As highlighted, many faculty members are working in silos when developing their own assessments which is highly time-consuming. Due to the high degree of overlap in content for many courses (especially the introductory ones), it is far more efficient to have assessments built from a common/ shared database as opposed to any number of similar but discrete databases. The second advantage is quality and relevance. The content of many universities courses is obsolete compared to what is going on in that industry today, due to the lack of review from professionals currently working there. By incorporating

their input and updating the database, all universities can benefit from such quality assurance. The third, is that despite the huge overlap in content for many courses, differences in the style and format of the assessments resulting from individual idiosyncrasies of instructors make it nearly impossible to compare results outside of one university and derive insights. Standardisation enables cross-institution comparison and benchmarking, offering valuable insights for faculty and university management alike.

This paper has introduced the 'smarter learning' initiative, spotlighted reforms required to achieve underpinning 'next-generation' forms of assessment and articulated feasible steps ahead. By adopting a design-oriented perspective, the paper has argued that learning is becoming more important, that advancing

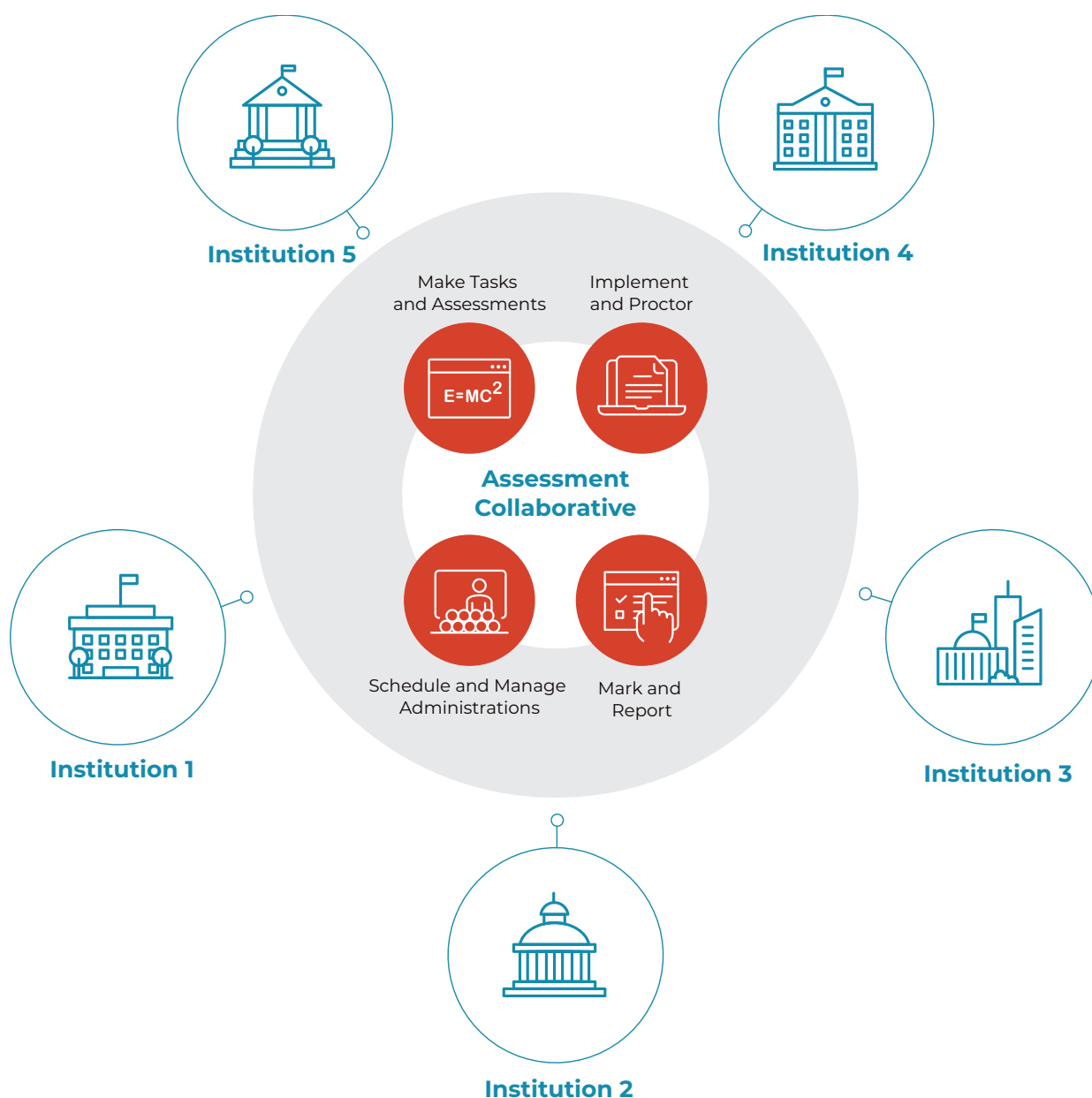


Figure 5: Potential structured collaboration model

learning hinges on assessment reform, and that such innovation will yield major productivity advances for higher education and broader communities. Figure 5 diagrams this idea with reference to the step-change developments discussed above.

Figure 6: Journey to Smarter Learning

How such progress unfolds hinges on a range of factors, many with no particular relationship to higher education assessment. Coates and Lennon (2014) identified a number of spurs to development, including quality contexts, disciplinary engagement, effective academic leadership, and perceptions of value and relevance. Coates (2020) further integrated these affordances in a broader review of change blockers and identified the 'value-creating constellation' depicted in this paper. The accelerated shift to hybrid education

since 2020 has augmented the need and opportunity for reform.

The year 2020 marks an inflection point when assessment became harder for higher education institutions than students. Every day, hundreds of millions of people in the world's 20,000-plus institutions engage in unproductive assessment of learning. This costs time and money, hinders learning, and squanders the capacity for higher education to prove its social, economic and professional contribution. This paper articulates the smarter learning initiative, which rests on next-generation assessment creating step-change value for education by raising standards and reducing costs. Smarter learning makes step-change advance possible, which enhances the sustainability and prosperity of higher education.



Figure 6: Journey to Smarter Learning

References

- American Educational Research Association, American Psychological Association & National Council for Measurement in Education (AERA, APA & NCME) (2015). Standards for Educational and Psychological Testing. Washington DC: AERA.
- Bennett, R. (2015). The Changing Nature of Educational Assessment. *Review of Research in Education*, 39(1), 370-407.
- CALOHEE (2021). Measuring and Comparing Achievements of Learning Outcomes in Higher Education in Europe. Accessed from: <https://www.calohee.eu>.
- Cantwell, B. Coates, H. & King, R. (Eds.) (2018). Handbook on the Politics of Higher Education. Cheltenham: Edward Elgar.
- Coates, H. & Lennon, M. C. (2014). Propelling the field: Insights, trends and prospects. In H. Coates (Ed.), Higher education learning outcomes assessment: International perspectives: 295–312. Frankfurt: Peter Lang.
- Coates, H. & Richardson, S. (2012). An international assessment of bachelor degree graduates' learning outcomes. *Higher Education Management and Policy*, 23(3), 51-69.
- Coates, H. (2016). Assessing student learning outcomes internationally: Insights and frontiers. *Assessment and Evaluation in Higher Education* 41(5), 662-676.
- Coates, H. (2017). *The Market for Learning: Leading transparent higher education*. Dordrecht: Springer.
- Coates, H. (2018). *Next Generation Assessment*. Melbourne: Online Education Services.
- Coates, H. (2020). *Higher Education Design: Big deal partnerships, technologies and capabilities*. Singapore: Palgrave Macmillan.
- Coates, H. (Ed.) (2014). *Higher Education Learning Outcomes Assessment*. Frankfurt: Peter Lang.
- Coates, H., Xie, Z. & Hong, X. (2020). Engaging transformed fundamentals to design global hybrid higher education. *Studies in Higher Education*.
- Edwards, D., Wilkinson, D., Canny, B., Pearce, P. & Coates, H. (2014). Developing outcomes assessments for collaborative, cross-institutional benchmarking: Progress of the Australian Medical Assessment Collaboration. *Medical Teacher*, 36(2), 139-147.
- Genix Ventures (2021). *Examina+*. Accessed from: <https://examinaplus.com>.
- Hazelkorn, E., Coates, H. & McCormick, A. (Eds.) (2018). *Research Handbook on Quality, Performance and Accountability in Higher Education*. Cheltenham: Edward Elgar.
- Liu, L, Wong, E. & Coates, H. (2019). Exploration on the reform of online course operating mode in Chinese universities: Inspiration from OPM provider-university cooperation model in Western countries. *Distance Education and Online Learning*, 1, 122-28.
- Mislevy, R.J., Almond, R.G. & Lukas, J.F. (2003). *A Brief Introduction to Evidence-centered Design*. Princeton: Educational Testing Service.
- Richardson, S. & Coates, H. (2014). Essential foundations for establishing equivalence in cross-national higher education assessment. *Higher Education*, 68(6), 825-836.
- Wong, E. (2019). *New Insights into Contemporary Higher Education in the US: Framework of Online Program Management (OPM) providers and its implication for China*. Beijing: Tsinghua University.