



The Centre
for Education
& Youth

Making Waves

A better future
for assessment

December 2020



Pearson

Loic Menzies | Abi Angus | Sam Baars | Will Millard
Bart Shaw | Ellie Mulcahy | Kate Bowen-Viner
Alix Robertson | Iesha Small | Anna Trethewey





Pearson is the world's learning company with more than 22,500 employees operating in 70 countries. We provide content, assessment and digital services to learners, educational institutions, employers, governments and other partners globally.

We are committed to helping equip learners with the skills they need to enhance their employability prospects and to succeed in the changing world of work. We believe that wherever learning flourishes so do people.

Find out more about Pearson at www.pearson.com/uk.



The Centre for Education and Youth (CfEY) is a 'think and action-tank'. We believe society should ensure all children and young people receive the support they need to make a fulfilling transition to adulthood.

We provide the evidence and support policy makers and practitioners need to support young people.

We use our timely and rigorous research to get under the skin of issues affecting young people in order to shape the public debate, advise the sector and campaign on topical issues.

We have a particular interest in issues affecting marginalised young people.

Find out more at <https://cfey.org/> or hello@cfey.org.

Contents

List of abbreviations	5	<i>2.1.3 Policy-level drivers of innovation</i>	17	3 The wave makers	24
Foreword	7	<i>2.1.4 Workforce drivers of innovation</i>	17	3.1 Innovation 1: Heathfield Community College Assessment Innovation Team	25
1 Introduction and methodology	8	2.2 How do people shape their solution and decide what to do?	18	3.2 Innovation 2: Isaac Physics at Rickmansworth School	39
1.1 Why Making Waves?	9	2.3 What approaches are being taken?	18	3.3 Innovation 3: ImpactEd at Bengeworth Primary Academy	43
1.2 The Making Waves story	10	<i>2.3.1 Improving the way assessment information is used to provide feedback and inform teaching</i>	19	3.4 Innovation 4: Shine at Eltham Hill School	55
1.3 Methodology	10	<i>2.3.2 Providing more meaningful or valid information</i>	20	3.5 Innovation 5: Eedi at the Academies Enterprise Trust (AET)	62
1.3.1 Recruitment	10	<i>2.3.3 Maintaining and enhancing teacher and pupil autonomy</i>	20	3.6 Innovation 6: Laser conversations at the Midland Academies Trust	73
1.3.2 Selection	10	2.4 How do people put their ideas into practice?	20	3.7 Innovation 7: Curriculum and assessment reform in British Columbia, Canada	81
1.3.3 Research framework	11	<i>2.4.1 Dissemination mechanisms</i>	20	3.8 Innovations 8 and 9: Online assessment resource banks in New Zealand and Victoria, Australia	100
1.3.4 Fieldwork	12	<i>2.4.2 Flexibility and rigidity</i>	21		
1.3.5 Analysis and synthesis	12	<i>2.4.3 External partnerships</i>	21		
1.3.6 The advisory group	12	<i>2.4.4 Facilitating factors and barriers to success</i>	22		
2 What have we learned?	14	2.5 Reviewing and adapting	23	4 Conclusion	114
2.1 Why are people making waves in assessment?	15			Appendix: Research framework	118
2.1.1 Pedagogical drivers of innovation	15				
2.1.2 Cultural drivers of innovation	16				

List of abbreviations

ACER	Australian Council for Educational Research	MAT	Multi-academy trust
AET	Academies Enterprise Trust	NAPLAN	National Assessment Program – Literacy and Numeracy
AIT	Assessment Innovation Team	NDP	New Democratic Party
ARB	Assessment resource bank	NFER	National Foundation for Educational Research
ASCL	Association of School and College Leaders	NZCER	New Zealand Council for Educational Research
BC	British Columbia	OECD	Organisation for Economic Co-operation and Development
BCTF	British Columbia Teachers’ Federation	PAT	Progressive Achievement Test
CfEY	Centre for Education and Youth	PACT	Progress and Consistency Tool
CPD	Continuing professional development	PISA	Programme for International Student Assessment
CPL	Collaborative professional learning	SEND	Special educational needs and disabilities
CSP	Classroom support professional	SIMS	School Information Management System
DAL	Digital Assessment Library	SLT	Senior leadership team
EAL	English as an additional language	SSAT	Schools, students and teachers network
EEF	Education Endowment Foundation	TLR	Teaching and Learning Responsibility
FNESC	First Nations Education Steering Committee	VCAA	Victorian Curriculum and Assessment Authority
GCSE	General Certificate of Secondary Education	VCE	Victorian Certificate of Education
GELP	Global Education Leaders’ Programme		
GLA	Greater London Authority		
HoD	Head of Department		
ICT	Information and communications technology		
INSET	In-service training		
IT	Information technology		

Foreword

Pearson and The Centre for Education and Youth (CfEY) are delighted to have once again renewed our partnership, through the Making Waves study.

Making Waves is an important piece of research that is being published at an historic time for education assessment. The Covid-19 pandemic and the closure of schools in March 2020 resulted in the government taking the radical decision to cancel A Level and GCSE examinations. Instead, students' public examination grades were calculated through centre (or teacher) assessed grades.

Many countries already rely much more heavily than the UK on teacher assessment at key moments in learners' educational careers. And, whilst it took a global pandemic to force such a radical shift in this country, the merits of different types of assessment and their benefit to learner progress has long been debated. Never fully resolved, the debate has resulted in a myriad of assessment theories and practices but as Pearson and CfEY's 2017 report 'Testing the Waters' showed, this has left teachers feeling overwhelmed by

excessive workload and a feeling that support for learners' progress has been pushed to the side. Teachers and school leaders report that this has been exacerbated by an absence of sufficient training on assessment, whether as part of initial teacher training or through continuing professional development.

'Testing the Water' showed that workload is one of the factors that stand in the way of staff taking control of and developing assessment practices. However, as schools reopen after lockdown there is an opportunity for colleagues to join the practitioners featured in this report in re-evaluating school practices that fail to contribute to learning. Many schools are looking for alternative ways of capturing progress, and deliberate refining of practices has the potential to usher in a better future for assessment.

This report encourages colleagues to wrestle with the seemingly intractable problems of assessment in education. Pearson has always been committed to developing teachers' assessment expertise and ensuring that the

learner is front and centre. It is therefore encouraging that the 'Making Waves' project has found that teachers are taking the thorny issue of assessment into their own hands and trialling different approaches and models.

We hope that the 'wave makers' in this report empower educators to draw on their professionalism and experience in order to experiment with new forms of assessment. We encourage colleagues to draw on Making Waves' key findings regarding the vital role of taking a structured approach in doing so.

The etymological root of "assess" reaches to the idea of 'sitting by' someone as they learn. At its heart, assessment is being, metaphorically, 'alongside the pupil or student'. What follows from this is on one hand the identification of misconceptions and celebration of insights, and on the other, making judgements about 'standards'. 'Making Waves' provides important insights from practitioners and academics about moving discussions about both of these priorities to the next level.



Marry Myatt

Chair of the Board

The Centre for
Education and Youth



Rod Bristow

President, Pearson UK &
Global Online Learning

Pearson

1

Introduction and
methodology



1.1

Why Making Waves?

Workload in schools has spiralled out of control and this is pushing teachers out of the profession at an unprecedented rate. Teachers do not feel confident about their assessment expertise and this can stand in the way of them exercising their professional autonomy.

However, pockets of innovation are springing up in response to these two challenges and our previous research¹ showed that professionals were developing new approaches that could give rise to a better future for assessment. These innovations need to be properly studied so that lessons on two key questions can be learned, and good practice scaled up.

These questions are:

1. How can we reduce teachers' assessment workload?
2. How can we make sure teachers have the assessment expertise they need?

¹ Millard, W., Small, I. and Menzies, L. (2017) Testing the Water: How assessment can underpin, not undermine, great teaching, London: LKMco and Pearson, available at: www.lkmco.org/wp-content/uploads/2017/11/Testing-the-Water-Final-Report-WEB.pdf

Making Waves



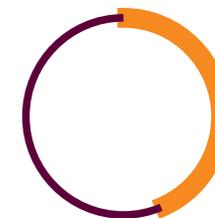
Only one third of classroom teachers feel 'very confident' conducting assessment as part of their day-to-day teaching.



The **need for training** is greatest at the chalkface: classroom teachers are less likely than their more experienced colleagues to have access to ongoing professional development in assessment



One in five classroom teachers would not know where to look for information on assessment if they needed it.



Under half of teachers received training in assessment as part of their initial teacher training and teachers access to assessment training over the course of their careers is far too limited

The Making Waves Project was designed to answer these questions in a consultative, deliberative manner through partnership with the sector.

The Centre for Education and Youth (CfEY) and Pearson believe that the answers to

many of the challenges in the education sector are in the minds and classrooms of professionals. We therefore embarked together on a mission to find out what we can learn from the teachers, schools and countries that are already 'making waves' win assessment.

1.2

The Making Waves story

Christmas 2017: The CfEY and Pearson publish *Testing the Waters* – a landmark report setting out the biggest challenges for assessment in England. The report brings together the views of thousands of teachers, pupils and parents as well as a number of international studies to suggest potential ways forward.

Easter 2018: We decide to focus on two of the biggest challenges to emerge from *Testing the Waters*, namely workload and teacher expertise.

September 2018: We launch our ‘Are you Making Waves?’ campaign, asking teachers, schools and school groups to share what they are doing to tackle these two key issues.

November 2018: Having received nominations for over 40 different innovations, our expert advisory group helped select 10 innovations or ‘waves’ for us to study over the course of the next year.

Spring 2019: The first round of fieldwork begins.

Autumn 2019: The third and final round of fieldwork concludes and final analysis begins.



1.3

Methodology

1.3.1 Recruitment

Potential innovations were ‘crowdsourced’ by disseminating information about the study and speaking to experts in the field, with links to a short survey asking educational institutions and teachers for information about their innovation.

The crowdsourcing campaign was promoted through the following networks and publications: Schools Week, Teacher Toolkit, NAHT, the Association of School and College Leaders (ASCL), NASUWT, BAMEed, WomenEd, the HeadTeachers RoundTable, Challenge

Partners, SSAT (the schools, students and teachers network), the Education and Training Foundation, the Education Endowment Foundation (EEF), Pearson and CfEY.

Crowd sourced responses were supplemented with international insight from Alex Beard and the International Education Policy Community of Practice (Teach For All), Dylan William, Tim Oates and Sally Brown (Cambridge Assessment), Joe Hallgarten (Education Development Trust), Sam Freedman (Ark International), Ben Durbin and Lesley Duff (National Foundation for Educational Research – NFER), Amelia Peterson (Harvard University) and Lucy Crehan.

1.3.2 Selection

The recruitment phase yielded a long list of 44 innovations, which were summarised and shortlisted based on the following four criteria:

1. Relevance: The proposed ‘wave’ is clearly focused on addressing the workload associated with assessment and/or teachers’ assessment expertise (their knowledge, understanding or skill in relation to assessment).

2. Newness: The ‘wave’ has recently been implemented (or is about to be implemented), allowing us to study the process of implementing it and the changes that are taking place.

3. Plausibility: There is a plausible evidence base or rationale for how the ‘wave’ will lead to the intended change. In particular, is there a considered rationale for the approach, that is, a clear answer to the question ‘Why do you think this will work?’ Does this answer clearly build on or go against an established evidence base?

4. Balance: The ‘wave’ allows us to study a good balanced basket of innovations – that is, it is not too similar to/takes a different approach from other waves.

The advisory group then met to review the shortlist. The aim was to secure at least

two innovations at each ‘level’ (individual classroom or practitioner, school, group of schools and country or jurisdiction). This yielded a final list of ten innovations. Some of these were eventually unable to participate or had to leave the study over the course of the year (for example, where a school’s circumstances or leadership changed or where an individual teacher left the school or profession). In this scenario, where possible, similar alternatives from the shortlist were selected as substitutes. This process led us to focus on nine innovations which we describe in section 3 as ‘The wave makers’.

- **Innovation 1:** Heathfield Community College Assessment Innovation Team
- **Innovation 2:** Isaac Physics at Rickmansworth School
- **Innovation 3:** ImpactEd at Bengeworth CE Academy
- **Innovation 4:** Shine at Eltham Hill School
- **Innovation 5:** Eedi at the Academies Enterprise Trust (AET)
- **Innovation 6:** Laser conversations at the Midland Academies Trust
- **Innovation 7:** Curriculum and assessment reform in British Columbia, Canada
- **Innovations 8 and 9:** Online assessment resource banks in New Zealand and Victoria, Australia

1.3.3 Research framework

Since the study focuses on the emergence and implementation of new approaches, we designed a framework based around a theoretical framework involving five elements.² This framework is typically used to understand policy processes but applying it to the different tiers of innovation (teacher/school/school group/state) provided a unifying structure.

The five elements were:

1. Agenda setting: What were/are the ‘pain points’ that this innovation sought to respond to? Who and what influenced this agenda? Why was it considered an important agenda to pursue?

2. Formulation: How was a solution arrived at? Who and what fed into shaping it?

3. Decision making: What solution was selected? Who was involved in deciding on this approach to addressing the pain points? Why were other solutions rejected and this approach preferred?

4. Implementation: What mechanisms have been used to implement the approach?

² These are drawn from the key stages in the policy process (Howlett, M. and Giest, S., 2013, ‘The policy-making process’, in Araral, E., Fritzen, S., Howlett, M., Ramesh, M. and Wu, X. (eds) Routledge Handbook of Public Policy, London: Routledge).

Who are the different actors involved? To what extent is the reified solution consistent with the planned/intended solution? What obstacles had to be overcome as part of implementation? What conditions contributed to or detracted from effective implementation?

5. Evaluation: How is success or failure being characterised? What if any new success criteria is the innovation being held to, compared to its originally stated goals? What approaches are being used to evaluate the innovation's success? Is the innovation believed to be achieving its intended objectives?

1.3.4 Fieldwork

In most cases, fieldwork took place through three, in-person visits over the course of the year, one each term. However, in some cases, flexibility was needed – for example where a case study (like Rickmansworth School) was added at a later stage due to attrition or, in the case of New Zealand and Victoria, Australia where it was considered better to conduct two visits combining the two jurisdictions, one at the start of 2019 and one in the autumn of 2019.

Each 'mini-study' was structured around the same framework (see the Appendix). The first visit generally focused on the early

stages of the policy cycle (agenda setting and formulation), while later visits looked in more detail at the subsequent stages (such as implementation and evaluation).

Field researchers used the overarching framework to tailor semi-structured interview scripts and focus-group discussion guides to each setting. In doing this they took into account:

- institutional circumstances and practicalities
- who the actor was, for example whether the focus was on an individual practitioner, an overarching multi-academy trust (MAT) or a set of policy makers
- what the innovation was and what it sought to achieve
- emerging themes over the course of the year (such as identified challenges).

Wherever possible, pupils were included among the research participants. Interviews and focus groups were recorded and transcribed where practical. Attributed quotes were checked with participants.

1.3.5 Analysis and synthesis

Fieldworkers wrote up their studies using a common structure based on the five-stage cycle. The lead author then conducted

thematic analysis, identifying and combining recurring themes and linking these to quotes and examples. This provided the analysis that is presented in section 2: 'What have we learned?'

1.3.1 The Advisory Group

Membership of the advisory group evolved over the course of the project as members' roles and availability changed. Participants met four times to review emerging themes, identify emerging questions and ensure lessons arising would be useful to a range of audiences and stakeholders. The following individuals participated:

1. **Eleanor Andressen**, Senior Researcher Pearson
2. **Loic Menzies**, Chief Exec, The Centre for Education and Youth
3. **Debra Rutley**, Head of Aspire Alternative Provision
4. **Simon Yates**, Special School Head teacher
5. **Amelia Walker**, National Director of Strategy for Ormiston Academies Trust
6. **James Zuccollo**, Director of School Workforce, Education Policy Institute

7. **Nick Brook**, Deputy General Secretary of the NAHT
8. **Kate Chhatwal**, Chief Executive, Challenge Partners
9. **Mark Lehain**, Former Director of Parents and Teachers for Excellence
10. **Darren Northcott**, NASUWT
11. **Allana Gay**, Deputy Head teacher
12. **Leora Cruddas**, Chief Exec, Confederation of School trusts
13. **Julian Astle**, Former Head of Education at the Royal Society of Arts
14. **Tom Middlehurst**, Head of Policy and Public Affairs, SSAT
15. **Olly Wimborne**, Senior Researcher, Royal Society of Arts
16. **Clare Sealy**, Former Head, St Matthias Primary, Tower Hamlets.



2

What have we
learned?



A vanguard of teachers, schools, groups of schools and countries is working towards a better future for assessment

We have spent a year studying their journeys.

The road has not always been easy, and the story that emerges is not one of simple, transferable solutions. However, we hope that by telling these wave makers' stories we will help educationalists in schools and governments to rediscover the professional agency needed to take back control of assessment and ensure it better serves pupils' needs.

We hope the challenges taken on by the teachers and educators we met will resonate for professionals across the system, and provide a taster of some of the avenues they might pursue when tackling assessment challenges.

We argue that this challenge involves not just finding 'the right solution' off the shelf, but also planning implementation carefully, reviewing progress regularly and adapting where necessary.

2.1

Why are people making waves in assessment?

Innovation does not happen without reason, so the first stage in understanding new approaches to assessment is to pin down the reasons for moving beyond the status quo.

The innovations we studied responded to multiple drivers. Some of the time these were framed in positive ways – a desire to find new and better pedagogical models or to empower students to learn independently – but in most cases they were more responsive, seeking to tackle perceived pain points, such as an unmanageable workload.

The four main drivers of innovation were pedagogical, cultural, policy-based and workforce-driven.

2.1.1 Pedagogical drivers of innovation

In the countries we studied outside of the UK, assessment innovation was often seen as part of a shift towards new or alternative pedagogical approaches – for example, developing a cross-curricula approach that prioritises independent learning and 'holistic' non-academic outcomes.

“This is not about curriculum and disciplines being the end destination. The end destination is the process of becoming as a human being. It’s about human development that’s the end destination. The confidence around communication, around critical thinking, creative thinking, around social and emotional wellbeing, around personal identity. That’s our destination.”

Maureen Dockendorf,

*Superintendent of Literacy and Numeracy,
Ministry of Education, British Columbia*

This was often linked to dissatisfaction with summative assessment, which many feel to be insufficiently focused on monitoring progression and mapping out next steps. This is a key agenda in Victoria, Australia and in New Zealand.

In England, teachers, schools and MATs are often dissatisfied with the quality of assessment. Many want to see a closer link between the information that assessment provides and the actions it prompts. This makes some of the innovation that is currently taking place a natural continuation of longstanding efforts to move towards 'assessment for learning'.¹

¹ Black, P. and Wiliam, D. (2005) Inside the Black Box: Raising standards through classroom assessment, London: Granada Learning.

This led some innovators to develop tools for rapid feedback on pupils' misconceptions or what they had and had not been remembered, so that they could reshape their teaching based on the information they gleaned. This challenge was partly linked to teacher expertise² in that some teachers were said to lack the subject knowledge, or assessment expertise, needed to optimise the relationship between assessment and the next steps for their teaching.

“Because children give you the wrong answer but they think it’s the right one... it’s having the subject knowledge as a teacher to think, ‘Why do you think that’s right? What’s going on in your head? What’s your misconception that we need to unpick?’ And that is where the poor subject knowledge [from teachers] means that that child doesn’t necessarily move forward.”

Maths lead

Lack of assessment expertise combined with extensive data collection could lead to what was

² Millard, W., Small, I. and Menzies, L. (2017) Testing the Water: How assessment can underpin, not undermine, great teaching, London: LKMco and Pearson, available at: www.lkmco.org/wp-content/uploads/2017/11/Testing-the-Water-Final-Report-WEB.pdf

described in the Midlands as “data for data’s sake” and in Victoria as “a lot of time (being) wasted ‘weighing the pig’ and producing data that is not reliable, valid or comparable”.

Of course, none of the innovations we studied are the first attempts to solve these problems. In fact, rejection of previous failed initiatives often spawned a desire for something new and drove innovation.

Typical reasons for failure included several of those explored in greater length in this report’s predecessor, Testing the Water,³ such as:

- excessive cost
- labour-intensive approaches that provided limited information of value
- approaches that tried to serve multiple competing functions (such as predicting performance as well as guiding teaching and intervention).

2.1.2 Cultural drivers of innovation

Innovators were keen to reflect their classroom, school or country’s culture in their approach to assessment.

³ Millard, W., Small, I. and Menzies, L. (2017) Testing the Water: How assessment can underpin, not undermine, great teaching, London: LKMco and Pearson, available at: www.lkmco.org/wp-content/uploads/2017/11/Testing-the-Water-Final-Report-WEB.pdf

Many were looking for a more collaborative approach with more pupil or parent involvement, or with teachers and managers working together.

The innovators we spoke to often valued inclusion highly, and wanted to ensure their approach to assessment reflected this. Sometimes this was based on an overall desire to ensure that assessment helped identify and respond to pupils’ individual needs rather than being a tool for accountability. At other times, teachers wanted to ensure assessment took into account non-academic priorities. A desire to respond to specific groups’ needs was also a key consideration, for example indigenous groups in British Columbia in Canada, Victoria in Australia and New Zealand, or pupils with special educational needs and disabilities (SEND).

Some contexts provide particularly fertile ground for innovation. This could be seen in schools or countries that reported an overall interest in, or appetite for, innovation in and of itself. Some schools, for example, described themselves as particularly engaged in research, and policy makers in British Columbia referred to the province’s desire to be at the cutting edge of educational innovation and reform.

2.1.3 Policy-level drivers of innovation

National shifts in policy can drive innovation at school level; The Assessment Innovation Team at Heathfield Community College was partly set up in response to the government's decision to move to 'assessment without levels' whilst the approach taken by the college's Biology department was shaped by a perceived increase in the rigour of the curriculum. Meanwhile, changes in the Ofsted framework, SEND code of practice and 'national agenda' influenced Bengeworth CE Academy's priorities.

"[Previously] I don't think the focus was on inclusion as it is now in the national agenda. I think it's really become more pivotal... inclusion drives the agenda of the school, its CPD [continuing professional development], its vision and its values should be inclusive... That's become high importance in Ofsted and other branches of national agendas."

Policy drivers are also linked to school funding, which can push schools to pursue value for money.

2.1.4 Workforce drivers of innovation

Our 2017 report, Testing the Water,⁴ showed that assessment is often linked to excessive workload, and many of the innovations we studied responded to an urgent imperative to reduce teacher workload. In other cases, a key design principle was that any change should not add to workload. Innovators also sometimes distinguished between 'worthwhile' and 'less worthwhile' workload, arguing that what mattered was ensuring teachers spent less time on unproductive or meaningless activity. This sentiment was echoed by our advisory group.

"In the long term, it's going to alleviate the workload of teachers because they're producing those provision maps and going back into them afterwards and collating them. So, [the platform we use provides] a dropdown menu and it will produce it for you."

Head of Inclusion

4 Millard, W., Small, I. and Menzies, L. (2017) Testing the Water: How assessment can underpin, not undermine, great teaching, London: LKMco and Pearson, available at: www.lkmco.org/wp-content/uploads/2017/11/Testing-the-Water-Final-Report-WEB.pdf

"[The innovation needed to be something] that I could embed very easily into my practice and not have to then spend any extra time doing anything. I think anything that is going to be effective, in the classroom, does need to consider that because you can't maintain it otherwise."

Jess Gillespie, Heathfield

Additionally, some schools and MATs selected their approach in response to struggles in securing the workforce they needed, particularly in subjects like maths and physics. Technological solutions might therefore be one way around the assessment challenges staffing difficulties can cause.

"The main thing was the lack of stability in the teaching staff, because there were lots of changes, you would have a lot of people coming in, and you wouldn't know what assessment had taken place, you wouldn't know how well the students had done. You had no real overview. So, from leading a department point of view, I just needed to get a handle on good feedback."

Head of Physics

2.2

How do people shape their solution and decide what to do?

Once innovators spot an issue with assessment, they formulate their response and decide on a solution. A simple, rational model of decision making would be to imagine that a range of different options are assembled, weighed up and the best bet selected. However, this is not what happens in practice.

A range of individuals are involved in shaping and making decisions and they have contrasting priorities, different lenses to look at issues through and limited time. Describing their decisions as ‘non-rational’ would be unfair, but it is clear that their choices are not always a case of simple optimisation.

Alternatives are not always considered and weighed up. When teachers, schools and countries innovate in assessment, they often proceed, step-wise, from one initiative to another, either building on the previous approach and taking it to the next level, or trying and rejecting different approaches sequentially. This is not necessarily a bad thing. In many cases, innovators identified an approach that was already working well and

extended it, or applied it in different contexts or cases. Examples of this include building an ‘Assessment Innovation Team’ based on a school’s existing successful use of Innovation Teams focused on other themes; or identifying a practice within one school in a MAT and implementing it across different settings, as the Midland Academies Trust did. On the other hand, the curriculum lead in one MAT described an approach designed to help weigh up alternatives through “thoughtful disagreement”:

“We thrashed it out with the team. We always encourage absolute open debate and disagreement. We call it thoughtful disagreement... we ask people to play devil’s advocate... it’s not about the loudest voice at winning the arguments, it’s about the most compelling reasoning and logic.”

MAT curriculum lead

Research and learning from others also played a role in directing innovators to certain solutions, but this was not based on extensive searches through a broad range of literature. Instead, certain high-profile researchers like Daisy Christodoulou, Dylan Wiliam and Becky Allen were repeatedly mentioned, while others drew ideas from their close peers.

Context shapes options in many ways, making some innovations opportunistic or ‘bounding’ the options that can be rationally selected. Online assessment platforms in Victoria, Australia and New Zealand, for example, were shaped by the status quo, including who owns data, platforms and tools. Meanwhile in British Columbia, some argued that decisions were limited by the ‘Overton Window’, that is, the range of policies that are politically acceptable at any one time.

Ultimately, decisions are made by individuals and groups of individuals. At Heathfield Community College, the fact that the Head of Geography was also the school’s research lead influenced their approach, while the Photography teacher’s background as a sports coach shaped theirs. This means that who has power – in terms of both their degree of autonomy and their background and beliefs – shapes decisions.

2.3

What approaches are being taken?

The innovations we studied were chosen from a pool of around 40 different initiatives. These clustered around four broad and overlapping approaches, or types of innovation:

1. **Reduced or simplified approaches to marking**
2. **Education technology**
3. **New approaches to assessing non-academic or ‘soft’ skills**
4. **Programmes of training and development seeking to enhance teachers’ assessment expertise.**

As we looked at our selected innovations more closely, it became clear that the three underlying themes behind assessment innovation were:

1. **Improving the way assessment information is used to provide feedback and inform teaching**
2. **Providing more meaningful or valid information**
3. **Maximising teacher and pupil autonomy.**

2.3.1 Improving the way assessment information is used to provide feedback and inform teaching

The innovations we studied did not focus on accountability or qualifications; instead, they sought to improve teaching and learning.

The first way they tried to do this was by shortening the feedback loop between gathering assessment information and making adjustments to teaching. This often involved combining technology and homework.

Gaining immediate information about misconceptions was often what teachers found most valuable about these approaches, since it allowed them to respond straight away.

“What we’re trying to do is basically allow teachers to be better informed so they can provide better instruction, allow students to understand – not just that they’re wrong, but why they’re wrong, and for parents to understand where their child is struggling and how they can support them.”

Eedi central team

“I keep a continual eye on the information that I get from the website about how many questions they’ve attempted, and I use this as a way of just keeping an eye on how much they’re doing in their own time, independent learning at home, and also as a way of seeing whether or not they’ve had particular issues with a certain style of question.”

Head of Physics

On the other hand, homework completion rates and technology cannot always be depended on so relying on homework and technology can mean teachers do not get the information they need to shape their teaching. Some schools within the Academies Enterprise Trust (AET) have therefore diverged from the initial plan by conducting assessments in class or on paper. The Biology teacher within the Assessment Innovation Team at Heathfield have integrated assessment tools into lesson starters as well as homework.

Several teachers who submitted ideas to our crowdsourcing exercise described developing Google Form surveys that allow them to get auto-marked, at-a-glance information about their class’s understanding through starter activities. This demonstrates how individual

teachers can take the initiative in developing solutions, even where they do not have access to commercial platforms. In contrast, in New Zealand and Victoria, Australia, it is the state that has taken the initiative in developing assessment platforms. These provide feedback that can flag pupils' misconceptions as well as mapping out progress, while informing summative assessment. These roles are particularly important in these countries' contexts where there is a loose and un-prescriptive curriculum which leads some teachers to call for further guidance.

2.3.2 Providing more meaningful or valid information

Not everyone is convinced of the value of assessment that only includes academic outcomes. The most radical manifestation of this agenda which we studied was in British Columbia. There, pupils self-assess in three Core Competencies, and traditional assessment has largely been done away with.

Within England, several innovators introduced initiatives that extended what was assessed and how, but did so alongside more traditional approaches. Bengeworth CE Academy did this using ImpactEd – a platform that provides tools for assessing non-academic outcomes and for

benchmarking progress against these. Meanwhile, Rosie Osborne, Lead Practitioner at Eltham Hill School, wanted to give feedback to pupils in her after-school group in a way that has real-world validity and which boosts pupils' confidence. This is why she uses a non-teacher (a professional artist) to provide pupils with informal feedback.

2.3.3 Maintaining and enhancing teacher and pupil autonomy

Assessment can often be a top-down affair, and teachers and pupils sometimes resent prescriptive or rigid systems. Many of the innovations we studied were explicitly designed to avoid this. The photography department at Heathfield Community College decided to use exemplar work, giving pupils and teachers guidance on the features of high-quality work – without overly shaping what they produce. Meanwhile, the Assessment Innovation Team model at Heathfield is itself designed to promote and maintain teacher autonomy by giving individual teachers and departments the power to shape assessment according to their priorities and disciplines.

Assessment banks like those used by the online platform Eedi and the assessment resource banks (ARBs) in New Zealand are

another way of maintaining teacher autonomy, in that teachers have the freedom to select questions themselves. They are therefore designed to support teachers' use of high-quality assessment materials, without dictating teachers' focus.

2.4 How do people put their ideas into practice?

Implementing an innovation involves taking an idea, putting it into practice and (in most cases) disseminating or rolling it out. Approaches to doing this vary depending on how the initiative is taken beyond the initiator, how much flexibility and adaptation there is along the way, and the role of external partners.

2.4.1 Dissemination mechanisms

Most approaches to dissemination can be characterised as a form of 'cascading', in other words, they start with an initial node – which might be an individual, group or organisation – and are then widened out from there. Selecting and deploying an approach to cascading is therefore a critical element of implementation.

Both of the MATs we studied had spread an approach out from a central team to their schools. In the larger MAT, the central team disseminated their new approach through conferences, while in the smaller MAT, dissemination worked in both directions, with the ‘laser conversation’ approach originating in one school before being adopted by the central team and being spread to other schools. The context of being a small MAT made this easier and much of the dissemination happened informally, although all schools were expected to follow a set timetable and heads were required to use the new approach in order to feed information to the central team.

As we will see below, the degree of flexibility given to schools as part of this has important implications and, in both cases, the MATs have since provided further guidance and direction or plan to provide this in the future.

Where innovations were introduced at school level, cascading happened between departments (as in the case of Heathfield), while at state level (in British Columbia) cascading took place from state to districts and in turn to schools. Flexibility was built into British Columbia’s model with districts’ approaches differing in a number of ways. This raises the important question of flexibility versus rigidity - another key factor in implementation.

2.4.2 Flexibility and rigidity

Innovators did not tend to take their idea and roll it out as a finished product, or expect it to be implemented with total fidelity to the original plan. In some cases, flexibility was built in through piloting, for example Bengeworth CE Academy trialed ImpactEd with a small number of academic and non-academic interventions before adopting it more fully. In other cases, flexibility has been a more organic process of evolution and adaptation, with challenges identified throughout and modifications made in response. While this approach may be well suited to developing a new approach, it can make it harder to ‘test’ the effectiveness of an approach, as the Biology teacher at Heathfield Community College explained:

“You’ve got to have a bit of time to measure. So I only did it for a term, which is an extremely small amount of time, but you’ve then got to have the consistency and the patience really, to just keep with that, across a year.”

Furthermore, a flexible approach can mean that innovations stray from their original intentions and innovators therefore need to identify the

‘golden core’ of their approach so they can establish non-negotiables.

2.4.3 External partnerships

Innovators rarely work on their own. Even individual teachers implementing small-scale initiatives drew on expertise and resources from beyond the school, whether a professional artist in one case or an established technological product combined with a network of other teachers brought together through a WhatsApp group in the case of Rickmansworth School.

In some cases, external partners provided a tool or platform whereas in others they provided expertise and training, or a combination of both.

The Victorian and New Zealand governments used external specialist agencies to develop their assessment platforms and these agencies also used their high-level expertise to train and develop teachers so that they could develop a high-quality offer.

AET and Bengeworth CE Academy both emphasised the importance of a close relationship with their external provider, since this allowed them to co-develop the platforms, adapting and tailoring them to needs and responding quickly to challenges. It also meant

they were able to secure significant support and training, which helped overcome many of the challenges involved in implementing a technological solution. Meanwhile, Heathfield Community College sourced expertise from a university and – much like the training provided by the New Zealand Council for Educational Research (NZCER) and the Victorian Curriculum and Assessment Authority (VCAA) – this empowered teachers by giving them the assessment expertise they needed to become assessment innovators.

2.4.4 Facilitating factors and barriers to success

As explored above, effective dissemination, getting the right balance between flexibility and standardisation, relationships with external providers and the availability of training can all affect how successfully an innovation is implemented. Beyond these factors, a number of practical considerations can also contribute to or detract from successful implementation, as can the degree of alignment between stakeholders.

Practicalities

Innovators had to overcome a number of practical issues such as homework completion rates and access to technology. Additionally,

limited assessment expertise or excessive workload often increased individuals' resistance to an innovation and made it harder to implement the innovation successfully.

Innovation can also be expensive, and it was clear that several initiatives were highly vulnerable to changes in resourcing. Ultimately, if an innovation is to be successful, it needs to be adequately resourced – in terms of funding, training and freeing up teachers' time for an extended period. It is for this reason that AET now plans to award a Teaching and Learning Responsibility (TLR) payment for Eedi champions and why the Midland Academies Trust is keen to find time-savings in other areas (such as 'tick and flick' marking) to free up time for laser conversations.

Alignment

Securing and maintaining backing from a coalition of supporters makes it easier to ensure an innovation continues to be prioritised by all those involved. Developing and making a strong case for an innovation is therefore an important part of implementation.

In some cases, innovators did this by demonstrating impact; Rosie Osborne (Lead Practitioner at Eltham Hill School) reported that senior leaders trusted her to run her

innovation because they could see that students were more engaged in their learning thanks to their after-school projects. She created the conditions for this recognition by maintaining an open-door policy so that senior leaders could drop into sessions to see her work in action. Similarly, results from a mini control-group trial are making it easier for the Biology teacher at Heathfield Community College to persuade others of her approach. Her decisions about evaluation are therefore contributing to successful implementation. Meanwhile at Bengeworth CE Academy, the leadership team considered discontinuing their use of ImpactEd rather than rolling it out, but because Rachel Seneque (the Head of Inclusion) was able to show long-term potential for financial savings, they eventually backed deeper and broader implementation.

As we saw earlier, innovations often gain traction when 'the stars align' and several different agendas, held by different individuals, come together, or when an innovation fits particularly well with a specific context. Innovators therefore found that securing buy-in from leaders at different levels as well as from parents and pupils could oil the wheels of implementation. In contrast, implementation often suffered when one group struggled to understand the innovation or where sceptics could not be persuaded of its value.

Sometimes alignment was created between different, otherwise unrelated pedagogical agendas such as parental engagement, raising attainment, curriculum and improved assessment. In others it arose from complementary political agendas, something that could require lobbying or adapting to a changing political context in order to maintain long-term support, as happened for a time with the ARBs in New Zealand.

In contrast, problems can arise where there are clashes between different agendas or individuals as well as where remits are overlapping or confused. It is also harder to maintain alignment when there are frequent changes of staff.

2.5 Reviewing and adapting

We did not generally find a tight link between the impact innovators originally intended to have and the indicators of success that they referred to when describing their approaches to evaluation. Notions of success included:

- improved academic results and progression
- pupil and parental engagement
- uptake and usage of a platform or tool
- reduced workload or time-savings

- improved use of evidence
- changes in pedagogy – in particular, better use of assessment information
- improved non-cognitive skills
- shifts in mindsets and educational approach
- inclusion
- better relationships between teachers and pupils.

Apart from the biology department's Assessment Innovation Team at Heathfield, very few of the innovators we studied evaluated their initiatives formally.

“I’m a scientist, so maybe it was just my natural instinct to think, ‘well actually I do need something to compare this to’ but I was just in a really fortunate position, that I had a natural control in the sense that I had the two parallel groups.”

Jess Gillespie,
science mini project

It might be argued that this means initiatives will not lead to their intended changes and that this will go unnoticed. On the other hand, many initiatives were at an early stage where processes and approaches were still being refined. This meant that rapid and flexible adaptation was often a greater priority than

a formal evaluation, which innovators feared might create rigidity prematurely. Gathering feedback was therefore considered a more pressing priority and many innovators favoured informal assessments of whether pupil work and engagement were improving through - what Heathfield's Photography teacher called - “over-the-shoulder assessment”.

Others decided that usage of a tool was a good proxy for whether an initiative was effective. For example, in New Zealand, one interviewee argued that the ARBs did not need to be evaluated because their popularity was sufficient evidence of success. In large MATs like AET, gauging success required triangulation between various different sources of information. These included a usage tracker, book scrutiny by regional curriculum leads, as well as upward cascading of feedback from schools' maths leads via regional leads all the way up to the central team through the school engagement lead. There is clearly a compromise needed between flexibility and rigorous approaches to gauging efficacy, but Heathfield's approach – with small-scale innovation taking place at departmental level, with planned but flexible evaluation – may offer a model that balances these two competing priorities.

3

The wave makers



3.1

Innovation 1: Heathfield Community College Assessment Innovation Team

3.1.1 Introducing the innovation

Heathfield Community College is a local authority maintained secondary school in rural East Sussex, catering for around 1,500 pupils in Years 7 through to 11. The college also has a sixth form with around 400 students. ‘Innovation Teams’ are a central feature of how Heathfield delivers its school improvement plan and develops its staff. Teachers are encouraged to participate in fortnightly working groups that focus on specific areas of the school improvement plan. The Assessment Innovation Team in this study is one of these working groups.

3.1.2 Agenda setting

A consultation with teachers at Heathfield revealed gaps in teachers’ knowledge and

confidence relating to assessment, and it was evident that practice was often driven by habit rather than evidence of best practice. Information from work scrutiny also suggested that workload was an issue, and that teachers were spending time on assessment that was not always delivering value due to poor validity, accuracy or utility for students.

Experimentation with Flightpaths showing pupils’ current and anticipated progress trajectories had not provided a solution to these issues. Caroline Barlow, the school’s Head, and Tom Flower, the Deputy Head, therefore identified assessment as a priority area in the school’s improvement plan and decided to take targeted action on assessment by establishing an Assessment Innovation Team (AIT).

The AIT set out to tackle four main ‘pain points’:

1. The quality and consistency of the data underlying judgements about pupil progress
2. The consistency and utility of processes for capturing and reporting on assessment data
3. The extent to which assessment was driving effective feedback, and therefore supporting pupils’ learning
4. Teacher workload: work scrutiny had revealed overly burdensome assessment and feedback policies.

Different AIT ‘mini projects’ have focused on different combinations of pain points. For example, in biology, teacher and member of the AIT – Jess Gilespie – was keen to find techniques that would improve teaching and learning without increasing workload, because she felt that if an assessment innovation increased her workload then it would be more difficult to maintain and embed. Although Jess’s primary focus was to improve teaching and learning, she acknowledged that workload was an important consideration:

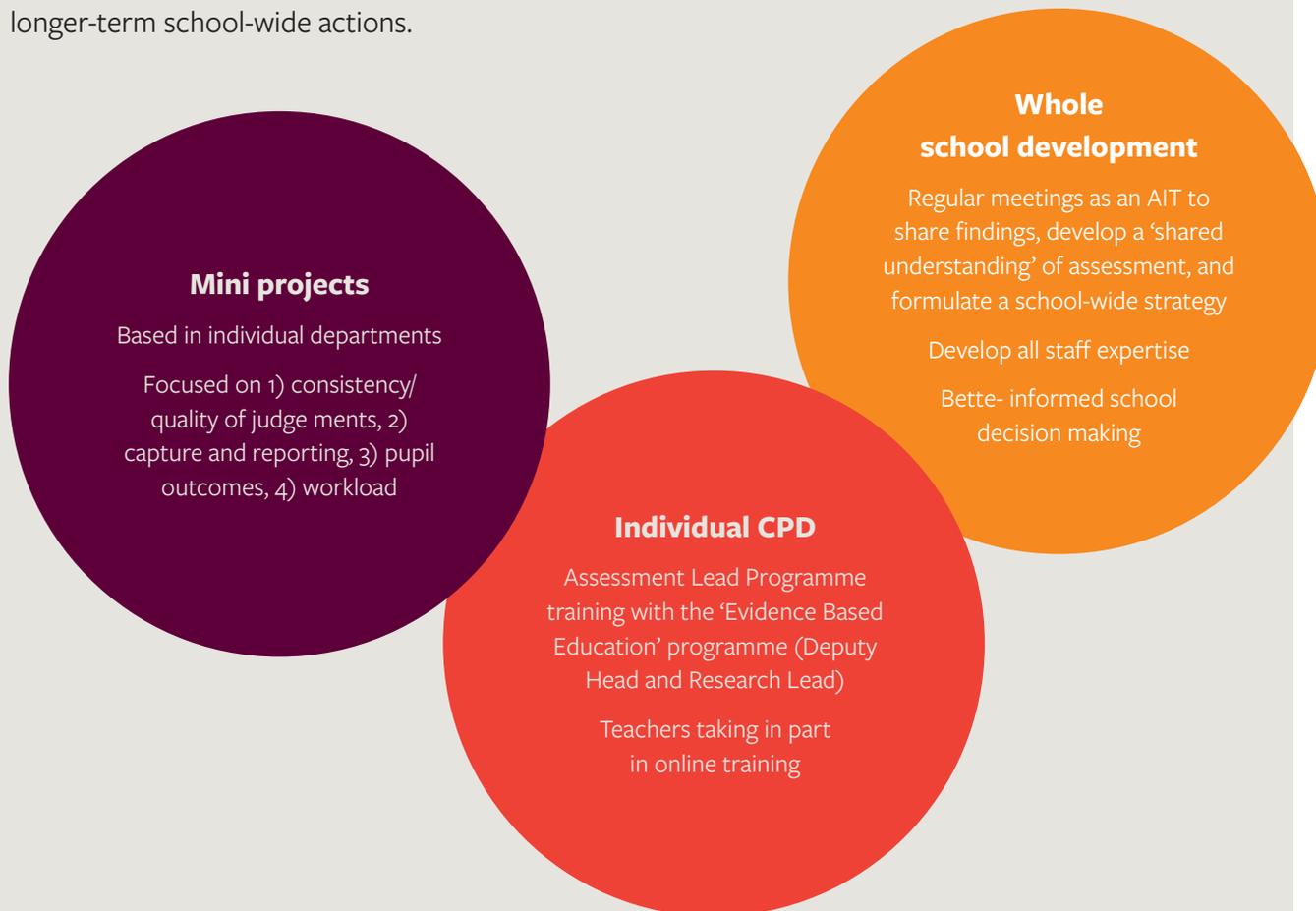
“[The innovation] would be something that I could embed very easily into my practice and not have to then spend any extra time doing anything. I think anything that is going to be effective, in the classroom, does need to consider that because you can’t maintain it otherwise.”

Jess Gilespie,
science mini project

3.1.3 Formulation

Heathfield has used Innovation Teams before to address other areas identified in the college improvement plan such as boys' achievement and literacy. Innovation Teams typically have a two to three year life-span. They function as action learning sets, with reviews of the evidence base followed by pilot work and a review, which then leads to the formulation of longer-term school-wide actions.

The overall aim of the AIT is to make sure there is sufficient capacity and expertise in the school to improve assessment and make it more effective. It is headed up by the Deputy Head and the school's research lead, and consists of:



The school's research lead drew on evidence from key thinkers like Daisy Christodoulou, Dylan Wiliam and Becky Allen to identify some principles for a more effective approach to assessment across the school, and these features of effective assessment practice fed into the design of the mini projects.

3.1.4 Decision making

Why and how was the decision made to go with the AIT?

Innovation Teams are an established structure for driving school improvement at Heathfield. Therefore, once assessment was identified as a priority area in the school's improvement plan, the Innovation Team approach seemed a logical, tried-and-trusted route to follow. While the decision to take this approach was centralised within the senior leadership team (SLT), the Innovation Team model involves a high degree of staff involvement, with teachers from across the school being invited to participate. For instance, individual members of staff were given a significant degree of autonomy over the formulation of their mini projects.

Mini projects

The AIT approach involves a series of different 'sub-innovations' within different subject areas.

Here, we explore the origins and development of some of those mini projects.

Photography: improving the consistency of teacher assessment

Agenda setting

- Results in photography not consistent between teachers
- Difficulty moderating in an arts subject
- Deputy Head wanted a creative subject represented on the AIT

Formulation

- Informed by blogs, photography teacher Darren's personal experience of 'what works' in verbal formative feedback and in sports coaching
- More formal approaches, inc. stamps to show "verbal feedback given," were unworkable
- Combined with using student exemplar work and templates

Decision making

- Darren's approach (student-centred, conversational) "seen to work" as he secures the best results
- Darren and his team jointly agreed the key areas for improvement in students' work, which formed the basis of the project
- The entire team shared a desire to standardise their assessment and feedback practices and were keen to be involved

Geography and history: retrieval/quizzing

Agenda setting

- Head of Department wanted assessment in geography to be more diagnostic and more closely tied to the curriculum
- Belief that good assessment will help disadvantaged pupils by improving working memory deficits

Formulation

- Based on evidence e.g. ResearchEd, Becky Allen to decide which techniques to try
- Also trialling other techniques experimentally e.g. rank order assessment

Decision making

- Head of Department is the school's Research Lead, and had started a programme of innovation in assesment before the AIT kicked in
- Students are aware they're part of a trial and understand why this is taking place

English: comparative judgement

Agenda setting

- English marking is subjective, department wanted to know that their marking was both accurate and useful
- Marking takes a lot of time, desire to reduce teacher workload
- There had been some cases where exams had not been marked accurately and the school appealed the grades given

Formulation

- Trialled different methods and read a range of research on assessment before settling on specific innovation
- Attended a seminar and trialled the 'no more marking' software and saw it as 'the future'

Decision making

- Tried using comparative judgement to mark mock papers and it was felt to have worked, but was time-consuming to 'clean' and prepare the data
- Decision was taken by the department to use comparative judgement in future, but not using the system/software. Instead to manually compare whole paper scripts in different pairwise combinations

Science: interleaving

Agenda setting

- New science GCSE course requires more depth of knowledge than previous course
- Department given marking template last year, resulting in confusion over what was being marked and for what purpose

Formulation

- Considered using retrieval practice (and is using this a little) but wanted to try something different to other members of the AIT
- Relevant and current research on interleaving in STEM (science, technology, engineering and maths) subjects was easier to find

Decision making

- A relatively junior teacher has been leading this innovation, with support from the school's Research Lead
- The teacher leading this innovation has been doing so relatively independently from the rest of their department
- The absence of any significant additional workload was a key deciding factor in selecting this innovation

3.1.5 Implementation

Implementing the AIT

The AIT was led by the Deputy Head and the school's research lead, who undertook the Assessment Lead Programme delivered by Evidence Based Education. The Assessment Lead Programme is a year-long programme of continuing professional development (CPD), which supports teachers to develop their assessment knowledge, skills and



confidence, apply this learning in their school, and lead other teachers to do the same.

Teachers at Heathfield were given an open invitation to join the AIT, which met fortnightly throughout the school year during timetabled hours. With support from the Deputy Head and Research Lead, teachers identified their mini projects and developed plans for evaluating and disseminating their work, for example through in-service training (INSET) days. They were also provided with online assessment training.

Access to expertise and support from senior staff within the AIT was seen as vital, particularly in the planning stage:

“If you can have discussions with other staff and people like Mark, who is a research lead, or people like Tom [Deputy Head] who are in those positions of expertise... then you can plan something that is doable, manageable and actually can really change the way that you do things.”

Jess Gillespie,
science mini project

The AIT's work concluded at the end of the academic year and the team's work now sits within individual departments to take forward.

Comparing innovations

We have focused on two mini projects in detail: using exemplar work and in-person feedback in photography; and interleaving homework in biology. The projects overlap in some areas. For instance, both have some degree of focus on raising attainment, and both appear to have had a positive impact on their intended outcomes. However, the projects differ in other ways. For instance, the biology project was driven by a single member of staff, while the photography project was delivered collaboratively by a team. Meanwhile, the biology project has been formally evaluated using an experimental research design, whereas outcomes of the photography project have been assessed on a more anecdotal basis.

The following table compares the two projects' implementation side-by-side. This illustrates the different ways assessment innovations can be implemented, even when they sit within the same overall programme within the same school. It also exemplifies how teachers can tailor innovations to their own priorities, expertise and subject areas, when they are given the freedom and support to do so.

Phase	Photography: using exemplar work and in-person feedback	Biology: interleaving homework
Initial implementation	<ul style="list-style-type: none"> ● Realised that creative work can still have a core of non-negotiable ‘quality markers’ ● Began by using the best student work to identify these non-negotiables for presentation (e.g. digital sketchbooks); what does it need to look like in order to get a Grade 9? ● Defined a set of criteria that teachers and students can agree on, but left room for creative ‘flex’ ● Developed resources/guidance for students based almost entirely on examples of excellent student work, and some of the teachers’ own work 	<ul style="list-style-type: none"> ● Designed as a trial in two parallel Year 9 classes ● One group has been set homework that is related to the topic they have just studied; the other group has homework based on content they covered two or three topics back ● Question bank software has allowed the teacher to set homework of equivalent difficulty, from the same exam board, even though they are on different topics
Evolution	<ul style="list-style-type: none"> ● Went on to create additional resources to address other areas of work ● Used the resources to cover common questions, reduce the amount of verbal feedback required, and to allow more time to be spent giving face-to-face feedback on questions that are specific to each student’s work 	<ul style="list-style-type: none"> ● Focus and approach has not changed since the outset
Facilitating factors	<ul style="list-style-type: none"> ● The whole department is collaborating on it ● The department was already quite far along in developing new schemes of work, so teachers had these schemes of work to work from ● Students have been supportive and a supportive Head of Department means there is commitment to the project 	<ul style="list-style-type: none"> ● Has not required any additional staff, time or resources. One teacher has implemented the project with some support from the AIT, particularly around evaluation ● Conducting this mini project under the banner of the AIT has given it more clout; staff are more likely to listen to the findings as a result
Barriers	<ul style="list-style-type: none"> ● It is a creative subject, and the departmental team developing the resources are mindful of the need to avoid them becoming too prescriptive 	<ul style="list-style-type: none"> ● Students were not initially very receptive, but this was predicted and has not proven to be a barrier ● Rolling out across the department in future might encounter barriers because it involves ‘doing things differently’ and some longer-serving teachers may not be keen to change their practice

There has been push-back against innovation from some quarters due to the difficulties of changing established practice. As the Deputy Head put it, “once you let the assessment genie out of the bottle, it’s hard to put it away”. In other words, old habits and received wisdom in assessment die hard and asking staff to revisit, or change, what they have done for a long time can be difficult. Despite this, overall the Deputy Head feels that staff have been open-minded about the AIT and happy to acknowledge gaps in their knowledge and practice.

In the English department, the Head of English experienced push-back against the work involved in rank-order assessment. Having trialled comparative judgement to mark mock papers, he and his team felt the approach increased the reliability and validity of their assessment judgements, but that these benefits were outweighed by the increased workload involved in systematically comparing a large volume of work samples, and preparing students’ work samples for analysis. The department decided on a compromise solution whereby whole student scripts would be compared, in different pairwise combinations without using software to compare individual excerpts of work. This retained the underlying logic of rank-order assessment, but applied the principles of the approach in a less labour-intensive way. The Head of English feels this

approach is still helping his team to deliver more reliable assessment judgements.

3.1.6 Evaluation

Focusing once again on the biology and photography innovations, numerous differences emerge in relation to evaluation. Broadly speaking, the biology project took a more structured, data-driven approach to evaluation. Meanwhile, evaluation of the photography project took into account teacher workload and student engagement as well as attainment, whereas the biology project primarily focused on raising attainment.

“If there was some sort of quantitative measure that we could look to, then that was something that we used as a key means of evaluating its impact. And other than that we were very happy with any qualitative evaluation comments, either by members of staff or by students.”

Tom Flower,
Deputy Head

Biology

The biology project was evaluated using a control-group design. There were two parallel

groups in Year 10 that were a close match in terms of ability and who were being taught the same schemes and lessons by the same teacher. One group was therefore taught using the innovation and one was used as a control group to assess the impact:

“I’m a scientist, so maybe it was just my natural instinct to think ‘well actually I do need something to compare this to’ but I was just in a really fortunate position, that I had a natural control in the sense that I had the two parallel groups.”

Jess Gillespie,
science mini project

Jess is keen to explore additional ways of measuring and evaluating impact and there is an assessment the students will sit at the end of the year, reflecting on the whole year’s teaching. She believes this will be a useful way to see what impact the innovation has had.

After seeing the positive impact that the innovation is having, Jess is considering introducing the innovation with the other class, no longer using them as a comparison group, exemplifying a common ethical consideration when running control-group trials:



“[We’re] not so keen to carry it [the trial] on because it was having such a positive impact on the other group. I thought actually knowing that it’s working, I don’t want to then, for another term, hold back that other class.”

Jess Gillespie,
science mini project

Photography

The photography innovation was assessed in a less formal manner, relying on continuous conversations with, and observation of, students. This involved photography teacher Darren Causton assessing pupils’ work “over the shoulder”. There are no standardised, academic tests in photography, so it is harder to assess students’ work using metrics. Furthermore, Darren’s discursive approach to feedback allows him to assess the quality of students’ work and give direct feedback on how it could be improved. It also allows him to “take the pulse of the room” – to see if students are enjoying the work and feeling motivated by the topic. Darren used this type of student feedback to shape the templates he put together to support students with their portfolios; allowing the resources to focus on the areas students found most challenging. Occasionally, Darren undertakes a more structured student

voice exercise, to gauge student feeling on a particular component of the course.

Last year, Darren and his team developed templates to help students with the research pages in their portfolios. They are now working on a guide to taking good photographs. Darren acknowledges that photography is a creative process but also feels that the process can be directed – there are some non-negotiables that all students have to grasp.

“It can be very difficult to get the marks under [the new] very specific guidelines. So I think what we’ve tried to do in a creative subject is apply a formula or a set of rules and some of those are non-negotiable which seem for a creative subject quite controlling but if you initially understand these non-negotiables and you apply this key formula then success is going to be a lot easier to come by.”

Darren Causton,
photography mini project

As well as acting as a teaching resource, this template will also act as an assessment resource, in the same way as the research guide.

Comparing approaches to evaluation and assessments of impact

The table below provides a summary of the different ways in which impact is being evaluated in the two mini projects.

	Photography	Biology
How is success being characterised?	<ul style="list-style-type: none"> Improved grades Allowing teachers to make more effective use of the time they spend giving feedback Developing students' love and enjoyment of photography: students feel positive about their work No specific indicators developed for each of these desired outcomes 	<ul style="list-style-type: none"> Greater-than-expected progress and attainment
What does evaluation look like?	<ul style="list-style-type: none"> The department is hoping to see the project's impact in three ways: <ol style="list-style-type: none"> Visual comparisons – does the work look better than it did before? Data – student progress and attainment, taken from formal assessment; for example, they compared students' marks on an existing coursework project (before the assessment initiative), with those on a recent exam project (after the initiative) Student confidence and independence in the creative process No formal plan for evaluation beyond "discussion and making notes" 	<ul style="list-style-type: none"> Designed as a trial in two parallel Year 9 classes Pupils from both classes sat a common baseline test on Key Stage 3 material After 1.5 terms, pupils sat a progress test – this was compared with their baseline test score to see distance travelled They then looked at predicted Key Stage 4 grades to see if their progress outstripped their predicted progress Study design controls for teacher, class size and predicted attainment Plan to add some student voice to the evaluation next year
Unexpected benefits	<ul style="list-style-type: none"> Some students report early and rapid gains Impact is being felt most strongly among middle-attaining pupils who might not have sought these resources out for themselves 	<ul style="list-style-type: none"> For Jess as a teacher, it has been an opportunity to reflect critically on her own practice, and to experiment with different ways of teaching
Downsides	<ul style="list-style-type: none"> Some students feel constrained/overly directed, but Darren feels they probably would anyway, due to the drive within the curriculum to revisit completed work and continuously improve it 	<ul style="list-style-type: none"> It has added somewhat to Jess's workload, and this has sometimes been difficult to sustain alongside her other responsibilities

Photography

How effective has it been?

- Judged “very successful” based on feedback and work seen
- Believed to have had positive impact on grades, students’ understanding of what they are being assessed on, teachers’ effectiveness, quality of lesson resources, students’ positivity about their work, and their grades
- Currently these claims are based on anecdotal evidence
- Awaiting good results in the form of ‘hard data’ on results day in the summer

Longer-term impact

- Resources are now written into next year’s schemes of work in order to embed the approach
- Good results will make the course more appealing/boost the department’s reputation
- The resources will make it easier for teachers to communicate what is required of students in terms of presentation and research
- In the long run, having these core resources, based on the mark scheme, will give teachers greater flexibility to support individual students in different ways. By covering basic common queries, the resources will allow teachers to focus their face-to-face feedback on more personal, individual elements of students’ work. This, in turn, supports Darren’s focus on getting to know individual students and tailoring support on the basis of that knowledge
- This approach has been transferred to other creative subjects, such as art and design

Biology

- The interleaved class have made considerably more progress than the control-group class in 1.5 terms

- Jess has shared what she is doing with her department, and more widely through whole-school collaborative professional learning (CPL)
- The mini project in biology has contributed to the AIT’s overall momentum by making teachers’ practice more research-informed:

Professional impact

The assessment innovations in photography and biology were both felt to have impacted on the professional practice of the teachers leading the innovations.

Biology

Taking part in the AIT has encouraged Jess to think about assessment as a tool for teaching, rather than a separate activity. Jess has built the principle of interleaving into her teaching, beyond assessment activities such as quizzes and marked homework. For instance, she is increasingly making use of starter activities in lessons that revisit material from previous weeks, terms and even years.

“I’m more reflective now on the purpose of it and the point of it. So if I’ve got, say, an assessment that I’m giving my Year 7s on a piece of work, that I’m going to mark and give them feedback on, I’m much more interested to know, actually is that a useful assessment? What am I actually getting from it and the feedback that I then give to them, is that actually useful?”

Jess Gillespie, science mini project

Jess has found that experimenting with one assessment technique – interleaving – has acted as a springboard for experimenting with other techniques, such as using ‘hinge’ questions. These are diagnostic tools used when students reach a “hinge” point where teachers need to know whether or not students have grasped a key concept before moving on to the next idea.

Students’ responses provide the teacher with valuable evidence about what their students know, don’t know and “need to do next” There is also some evidence that the innovation has had a spill-over effect on other members of the science department at Heathfield. Jess has presented her innovation’s results to colleagues in the department, and since then they have made further enquiries and started using interleaving when setting revision homework for their Year 11 classes. Next, Jess plans to share the approach with teachers from other schools:

“I’ve been asked to present on it at... INSETs, [where] we get all the local schools together. So I’ve been asked to do something on it to the science departments, so I guess [we’re] rolling it out really and getting the word out of how simple it is but how effective it seems.”

Jess Gillespie, science mini project

Photography

Darren feels that the innovation has encouraged him to think about how he gives feedback to students; rather than assessing students at the end of a topic and then supporting them with areas of weakness, Darren uses his templates to deal with problems pre-emptively. The process has reinforced in Darren’s mind that assessment is mainly for pupils’ benefit: assessment is the basis of robust feedback, and that feedback will help students to improve specific pieces of work for examination, but also help them to develop broader skills in their subject area. Overall, running the innovation has provided Darren with an opportunity to reflect critically on who assessment is for.

Despite the additional workload involved in creating the templates, they have allowed Darren to provide more structured feedback and to pre-empt common questions. This has, in turn, reduced his workload during lessons. This is particularly the case during the initial stages of the course, where the templates allow students to solve problems and develop their skills more independently.

“The workload lesson to lesson is definitely improved because you’ve given them a starter for what they need to do to make that work successful from the beginning... initially obviously creating these resources takes a little bit of time but in the long run, yes I think it’s had a positive impact on managing the lesson and the workload within the teaching period.”

Darren Causton,
photography mini project

Darren feels the resources have also had a positive impact on other teachers’ in the department’s workload. Making the assessment criteria and mark scheme clearer, and sharing this with students through exemplar work and templates, has reduced the amount of time teachers spend repeating common themes in their feedback. Teachers can instead focus their efforts on providing students with more individualised feedback.

Across the AIT

Beyond specific mini projects, it is clear that the wider Innovation Team-based approach to assessment innovation at Heathfield has yielded professional development benefits for staff. Jess, Darren and Tom all feel that

the AIT had encouraged staff to reflect critically with their own practice and that of their colleagues, and to consider changes to their practice based on research evidence.

“I think it’s made me read a lot more research, it’s made me engage with a lot more of what gurus, if you like, the people that know assessment, are saying.”

Jess Gillespie, *science mini project*

The AIT mini projects have enjoyed a high profile, are regularly communicated to all staff, and feature in school-wide CPL day. This is partly down to the fact that the AIT has been anchored in Heathfield’s school improvement plan, and that it benefited from the input of senior leaders,.

Pupil impact

Biology

The interleaved class have made considerably more progress than the control-group class in 1.5 terms, based on a mid-year summative assessment. Jess is looking forward to using end-of-year exams to assess impact on attainment after a year of teaching using the innovation.

“Improving outcomes for assessment just by simply tweaking something like homework was the desired solution, and that seems to so far have been the case.”

Jess Gillespie, *science mini project*

Photography

Students have historically achieved strong results in photography at Heathfield. Darren feels the new resources he has designed to structure his teaching, feedback and formative assessment have nonetheless contributed positively to students’ coursework and final grades, although this is not based on a formal evaluation of impact.

During the focus group we conducted, students explained how the formative assessment and feedback they receive in photography differs from what they experience in other subjects. They felt that Darren’s approach had two key components, which worked well together:

1. Giving initial, written feedback based on template work and mark schemes to identify areas for improvement
2. Following this up with more detailed verbal feedback, to help students interpret and understand the written feedback, and make improvements that fit with their own creative objectives.

Students felt the verbal component of Darren's approach made his feedback more detailed, descriptive and individualised. This allows them to respond to gaps or weaknesses in their work that would lose them marks in their final assessment, without sacrificing creative freedom. Crucially, Darren understands that the objective quality markers provided by a mark scheme and exemplar work need to be translated into a different action plan for each student, so their work remains personal to them. As one student explained:

"I feel like if he didn't have any connection with us and he didn't really know us as a person, it would be much harder for him to understand where we're coming from... he knows exactly what we want to do. It helps with the examples and ideas of other people's work."

Photography student

Replicability and transferability

Biology

Jess believes her approach to interleaving homework is easily replicable for different teachers and subject areas, as it simply involves reordering teaching and assessment

topics so that students revisit older topics. The availability of standardised question banks can help to reduce the workload involved in constructing new assessments.

Teachers would need to ensure they continue with the innovation for long enough to cover all topics and see the impact.

Jess also described how a school-wide structure such as the AIT helps in cascading learning from innovation to other teachers, and argues that it also provides high-level backing which encourages other teachers to listen and experiment.

The Head of the science department has been encouraging and supporting staff to experiment with the approach that Jess has piloted. She believes that this is a better approach than if she tried to drive the change herself:

"Some people have done it their way for a very long time and don't necessarily find themselves very open to changing what they do. So I think support in that sense, from the Head of Department, is just crucial to... get you across that awkward blurred line of 'this is something I'd like you to try' versus 'this is something we are all going to do:'"

Photography

Last year, Darren fed-back about the work he was doing in photography to other Heads of Department at Heathfield. Other staff were interested in his approach but it is not clear that other departments have implemented any of these approaches as a result. Darren voiced scepticism about the transferability of key elements of his approach, such as providing continuous 'over-the-shoulder' assessments:

"I always feel like we're in a really lucky position in the way that we can work with the students because it is very unusual now... I can assess students very quickly over the shoulder. I don't have to read an essay or go through the workings out and the answers as in maths; everything takes a bit longer in those kind of subjects."

Darren Causton,
photography mini project

Despite this, Darren feels that a cornerstone of his approach – delivering feedback that is personalised, and based on an understanding of each individual student – should be transferable to all subject areas. Meanwhile, his approach to using templates has now been transferred from photography to

art and design, with Darren supporting colleagues to adapt the system and resources that have worked well in his classroom.

Across the AIT

Within the AIT more widely, teachers running mini projects have shared their work with staff across the school in meetings and colleagues have shown interest in their findings. The school supports staff to write blog posts and tips that are shared across the wider team. The approach seems to be softer than a school-wide rollout of individual innovations, instead allowing teachers to access and take forward ideas they think are relevant.

This dissemination of learning has had success, with the Deputy Head reporting that, across the school, staff knowledge of assessment has improved, resulting in more consistent practice:

“It was the case until recently that not necessarily all the staff were using the same assessments, which is problematic if you’re trying to compare how students have done. So I think we’ve got better at that.”

Tom Flower, Deputy Head

While the AIT has now concluded its work, assessment-related objectives remain within the school development plan:

“It’s absolutely embedded in our improvement plan every year now, that we’ve got something relating to improving assessment that is there every year. So I think that’s its legacy really.”

Tom Flower, Deputy Head

Teachers who did not take part in mini projects as part of the work of the AIT are therefore nonetheless experimenting with new assessment approaches, overseen by senior leaders. Tom is currently working with a group of 20 staff from different curriculum areas to develop their understanding of diagnostic questioning, and to experiment with hinge questioning.

Next steps

Jess and Darren identified four main recommendations for teachers considering piloting an assessment innovation:

1. Plan carefully, making sure your innovation will not impact on your workload, as this makes it easier to maintain.

2. Ensure your innovation has SLT backing and that they:
 - a. have put structures in place to support you, for instance with research and evaluation expertise, providing additional non-teaching time to plan and evaluate, and establishing forums to discuss and share practice with other teachers
 - b. see your innovation as part of a wider strategic focus and are willing to give the innovation adequate time to bed in and achieve results, for example by mapping it to a theme in the school’s development plan.
3. Discuss your work with other teachers so you do not miss out on opportunities to share your experiences and receive supportive feedback.
4. ‘Innovation’ does not have to mean a radical departure from existing practice; it might just mean focusing more closely on what is already working well, codifying practice so others can adopt it, and establishing systems for teachers to share and discuss what they’re doing.

The Deputy Head also provided five main pieces of advice to other school leaders:

1. Make sure your own knowledge of assessment is solid first; get the sequencing of training right.

2. Use the SLT's solid knowledge of assessment to audit where assessment practice is strongest and weakest in the school.
3. Give innovation the time it needs to bear fruit and for impact to be evidenced meaningfully.
4. Be proportionate when you are evaluating (for example, make use of qualitative as well as quantitative evidence).
5. Be aware of your biases and be open-minded in your expectations of particular innovations or practices.



3.2 Innovation 2: Isaac Physics at Rickmansworth School

3.2.1 Introducing the innovation

Rickmansworth School is a secondary school for 11- to 18-year-olds in Hertfordshire. It is a partially selective academy, admitting a small proportion of pupils based on high performance in music at the age of 11.

The 'Isaac Physics assessment programme' is an online assessment tool available to UK schools and pupils. It assesses pupils' performance in physics. The tool stores thousands of questions related to different Key Stages and topics in physics. Pupils complete questions and the programme generates details about their performance and progress.

Rickmansworth School is using the tool with GCSE and A-Level pupils in order to improve pupil performance in physics and to reduce teacher workload. It is being used to:

- help teachers assess GCSE and A-Level pupils' progress throughout the academic year

- set pupils' homework and revision tasks
- encourage pupils to take responsibility for their own revision.

Pupils are given a 'block of questions' to complete and teachers monitor their progress.

"I keep a continual eye on the information that I get from the website about how many questions they've attempted, and I use this as a way of just keeping an eye on how much they're doing in their own time, independent learning at home, and also as a way of seeing whether or not they've had particular issues with a certain style of question."

Head of Physics

3.2.2 Agenda setting

What were/are the 'pain points' that this innovation sought to respond to?

The Head of Physics at the school introduced the idea for the innovation shortly after joining the school. He has a considerable degree of autonomy in the school and senior leaders have had limited involvement. He had previous experience with 'Isaac Physics'

and felt that the assessment tool would be a useful way of responding to three ‘pain points’ he recognised in the department:

1. Teacher absence and teacher recruitment problems
2. The need to improve pupil outcomes
3. Lack of parental engagement.

Teacher shortages

1. Amidst a national shortage of science teachers with a physics specialism, the school was struggling to recruit specialist staff in the subject. Furthermore, a number of physics teachers had been on long-term sick leave because of ill-health, including the Head of Physics himself. This means that some staff teaching physics were supply teachers or not subject specialists. Therefore, the Head of Physics decided to use the online assessment tool to help non-subject specialists to assess pupils and improve the consistency of assessment. It also allows him to review pupils’ assessment data in periods where he is absent from school due to ill-health.

“The main thing was the lack of stability in the teaching staff, because there were lots of changes, you would have a lot of people coming in, and you wouldn’t know what assessment had taken place, you wouldn’t know how well the students had done. You had no real overview. So, from leading a department point of view, I just needed to get a handle on good feedback.”

Head of Physics

Pupil performance

2. ‘Isaac Physics’ also responds to pupil underperformance. The Head of Physics explained that some pupils “did not do well on their end-of-year exams”. The online assessment tool is therefore used as a revision programme for underperforming pupils to complete over the summer holidays.

Parental engagement

3. Although parental engagement was not problematic before the innovation began, the Head of Physics wanted to harness the power of parental engagement to enhance pupils’ progress and their outcomes. ‘Isaac Physics’ allows parents to access data about their child’s performance and helps

them to understand what topics pupils are studying and what they are struggling with. This helps them to understand how they can support their children with physics.

3.2.3 Formulation

How was the solution arrived at?

The solution was developed in two stages:

1. **The Head of Physics assessed the areas for improvement within the department when he joined the school.** He then drew on his previous experience of using ‘Isaac Physics’ to develop a plan about how this could be used to respond to the ‘pain points’ he had identified.
2. **The Head of Physics attended a course run by the developers of ‘Isaac Physics’.** This course helped him to familiarise himself with the tool and design a plan for how it could be used and implemented to tackle issues in physics within the school.

Who and what fed into shaping it?

The Head of Physics shaped, and continues to shape how ‘Isaac Physics’ is used in the department and the training he has received has helped him make decisions. On the ‘Isaac Physics’ course he met the platform designers and other teachers using the

tool and informal discussions with these people helped him shape the innovation:

“I went on a week-long course and became, if you like, au fait with all the various aspects and also meeting the people behind it was very useful and that’s one of the things that I think about having remote assessment tools or learning platforms is knowing that you’re going to have good communication with the people behind it so you can query if there’s problems, if there’s solutions you don’t understand, if there’s new things you want to add.”

Head of Physics

New members of staff have recently joined the department. Although the agenda is led by the Head of Physics, new teachers are starting to shape the way the tool is used. For example, one physics teacher has suggested using the tool in competitions between students to improve enthusiasm for the subject. While this plan has not yet come to fruition, the Head of Physics is taking teachers’ feedback into account while thinking about the future of ‘Isaac Physics’ in the department.

3.2.4 Decision making

Who was involved in deciding on this approach to addressing the pain points?

The Head of Physics was the driving force in deciding to use the ‘Isaac Physics’ approach. Other teachers in the department support using the tool, largely because of the tool’s reputation. As one teacher explained:

“The reason why we found it so useful is because it not only offers the students instant feedback, but more importantly I think it’s quite challenging. Other online platforms that we’ve tried in the past often have been quite simple and ask quite closed questions, but the opportunity within physics and with chemistry to a certain extent as well, is the mathematical side, and the challenge that that provides to our quite high-achieving students is actually really, really valuable.”

Head of Physics

The Head of Physics did not explore other solutions. However, the science department has previously used other online

assessment tools and rejected them because questions were not challenging enough.

3.2.5 Implementation

What mechanisms have been used to implement the approach?

Implementation took place in four steps:

1. **Attending a training course:** The Head of Physics had detailed conversations with the platform’s designers, and other teachers on the course, to learn more about the platform’s features and how to use it most effectively.
2. **Trialling the platform with sixth-form classes:** The Head of Physics piloted the programme with his sixth-form classes because the platform is relevant to the A-Level course.
3. **Rolling out the platform to Year 10 and Year 11:** After seeing ‘Isaac Physics’ work well with sixth-formers, the Head of Department introduced his Year 10s and 11s to the platform.
4. **Introducing the platform to other staff:** The Head of Department gave short training sessions to other staff. He demonstrated how to use the platform and asked teachers

to complete questions as if they were students. He believes this helped them to become familiar with the platform.

Not all staff have been using the platform consistently. For example, some new members of staff have not been collecting data on student progress in the same way. The Head of Department is therefore offering additional training to these staff. He is also having informal conversations with staff to understand any hurdles they are facing.

To what extent is the reified solution consistent with the planned solution?

The implemented approach is largely consistent with the intended plan. GCSE and A-Level pupils complete questions and receive feedback from 'Isaac Physics'. Teachers also find the tool useful for assessing pupils.

However, some pupils cheat when answering online questions. Other pupils simply answer a set of questions the day before a deadline rather than more frequently as was initially intended. This limits the innovation's ability to support pupils' progress.

Some pupils also find that the program's feedback is not helpful and does not help them to progress. As one pupil explained:

"They have hints and tips on each question, but for some questions, it might just be 'Oh no, I've used the wrong equation' or something like that, where it would have been really useful just to be like, 'Have you used this?' If you get it wrong after a certain amount of tries, it stops you from just redoing it until you get the right answer."

Pupil

Furthermore, as noted above, not all staff have been using the platform consistently and some new members of staff have not been collecting data on student progress in the same way that the Head of Department does. The Head of Department is therefore having informal conversations with staff to understand their experience of using the platform so that he can adapt and add to training. However, teachers have not highlighted any particular difficulties.

What conditions contributed to or detracted from effective implementation?

Training for the Head of Physics contributed to effective implementation because familiarity meant he could plan how to implement 'Isaac Physics' in his own department. The opportunity to meet other professionals using the platform was also valuable and he has remained part of a community of 'Isaac Physics'

teachers who use a WhatsApp group to share ideas about how to use the platform effectively:

"We have a little WhatsApp group and occasionally somebody will say, 'How did you get the answer to this one?' and we will have a conversation about the best approach to the question... it's a great learning forum for us as well."

Head of Physics

Having new members of staff join the physics department has made implementation slightly more challenging because these staff need to be trained up and monitored to ensure they are using the platform consistently.

3.2.6 Evaluation

How is success or failure being characterised?

The physics department currently define success as:

- pupils completing assessments as revision
- pupils making progress in physics.

What approaches are being used to evaluate the innovation's success?

Teachers of physics, and the Head of Physics, are evaluating the innovation's success using

quantitative data from the 'Isaac Physics' program. They are triangulating this with pupil progress data and informal comments from pupils. Teachers are using this to understand:

- how many questions pupils have attempted to answer
- which topic areas pupils are struggling with
- areas where pupils are making progress
- whether or not success in GCSE mock exams correlates with answering a large number of 'Isaac Physics' questions.

The Head of Physics, and other teachers in the department, will use this information to determine how much teaching and learning time they should dedicate to 'Isaac Physics'. This will feed into the department's strategy for the next academic year as the Head of Physics explained:

"It's sort of a cost-benefit analysis of how much time and effort we devote to it, and what we feel the returns are. So, we feel it's giving us positive returns at the moment, but then we've got to decide how much time we actually devote."

Head of Physics

3.3 Innovation 3: ImpactEd at Bengeworth CE Academy

3.3.1 Introducing the innovation

Bengeworth CE Academy is an outstanding three-form entry primary school in Worcestershire, with just over 500 pupils. It took over a second school in the area in 2018 so is now part of a small MAT. Approximately 15% of pupils at Bengeworth have a special educational need (SEN) and around 30% speak English as an additional language. The high proportion of pupils with additional needs at the school has led Bengeworth to focus on creating an inclusive culture.

Although interventions also provide academic support for pupils, Bengeworth's programme of interventions places great emphasis on nurture and support with non-cognitive skills. The leadership team are keen for their work to be evidence-informed and scrutinised for effectiveness. Rachel, Bengeworth's Head

of Inclusion, explained the origins of this focus on inclusion and evidence:

"I think that since 2014 when the code of practice was brought in and new recommendations were made, there has been a real change and a shift towards partnership work and creating a truly inclusive culture. We've done a huge amount for inclusion. I've worked nationally, with the local authority, developed myself as a local leader of SEND... Historically, interventions wouldn't have necessarily been recorded in a systematic, robust way... there was a need to move towards evidence-based ones, that showed clear outcomes for pupils."

The innovation we studied involves Bengeworth's partnership with the organisation ImpactEd. ImpactEd provides schools with a data platform, which allows schools to input progress data and make comparisons across pupils and across other schools. ImpactEd provides schools with information about evidence on different interventions they might wish to use while building the evidence base for interventions using data from across schools.

ImpactEd offers a variety of scales and validated tools to measure progress against different outcomes. Bengeworth has worked with ImpactEd to build the collection of tools on the platform that measure non-cognitive skills such as metacognition, self-awareness and resilience. These include both validated tools and Bengeworth's own tools relating to tailor-made interventions.

Bengeworth's use of ImpactEd originally focused on measuring the impact of non-academic interventions, and this study focuses on this element, although as it evolved, Bengeworth began using the platform to record children's progress in all intervention programmes. The school aims to use ImpactEd to identify where and why interventions are or are not working so it can make evidence-based decisions.

3.3.2 Agenda setting

What were/are the 'pain points' that this innovation sought to address?

Staff at Bengeworth feel that children who are struggling need support to develop the non-cognitive skills that then allow them to access learning and make progress in academic areas, as Rachel, the inclusion lead, explains:

"If you don't get nurture-based interventions correct, then there's no point in doing any cognitive-based interventions to improve academic outcomes because... if a child isn't mentally healthy then they are not able to take on new learning."

Therefore, leaders felt that a focus on evidencing non-academic interventions had to come before, or at least be developed in tandem with, evidencing academic interventions. The decision to use ImpactEd to measure the progress pupils make in non-academic interventions was borne of a number of pain points, primarily:

- a lack of evidence (or lack of accessible evidence) about which interventions develop pupils' non-cognitive skills, resulting in decisions about interventions being made without an evidence base,
- high staff workload and inefficient systems for recording evidence.

Leaders felt there was a lack of evidence about which non-academic interventions were effective, both in their school and across the sector. The potential impact of this lack of evidence was threefold:

1. Ineffective interventions continued for too long.

2. It was more difficult to secure funds for and resource effective interventions.
3. Teachers and classroom support professionals (CSPs) were not able to choose and implement interventions based on evidence, leading to low teacher confidence and a lack of accountability for the effectiveness of non-cognitive interventions.

Leaders and staff commented on a tendency in primary schools to allow interventions that are 'well intentioned' but ineffective to continue. However, they felt that this culture needed to change and that a move towards robust, measured programmes that support better outcomes was required. Leaders felt that Bengeworth should lead that change.

Bengeworth staff acknowledge that the system they have used in the past meant that interventions' effectiveness was not scrutinised sufficiently. This old system involved rating interventions as red, amber or green by making notes on each child's progress in a paper file and manually inputting data into Excel in order to carry out analysis and discussing progress during half-termly pupil progress meetings. As a result, some interventions that were not effective were left to go on too long.

On the other hand, where interventions are effective, it can be difficult to secure

funding and resources without sufficient evidence of their impact. This was a particular problem for interventions with non-academic outcomes, even though these were sometimes transformative for pupils with very high levels of need.

Previously, decisions about interventions were made based on ‘the status quo’ or a specific staff member’s knowledge or experience, rather than on the evidence of what was most likely to be effective for the pupil. The Head of Inclusion anticipated that using ImpactEd would make teachers more confident in using assessment data and evidence to select interventions, and would make teachers more accountable for the interventions they choose for pupils:

“My vision is that it would improve teacher confidence in terms of making them more accountable for the measures that they’re putting in... and monitoring it more rigorously. From my perspective, we’re already doing that as a school but it’s the leadership that are doing it... and we’re filtering that information back down to staff and empowering them.”

Additionally, although the previous system of recording interventions was ineffective, it nonetheless involved a high

workload for teachers and CSPs. This pain point was therefore an important driver behind the intervention.

Why was it considered an important agenda to pursue?

The school’s inclusive culture and values partly drove the school’s decision to adopt ImpactEd and this culture was, in turn, driven by the high proportion of Bengeworth’s pupils who have additional needs and Ofsted’s increased focus on inclusion.

“[Previously,] I don’t think the focus was on inclusion as it is now in the national agenda. I think it’s really become more pivotal... inclusion drives the agenda of the school, its CPD, its vision and its values should be inclusive... That’s become high importance in Ofsted and other branches of national agendas.”

Rachel, Head of Inclusion

Inclusive practice requires interventions that support pupils in developing their non-academic skills (as well as academic skills) and the leadership team felt that ImpactEd would allow teachers and CSPs to make intervention decisions based on evidence, including both evidence of the

previous impact of the intervention in the school and the wider evidence base.

Prioritising value for money and evidence was also linked to the climate of reduced funding as this has affected which interventions continue and also shaped the school’s choice of system. The use of other systems was stopped due to their cost and ineffectiveness; however, Rachel, the inclusion lead, argued for an investment in ImpactEd, believing it would save money by ensuring that only effective interventions were continued.

3.3.3 Formulation

How was a solution arrived at?

The senior leadership team at Bengeworth first heard about ImpactEd through a Challenge Partners conference. They were already aware of the workload associated with the existing intervention tracking system but had not yet found a solution. Bengeworth’s leaders feel that partnerships such as Challenge Partners have been useful in pointing the way towards ‘tested good practice’.

Who and what fed into shaping it?

Bengeworth became one of 10 Challenge Partner schools that formed a ‘co-design group’ for ImpactEd. These schools were very early adopters of the platform and

therefore fed into shaping ImpactEd's practice and offer to schools.

Rachel, Bengeworth's Head of Inclusion, was key in driving the innovation forward during the first stage of implementation. Over the 2018–19 academic year, Bengeworth piloted ImpactEd with some academic and non-academic interventions before deciding to roll it out across the school with a focus on non-academic interventions in 2019–20.

As it was part of the small group of schools that received early access to the platform to help shape its development, Bengeworth received considerable tailored support from ImpactEd during its pilot phase. ImpactEd acted on Bengeworth's suggested updates and changes to the platform. Although this piloting phase may not be needed by all schools as the innovation is scaled, Bengeworth's experience highlighted the importance of a tailored approach. Bengeworth's leaders highlighted that without ImpactEd's support they would not have been able to make the innovation work for them.

During the pilot phase, Natalie Snowdon, Bengeworth's data and assessment lead, shaped the innovation by working with ImpactEd to decide which interventions and measures would be included on the platform. This involved trialling different impact

measurement tools, some of which were found to be ineffective or inappropriate for Bengeworth's in-house interventions. These measures were then removed and Natalie worked with ImpactEd to identify different validated measures or Bengeworth-made measures to better fit intervention outcomes.

3.3.4 Decision making

Who was involved in deciding on this approach to addressing the pain points?

The decision to use ImpactEd was primarily made by the senior leadership team. The Head of Inclusion then drove the innovation forward throughout the pilot phase, with support from the Headteacher, the data and assessment lead and the senior leadership team.

As the innovation was rolled out across the school in the autumn term of 2019, all staff contributed to the decision to expand the use of ImpactEd to track the impact of all interventions, both academic and non-academic.

Why were other solutions rejected and this approach preferred?

Other approaches to assessing interventions' impact on non-academic skills, beyond the systems in place beforehand, were not considered. The previous systems involved

a combination of tracking on the School Information Management System (SIMS) and RAG-rating interventions, but this was not felt to be working effectively.

Leaders felt that sourcing solutions through collaborations such as Challenge Partners meant that they could be more efficient by choosing solutions that partner schools have already found to be effective. Also, given that Bengeworth was one of the co-design pilot schools for ImpactEd, the school has concentrated on shaping and refining the tool rather than comparing it to other potential solutions.



3.3.5 Implementation

What mechanisms have been used to implement the approach?

Bengeworth has implemented this innovation in four stages:

1

Preparing the platform

When?

2018-19

Who?

Head of Inclusion and Data and Assessment Lead

What?

Working alongside ImpactEd to input interventions and corresponding measurement tools and scales onto the platform

2

Small-scale pilot

When?

Spring and summer term 2019

Who?

Head of Inclusion and Data and Assessment Lead

What?

Trial of the platform to assess a few interventions on a small scale

3

Refining the approach and preparing for rollout

When?

Summer term 2019

Who?

Head of Inclusion and Data and Assessment lead

What?

Adjusted measurements and scales to fit Bengeworth's interventions. Prepared whole-staff training

4

Rolling out the platform to all staff

When?

Autumn term 2019

Who?

All teachers, followed by all CSPs

What?

All staff received training on the platform form either ImpactEd or the Head of Inclusion. All staff then used the platform to assess impact of some interventions

Who are the different actors involved?

In the first stage, the main people involved in the innovation were the leadership team, primarily Rachel (the Head of Inclusion) and Natalie (the data and assessment lead) as they worked with Ella Knight, Partnerships Manager at ImpactEd to ready the platform and the tools for rollout across the school.

The next phase of implementation – initial rollout – involved all teachers and some lead CSPs who were trained to use ImpactEd. They began to use the platform to input data on interventions in the autumn term in 2019. Towards the end of the autumn term, the lead CSPs began to train other CSPs to use the platform. This stage of the rollout was key in reducing the duplication of systems whereby CSPs were using some of the old tracking systems, such as RAG-rating, while teachers were inputting data onto ImpactEd. Given that most interventions are delivered by CSPs, it was felt that to truly embed the system, they would need to use ImpactEd and take some responsibility for tracking data, though ultimately the responsibility for monitoring the effectiveness of interventions and making decisions will remain with teachers.

To what extent is the reified solution consistent with the planned solution?

In terms of the platform and the implementation plan, the reified solution was consistent with the intended solution throughout planning, piloting and implementation. However, there were two differences observed between the initial plan during the set-up phase and the reality of implementation:

1. The platform was used to measure the impact of both non-cognitive and academic-focused outcomes, rather than the original plan of mainly focusing on the impact of non-cognitive interventions.
2. Many measurement tools that have been used were developed in-house by Bengeworth rather than primarily using validated tools and scales.

i) Measuring the impact of academic and non-academic interventions

The decision to broaden the focus from primarily measuring the impact of non-cognitive interventions, to include academic interventions, was taken for two main reasons: first, the school felt there was a need for greater consistency; and second, it was felt to be more straightforward to measure impact on academic outcomes.

This has created greater consistency and may have increased buy-in from staff as they prefer using a single system. However, it may also have reduced the focus on non-cognitive outcomes. It also highlights the difficulty of measuring progress in non-cognitive skills.

ii) Using tailor-made measures

While ImpactEd is designed to allow schools to use validated measures, many, though not all, of the tools the school is using have been developed by the school alongside their own interventions.

This flexibility and the responsiveness of ImpactEd in uploading new measures are viewed positively by leaders and staff. Staff also felt strongly that use of tailor-made interventions (especially in lower Key Stages) was a crucial element of good practice.

However, this change also hampers one of ImpactEd's main aims: allowing schools to compare the impact of their interventions with nationwide datasets. During the initial visits to Bengeworth, the Head of Inclusion highlighted the need to base intervention decisions on evidence and compare Bengeworth's outcomes with those seen in other schools as a major benefit of ImpactEd. This will not be possible if the interventions and measures are school specific. In some

senses this suggests that ImpactEd may be being used with less rigour than intended, as wider datasets cannot inform decision making.

On the other hand, ImpactEd is providing an opportunity for the school to monitor the impact of its interventions more closely when there is no prior evidence regarding their effectiveness (unlike some bought-in interventions that have an established evidence base). Therefore, using ImpactEd to some extent ensures decisions are becoming more rigorous. There was also evidence that this approach to using ImpactEd had fed into decisions:

“There are really two avenues you can go down with this. We’ll look at this data and in terms of the children that were in this top-dog intervention, for those children it’s made a difference to their confidence in class... so we are just going to monitor them in class now. But for this child, this hasn’t worked, perhaps it wasn’t well suited, it wasn’t right, so what is better? What will work? And we will give them that next. All the while, it gives the overview of ‘Generally is this intervention really doing what we wanted it to do when we designed it?’”

Teacher

Teachers explained that while they were, to an extent, using measures and tracking pupils’ progress in interventions before using ImpactEd, the time taken with data entry and analysis meant far less time was available for evaluative conversations and decisions about pupils’ next steps. Therefore, they felt ImpactEd was allowing them to scrutinise Bengeworth’s tailor-made interventions more closely and make better-informed decisions for pupils.

It is also worth noting that the process of adding new measures exemplifies ImpactEd’s intense and ongoing involvement in helping Bengeworth set up the platform. This was a crucial facilitating factor behind success. Bengeworth’s leaders praised ImpactEd’s responsiveness: often ImpactEd would respond to Rachel’s request for additions to be made to the platform within a day. At the start of the study, ImpactEd had not decided whether it would provide this type of intensive support to all schools in future; however, findings from this study and of ImpactEd’s own review of the pilot demonstrate that it is a key element of success. Therefore, by the end of this study ImpactEd had decided to put partnership at the heart of its model and continue with an approach that tailors support to schools. In future, there may be some challenges to scaling this level of support but ImpactEd recognises the importance of this part of its model.

What obstacles had to be overcome as part of implementation?

Over the course of the pilot year and the whole-school rollout, Bengeworth faced several obstacles in setting up ImpactEd. However, most were resolved with relative ease. For example, identifying problems with the platform or the tool or needing to upload a new tool could have represented potential barriers to success; however, changes were quickly made and were therefore no longer considered obstacles.

In the school year 2018–19, during a review of funding and systems, the leadership team considered discontinuing their use of ImpactEd, rather than rolling it out. This decision would have been entirely driven by funding constraints rather than issues with the innovation. However, Rachel made a strong case for continuing with the innovation and the school decided to invest in the system in the long term.

During the set-up phase, the leadership team experienced some obstacles, mainly relating to time. During the implementation phase, the main obstacles were technical.

Setting up interventions and tools on the platform was time-consuming. Although the leadership team responsible for the set-up

managed to get the platform ready for the September 2019 rollout, this issue highlights that any school wishing to use the innovation needs to provide a considerable upfront time investment from leaders to secure long-run reductions in workload. Yet this does seem to pay off. Whilst in September 2019 teachers reported having to spend time getting to grips with the platform, as early as December 2019, they began to experience a reduction in data-entry and they anticipated that this would continue.

ImpactEd believes platform development since the pilot started means set-up is now a smoother and more straightforward process. However, an initial time investment is still needed to save time in the long term.

During training in September 2019, staff experienced some technical issues. The training was delivered via video link by ImpactEd; however, staff experiencing technical issues or confusion were not able to benefit from the training because they were not able to ask questions. Rachel addressed this by

delivering in-house training which was felt to be more effective. Though this was an in-school technical issue, it demonstrates the potential challenges of online training.

While this obstacle was overcome, a leader with a good understanding of the platform and capacity to deliver training was needed. If other schools wish to roll out the innovation without a long pilot phase, they will not have this in-house expertise and ImpactEd must consider how to deliver effective training for those schools.

What conditions contributed to or detracted from effective implementation?

During the pilot phase the main factors that contributed to effective set-up were:

- **leaders having dedicated time** to prepare the innovation to be rolled out across the school – although time was limited, Rachel and Natalie (the data and assessment lead) were given time to develop the platform
- **having on-hand support from ImpactEd** to shape the platform according to the school's needs – this continued to be a key factor of success during rollout
- **having a member of the senior leadership team leading the innovation.**



As the innovation was implemented and ImpactEd was rolled out across the school, other key factors for success emerged:

- **Ongoing and responsive support from ImpactEd was**, as in the initial phase, highly important. On a number of occasions, Rachel and Natalie asked ImpactEd to make changes to the platform in response to staff feedback. In December 2019, Rachel reflected on how the rollout could have gone differently had this support not been available:

“It only takes a few times for people to have problems that aren’t dealt with for them to lose interest. Teachers have a lot to think about and if something doesn’t work for them it’s difficult to motivate them to use it.”

- **The availability of an in-school expert** to deliver training and support to staff was crucial. Rachel and Natalie’s understanding of the system, having used it for the whole of the previous year, allowed them to support staff. Rachel created guidebooks for staff, which also ensured that new staff in the future would be able to use the platform. Leaders highlighted that the

Special Educational Needs Coordinator (SENCO) was the ideal person to do this.

- **A culture of innovation and embracing change** was also seen as helpful, though not necessarily crucial, to ensuring success. There was some disagreement among different staff members as to whether ‘any school could do it’ or whether a culture of innovation, especially when it comes to new technology, was needed in order to implement ImpactEd successfully.
- **Strong leadership and a strategic vision for the innovation** was identified as a key factor in ensuring the innovation was used consistently. Natalie, the data and assessment lead, felt that the system would be vulnerable to under-use and could be treated simply as a data-entry platform rather than as a way of comparing school impact to benchmarks and national comparisons and as a way of examining the evidence base to support intervention decisions. She felt that the set-up and implementation of the innovation had been successful because leaders had identified the area of need, wanted to solve the issue without increasing workload and had driven the change forward throughout:

“When you introduce anything new to staff, with teacher workload being high on the agenda and the Ofsted framework, you are mindful of that... particularly with data systems, there can be that preconception that it’s going to be hard and really difficult, especially with IT. But it hasn’t been like that because of the way it was delivered and because they can see the impact of it already because we are getting data.”

Rachel and Natalie also highlighted that further development of ImpactEd is needed to ensure long-term use and success, specifically with regard to provision mapping. Currently, ImpactEd does not have a feature to report a summary of all interventions that a single child, over the course of their time at school, has received. Schools are required to produce provision maps for children with special educational needs, which summarise this support and its outcomes. Natalie explained:

“Now we are using the program we do find we want it to do a little bit more... to show these are all the interventions a single child has had, these are the results, what is the impact.”

3.3.6 Evaluation

How is success or failure being characterised?

The intention behind using ImpactEd is to reduce teacher workload and improve the evidence used to make decisions about interventions. Originally, leaders envisioned that achieving these goals would ultimately lead to pupils developing their non-cognitive skills through effective interventions, as The Head of Inclusion explained:

“Success will look like, every pupil is making the most progress that they can and accessing the correct intervention at the right time, ensuring better outcomes and that will be being closely monitored. It won’t look like children have been in interventions that have made no real impact to their performance, whether that be pastorally or academically.”

This vision was then extended to include improving pupil progress in academic interventions too as the school extended its use of the platform.

During the implementation phase, Bengeworth’s senior leadership team used

the following three key criteria to assess whether the rollout was successful:

1. All staff understand ImpactEd’s purpose and potential benefits.
2. All staff involved in running interventions input pupil data into the system consistently.
3. All staff involved in making decisions about interventions use ImpactEd data to do so.

Leaders also anticipated that once all interventions were on the platform, if staff met the above criteria, the innovation would result in reduced workload, despite an initial increase in workload as teachers familiarised themselves with a new system. The previously used process of RAG-rating interventions – making notes on each child’s progress in a paper file and manually inputting data into Excel in order to carry out analysis – was extremely time-consuming. Leaders felt ImpactEd would save time, particularly with data entry and analysis.

In addition, as part of the initial plan to focus on non-cognitive interventions, one success criterion was that staff would view these interventions and the need to evidence their impact as on a ‘par’ with academic interventions.

What approaches are being used to evaluate the innovation’s success?

Before implementation, although Bengeworth’s leadership team had a clear plan to train and support all staff, as well as clear objectives in terms of what would constitute success during the implementation phase, they did not plan to formally evaluate the innovation’s success beyond gathering staff feedback and reviewing progress. However, collecting staff feedback appeared to be an effective way of monitoring the innovation during implementation. Feedback from staff during our research interviews was consistent, and positive, and leaders demonstrated that they were aware of attitudes to ImpactEd. Rachel and Natalie also monitored the data that was being input in order to check that the second ‘success criterion’ – consistent data entry – was also being met.

Is there evidence of success or failure?

Interviews with teachers about their experiences and discussions with leaders who had monitored the intervention’s impact suggested it had been successful.

i) Consistent use and decision making

All staff we spoke to had a consistent understanding of the potential benefits of

ImpactEd and its intended use. All teachers had received training and were using the system by December 2019. Lead CSPs had also done so and were preparing to disseminate this training to other CSPs in the near future.

ImpactEd had fed into decisions about which interventions would be run. Data collected during the pilot phase had supported discussions about reducing the number of interventions in Key Stage 1, which staff had felt was too high:

“We were reflective enough to know through professional conversations and timetabling that we were struggling to fit the interventions in. What ImpactEd did was give us the data result to support our discussion and say we need to not put so many in because the data is showing that hasn’t had as much impact as it could have had.”

Year One Teacher

Since the whole-school rollout, most interventions had not reached their endpoint but, where they had, teachers had entered data and considered the analysis results. Teachers were clear that they could, and would, use this data to support decisions about which children

would receive interventions in the new-year. However, most teachers felt that they did not yet have sufficient data to make decisions about which interventions would be run, or stopped. This may have been influenced in part by a lack of urgency as all interventions showed an overall trend of positive progress. Where some pupils had not made progress as part of an intervention, despite the group making progress overall, it was generally considered that the pupil should receive the intervention again before the decision was made to try something different. Teachers highlighted that one term was a short amount of time in

which to expect pupils to make progress; they planned to review the interventions they were running in spring. While this may be wise, and some interventions such as ‘Talk About’, which lasts for a year, are planned to last for longer than a term, the decision to review intervention impact after three or four terms stands at odds with leaders’ original intentions to make quicker decisions about what is and is not working.

Furthermore, there was not yet evidence of teachers comparing their pupils’ data with wider datasets, often because the measures were tailor-made for Bengeworth and therefore



there was no evidence base available. Staff were not concerned about this: they felt that the need for tailor-made interventions outweighed the need for cross-school comparison.

Nevertheless, this demonstrates a move away from the original vision of what success with ImpactEd would look like. Leaders felt that control groups and comparisons would be the next step in the implementation as it had not happened as quickly as originally planned. Leaders recognised that where they were using their own measures they would have to develop their own control groups internally.

ii) Non-cognitive intervention impact and inclusion

It was clear that staff did not view ImpactEd as a tool specifically for non-cognitive impact measurement. Teachers referred equally to academic interventions and non-cognitive interventions. Therefore, while non-academic interventions retained importance and the school continued its culture of inclusion, ImpactEd was not seen as an innovation focused on evidencing non-cognitive interventions.

iii) Teacher workload

There was a clear positive impact on teachers' workload by the end of the autumn term in 2019, three months after

rollout, which was sooner than leaders expected. Teachers found it 'quick and easy' to enter data and reported that the analysis outputs were easy to understand:

"It's much easier than what we were doing before, you just put it in, click a few buttons and it gives it to you and then it allows you to be more forensic."

Staff anticipated that workload would be further reduced once they were able to stop using duplicate systems as some interventions were still being rag-rated as before while the changeover to ImpactEd was still in progress.

iv) Leaders' knowledge and workload

ImpactEd has also had a positive impact on leaders' ability to understand what is being delivered across the school, to pick up on instances where things are not being delivered and to monitor the impact of all interventions, forming a holistic picture. Rachel feels this has been achieved while reducing her workload:

"From my point of view and strategically across the school to be able to see what's being done has been invaluable... At a glance I can have a holistic view of what everybody has been doing."

What are the medium- or long-term impacts expected to be?

In the longer term, leaders at Bengeworth feel ImpactEd will allow them to know which interventions best achieve their intended outcomes and that this will make Bengeworth's practice more effective. Although teachers had not used ImpactEd to make decisions to cut or change interventions based on the platform yet, the pilot had helped to support the decision not to do so, and most felt that ImpactEd would support future decision making.

Staff felt that ImpactEd would improve CSPs' confidence in their assessments and teachers' confidence in their decision-making about interventions. In turn, leaders and staff felt that the main underlying purpose of using ImpactEd, like all of their practice, was to have a greater impact on pupils' progress and they believed that, by supporting better decision making, ImpactEd would have a positive impact on pupils.

3.4

Innovation 4: Shine at Eltham Hill School

3.4.1 Introducing the innovation

This innovation, 'Shine', takes place at Eltham Hill School, a single-sex local authority secondary school in Greenwich, London, with an intake of approximately 1,100 pupils. Shine is in its third year and it targets the new intake of Year 7s, while still retaining older pupils who have been attending regularly since they started at the school.

Shine is an after-school extra-curricular activity that is aimed specifically at pupil premium pupils; pupils with special educational needs; and pupils with mental health issues, although access is open to all pupils. It takes place in an information and communications technology (ICT) room so that all pupils have access to a computer and email.

The innovation involves the pupils creating memes and sending them to an artist offsite via email through their teacher during the session. The artist gives live feedback in the form of direct comments on the work or changes to the original meme.

The pupils use this feedback to revisit and develop their work, creating a new meme to send back to the artist. As this is an extra-curricular activity, it is not linked to any particular subject or curriculum, although its closest links are to art and English. It has been funded by The Shine Trust and the Stepping Stones Fund from the Greater London Authority (GLA).

Last year, pupils' work was displayed in art galleries such as the Hoxton Gallery, Vinyl Deptford and Lewisham Arthouse. The lead teacher organised for the pupils to be on the reception desk at the Lewisham Arthouse for a week and for Saturday trips to the galleries for the pupils and their families to view the work.

3.4.2 Agenda setting

What were/are the 'pain points' that this innovation sought to respond to?

The innovation was set up to engage the target pupil group in their learning and to build their love of school, or as the teacher who set up the project, Rosie Osborne, puts it, "working with students who can't access the curriculum in traditional ways".

Rosie believes traditional forms of assessment can be a barrier to access – and this is

why the innovation takes such a different approach. Rosie wanted to ensure pupils received feedback in a way that encouraged them and built confidence, hence using real-life feedback from an artist:

"All of the assessment is meant to be about giving them feedback that will help them to improve, but in a way that is not restrictive."

Lead Practitioner

Historically, the school has had difficulty engaging pupil premium eligible students and those with special educational needs; their attendance was poor and there was a wide gap between outcomes for this group and the rest of the pupils at the school. The school had noticed particular difficulties during transition from primary school, so they targeted Year 7 pupils for Shine initially.

Rosie, an experienced teacher and one of the longest-standing members of staff, was appointed Lead Practitioner in the school, a role that allows her to pilot a series of extra-curricular groups that target pupil premium students, those with special educational needs and those with mental health difficulties, to help them re-engage with school.

Rosie wanted to improve the pupils' confidence and willingness to open up, ensuring they felt free to express themselves in all lessons. Rosie therefore set up the innovation to support the target group by helping them to be creative and to ensure they were assessed in a way that was less "judgemental."

It should therefore be noted that while this project uses innovative assessment methods, the key aim of the project is to develop non-cognitive skills and engagement with school.

Who and what influenced this agenda?

Rosie has driven the innovation. She is an English and drama teacher as well as the school's Lead Practitioner on the leadership team. An important part of her role involves developing creative projects for pupils who find it hard to engage with the curriculum, and to seek additional funding for such projects. She developed and leads the innovation with an existing artist friend.

Her experience with Creative Partnerships earlier in her career led her to believe that pupils in receipt of the pupil premium lacked confidence and did not always engage with school. Her motivation for starting this project was to address these two areas:

"What we called the D-C borderline students at GCSE level, which is now four or five, is often based around that self-belief, and also things around attendance and behaviour, and literacy and oracy, reading and writing... So, I think a lot of those projects that I do address those issues, but in a really fun way that they really get engaged with, and music and art and writing, but in a really non-pressured environment."

Lead Practitioner

Rosie was appointed Lead Practitioner at the school due to the successful work she has done in the past on creative projects. For example, she previously devised a project that involved pupils working with writers from the Royal Court Theatre in London to craft their own plays that were performed. The pupils involved in this project became more engaged and confident in the rest of their school life, so the leadership team wanted to formalise her role to give her the time to lead on future creative projects.

The senior leadership team are committed to not becoming an 'exams factory' and to offering something broader than just the taught curriculum. They therefore prioritise projects like this, which broaden the curriculum experience. They have taken time to articulate their vision and aims as a school.

Rosie reports that senior leaders trust her because the students she works with are often more engaged in their learning due to the after-school projects. She has an open-door policy and senior leaders occasionally drop in to sessions. They have also shared their transition interventions, of which this is a part, with a wider audience at a national conference held by SSAT (the Schools, Students and Teachers network).

Why was it considered an important agenda to pursue?

One senior leader talked about the importance of articulating your vision as a school and then prioritising projects that help to achieve that vision. Eltham Hill's vision is to extend its offer beyond the taught curriculum so that students feel valued and that they belong.

Furthermore, many of the tasks in English statutory assessments ask you to write for an specific intended audience and Rosie felt that students found this hard when they had no actual input from the "outside world". She therefore came to believe that it was important for pupils to have engagement and feedback from the outside world.

Rosie also believed that the target group of students would benefit from having time outside lessons with set criteria, where they could enjoy "freedom outside lesson

time, [which] opens it up to some really exciting and creative responses”. She felt that if they had the opportunity to feel freer in a school environment, even outside of normal lessons, then they would feel more engaged with school overall.

This innovation is part of a set of extra-curricular solutions Rosie is working on, all partly funded by the GLA’s Stepping Stones Fund and The Shine Trust. One of the other projects, for example, targets a group of pupil premium pupils to create soundscapes with an artist-in-residence, where they have gained informal feedback on their work to make improvements. Rosie feels it is important that this type of innovation happens as an extra-curricular activity to give teachers the freedom to take risks:

“I love experimenting with things, and in that extra-curricular environment, you have the ability and you have that space to do it, because it can only add to their experience. If it’s in lessons all the time, you never quite know what the impact’s going to be, and then it’s a bit of a danger, but with this, you really feel like you’ve got the freedom.”

Lead Practitioner

3.4.3 Formulation

How was a solution arrived at?

Previously, in an attempt to secure ‘real-world’ feedback in English lessons, Rosie asked students to write letters to authors in the hope they would respond, but they rarely did.

After delivering her project with the Royal Court Theatre, she decided to create a new project where students could gain instant feedback on their work from someone from the ‘real world’, something that the senior leadership team supported.

Rosie had the idea of getting an artist involved after talking with her networks outside of school; one of her artist friends, Jeremy Gluck, offered to volunteer for a pilot. He reflects on his work with the school in the Master’s degree he is undertaking and he is planning to undertake a PhD to explore how schools and artists collaborate to improve young people’s mental health. Rosie secured funding from Shine and the Stepping Stones Fund for a yearly project. This year she is running this project and another, where students create soundscapes with artists who are paid for their time.

Rosie initially tested the solution with a small group of students before seeing the potential for it to work with a larger group on a regular

basis. She was working with a small group doing drama, creative writing and art and when they started to create memes, she got in touch with another artist friend of hers, to experiment with assessment through live feedback.

One senior leader said that the intervention was influenced by national evidence on poorer outcomes for pupil premium students and research by colleagues conducting studies for a Master’s degree. This research suggested that the blockers the target group were experiencing – such as poor organisation and a lack of support at home – could be addressed by increasing their sense of belonging at school:

“Some of the work that Rosie is leading and what happens in that group is those children have got almost a sense of family, a sense of kinship, it’s developing those friendship skills and all the other things that come alongside that as part of the benefit.”

Senior leader

Who and what fed into shaping it?

Both Rosie and the artist were keen to provide pupils with feedback in a way that did not constrain their creativity. They

therefore experimented with the best way to give students live feedback. Initially, Rosie sent the memes to the artist after the session, but it was decided that him offering feedback during the session was a better way of engaging the students, as they enjoyed receiving an immediate response.

3.4.4 Decision making

Who was involved in deciding on this approach to addressing the pain points?

Rosie, as Lead Practitioner, was the driving force behind the initiative, with support from the artist. The artist said how important it was to trust the professional in certain instances, for example around how safeguarding would work and which students to select for the intervention. The senior leadership team gave Rosie time, autonomy and the creative freedom to develop the project.

Why were other solutions rejected and this approach preferred?

Other solutions were not considered or weighed up against each other; instead, the project grew organically. Initially, the project just involved a group of Year 7 pupils, but now, some of those older pupils have stayed as the project has grown each year

to include the new cohort of Year 7 pupils. Some of these older pupils take on the role of informal mentors, supporting the younger students with their work. They also now use internationally significant dates to inspire the students to create themed memes, such as International Women's Day.

3.4.5 Implementation

What mechanisms have been used to implement the approach?

Shine is an after-school club:

- The students work in the ICT room to create a meme based on stimulus material that Rosie provides.
- When they feel their work is ready, they email it to Rosie, who emails it on to the artist.
- The artist offers live feedback in an email back to Rosie, who then shares it with all the students on an interactive whiteboard.

Who are the different actors involved?

Both Rosie and a senior leader highlighted the importance of clear safeguarding processes to ensure that students are carrying out the work in a safe environment, especially as

students sometimes share sensitive material. This means an experienced professional with links with other professionals in the school who have pastoral and safeguarding responsibilities needs to mediate the relationship between the students and the artist.

“You’ve got to have somebody who knows what they’re doing, because of the safeguarding aspect of it. Because it’s going through a third party, you’ve got to be there to mediate it, you can’t just create a link with somebody that you don’t know. The person from the outside sends the message to you, then you send it back to the student. Yes, so that can be problematic. But once you’ve got that set up it’s fine.”

Lead Practitioner

Rosie also includes the parents and carers so that the young people have permission to stay after school. For some students in particular, she tailors her communication to maximise the chances of them engaging with the project:

“I contact one student’s mum regularly. I just wrote to her last week and she came along to the exhibition. She was so open to it, and she was so grateful that she [her daughter] had been kind of caught, and channelled, in a way, that meant that she stayed in school, because there was a danger that she would have left school.”

Lead Practitioner

The partnerships with the gallery developed because she took the initiative and contacted the galleries directly. She was surprised at how quickly they agreed to exhibit the work and she collaborated with the art department in school to help the students mount their work ready to exhibit.

What obstacles had to be overcome as part of implementation?

Technological and safeguarding constraints mean that Rosie has to submit student work via email. There do not tend to be technological issues, such as logins not working, but if there are, Rosie has access to the school’s technology support team. If she is unable to attend a session due to CPD, the session does not run that week.

As it is a drop-in session, Rosie does not know which students will attend from week to week. To overcome this, she makes sure that each session is self-contained and that participation is not dependent on having attended before. Rosie has worked closely with the member of staff who oversees the summer school Eltham runs with students going into Year 7 so they can identify the students who might benefit most from the innovation.

3.4.6 Evaluation

How is success or failure being characterised?

Because this project is not part of a particular subject or syllabus, there are no specific attainment or progress measures. Instead, the innovation’s success or failure is gauged according to four loose goals and these have remained the same since the initiative began:

1. Increase students’ confidence and ability to take feedback positively.
2. Increase students’ engagement with school.
3. Improve students’ attendance and behaviour.
4. Improve students’ literacy skills.

What approaches are being used to evaluate the innovation’s success?

Rosie uses student attendance, progress and attainment data to monitor the intervention’s success and discusses students’ progress with Heads of Year and subject teachers in English, drama and art – as these are the subjects most closely linked to the programme.

It should be noted that many of the students who attend are in Year 7, so while they form part of a group who have historically not engaged with their learning, it is hard to ascertain what impact this intervention has on these individuals as they start the programme as soon as they enter the school. However, Rosie reports that they improve over the course of the year and as they spend longer on the programme. During this academic year, the whole of the Stepping Stones project, of which this is a part, is being evaluated by the independent evaluator.

Rosie has to provide Stepping Stones and Shine with impact data each half-term, in the form of a narrative report with details of attendance, demographics of students who attend, attitudes to learning in the form of behaviour and exclusions and attainment.

One senior leader reported that while attainment was important, it was also some of

the more intangible outcomes that were most valuable – for example, re-engaging students who were at risk of exclusion with their learning – and where they saw the most difference.

Is the innovation believed to be achieving its intended objectives?

All those interviewed, including a student, felt that the initiative is achieving its aim of giving students – especially those in the target group – a chance to receive feedback in a non-pressured, freeing environment to develop their confidence.

Rosie particularly highlighted one student with special educational needs who had benefited from the opportunity to be creative:

“It was more a creative element, and having the freedom to explore your feelings, but without the pressure of being judged against an exam criteria. So, that was what was really interesting for me, especially for students with special needs, in terms of some students who never, ever spoke in lessons were suddenly free. Like [student] was totally free to suddenly express herself, in a way, and just flew with it. And still, after three years, is still flying with it.”

Lead Practitioner

The student we spoke to explained differences between how she receives feedback through Shine, compared to how her work is normally assessed:

“It’s got round edges, so I mean there’s no like box which can be filled in or fitted into, it’s all just one wriggly shape with basic outlines... [It was] so much more helpful because then it means there isn’t a pressure of having these targets to reach... It honestly took everything away from normal feedback, which I hated, and just put everything I loved about arts, different arts and communication into a just wonderful thing.”

Student

This student had also created an alter ego to interact with the artist, through which she felt she could express herself and receive feedback in a more creative way. She talked about how working with an artist outside of school has helped her self-esteem and ability to communicate her special educational needs to other pupils, teachers and others outside of school:

“It boosts your self-confidence, definitely, because it’s an outside opinion coming in and giving you that positive feedback, where they’re not influenced by anything, they’re just literally seeing your work... When I’m there [in Shine], it’s definitely lifted a lot of my anxiety because school is an anxiety-ridden place for me. It’s lifted a lot of anxiety, like because for most of Year 7, nobody except teachers and stuff in the school knew about... different disabilities... But then doing Shine, because that was a place where nobody really knew me, so I could just completely just go in with a new coat of paint... since I came into Year 8, I’d say, I’ve definitely been more open about my disabilities and stuff, because it’s part of me. Why should I bother trying to hide it or cover it up?”

Student

Rosie reported that an additional benefit was finding out the interests and additional needs of the students so that future projects, such as one on gender equality, could be designed around them. The artist also talked about how much he had learned from the students, such as how quick-witted and open they were. Finally, one senior leader

advocated working with external professionals as it was a way for teachers to develop their expertise and form links in the community.

While the innovation was not primarily focused on student attainment, in terms of whole-school data, Rosie also said that there had been a positive impact on the target group's results in English and drama, with 97% achieving better-than-expected progress in the first year and 100% in the second. Attendance for the target group also improved, especially with those who previously had poor focus and engagement at the school; the Head of Year 7 reported fewer issues with this target group than they have had with similar students in the target group in previous years.

As Rosie develops future iterations of the projects, she is committed to using external professionals:

"I love the idea of using professionals again because that raises the aspiration and also it gives the students the idea that they can do it; getting that instant feedback and also the possibility of exhibiting or sharing in a wider public sphere, which is great for them... then they can say, 'I've been in a gallery in Hoxton and I can do it, I've worked with a professional musician, I have created a garage band track, I've performed as a DJ, I've worked with a professional dancer, I've worked with a professional artist.' But I love the way that they don't see that as a hierarchy; they will talk to these professionals on a level and they are not frightened to."

Lead Practitioner

As yet, no other teachers have tried this approach in an after-school club or in their lessons, although Rosie feels it would be a simple and valuable process.

If another school were considering developing this type of innovation, the senior leader interviewed would recommend:

- pinning down your values as a school first, and considering how this project fits
- finding the pioneers on the staff team and giving them time, encouragement and the creative freedom to try something different
- considering what the rest of the sector can learn from your project.

The Lead Practitioner would recommend:

- using experienced staff who have a good understanding of safeguarding processes
- working with other professionals in the school, such as pastoral and safeguarding leads and technology support
- developing a network of external professionals who can offer 'real-world' feedback
- identifying how the students' interests and additional needs can feed into future innovations.



3.5

Innovation 5: Eedi at the Academies Enterprise Trust (AET)

3.5.1 Introducing the innovation

Approximately 60 primary, secondary and special schools in the Academies Enterprise Trust (AET) – the largest MAT in England – have been using an online maths platform called ‘Eedi’. It is being used across different age groups up to Year 8, and additionally with Year 11.

Eedi consists of multiple-choice, low-stakes, formative assessment tests. The questions are designed to be diagnostic, with each potential answer specially designed to reveal a common misconception or mistake for that topic.

For this study we focused on two primary schools, both serving deprived, one in Birmingham (400 pupils) and one in Burton on Trent (593 pupils). Eedi is being used in every year group from Year 1 to Year 6 in both schools.

In our first visit, in the autumn term of 2018, AET’s national curriculum lead set out the trust’s expectations of how Eedi might work in practice. In the trust’s schools, teachers teach a topic and then pupils complete the test on the platform as homework. The platform marks the test automatically, but also gives feedback on each pupil’s misconceptions so they understand why they made a mistake, not just what the mistake was. The diagnostic questions reveal pupils’ common mistakes to teachers so that they can re-teach or explain the mistakes. Pupils take a test again three weeks later and this is intended to show whether the topic has been embedded in their longer-term memory and to provide further information for teachers on pupil understanding. The Eedi team also hope that the platform will be used by parents, although this is not yet a part of AET’s use of Eedi:

“What we’re trying to do is basically allow teachers to be more informed so they can provide better instruction, allow students to understand, not just that they’re wrong, but why they’re wrong, and for parents to understand where their child is struggling and how they can support them.”

Eedi central team

In theory, teachers are able to retain autonomy in choosing their own questions from the question bank to fit the topics they teach. The diagnostic question bank and the feedback that the platform provides were designed by an experienced maths teacher who is Secondary Maths Advisor to the Times Educational Supplement who co-created ‘diagnosticquestions.com’, a free website hosting the world’s largest collection of high-quality diagnostic multiple-choice questions and now part of Eedi.

In AET schools, including those we visited, Eedi is used to supplement other forms of assessment and to triangulate the data they generate. Stop the Clock assessments, produced by AET and used across AET schools at the end of each unit, assess pupils on the different elements of the maths curriculum and their skills in maths. Questions cover the same content as Eedi (relating to the previous unit’s scheme of learning, which is centrally planned by AET). Questions are not multiple choice as they are with Eedi, but are similar in style and the results are analysed by AET. AET then feedback pupil data from Stop the Clock to schools, to allow teachers to understand pupils’ progress in comparison to their age-related level.

Our case study schools both used Eedi as well as Stop the Clock. In School One, Eedi was used as an interim assessment for each unit. The school attempted to use results from Eedi to target interventions before pupils are assessed again on Stop the Clock. In theory this would have allowed the school to act swiftly to support pupil progress in maths, rather than relying on termly data. However, this was not the experience of School One, which decided to stop using Eedi during the summer term of 2019.

School Two used Eedi for regular homework. This homework was used formatively to help ensure that lesson planning targeted pupils' misconceptions. The school also used Eedi at the start and end of each unit of work as a means of assessing progress. The school also used teacher assessments, which are moderated internally by school staff, and used to triangulate with the more formal termly assessment that AET produces.

3.5.2 Agenda setting

What were/are the 'pain points' that this innovation sought to address?

AET's deployment of Eedi emerged from a combination of different pain points experienced by different parties.

i) Eedi team

The Eedi team themselves were heavily influenced by research on formative assessment and memory, especially the work of Robert Bjork and Dylan Wiliam.

They felt that their platform could address a number of 'pain points' that they identified in schools across the country. These were:

- **Workload:** Eedi seeks to save teachers time by automating marking and producing reports and analyses of pupils' understanding.
- **Weaknesses in formative assessment:** Eedi aims to make it easier for teachers to understand why pupils make mistakes, and what they need to practise or know.
- **A need for more precise reporting to parents:** The Eedi team believed that parents would find detailed analysis of what a pupil can and cannot do more useful and meaningful than generic pupil reports.
- **Weaknesses in pupil self-assessment:** Eedi aims to help pupils understand why they got something wrong, not just that they got something wrong.

The co-founder of the Eedi team had originally acquired the diagnostic question bank for online tutoring, but realised that schools were

vital in order to benefit as many pupils as possible. Teachers needed something more structured than a question bank; specifically, they needed to be able to quickly match questions to existing schemes of work.

ii) AET's central team

AET's central team felt that assessment across their schools was not precise enough in primary and Key Stage 3. This was the case across an estimated 80% of schools across AET. Assessment had tended to be based on learning objectives for individual lessons rather than the granular understanding of the steps needed to understand a concept in maths. AET wanted granular understanding of what it calls "fluency steps" to be part of maths assessment practices in every school across the trust.

iii) The case study schools

The two case study schools had different reasons for using Eedi. At School One, the Headteacher reported that, having been recently appointed, their immediate priority was to improve teaching and learning following a disappointing Ofsted report, which precipitated a change in the senior leadership team (SLT). Interviewees explained that previously teachers held low expectations and lacked subject knowledge and assessment expertise in maths and some other subjects.

The maths lead on the new SLT felt that pupils lacked knowledge about how to use mathematical skills and concepts. This led to difficulties understanding why they got questions wrong, a problem exacerbated by weaknesses in maths teaching in the early years and gaps in teacher knowledge in relation to key mathematical concepts, particularly when it came to content that was covered in other years or Key Stages. This view was shared by the Headteacher:

“[The Head of Maths] and I haven’t been here very long, and that’s the kind of legacy that we’ve picked up, and it’s a challenge in most areas of the curriculum but it is also a challenge that’s specific to maths.”

Headteacher: case study school

At School Two, interviewees explained that the school itself was not actively looking for a solution to any issues with maths assessment. However, once AET made schools aware of Eedi, the two maths leads at School Two decided to incorporate Eedi into their practice. They felt that Eedi promised three main benefits:

1. It was an opportunity to reduce workload.
2. The platform was already linked to their existing schemes of learning.
3. Eedi was free, in contrast to the maths homework package they had previously used.

Of these, workload was considered to be most important. School Two’s maths leads felt that Eedi saved time for teachers. Instead of teachers having to match homework questions to their schemes of learning and lesson content, and then print out questions for pupils to take home, Eedi created questions that were pre-matched to lesson content and were accessible online. They were also excited by the potential for Eedi to reduce marking, as the platform records and shows teachers which pupils have completed each homework, as well as each pupil’s score. The maths leads explained that workload was something that was considered in every decision made in the school.

Thus, while the decision to use Eedi in AET schools was taken by the AET central team, the innovation tallied with priorities in each school. For one of our two case study schools it tallied with the SLT priority to improve the quality of maths teaching and assessment, while in the other it tallied with a focus on reducing workload. There was therefore alignment between the central and school teams.

Why was it considered an important agenda to pursue?

The AET core curriculum team prioritised improving the quality of formative assessment in order to improve the quality of teaching and learning and help pupils make more progress in maths. The majority of schools in AET are ‘sponsor academies’, which had gone into special measures, or been judged to require significant improvement. AET recognised a weakness in maths outcomes across the majority of its schools, and was looking for changes that would support pupil progress in maths.

Eedi’s team believed that diagnostic questions were a powerful tool to improve formative assessment. They were also keen to help parents understand how their child is doing. The team believed that they could support pupils by helping parents understand their children’s progress in maths.

The parental element of Eedi was also influenced by the need to monetise the platform, without using schools as the source of revenue. The suggestion had been to provide parents with some elements for free and others as part of a paid subscription otherwise known as a freemium model.

Interest in diagnostic questions as a linchpin of formative assessment and the desire to engage with parents meant that Eedi and AET had shared priorities. Schools also shared a belief that the diagnostic questions would be useful

Schools within AET were introduced to Eedi through regional conferences, and were expected (although not required) to try to make use of the platform. Not all schools across the MAT chose to use the platform, but the two schools we visited saw alignment between Eedi and their own agendas. For School One, Eedi appeared to offer a means of upgrading teachers' knowledge and understanding of assessing progress in maths. As the maths leads explained, Eedi allowed teachers to gain an:

“Children give you the wrong answer but they think it’s the right one. Nine times out of ten they think it’s the right answer, and it’s having the subject knowledge as a teacher to think, ‘Why do you think that’s right? What’s going on in your head? What’s your misconception that we need to unpick?’ And that is where the poor subject knowledge [from teachers] means that that child doesn’t necessarily move forward.”

Maths lead: School One

For School Two, maths leads did not feel that teachers needed support in the same way, but they wanted Eedi to strengthen formative assessment at the school and improve planning.

3.5.3 Formulation

How was a solution arrived at?

The AET core curriculum team took the decision to use Eedi. They took this decision after first deciding to develop a bank of diagnostic questions to use in schools across the MAT. For AET, a key reason for using Eedi was that it offered a simpler, less time-consuming solution to the other options the team had considered. Initially, the curriculum team considered creating their own diagnostic assessment resource, and went so far as to begin writing their own multiple-choice assessment questions. At the same time, the team recognised that many teachers were using a platform to write, print and administer their own diagnostic questions. After meeting the co-founder of Eedi, however, the central team realised that Eedi could offer a ready-made solution, which offered an accessible platform that might cut teacher workload by reducing the need to administer and mark tests.

Who and what fed into shaping it?

AET has a team of maths specialists and external experts who work on its maths

curriculum. This team worked with Eedi alongside the MAT’s core curriculum team to write questions that fitted with the MAT’s maths curriculum. At the point when AET decided to use Eedi (2016/17), Eedi was able to provide a bank of diagnostic questions, but not the automated analysis and feedback that are now a key feature of the platform. Eedi thus evolved as the AET curriculum team worked with the Eedi team over the course of the next year (2017/18) in order to shape a solution that allowed teachers to understand pupils’ misconceptions across all AET schemes of work in maths.

In order to match the case study schools’ needs schools’ maths leads have refined the deployment of Eedi by working with classroom teachers to adapt how pupils access the platform. In some cases, questions have been printed out when there is no access to IT, or when questions need to be read out to Early Years pupils.

3.5.4 Decision making

Who was involved in deciding on this approach?

The decision to use Eedi in AET schools was taken by AET’s core curriculum team. This team made the decision to address a particular

pain point (assessment in maths), and to use Eedi as a means of solving this problem. As we will see below, this team also made decisions around implementation and evaluation.

The AET curriculum team have an agreed way of working. This team is made up of a national lead (who we interviewed) and five regional leads. The team meet fortnightly to lead strategy in a number of areas across the MAT (including assessment). The team's decision making approach is described as "thoughtful disagreement". This involves the merit of each idea being evaluated and decisions being taken collectively, rather than being taken by the most senior member of the team. Team members are therefore encouraged to put forward deliberately contrasting views, in order to test the merit of a particular idea or decision.

The AET curriculum lead described the process in the following terms:

"We thrashed it out with the team, we always encourage absolute open debate and disagreement. We call it thoughtful disagreement... we ask people to play devil's advocate... it's not about the loudest voice at winning the arguments, it's about the most compelling reasoning and logic. And it doesn't matter how deep we went in our conversation, it just made sense to run with this, because of three reasons: it's clear in its precision, it's easy to use and we can embed our own content."

AET National Curriculum Leader

Decisions around platform functionality and design were taken by the Eedi team. These decisions needed to combine business and pedagogical perspectives – in particular by deploying research expertise to understand how best to address pain points around formative assessment and deciding how to develop a solution that was financially sustainable. For the Eedi team, the question of how to monetise the platform was a particularly thorny issue. The team decided from the very start to try to stay away from charging individual teachers. With AET, the team explored four options for monetisation before deciding to offer the platform to AET for free:

- **Charging each school a small amount:** Eedi and AET decided against this option due to pressures on school budgets.
- **Charging parents:** Eedi is keen to explore this in future but reports resistance to this approach from schools. The platform is currently only designed to be used by teachers, but Eedi is working on a parent tool, which it may charge for.
- **Charging AET:** In future, Eedi might consider charging a fee for MATs. This would require Eedi to build an interface that was useful to central teams within a Trust, for example giving insights into strengths and weaknesses of learning



across schools within a Trust. With AET, this option had not been explored as this interface did not exist, even before raising the question of whether AET might have been willing or able to pay for this service.

- **Charging teachers for resources and CPD:** Another possible future direction for Eedi might be to charge teachers for insights into strengths and weakness of their teaching (for example highlighting which topics an individual teacher’s pupils struggle with compared to what might be expected). Charging fees for insights to help teachers personalise their CPD would allow the platform to remain free for its original purpose as an assessment platform.

Eedi’s Chief Executive Officer (CEO) reported that these options had been explored informally, over time, and in discussion with the AET curriculum team.

In schools, maths leads and the SLT made the decision whether or not Eedi would be used in their school. Maths leads took decisions about precisely how assessments would be administered – for example, whether the assessment was conducted in class or at home, on a computer or on paper, at which exact point in the unit (guided by the directions given by AET) and with which classes.

Classroom teachers appeared to have limited decision-making power (other than those who were also maths leads). However, teachers were able to decide on how they would adapt their planning and teaching as a result of the feedback they received from the Eedi assessments.

Why were other solutions rejected and this approach preferred?

AET’s core curriculum team decided to pursue Eedi because they had worked with a member of the Eedi team in the past on maths assessment. The fact that they could work with Eedi to design their own questions tying into the MAT’s curriculum was particularly important as the AET national leader for curriculum explained:

“With Eedi we could write our own assessments and just embed them into their system. That was invaluable, because we could write them against our own schemes of learning, whereas other products we couldn’t.”

AET National Curriculum Leader

Case study School One believed AET’s decision to use Eedi fitted with its own need for a

solution to problems with assessment in the school and a sense that it did not have anything else and needed something quickly:

“Judgements were not secure in most classrooms. We [the new SLT] came in the middle of AET working on a different way of primary schools assessing in reading, writing and maths anyway, so we didn’t seek Eedi out but we knew we needed a solution and when AET said, ‘This is what we would like schools to do,’ we jumped on that quite quickly because we didn’t have anything else. So it wasn’t a switch from one thing to another. We said, ‘Okay, that’s the one that AET wants us to use and we need something really quickly. Let’s get that in.’ So we were quite proactive in pushing that forwards then.”

Senior leader: School One

Case study School Two already used a different online maths homework tool. For this school the decision to replace the previous platform with Eedi hinged on two factors:

1. **Cost:** Eedi was free to use while the previous platform carried a fee.
2. **Workload:** The school reported that

the previous platform required printing, marking, planning which assignment to use for a particular homework, and choosing questions to match schemes of work. In Eedi, question choice is automated and linked to AET's curricula, and feedback is also automated.

For the Eedi team, an important design principle was to build a platform that is free for schools to use and which prioritises simplicity. The Eedi team felt that these features would give them a competitive advantage in a crowded market for online formative assessment solutions. They argued that in contrast, other educational technology (Edtech) assessment solutions focused on keeping ahead of competitors by developing more complex technology. Instead, Eedi kept the technology simple and focused on bridging the gap to existing practice. This means that Eedi sought to develop a platform that was easily adaptable and could fit into teachers' current ways of working.

The decision to work with AET, a large MAT, was seen as a risk. Eedi's approach to marketing had previously been to work with teachers directly. For example, a teacher would find Eedi through their own research, and as a result would already buy into the aims and methods used by the platform. Working with

a large MAT was more top-down, with senior leaders or MAT central teams choosing the platform, and attempting to apply it across classrooms. For Eedi's CEO, this was an attempt to try a new strategy to increase uptake.

Eedi's CEO did not see this decision as totally successful. While there was enthusiasm at MAT level, Eedi believed that this had not always translated to enthusiasm at classroom level. This was borne out by the experience of our case study School One. Where there had been successes in the uptake of Eedi in AET schools, this was seen as being driven by one teacher within the school who had bought into the idea and taken steps to make it work, and shared that with other teachers.

3.5.5 Implementation

What mechanisms have been used to implement the approach?

For AET, it was important not to coerce schools into using the platform. Once the AET central team had decided to use Eedi across their schools, they worked with maths leads in schools across the MAT to explain what the solution was, and how it would benefit maths assessment:

“We pushed it out to heads of departments, because by that point we were clear with our narrative and, therefore, the way that we presented it, it would be difficult to say we don't want to use this, because it was so powerful.”

AET National Curriculum Leader

AET organised regional 'Eedi conferences' at which maths leads were shown how to use the platform. Schools were encouraged to use Eedi, through the regional conferences, offering support to schools that expressed an interest, and highlighting successes in schools with the highest usage of Eedi through a series of case studies. These case studies were in development at the time of our research. In addition, in one of our case study schools a trainer came in to show individual members of staff how to use Eedi.

AET recognised that Eedi formed one part of a bigger picture of improvement for many schools. AET has a series of minimum expectations known as 'non-negotiables' in terms of how it expects schools to operate. These minimum expectations are based around various aspects of schools' practice, such as attendance, financial systems and basic principles of teaching and learning. While

Eedi, as a formative assessment model, fits into one aspect of those, the ‘non-negotiables’ are not so granular that they include Eedi.

AET is planning to create a series of Eedi champions with leadership responsibilities, with additional pay attached (a TLR payment). This is a response to a perception at AET that its approach to implementation had been too ‘laissez-faire’, and that building in some accountability for schools’ use of Eedi, through TLRs and appraisals on the implementation of Eedi in each school, would help drive uptake.

Who are the different actors involved?

Implementation flowed from the core curriculum team, through maths leaders in schools, to classroom teachers. In one school, the key member of staff driving the Eedi initiative was a senior leader who was also the maths lead. In the other, a pair of middle leaders shared the role of maths lead, and drove implementation. In both schools, individual classroom teachers were then responsible for direct implementation with pupils.

Meanwhile, Eedi’s team include a range of individuals, ranging from those involved in business strategy, to those developing questions, building the software or building relationships and offering support to schools.

Broadly, there was one individual, the CEO, who shaped the business model and overall strategy. Another individual designed the platform. Eedi also has an individual who provides liaison with schools and MATs. Beyond this central team, there is also a team providing training for groups of schools, individual schools and individual teachers.

To what extent is the reified solution consistent with the planned/intended solution?

For Eedi, the platform itself has remained similar to its plans. The CEO reported that the experience of working with AET had led to approaches from other MATs, where implementation would proceed differently. In those MATs, the Eedi CEO described a more prescriptive approach, with Eedi being used as a formal “non-negotiable” assessment. Eedi is also working on developing assessments that do not require technology. One of these developments is taking place through pilots in India in schools where pupils do not have access to computers. This may lead to changes in the UK too, as Eedi attempts to find ways teachers can record data from assessments that take place in class, without computers. Eedi’s CEO estimates that:

“85%, possibly 90%, of our users use the content in class by projecting it on a whiteboard and then basically telling the kids to use their hands to respond or whiteboards to respond.”

Eedi central team

As a result, Eedi wants to remove the technology from the solution so that schools have an option to avoid the “disruption” that can occur with the use of computers in schools, such as lost passwords or lack of up-to-date IT infrastructure.

Across AET, the solution has remained consistent, although there are planned changes to the way in which it will be implemented in 2020. AET has also been flexible in responding to teachers’ use of the platform. Originally intended for use as homework, with teachers using feedback to adapt future lesson plans, the AET curriculum lead now describes a “best practice” model whereby pupils take the assessment independently in class, and are then given instant feedback on common misconceptions.

One way that the case study schools’ implementation of Eedi deviated from the original intention was that they now use the platform with pupils in Key Stage 1 as well as

those in Key Stage 2. Eedi was designed for use with older pupils (Key Stage 2). It was not clear what had led schools to make the decision to use Eedi with younger pupils, and in the case of one of the case study schools, this led to problems, set out below.

In the case study schools, there had been contrasting trajectories for the use of Eedi during our fieldwork, indeed one school has decided not to continue using Eedi. For this school, the decision was based on the difficulties it faced persuading teachers of the tool's value. In this case teachers were believed to be unable to identify pupils' misconceptions themselves, and while the platform highlighted where pupils made mistakes, teachers needed more guidance to understand how this related to pupils' understanding of topics. Logistical barriers also played a part. Teachers found that, with Key Stage 1 pupils in particular, their workload increased as assessments had to be printed out and inputted onto the platform or marked by hand.

What obstacles had to be overcome as part of implementation?

A number of challenges have had to be overcome. In particular:

1. Low homework completion rates: If pupils do not complete homework it makes a

teaching and learning strategy based on assessment through homework impossible. Both case study schools reported difficulties using Eedi for homework, albeit to different levels. For School One, at the stage of planning how best to use Eedi, the maths lead decided not to use the platform for homework. There were two reasons for this:

- pupils lacking technology at home
- Key Stage 1 pupils' difficulties in navigating the platform.

Instead, School One used Eedi in class, with paper printouts. While this allowed all pupils to access the assessment, it also created additional workload for teachers in printing the papers and marking pupils' answers.

For School Two, maths leads identified some difficulties linked to parental involvement at an early stage. Parents were used to the homework platform that the school used previously. The maths leads had to spend time inviting parents into school to show them Eedi, communicating with parents via letter to explain the change in homework system, and speaking about Eedi in parents' evenings. For a small number of pupils, non-completion was also an issue, but maths leads believed that this was to be expected in any homework system.

2. Teacher mistrust: Some teachers are sceptical about multiple-choice questions. The AET curriculum team identified this as a problem at the formulation stage. In part, the decision to have Eedi present at two AET national conferences was a response to this scepticism.
3. Teacher knowledge: As noted above, responding to misconceptions in one school. Whilst Eedi shows where pupils have misconceptions it does not provide feedback for pupils and teachers on what their misconceptions are based on. Teachers therefore need the knowledge to understand and correct misconceptions. In one school, not all teachers were secure enough in their subject knowledge for this to take place.

Both case study schools highlighted difficulties using the assessment with Key Stage 1 classes. Although the AET national curriculum lead told us that Key Stage 2 was the primary focus for Eedi, this message did not appear to have filtered down to our case study schools. For them, younger pupils struggled to navigate the platform, were more likely to guess at multiple-choice answers, and were less likely to complete Eedi assessments for homework. This was a major factor behind one school's decision to discontinue its use of Eedi in 2019.

3.5.6 Evaluation

How is success or failure being characterised?

For Eedi, success is characterised by uptake among schools. In the longer term, success would also involve monetising the platform.

For AET there are three success criteria:

1. **Use of Eedi across all schools:** Over the year, the AET curriculum team have modified this aim due to a slow take-up. Instead, they are now focusing on a small number of schools using the platform well, in order to provide case studies to encourage uptake across more AET schools. In those schools AET aims to see the majority of pupils in all schools complete 15 units a year, with follow-up assessments taking place three weeks after the initial assessment.
2. **Increased attainment as a result of using Eedi:** In schools using Eedi well, AET expects to see a 75% pass rate in the three-week follow-up assessment.
3. **Better learners:** Less tangibly, AET reports a success factor as “pupils who are better learners, because teaching is focused on their specific gaps”.

What approaches are being used to evaluate the innovation’s success?

AET is part way through its second year of using the Eedi platform. There has not yet been any formal evaluation. Instead, AET has been gathering feedback from maths leads and triangulating with learning walks where leaders visit classrooms and through AET’s termly summative assessment. Eedi has also developed a usage tracker to identify which schools are making more and less use of the platform.

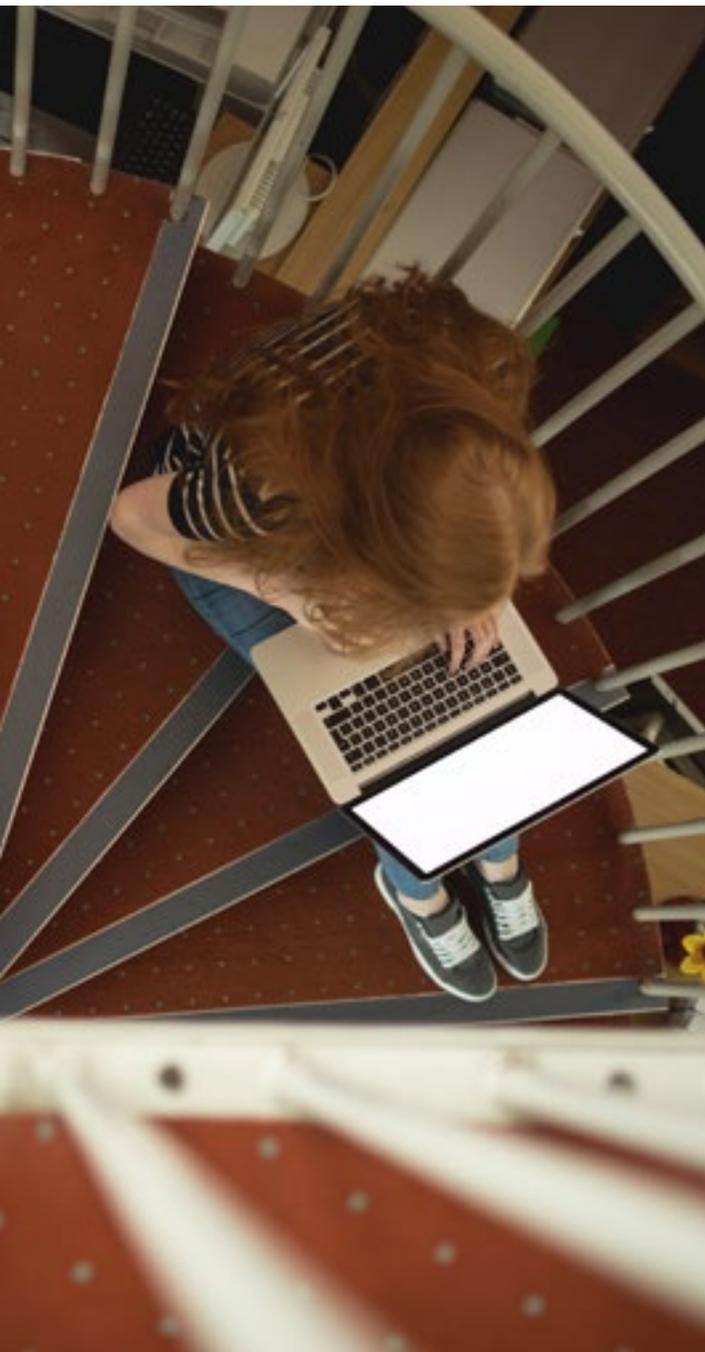
The AET central team are using a number of forms of evaluation:

- Regional curriculum leaders are looking at pupil books to track changes in the quality of learning.
- Regional leads also gather feedback from maths leads in schools (sometimes also called Eedi co-ordinators) on how they are using the platform, and whether they perceive positive impact.
- Once every half-term, this information is fed back to Eedi through Eedi’s lead on school engagement.
- The team are using the usage tracker, outlined above.
- The team are using summative data to show how maths attainment is changing across AET schools.

Lessons learned

1. **‘Bottom-up’ development of an innovation by maths teachers versus ‘top-down’ implementation:** Both Eedi and AET acknowledge that uptake and impact have been greatest where individual teachers have bought into the ideas behind the solution. For the case study schools, the ‘top-down’ approach has not resulted in success; one of the schools continues to use the platform, while the other has decided not to use Eedi in the immediate future.

In 2019/20, AET used an approach to implementation that was more ‘carrot’ than ‘stick’. Schools were encouraged to use Eedi, but there were no consequences for schools that chose not to use the platform, or that have different levels of use across different classrooms. It seems that this strategy chimes with the ‘ground-up’ approach taken thus far by the Eedi team. The approach allows individual teachers to decide whether or not the platform works for them. While this approach benefits Eedi, and the individual teacher, it creates a problem for the MAT. What do the central team do about schools that need most support with maths assessment, but which do not engage with solutions such as Eedi? In the future, it seems that AET intends to build in an



element of accountability for schools' use of Eedi by paying staff to take responsibility for how the platform is used in schools.

2. **Decision making at different levels:**

One feature of this innovation has been the involvement of several layers of decision makers. One layer has involved the Eedi team making decisions about the platform itself. This team decided on functionality, design and a strategy to market the platform to schools. The next layer is the MAT central team. They have some limited input into Eedi's decisions about functionality and design. The main focus of decision making here is to address pain points. School management – whether senior leaders or maths leaders, provide a third layer. This layer decides ultimately whether the platform will be used or not, and how it should be used in classrooms. Finally, teachers have limited decision-making power, but can feed back, indirectly regarding the design and functionality of the platform. This layer can also feedback on how successfully the solution is addressing the problem it was intended to tackle. Teachers can also make decisions about the teaching that occurs as a result of the feedback they receive from pupil assessments.

3. **Matching solutions to problems across several different schools:** At the AET/ trust level, assessment in maths was a common problem. AET's response was to make Eedi available to all schools. An alternative approach might have been to target specific schools for whom maths assessment was the most pressing issue and to tailor the solution around them. In this study, we found that, paradoxically, the school that reported the greatest problems with maths assessment did not choose to use the platform, whereas the school that appeared more confident with maths assessment found Eedi useful. For this school, the benefit of Eedi solved a different problem – that of teacher workload.

3.6

Innovation 6: Laser conversations at the Midland Academies Trust

3.6.1 Introducing the innovation

‘Laser conversations’ are an assessment practice used in all four of the secondary schools that are part of the Midland Academies Trust in the Midlands.

Laser conversations are intended to make assessment more meaningful. Rather than regularly entering pupil performance data onto a system, teachers have conversations with their line managers about how pupils are progressing and what steps they are taking to support pupils who are underperforming. The process involves classroom teachers, middle leaders and senior leaders, as set out in the diagram below.



Laser conversations are timetabled regularly for individual year groups. For example, in the autumn term, all Year 11 teachers might meet with their line manager to discuss their class. Information about pupil performance and progress can then be fed up to senior leaders and trustees.

Laser conversations are intended to be a more pupil-focused assessment practice compared to traditional data collections.

“It’s literally you sit down with whoever you’re working with and they talk about where your kids are at. It’s uplifting, because you’re talking about your students and you’re talking about how they’re doing and their progress, so it’s less about the data and more about the student. It has the student at the heart of it.”

*Teacher,
Midland Academies Trust*

They are also intended to be a more accurate way of assessing pupil needs and deciding next steps, compared to traditional data collections.

“For example, let’s say they’re very clear about the changes that took place during the Industrial Revolution, but as a cohort they are less clear about the reasons that those changes occurred. We would then want to have a conversation with the teacher about what is being done about that. So, we would want to know that learning sequences are being adjusted or tweaked.”

*School senior leader,
Midland Academies Trust*

3.6.2 Agenda setting

What were/are the ‘pain points’ that this innovation sought to respond to?

Former Executive Principal of the trust, Ros McMullen, explained that:

“The first main reason was the workload – we were seeing a culture where teachers were inputting vast amounts of data, but we felt it was having little impact in the classroom.”

Former Executive Principal

Within this, three agendas can be drawn out and the relative emphasis of

these has varied over time and among different stakeholder groups:

1. Developing more valuable, active responses to assessment information
2. Increasing professional dialogue between teachers
3. Reducing teacher workload.

Previously, the central trust team collected lots of data about pupil progress at different points in the year through ‘data drops’. Like in many schools and MATs, this was time-consuming but had limited impact, since it did not necessarily lead to teachers adjusting their practice in response to assessment information. Both teachers and the MAT’s former Executive Principal recognised this and referred to it as “assessment for data’s sake”, “or assessment as an end itself”. One teacher also argued that it ended up interrupting the flow of teaching.

“Two years ago we had a huge amount of data but we were not taking that and doing something meaningful with it. Like lots of schools we were falling into the trap of generating lots of numbers... it should be about the students and teaching and what we can do about it.”

School improvement lead

Thus, the trust found that multiple data drops did not result in improvements in pupil progress or attainment despite creating workload. Its innovation was therefore designed to provide a springboard towards professional discussion and action.

The approach was also intended to ensure that peers came together to spot issues for pupils across the school since “none of that gets captured when individuals sit on their own with data”.

Who and what influenced this agenda?

Laser conversations were not the MAT’s first attempt to address assessment challenges. Previously, the central team had sought to improve pupil outcomes and reduce workload by asking teachers to keep ‘live data’ on the centralised computer system (that is, they asked them to update data whenever there was a change or development in pupil progress rather than in onerous half-termly data drops). They hoped this would help teachers use assessment data to improve teaching and learning:

“All data in the trust was live. So, as a teacher you would be constantly updating your data as you needed to, as it was relevant to you. You’d be asking yourself the question ‘Have they learnt what I taught, how do we think they’re doing, what do we need to do about it?’”

MAT central team

However, the MAT realised this approach was not having its intended effect because teachers did not keep data updated regularly and workload remained high.

“Teachers were keeping rich information in their mark books, for example, but asking them to go into another system and keep it continually updated was an increase in workload without a real impact on the quality of teaching.”

MAT central team

Furthermore, the central team noticed that this type of data collection gave teachers the wrong message – that assessments were about holding teachers to account rather than helping pupils to make progress.

The MAT therefore decided to abolish data drops altogether in the hope that this would support school leaders, middle leaders and classroom teachers to make improvements to assessment and teaching and learning:

“We decided that we would stop teachers inputting data completely into the system because this was getting in the way. We decided we would replace it with conversations which would improve the way our middle leaders were able to line manage and quality assure and improve the performance in their departments. It would also mean that the principals were more effectively able to challenge what was happening in each department and that at the end of that process, there would be data produced which was really meaningful.”

MAT central team

Why was it considered an important agenda to pursue?

Some central team members were particularly motivated by a desire to reduce teacher workload, particularly where this was unproductive. They believed that introducing laser conversations would

lead to a combination of reduced teacher workload, and more effective assessment.

The MAT also prioritised this new approach because it felt that traditional data collection was unreliable and unhelpful in supporting pupils to make progress. This was partly based on feedback from teachers about traditional data collections, and partly on pupil progress tracking, which showed that pupils were not progressing sufficiently despite the ‘live data’ initiative.

Additionally, the approach reflected the MAT’s culture, which is said to emphasise responding to specific pupils’ needs, rather than prioritising accountability measures.

3.6.3 Formulation

One school had started using laser conversations to review progress in certain year groups and the central team therefore decided to experiment by rolling the approach out to other schools and year groups.

The MAT’s school improvement lead argues that the opportunity to take promising practice from one school, learn from it and take it to other schools is one of the important benefits of being a small MAT.

3.6.4 Decision making

Who was involved in deciding on this approach to addressing the pain points?

Although the laser conversation model was first developed in one school, the central team were the driving force behind the MAT-wide approach, with the Executive Principal and the strategic lead for academy improvement playing leading roles.

The MAT’s trustees were initially cautious because they were used to the previous assessment system and wanted to keep track of pupil progress by reviewing data regularly. The MAT central team therefore had to work closely with trustees to show how laser conversations could produce more valid and useful pupil progress data.

Why were other solutions rejected and this approach preferred?

The real-time and live-data approaches were rejected because data was not up to date or accurate and did not result in improved outcomes for pupils – despite involving considerable workload. Rather than a range of different options being weighed up in parallel, the new approach was therefore selected, cascaded out and gradually refined.

3.6.5 Implementation

What mechanisms have been used to implement the approach?

Laser conversations have been implemented across the trust. Implementation has involved the central MAT team, senior leaders in the secondary schools and middle leaders.

Different parties are involved in laser conversations at different levels to allow for information and decision making to be shared between different levels of management in the school.

i) The central team role has been to:

- communicate the plan and provide instructions to teachers across the MAT – they did this by cascading information out via head teachers and emails
- create an annual schedule for laser conversations
- adjust instructions and timetables in response to staff feedback
- give all teaching staff and school leaders a guidebook about laser conversations.

ii) The role of senior leaders in schools has been to:

- explain the new approach to school staff during meetings

- invite middle leaders to training on laser conversations
- monitor middle leaders' and teachers' use of laser conversations through meetings with middle leaders to discuss the outcomes of the conversations.

Senior leaders' implementation methods helped to ensure that middle leaders understood the basic expectations for laser conversations in their school context. However, training differed slightly between schools, and implementation was not uniform and homogeneous across schools. For example, the instructions that senior leaders gave middle leaders left space to adapt the innovation to their own teams. This had important consequences, which we return to later.

iii) Middle leaders' role has been to:

- explain how the approach works to classroom teachers
- set expectations for how the conversations would work within their departments
- schedule the conversations in line with the MAT's timetable
- structure the conversations according to their preferences
- identify how pupil outcomes can be improved and produce action plans where necessary.

iv) Teachers' role has been to:

- conduct formative assessments of pupils' work
- use information from formative assessments to make judgements about how pupils are performing
- talk to middle leaders during laser conversations about which pupils are underperforming and what steps they are taking to improve outcomes.

In some cases, teachers sit down with their Head of Department (HoD) to discuss pupil performance without any performance data written down. In these cases, the HoD records data in a spreadsheet. In other cases, teachers enter data into the SIMS/a spreadsheet in advance (for example, "working above expected standard"). They then talk about their classes' data with middle leaders and develop an action plan to improve pupils' performance.

Middle leaders discuss conversation outcomes with senior school leaders and produce department-wide or school-wide action plans. Like with teacher to middle leader conversations, the structure of the discussion varies. In general, middle leaders come to meetings with data on all pupils in a year group (for example, Year 9). Senior leaders then ask middle leaders about what is going well for

pupils, areas for development and plans for improving pupil performance.

In turn, Heads and school leaders meet with MAT leaders to discuss findings and this is used to hold school leadership teams to account and to develop MAT-wide strategies for improving pupil performance. Trustees receive summaries of the information that has been gathered and scrutinise the steps being taken to improve outcomes.

To what extent is the reified solution consistent with the planned/intended solution?

The implemented approach is constantly evolving and has moved on since the original plan as a result of the flexibility individual schools and staff members were given regarding how they conduct the conversations. The MAT's school improvement lead points out that this is part and parcel of trying and developing something new.

The biggest divergence from the original plan has been that in many cases teachers are recording data in spreadsheets. Although the MAT's former Executive Principal emphasised to staff that "if it isn't making a difference to student progress and your effectiveness as a teacher, don't do it", the MAT now believes that some teachers and leaders prefer to structure

laser conversations around recorded data and that it can prompt valuable conversations. On the one hand, this is time-consuming, with some teachers arguing that it is not a good use of time, which could be focused on working with pupils to improve their learning. However, as the MAT's school improvement lead explains:

“Whilst we initially envisaged the process would work better without any numbers at all, our teachers showed us that they already held information which was meaningful and would add value.”

MAT's school improvement lead

Thus, the MAT emphasises that its priority is now for data to be used as a stepping stone towards pupil-centred, professional conversations. As one middle leader put it:

“Everything starts with ‘What are we doing to improve individual pupil outcomes?’ And actually if we do that for each pupil the headline figures should improve as a result of the work that we’ve done.”

Middle leader

Our visits to schools suggested that as the approach had evolved, it had been hard to maintain a shared understanding across all schools and the central team regarding expectations for conversations, in particular whether recording data was appropriate or not. For this reason, the central MAT team have recently reviewed practice and have adjusted their guidelines. This continued iteration is seen as an important element of developing and refining the initiative. The new guidance has been written under the tagline ‘Accelerating Progress, Reducing Workload’, re-emphasising the dual focus on progress and teacher workload and pointing out that ‘we assess to improve our curriculum and our teaching’ and that data is a ‘by-product’ of this.

It goes on to state that ‘we understand that teachers will have their own records to help them assess and plan, and to aid their conversations; however, we do not ask nor expect any teacher to enter data on SIMS. If it doesn’t accelerate student progress and reduce teacher workload it isn’t worth doing’.

What obstacles had to be overcome as part of implementation?

The flexibility given to schools, leaders and teachers in terms of how they implemented laser conversations has had important

consequences and illustrates the trade-offs innovators need to make between flexibility – in order to preserve professional autonomy (as well as responsiveness to context) – and consistency.

Another important challenge has been the fact that laser conversations are sometimes time-consuming for middle leaders managing large staff teams. This may be one factor that prompted the return to recording data in advance of laser conversations since it allows middle leaders to ask classroom teachers about specific pupils who are underperforming. Some senior leaders have taken a similar approach by asking middle leaders to share data about pupil performance in a year group before the meeting.

Although this approach might be considered to be in tension with the agenda around reducing teachers’ workload, it remains consistent with one of the initiative’s initial goals – reducing ‘unhelpful’ or ‘wasteful’ workload – since the data has now become a prompt to action, rather than ‘data for data’s sake’. Furthermore, the MAT hopes that time can be re-allocated from less valuable, low-return workload such as ‘tick and flick marking’ because “that’s not what matters”, resulting in a net saving in workload.

What conditions contributed to or detracted from effective implementation?

Training for middle leaders helped ensure laser conversations were implemented across the trust. Training also provided space for middle leaders to discuss how they would implement the innovation within their teams.

The MAT believes that implementation has been made easier by the fact that it is small, because rather than designing an approach and then rolling it out, the approach has been more co-created, with best practice shared and adapted across schools.

However, as noted above, there is a trade-off between consistency and flexibility. The heterogeneous approach to how recorded data was used during laser conversations meant that some teachers reported being unclear about expectations and some did not see the reductions in workload that they had initially expected. This challenge may be linked to an underestimate of how time-consuming it would be to conduct all the conversations in the larger schools. This may have made it harder for staff to invest the necessary time in developing and implementing effective action plans and the MAT has therefore adjusted timelines for the conversations by spreading them out more.

3.6.6 Evaluation

How is success or failure being characterised?

As noted in section 3.6.2, the innovation's original goals were to:

- develop more valuable, active responses to assessment information
- increase professional dialogue between teachers
- reduce teacher workload.

The central trust team currently define success as:

- staff taking a pupil-centred approach to assessment
- reducing the amount of unnecessary, or 'pointless', work for teachers
- improving outcomes for pupils.

The innovation's goals have therefore evolved over time and, for now, there remains some variation in what goals different individuals emphasise. This leads to some variation in people's assessments of success.

What approaches are being used to evaluate the innovation's success?

The trust is gathering staff feedback by email and the central team believe that an important

benefit of being a small 'family' of schools is that it enables constant discussion and feedback. Because of this, and the fact that the MAT sees the innovation as a constantly evolving and iterative process, it has decided not to conduct a formal, structured evaluation. On the other hand, the decision not to conduct a systematic evaluation as part of the innovation may have made it harder to spot some of the emerging challenges early on.

Nonetheless, informal feedback has led the trust to make adjustments to the innovation, for example in relation to timetabling. Furthermore, it has reviewed guidance and set out the following principles for laser conversations.

What they are

What they are not

A reflection on our: <ul style="list-style-type: none"> • curriculum • delivery <ul style="list-style-type: none"> • organisation (of students) • resources. 	Data reporting
A focus on the progress students are making	A summative report based on current attainment
An opportunity to decide what needs to be done to accelerate progress	A definitive statement of expectation
An opportunity to share strategies and request advice	An opportunity to defend and justify
Assistance in making improvements to practice based on experience	A focus entirely based on targets and attainment
A focus on individuals, which provides the opportunity to identify patterns	A conversation purely about key groups of students
A professional dialogue that is supportive of teachers and students	An extension of performance management
A summary of the conversation will produce actions for the teacher/Director of Learning (DoL)	A judgement about the quality of teaching
A summary of the conversation will produce meaningful data for DoLs/the SLT	The data production, but the process which led to it

The MAT is also looking to roll out certain practices across schools in response to the strengths and weaknesses of different approaches. For example, following a pilot in maths and science, the MAT is now trialling the use of four quadrants to provide a consistent structure for conversations in certain subjects (see below). The latest guidance also includes examples of questions that staff at different levels might want to ask as part of the conversations.

	Working hard	Not working hard
Achieving highly		
Not achieving well		

3.7

Innovation 7: Curriculum and assessment reform in British Columbia, Canada

3.7.1 Introducing the innovation

Reforms to assessment and the curriculum are underway in the West-Canadian Province of British Columbia (BC). Instigated by the BC Ministry of Education, the reforms are now being enacted in school districts and individual schools.

In 2011, the Ministry of Education in BC began work on a sweeping set of curriculum and assessment reforms, culminating in 2015/16 with the launch of a new curriculum in Grades Kindergarten to 9, followed by a new curriculum for Grades 10 to 12 in 2017/18, and ongoing revisions to the assessment regime. This is the first time in BC's recent history that a comprehensive set of reforms to the curriculum have been designed and implemented.

BC's new curriculum and assessment is intended to better prepare students for life

in the 21st century. This includes placing three Core Competencies ('Communication', 'Thinking' and 'Personal and Social') at the centre of the curriculum model and these will underpin learning across all 'areas of learning' (subjects). With these Core Competencies as a foundation, the curriculum in each area of learning is structured around a Know-Do-Understand model, encompassing its:

- Content ('Know')
- Curricula Competencies ('Do', denoting skills students will develop during and across different grades)
- Big Ideas ('Understand', outlining principles and key concepts covered during a grade in an area of learning).

These reforms involve considerable changes to assessment, including:

- students self-assessing (and reporting) against the Core Competencies, at least once a year throughout their education
- the piloting of a proficiency scale ('emerging', 'developing', 'proficient' and 'extending'), for reporting students' achievements in each area of learning
- teacher assessment of the Curriculum Competencies
- Numeracy and Literacy Assessments replacing exams in Grades 10 and 12

- revised formats for Foundational Skills Assessments in Grades 4 and 7.

This study will particularly focus on student self-assessment in Core Competencies, acknowledging that this is only one aspect of a far wider range of reforms.

Students self-assess against the three Core Competencies – rather than knowledge. The competencies do not have to be taught explicitly; they underpin teaching and learning across the entire curriculum. The assessment approach is intended to:

- help students to self- and peer-assess and to draw students into the reporting process so that they take greater ownership over their learning
- increase students' engagement in their work
- prioritise formative assessment over summative tests
- promote positive language and 'strength-based' assessment, which describes what students can do, rather than grades
- ensure more students are able to graduate
- use a wider range of assessment methods (for example, shifting away from a reliance on quizzes and tests towards using observation, interviews, dialogue, e-portfolios and so on).

Province-level reforms are now being implemented in districts across BC, although individual teachers', schools' and districts' responses to the reforms vary considerably.

3.7.2 Agenda setting

What were/are the 'pain points' that this innovation sought to respond to?

The reforms respond to three main perceived pain points:

1. The need for pupils to develop competencies and skills, over and above developing lots of knowledge
2. The need for BC's education system to keep pace with other, leading education jurisdictions
3. The need to include First Nations communities.

i) Holistic development

BC's scores in international assessments, including the Programme for International Student Assessment (PISA), were world-class but policy makers and educators were concerned that students' excellent subject knowledge was not matched in terms of their personal and social development, nor their

ability to communicate effectively. Addressing this perceived shortfall and ensuring pupils have support to develop holistically has been a key focus for the reforms, which reposition the curriculum (and consequently assessment) as a means to an end.

"This is not about curriculum and disciplines being the end destination. The end destination is the process of 'becoming' as a human being. It's about human development, that's the end destination. The confidence around communication, around critical thinking, creative thinking, around social and emotional wellbeing, around personal identity – that's our destination. Curriculum? Of course, you can't critically think in a vacuum. That's the vehicle, not the end destination."

Maureen Dockendorf,

Superintendent of Literacy and Numeracy, Ministry of Education

Kim Schonert-Reichl, Professor in Human Development, Learning and Culture at the University of British Columbia's Faculty of Education, explained that another important dimension has been extending and formalising

the revised curriculum's role in helping pupils develop social responsibility.

ii) Keeping pace

The Ministry has stressed the need for BC to maintain pace with 'the 21st-century world', for example in its 2011 Education Plan, which set out the perceived need for a "more nimble and flexible [system] that can adapt more quickly to better meet the needs of 21st Century learners". Its revised curriculum focuses on 'education for the 21st century', making personalised learning, flexibility and choice key tenets. Many practitioners, administrators, academics and policy makers in BC are supportive of this rationale. For example, one District Assistant Superintendent⁵ argued that "the world is changing and education needs to change".

For the British Columbia Teachers' Federation (BCTF), the reforms could also help unify the curriculum, ensuring subject areas pull in the same direction. There was a sense that subject learning – especially at secondary level – had become too siloed, and that the reforms offered an opportunity to open up the curriculum so that it had greater overarching coherence.

⁵ A superintendent and their team are responsible for standards and budgets in a school district.

However, some are sceptical of these arguments. Glen Hansman, Immediate Past President of the BCTF said discourses pushing for modernisation had been around for decades:

“One should be sceptical of rationales like a rapidly changing world. The world is always changing. I could produce for you newspaper articles from this province from the 1960s that utilised similar sort of constructivist language around getting rid of siloed subject areas, students take charge of their learning and teachers being facilitators, rather than somebody at the front of the room sharing information.”

Glen Hansman, Immediate Past President, BCTF

iii) Inclusion of First Nations communities

Educators, policy makers and the wider public in BC are keen to improve the quality of education for, and about, indigenous communities. This has long been perceived as substandard across the province, and there is a fear that it will remain so unless curriculum and assessment reform places this issue front and centre. The motivation

to reform the curriculum and assessment came from concern that lower graduation rates among First Nations⁶ students were a result of a lack of engagement and support for them within the school system:

“What did they need for support? When we started to look into the data, it was around engagement, culture and climate, feeling welcome in school, seeing their culture reflected.”

Jennifer McCrea, Assistant Deputy Minister, Ministry of Education

In addition to the three reasons for reform outlined above, while most educators see teachers’ practice as a strength of the BC system, there is a feeling among educators that pedagogy needs to evolve. In particular, teachers’ assessment practice has been deemed by policy makers and many educational thought-leaders to be too dependent on a narrow range of practices that could not support new system-level priorities such as inclusion, the co-construction of learning and holistic development. Without wholesale reform of the curriculum and

6 ‘First Nations’ is a term encompassing many different indigenous groups. For more information, see: <https://www.welcomebc.ca/Choose-B-C/Explore-British-Columbia/B-C-First-Nations-Indigenous-People>

assessment regime, some argued that teachers would not be incentivised to develop their pedagogy. Specifically, they would not be incentivised to move away from more ‘traditional’ teaching methods.

Students’ self-assessment against the Core Competencies reflects these various rationales, placing greater emphasis on individual students’ needs and experiences.

Who and what influenced this agenda?

i) The philosophical foundations

The reforms’ foundations were laid in the 1988 ‘Sullivan Report’,⁷ written because of the ‘unprecedented challenges’ posed by shifting economic and technological trends. The Ministry of Education responded with Year 2000: A framework for learning, a plan for reforming education by enabling ‘learners to develop their individual potential and... contribute to a healthy society and a prosperous and sustainable economy’.⁸ It emphasised the need for personalised and inquiry-based learning, and for pupils

7 Royal Commission on Education (1988) A Legacy for Learners: Summary of findings, Victoria: Royal Commission on Education, available at: <https://crofsblogs.typepad.com/files/legacyforlearnerssummary.pdf>

8 Ministry of Education (1990) Year 2000: A framework for learning, Victoria: British Columbia Department of Education, available at: <https://files.eric.ed.gov/fulltext/ED357999.pdf>

to develop ‘the thoughtfulness, creativity, flexibility, motivation, and responsibility that characterize educated citizens’. However, many parents and teachers expressed concern about moves towards ‘anecdotal reporting’, and universities were unhappy with the extent to which students were prepared for their studies.⁹ Consequently, the New Democratic Party (NDP) abandoned many of the Year 2000 pledges when it took office in the early 1990s.

Nonetheless, influential educators and policy makers who developed the current curriculum and assessment reforms cite both the Sullivan and Year 2000 documents as influential in shaping their thinking. For example, Jordan Tinney, Superintendent in Surrey, the largest school district in BC, explained that the documents’ recommendations were “foundational” in shaping his and others’ views about education, and that many of the ideas were “dusted off” in 2011.

Maureen Dockendorf was seconded to the Ministry of Education to help engage school districts and teachers in the reforms, and explained that the Year 2000 text “set the tone” for the latest reforms:

⁹ See Sanford, K. and Hopper, T. (2019) *Educational Transformation: The BC story*, Victoria: Houlihan Publishers; Kilian, C. (2011) ‘The last time we had an education commission’, The Tyee, 9 March, available at: <https://thetyee.ca/Opinion/2011/03/09/LastEducationCommission>

“[The Year 2000 report] was just way ahead of itself and it was a desire to make public education meaningful to all learners. It was a desire to ensure that they were engaged, that they had a sense of agency, that they had voice. ... So the whole focus around social and emotional wellbeing, social responsibility, personal awareness began in the nineties.”

Maureen Dockendorf,
Superintendent of Literacy and Numeracy, Ministry of Education

ii) Progressive educational philosophy

Politicians, educationalists, academics and senior policy makers attribute the reforms to the growing popularity of progressive educational philosophy.

In the late 2000s, the ideas of John Abbott (author of *Overschooled but Undereducated*¹⁰) and other progressive educationalists were taking hold among BC politicians, policy makers and teachers. For example, former Minister of Education, George Abbott, recalls “considerable affection” for Ken Robinson’s arguments about creativity and individualised learning:

¹⁰ Abbott, J. (2009) *Overschooled but Undereducated*, London: Continuum International Publishing Group.

“The kind of Ken Robinson approach to revitalising education I think had significant currency in British Columbia for probably a couple of decades.”

George Abbott, former
Minister of Education

Notably, these ideas influenced Rod Allen, a teacher, Principal and Superintendent turned Associate Deputy Minister in the BC Education Ministry. Rod worked closely with Deputy Minister James Gorman (who was, at the time, concerned about his Grade 4 son’s lack of reading ability) and with Minister George Abbott to design and pass the curriculum and assessment reforms. Gorman identifies Allen as “the intellectual leader” behind the reforms.

Organisations such as the Global Education Leaders’ Programme (GELP) and the Canadian 21st Century Skills Consortium were also influential in shaping thinking underpinning the reforms (including at the Ministry), adding to a desire for radical transformation of the BC system. For example, Allen and Gorman asked Valerie Hannon, the founder of the UK’s Innovation Unit and co-founder of GELP, to speak to Cabinet in order to provide an ‘external voice’ and gravitas.

iii) Political timing and buy-in

While Gorman attributes what he sees as the reform's successes to "good manufacture and good design", he also believes it was "good luck and good timing, too". When Premier Gordon Campbell¹¹ asked Gorman to run the Ministry of Education, he told Gorman "your job is just to keep it quiet and keep it out of the news". After the 2006 election, Campbell reprimanded Gorman, saying – in Gorman's words – "why the hell are you not making substantive change to education?" Gorman replied saying the Ministry was ready to do so, and sought Campbell's support in securing the political buy-in from Cabinet:

"I said, 'Look, I've got people busting at the seams to make substantive change here', and I said, 'If you're interested in doing it, then you help me bring Cabinet along and we'll get there.' So we started this process and we must have made 15 trips to Cabinet, pitching how we were going to do it."

James Gorman, former Deputy Minister of Education

¹¹ The Premier is the head of the provincial government. In Canada, provincial governments wield a lot of power, with responsibility for education, health and other public services.



George Abbott entered the Ministry for the first time in 2010, and suggests that "when I arrived in 2010, 2011, the time was really right for this to move forward". While not an education expert, Abbott described feeling an immediate affinity with the work already underway within the Ministry. He and Gorman recall how Premier Campbell gave them the space to develop ideas, and political support to achieve buy-in from Cabinet. They achieved this by asking influential thinkers to present to Cabinet, and Gorman explained how speakers such as Valerie Hannon "did a beautiful job" of setting out the need for change.

Abbott lost the provincial government leadership battle to Christie Clark in 2011, but was given his pick of Ministries afterwards and an assurance that Clark would support his agenda for reform. Abbott returned – by choice

– to education. Clark was subsequently 'hands off' with education and, because the rationale for reform had already been communicated effectively to Cabinet, Abbott and his team were left to proceed unimpeded:

"[Clark] was quite happy to largely let me go in terms of what we were trying to do. And again, what we were trying to do was not unfamiliar to caucus and Cabinet because of some of the speakers that we had at different caucus sessions and so on. ... There wasn't any resistance from the Premier's office under Clark. ... I don't think she had any interest in battling... She might battle on other political issues, but she would not I think expend a lot of political capital on trying to influence educational reforms that clearly were well defined already and [well] perceived."

George Abbott, former Minister of Education

Perhaps ironically for a party associated by some with austerity and a reputation for squeezing public services, reforms in education took on a wide-reaching, ambitious and progressive bent. One senior policy maker within the Ministry of Education said ministers

seized the opportunity to reform education. This high-level political support was vital in the reforms gaining traction.

iv) Involving the British Columbia Teachers' Federation and other stakeholders

Policy makers and politicians acknowledged that, in theory, the reforms could have been designed without BCTF support. However, the Ministry's relationship with the BCTF was fundamentally important in both symbolic and practical terms. Relations between the Ministry and the BCTF were woefully poor when Abbott arrived. Abbott said that his first call when he arrived at the Ministry was to Susan Lambert, then President of the BCTF:

“The first call I made as an education minister - and I think it was in retrospect a really wise thing to do - was to make my first call as a minister to Susan Lambert, who was the President of the BCTF, just to say, ‘Hey, this may surprise you, but I wanted to make my first phone call as a minister to you. It’s no secret that our government has had a poor relationship with you for a long time. I’d like to see if we can change that.’”

George Abbott, former
Minister of Education

Abbott saw the reforms as an opportunity to build bridges with the BCTF, and he and others (including Rod Allen and James Gorman) believe that the curriculum and assessment reforms helped achieve this even during ongoing and acrimonious negotiations about other areas of education. Even though Lambert always remained “cool” towards the Ministry, dialogue about the reforms helped build trust with her deputies, both of whom later became BCTF Presidents and who speak favourably of the Ministry's work to involve the BCTF in the reforms.

According to George Abbott, other stakeholder groups also played a vital role in building support for the reforms, including the BC School Superintendents Association and the British Columbia School Trustees Association. Abbott spoke in person to these groups, attending meetings to update them on the Ministry's work:

“One of the things that I like[d] to do, particularly when attending their conferences, [was] give them lots of time. I would spend the whole day there. When they ask me to speak, I’d say, ‘I’ll take questions until you’re tired of [me].’”

George Abbott, former
Minister of Education

Abbott, Gorman and Allen (among others) all felt that Abbott's investment in communicating and building bridges with these key groups had been hugely important in building support for the reforms.

v) Sociocultural imperatives

The social, cultural and political debates taking place in BC about the rights of, and reparations owed to, indigenous peoples were significant influences on assessment and curriculum reform. The First Nations Education Steering Committee (FNEESC) was guaranteed seats on each of the Ministry's curriculum committees. Furthermore, from the early 2000s, the grassroots Networks of Inquiry and Indigenous Education (as it is now called) were exerting considerable influence over the teaching practice of a small (but growing and prominent) group of teachers who held sway over the education debate. The sorts of practices these networks supported aligned ideologically with the Ministry's emerging ideas. As Markus Baer, Director of Assessment and Reporting at the Ministry, put it, “we don't do anything without the involvement of our indigenous educator colleagues”.

While the commitment to students self-reporting against Core Competencies emerged from the curriculum and assessment drafting and revisions process undertaken by various

committees from 2011 onwards, the idea's foundations were laid during the 1990s and early 2000s. The research and writing of Caren Cameron, Anne Davies and Kathleen Gregory had been influential among teachers and academics, who themselves were later heavily involved in designing the revised curriculum. For example, Dr Leyton Schnellert sat on the initial Curriculum and Assessment Framework Advisory Group, and explained:

“We have a long history of student self-assessment in this province. ... we did that work extensively in the late 90s and early 2000s and all through the 2000s, and so formative assessment and student self-assessment was something we'd been working on extensively. So the idea of self-assessment for the Core Competencies built on a practice that we were familiar with.”

Dr Leyton Schnellert, Associate Professor, Faculty of Education, University of British Columbia

Why was it considered an important agenda to pursue?

Different individuals and organisations expressed different motivations for

being involved in the reforms, and highlighted a range of ways the reforms have been important. However, there was lots of overlap in what they said.

For example, some – including academics involved in teacher training, district administrators and teachers themselves – talked about how curriculum and assessment reforms had the potential to reshape teaching practice by:

- raising students' awareness of how they learn, as well as what they learn
- giving students a greater role in their own learning
- encouraging students to develop the skills and capabilities associated with the Core Competencies (in turn enabling them to flourish at school and in life)
- repositioning teachers as facilitators of learning, rather than 'lecturers'
- encouraging collaboration between teachers, within and across departments and year groups
- increasing the inclusion of vulnerable and marginalised groups and in particular First Nations students.

Assessment reform, particularly the increased role of self-assessment against the Core Competencies, is seen as a way of embedding

these shifts in pedagogy. Importantly, it is hoped that all teachers – rather than an enthusiastic minority – will embed these approaches in their teaching. Sharon Jeroski, a former advisor at the Ministry and key architect of the Core Competencies curriculum approach, explained that self-assessment could ensure teachers were not overburdened while also gathering more valid information:

“We kept coming up against the issue of how secondary teachers, who already were feeling overburdened, were going to be able to somehow develop a reporting system for Core Comps for students who had as many as eight different teachers. We kept coming back to the fact that the only person in the process who had enough knowledge of all the various pieces to summarise and report on them was the student, and if somehow, we could find a way to make self-assessment the guiding process, there was a hope it would all work.”

Sharon Jeroski, former advisor to the Ministry of Education

The reforms also continue to be seen as important in reducing teachers', parents' and students' reliance on summative grades

and percentages. The hope is that formative feedback and self-reflection will gain greater prominence, replacing the current preference for summative judgements. It is hoped this will have a knock-on impact for students' wellbeing, and, particularly at secondary level, reduce the 'obsession' with summative judgements. In turn, policy makers, teachers, union officials and academics believe the shift towards formative feedback and greater self-agency among students has the potential to make assessment more inclusive. A number of contributors, including policy makers, practitioners and academics, felt the grade- and percentage-reliant system placed vulnerable and marginalised students – often including members of indigenous communities – at a disadvantage. This is because these groups often perform worse under such assessments, potentially driving disengagement.

Young people highlighted the stress that summative scores can cause:

“It’s really competitive. Someone will be like, ‘I got an 81.’ And other kids are like, ‘I got an 82 – hah!’ It’s really competitive. I also know some kids who [are] so stressed out about their marks that they can’t see anything past it.”

Secondary school student

Students' self-assessment against the Core Competencies also reflects the wider education reforms' overall goals of promoting:

- skills and competencies, alongside knowledge
- social responsibility
- inclusion
- the drawing in of a range of perspectives, including indigenous perspectives
- flexibility and autonomy for teachers and students
- student self-awareness and agency, with students taking ownership over and co-constructing their learning.

3.7.3 Formulation

How was a solution arrived at?

i) Key individuals within the Ministry

As described, above, a small team at the Ministry was instrumental in setting the vision for the reforms and ensuring they gained traction. George Abbott – Education Minister between 2010 and 2012 – was crucial politically, persuading his government's Cabinet to support the reforms. Within the Ministry, Nancy Walt – a senior policy maker – said that Deputy Minister James Gorman and Rod Allen worked closely with Abbott to shape

the reforms and initial approach. Equally, Allen, Gorman and Abbott (among others) all recognised the role that policy experts such as Walt played in moving the ideas into policy.

ii) Committees

Rod Allen emphasised the need for the Ministry to 'co-construct' the curriculum and subsequent assessment systems with a wide range of stakeholders. This involved establishing various committees to shape the curriculum principles and, later, subject-specific content and approaches to assessment. These committees involved individuals and organisations including:

- University-based academics
- The BCTF
- Indigenous communities
- Representatives from independent schools
- Superintendents.
- School administrators including Principals.
- Organisations including GELP and the Organisation for Economic Co-operation and Development (OECD).
- Employers and industry representatives.

Committees set ground rules, emphasising First Nations principles. Rod Allen explained that acknowledgement of (and adherence to) these principles was the “starting point for

everything”, but also helped build coalitions with indigenous communities. The first committee explored underlying principles that would go on to shape the curriculum content and ‘areas of learning’ while, later, committees had more specific remits.

The process was intended to be very open; for example, members of committees were not required to sign non-disclosure agreements, as can sometimes be the case with this sort of work. Instead, the Ministry encouraged committee members to talk to colleagues about what they were discussing and gather feedback for the next session. This was to build awareness and buy-in for the reforms, and also so that committee members could gather feedback from their peers and colleagues about the ideas.

One contributor noted the importance of ‘grassroot’ contributions to the process, saying “we would not have this curriculum” without such “deep partnership”.

The Ministry, and domestic and international experts also fed into the process including influential educationalists within the province including were brought into the Ministry to provide guidance on the reforms and shape the committees’ work. As the committees’ work progressed, the Ministry put draft versions of

the curriculum and assessment arrangements out for consultation. Early drafts had over eight million online hits.

There was generally said to be broad, philosophical alignment among committee members. For example, the co-chair of an assessment review group said:

“The Ministry had an agenda that it wanted to move forward, and that’s why [we] were asked to chair the committee in particular, because they thought we probably shared a political outlook on moving this agenda forward, around assessment.”

Blye Frank, Dean of the Faculty of Education, University of British Columbia

This shared ‘philosophical outlook’ meant that committee members often broadly agreed with one another. A member of the Ministry’s Curriculum and Assessment Framework Advisory Group, responsible for the initial design of the revised curriculum and assessment frameworks, explained the group did not start out from a ‘blank slate’. The Ministry selected people who could help build perceived best practice into the heart of the revised curriculum.

A senior policy maker with oversight of the different committees’ work said there had been two key disputes during the committees’ drafting phase. The first was in relation to the environment, and specifically the prominence (or perceived lack thereof) of content relating to climate change, and a clash with mining companies who felt their industry was not given due prominence. The second was in relation to maths, and concern that a cross-curricular approach would threaten standards. The former issue was resolved through raising the profile of climate change and discussions about industries’ roles in mitigating it in the subsequent iterations of the curriculum. The latter ‘died down’ in the face of broad consensus among other committee members.

The Advisory Group on Provincial Assessment submitted its final report in 2015, and recommended that self-evaluation be ‘embedded into the learning and assessment process’.¹²

The idea of student self-reporting against Core Competencies, emerged from this cycle of creating and revising curriculum and assessment framework drafts.

¹² Magnusson, K. and Frank, B. (2015) Advisory Group on Provincial Learning: The graduation learning years: Final report, available at: <https://curriculum.gov.bc.ca/sites/curriculum.gov.bc.ca/files/pdf/agpa-final-report.pdf>

iii) Research, Consultation, communication

Academics and policy makers working at or with the Ministry to design the reforms emphasise how important research was in shaping decisions about the curriculum and assessment. In particular, research published by the OECD, GELP and other international comparisons was influential.

However, some practitioners expressed scepticism about the presence and quality of research used to justify the reforms:

“I’ve searched and I’ve searched, [and] my opinion is that this has not been based on research; it’s not on peer-reviewed evidence. It’s based on philosophy or ideology.”

Secondary school teacher

Alongside the committees, districts (normally in collaboration with local branches of the BCTF) held information and feedback sessions with teachers, district administrators (including school leaders) and parents. Feedback from these sessions was fed up to the Ministry. However, this process of sharing feedback with the Ministry has been stronger in some districts than in others.



Universities were sometimes more sceptical, and had questions about how the revised assessment arrangements would affect admissions procedures and enable them to distinguish between applicants. However, some universities were more supportive than others and Jan Unwin, the Superintendent of Graduation and Transitions at the Ministry, was able to leverage support from certain universities to persuade those who were less supportive. Unwin also visited teacher training faculties at universities to speak to cohorts of new teachers, explaining: “Here’s the transformation, here’s why, here’s what we’re hearing, here’s what we’re doing, here’s what it looks like.”

Individual districts communicated about the reforms with teachers and schools,

parents and other local stakeholders including employers. This communication looked different in different districts, but included combinations of:

- meetings
- online communication, including emails and newsletters, and signposting draft content on the Ministry’s website
- paper communication (particularly with schools).

3.7.4 Decision making

Who was involved in deciding on this approach to the pain points?

Key decision makers tended to be ministers and their deputies, including:

- Rod Allen, the Assistant Deputy Minister of Education
- James Gorman, the Deputy Minister
- George Abbott, the Minister.

Part of the reason ministers played such a key role is that, despite the far-reaching scope of the reforms, only limited legislation was required. A senior policy maker explained that the key legislation relating to schools in BC (the School Act) is written at such a high level, it “doesn’t seem to need to be touched”. Consequently, the necessary legislation for the curriculum and assessment reforms could be passed as Ministers’ Orders, which “don’t have to go anywhere else”. She commented that legislation has been a sticking point in other jurisdictions attempting reform:

“They really get bogged down in this, with governments changing, or just [in] debate. We’re seeing this even in our neighbouring province, Alberta. This is playing out a little bit in Ontario. Lots of views [on] what curriculum should be, but we didn’t have that legislative angle at all. I think that surprises people.”

Nancy Walt, Executive Director,
Ministry of Education, British Columbia

With the political will and permissions in place from Cabinet, ministers could therefore unilaterally adapt and create the legislation required to implement the reforms. This was done for the majority of the reforms before the change in government in 2017, by which point sufficient legislative progress had been made to safeguard the bulk of the reforms. This is not the case for reforms to the Provincial Assessments at Grades 10 and 12. The legislation for this was not passed before the change in government, and this has proven a political sticking point since then, with the new provincial government scaling back on the initial promises to remove graduation assessments altogether.

Why were other solutions rejected and this approach preferred?

Some contributors during the committee phase advocated more ‘conservative’ suggestions prioritising knowledge and more didactic forms of teaching, and expressed concern about the emphasis in the new curriculum and assessment arrangements on skills and competencies, and student agency. In particular, they felt the reforms would pull the rug from beneath teachers’ feet, undermining existing good practice, and threatening standards. However, there was “critical mass” (as one Ministry of

Education Director put it) around the Core Competencies model, and a more skills focused curriculum and assessment regime).

A drawback of the high degree of consensus is that committees may have been vulnerable to ‘groupthink’. One contributor expressed concern that the ‘Overton Window’ – that is, the parameters for what it is and is not acceptable to say as an educator – have shifted towards more progressive views of education. This impacted on the reforms’ design, because more traditional views (regarding the importance of knowledge, memory and practice) were seen as outmoded. Some contributors argued that while those designing and advocating for the reforms talked about ‘consultation’ and ‘dialogue’, they in practice “just heard what they wanted to hear”. However, one senior policy maker argued that the curriculum still contains lots of content, but laid out in a different way.

Sometimes there were disputes in subject-specific committees about the content and structure of the ‘areas of learning’. Some felt the reforms were not going far enough, and essentially wanted to scrap subject labels altogether. Others expressed concern that the reforms were not applicable to their subject. For example, there was disagreement about the extent to which mathematics could be taught and assessed in a cross-curricular fashion.

Furthermore, one union representative suggested that “a lot of the stuff early on was sort of predetermined in terms of structure”, and because this pre-dated the BCTF’s involvement, many teachers were excluded from initial discussions. This mattered, because decisions were taken during these discussions regarding the nomenclature for the Core Competencies and the structure that would be used to present the revised curriculum. These provided the subsequent foundation for future committees’ work.

This generated some suspicion about how open the drafting process was. However, another commentator – a university academic involved in several of the committees – observed that some teachers who were resistant to the reforms had also complained during previous sets of reforms in the 2000s. She hypothesised that these teachers’ reservations stemmed more from a conservative outlook on reform and resistance to change, than an objection to the new reforms per se.

There has also been some disagreement with regards to Provincial Assessments. The government had previously announced that it would scrap these end-of-secondary assessments on the grounds that they ‘stifle creativity’ and result in teaching to the test. However, this decision was reversed in 2017

following the change in government. New Provincial Assessments (cross-curricular literacy and numeracy assessments at Grade 10, and a cross-curricula literacy exam at Grade 12) are now being phased in. The government decided to slow down the reform of Provincial Assessments in part because of concerns among parents (that is, voters) and universities. A senior policy maker within the Ministry explained this was “not the hill to die on” when far-reaching reforms to curriculum and assessment are underway across Kindergarten to Grade 9.

3.7.5 Implementation

What mechanisms have been used to implement the approach?

The Ministry is overseeing the assessment and reporting pilot, although day-to-day responsibility lies with individual teachers, schools and school districts.

Implementation is ongoing and involves:

- documentation, including curriculum documents and assessment guidance available on the Ministry’s website
- a Ministry-coordinated pilot of assessment and reporting arrangements for Kindergarten to Grade 9 in 13

different districts – this pilot has been extended by a year, to allow districts more time to embed changes

- district-level communication with schools, teachers and parents about the reforms
- individual teacher and school-level experimentation.

Implementation of differs between districts. Key differences include:

- **Timeframes.** Different districts have become involved in the pilot in different years, some as early as 2016.
- **Numbers of participating teachers and schools.** In some districts a high proportion of schools are involved. In others, the pilot only involves individual teachers and schools.
- **Approach.** Depending on the ‘pathway’ the district has selected, in some districts implementation is far more extensive than in others. For example, some districts have moved wholesale towards the new assessment and reporting frameworks across Kindergarten to Grade 9, while others are only adopting certain elements.
- **Communication.** Some districts are communicating more widely than others about the pilot and associated reforms. Some districts are focusing their communications

on teachers and administrators, while others are involving parents, students and employers too. Districts use different modes of communication, including in-person meetings, online communications including emails and videos, and paper communications including posters and letters.

Consequently, in some classrooms, students are self-assessing their learning against the competencies in most lessons; in others, this is a one-off activity at the end of term.

To what extent is the reified solution consistent with the planned/intended solution?

As yet it is unclear how the requirement for students to self-report each year will affect teachers' practice, and the extent to which teachers, schools and districts have been responding to the reforms and embedding the changes varies hugely.

Many are enthusiastic about the reforms and how they are playing out. Academics, teachers and administrators who engaged with the committees and consultation process are generally enthusiastic about the changes, saying the reforms represent a logical extension of work they were doing already. They believe the process for self-assessment against the

Core Competencies is appropriate as there is a lot of flexibility available to teachers and so the reform does not feel heavy-handed. Furthermore, teachers are still required to assess pupils in line with the Curriculum Competencies, providing pupils with feedback on their academic progress. However, even the teachers, schools and districts who wholeheartedly support the reforms, are finding their way, and implementation has been challenging.

What obstacles had to be overcome as part of implementation?

Interviewees talked about a range of challenges in relation to implementing recent education reforms including:

i) Geography

BC is a vast province, covering nearly 1,000,000 km² stretching up Canada's western coast. It has a population of around five million, and approximately half the population lives in the Vancouver metropolitan area. This means large areas are sparsely populated. This presents challenges for implementing and monitoring new policies such as the revised curriculum and assessment system. Some contributors observed that while implementing reforms is comparatively straightforward in urban and metropolitan areas, transport

links and other infrastructure challenges including broadband access can slow uptake in more sparsely populated and rural areas.

ii) Conceptions of assessment

The move away from allocating grades and percentages challenges many teachers' conceptions of assessment and not all teachers feel comfortable with the shift.

iii) Reliance on summative judgements

Many teachers, parents and students still prioritise summative judgements, including percentages and grades, over and above other forms of feedback encouraged by the reforms (including self-assessment). Even teachers supportive of the shift towards different forms of feedback acknowledged that parents' and students' views will take time to shift. Furthermore, parents' scepticism has been compounded in some districts by inconsistent reporting; some teachers provide narrative feedback, and others stick to grades. One student said her parents "hate" the new proficiency-scale-based system because they do not understand it.

iv) Teachers' confidence and knowledge

Many teachers are subject experts and experienced at creating and deploying

assessments in their subject areas. However not all teachers feel comfortable supporting other forms of assessment (including helping students self-assess against the Core Competencies). This challenge is compounded by difficulties in specific phases and subject areas as well as by time, with many teachers feeling they do not have time to give the reforms the attention needed to embed them effectively.

Many teachers are unsure how best to help students self-assess and report against the Core Competencies. The Ministry has published examples and guidance, and technology companies are offering solutions (for example, FreshGrade, an online reporting platform). However, this technological support has also proven problematic.

One teacher said he felt more resources were needed to help teachers (especially new teachers) to implement the changes.

While interviewees from virtually all schools and districts talked about struggling to update their practice, there is disagreement about how to respond to the challenge of a deficit in teachers' knowledge and confidence. Some interviewees said districts and schools do not have sufficient funds to support all teachers to shift their practice.

Others questioned whether offering support was worthwhile in the first place:

“I’m not going to set up educational systems for dinosaurs when mammals are thriving. They’re going to be extinct soon enough.”

Dr Kris Magnusson, Dean of Education, Simon Fraser University

v) Engaging teachers

Some teachers have been resistant to the reforms. One school Principal said that while some of her staff had whole-heartedly adopted the changes, other colleagues were “dragging their heels”. She believes that this stems from a combination of scepticism about the reforms themselves, and feeling the reforms are “top down”, and teachers not liking “being told what to do”.

This also stems from a sense among some teachers that the reforms are ideologically driven, rather than evidence-based. Some teachers requested more information from the Ministry about the reforms and felt their concerns had not been adequately addressed:

“We had this deputy minister [visit], [and] one of our colleagues asked, well, ‘Is there a robust, peer-reviewed research base that supports all these changes?’ [The minister...] clearly wasn’t ready for that question. He said, ‘Yes, there is, we’ll get back to you on that.’”

Secondary school teacher

A number of contributors less in favour of (or in some cases opposed to) the reforms explained a sense that alternative viewpoints were being shut out of the debate. For example, views that emphasise the centrality of subject-specific knowledge, memory and practice to learning, as opposed to cross-curricular and generic skills, not only felt their views are under-represented in the reforms, but that it is taboo to raise such concerns at all. This has alienated some teachers, administrators, parents and pupils with different views about what effective teaching looks like. Others counter that teachers' autonomy remains an important principle underpinning education in BC, and that knowledge and content remain important in the revised curriculum.

vi) Transferability across subjects

Most of the teachers and pupils we spoke to observed that the reforms impact on different



subject areas differently. For example, teachers explained that while English and social studies have been changed quite dramatically, fewer substantive changes have been noted in maths and science (particularly at secondary level).

vii) Curriculum design

While the volume of content in the written curriculum has in many areas been scaled back, it is still necessary to teach more than is specified in the key curriculum documentation. For example, one district administrator said that a teacher of Grade 6 social studies is now instructed to cover “huge ideas around global conflict and patterns of interconnection between developed countries”, “so, sure, it fits on one page now, but the ideas are massive”.

Giving another example, he explained that to understand electricity, pupils still need to know the same component content, whether or

not the curriculum specifies what this is. He said responsibility has shifted onto teachers to work out what this component content is:

“For example, in social studies, 12-year-olds explore ‘the anthropological origins of humans, human responses to particular geographical challenge including climate, landforms, features and characteristics of civilizations that lead to their rise and fall’. That whole learning outcome is stated in just one content bullet – but behind it is a great deal of specific ideas, vocabulary and skills. So the curriculum looks reduced, but the foundational knowledge needed to develop understanding doesn’t change, it’s just not as readily available.”

Neil Stephenson, Director of Learning Services in a school district

A lot of content is “buried in hyperlinks”, again meaning that while the curriculum itself is more succinct, teachers need to do more work to find out the content they are expected to cover. Indeed, despite a rhetorical emphasis on teacher autonomy and learner freedom through scaled-back curriculum content, some areas of the curriculum feel cluttered. For example, Neil Stephenson said that in Kindergarten:

“There are 28 big ideas (although these are not mandated), plus curricular and Core Competencies, plus you’re trying to develop literacy and self-regulation, right? How [can] teachers... develop and assess and design learning for 28 big ideas for a room full of six-year-olds? The curriculum requires a lot of interpretation and filtering to be useful by teachers. The end goals of the new curriculum are admirable; I do wish it was a little more streamlined to help teachers focus in on fewer essential learnings connected between grades and subject areas.”

Neil Stephenson, Director of Learning Services in a school district

Furthermore, there are some concerns about curriculum coherence since objectives

at different grade levels do not necessarily cohere with objectives lower and higher up the system. For example, a district administrator explained that Kindergarten outcomes in some subject areas do not align with those in Grade 1, which in turn do not link fully with those set out for Grade 2. In his words, the curriculum “does not hang together”.

viii) Inconsistency in assessment

While teachers and students said there had always been discrepancies in how individual teachers approached assessment, marking and reporting, students expressed concern that the new system would exaggerate these differences. This could mean teachers’ perceptions and biases exert greater influence over students’ outcomes:

“Each teacher teaches differently; they have different styles and different ways to share their knowledge with their students. Some classes might be harder than other classes, so for two students taking [the same class with] two different teachers, one might have 90% because that teacher is really easy or something, and there might be another teacher who is really hard... and the other student gets like a 60 or something.”

Secondary school student



Lots of commentators argued that the revised approaches to assessment (including the revamped Provincial Assessments) make them more accessible and, by extension, something a greater number of pupils can engage with. However, others (including teachers, parents and some pupils) suggested that even though the Provincial Assessments appear on students’ transcripts, they no longer need to be passed in order to graduate, which some pupils find demotivating. One parent felt the process of sitting an assessment that does not need to be passed “wastes times”.

Another issue is the modes of assessments themselves. For example, while gaming was in theory possible under the previous system, this risk is heightened especially with students self-reporting, and these assessments form part of their applications to universities or jobs. Glen Hansman noted that on several occasions

BCTF representatives raised concerns about students’ self-reporting at the secondary level, given the lack of exemplars from the Ministry.

Perhaps partly as a consequence (as well as because any new system takes time to get used to), there are concerns about certification and progression, particularly when it comes to university.

Despite self-reporting being used only up to Grade 9, some universities are concerned about interpreting and using the revised assessment outputs when offering prospective students places. We heard about some universities changing their entry criteria to make more of competency-based judgements about applicants’ suitability. However, students and teachers alike expressed concern that they do not know how universities in general will respond to the reforms:

“The thing that I was really concerned [about is that] they [the government] took off the Provincial Assessments and they just put the numeracy and the English assessments aside. We don’t know if the universities are looking at that. Now we’re just solely based on the teachers’ marks on us.”

Secondary school student

However, a policy maker at the Ministry of Education said that this will become clearer over time and welcomed universities' shift to a more 'broad-based' admissions process.

Some teachers and students suggested the reforms might 'move the goalposts' without tackling the underlying causes of stress when it comes to university applications:

"They look at your extra-curriculars on top of your academics, so that's why there's a lot of stress on us. If you want to get into UBC [University of British Columbia] there's so many things you have to qualify for."

Secondary school student

viii) Lack of clarity

Some teachers involved in the pilot have felt confusion or frustration during the pilot despite in many cases having self-selected into the pilot because they approve of the changes to the curriculum and assessment frameworks. This has resulted from confusing or contradictory guidelines about what was happening, including the emphasis on moving away from letter grades and a simultaneous requirement to 'convert' scale ratings. Others

cited an absence of communication about aspects of the pilot and felt frustrated that this is leading to duplication of effort.

"I clearly wasn't paying attention to some meeting somewhere obviously and I didn't realise that all of our proficiency scales would have to be converted to a letter grade at the end of the year. I didn't know that that would have to happen. So now at the end of the year we now have to take all those proficiencies and convert them to a letter grade."

Secondary school teacher

This is striking given that converting the proficiency scale results is not in fact a requirement.

Many of the challenges mentioned here are particularly pronounced at secondary level, where teachers are more accustomed to working in discrete subjects and reporting using summative grade or percentage judgements. Furthermore, teachers spend comparatively less time with their students at secondary level, which raises questions about their ability to accurately judge students' development under the competencies.

ix) Technology

Administrators, teachers and students talked about adopting electronic and online systems for assessment (such as sharing comments and photos of students' work), and reporting scale results and grades. There are teething issues, including a lack of alignment in how grades are reported on these platforms. One teacher said:

"The computer programme does not really fit the new assessment styles that we're trying to implement. So there were a lot of staggered starts and stops at the beginning of the year."

Secondary school teacher

There have also been challenges in getting teachers and parents to adopt the new systems.

x) Preparation for adult life

Students expressed concern about the fact they are 'guinea pigs' trying a new system out without being clear whether the new system is better. Some are optimistic and welcome the emphasis on the assessment of skills and competencies because they do not believe that simply learning content will help them flourish after school. However, others do not

believe the new assessment and reporting frameworks are beneficial, and one student even said she found the reforms “patronising”.

What conditions contributed to or detracted from effective implementation?

Interviewees highlighted several factors that have inhibited effective implementation:

- **The scale of the task:** The reforms involve revising an entire curriculum and assessment system which has been in place for years. This has created a huge volume of work and uncertainty for lots of teachers.
- **The substance of the reforms:** Some interviewees felt the curriculum and assessment reforms themselves are flawed.
- **Confusion:** Because of the scale of the reforms, some have experienced confusion and uncertainty about what needs to change and when. This has been particularly prevalent for the reporting pilot and for Provincial Assessments. With regards to the self-assessments, many are unclear how to embed these effectively, or what ‘good’ looks like.
- **Prioritisation:** Student self-reporting has achieved limited traction, in part because this is such a small part of a much wider set of changes, with many practitioners describing self-assessment as an ‘afterthought’.
- **Communication:** Many interviewees felt that the Ministry’s work was well-intentioned, but that sometimes the reforms were poorly communicated.
- **Funding:** There is not enough funding available to pay for cover and free up teachers’ time so that they can communicate, plan and review implementation.
- **Levers to influence practice:** BC does not have some of the ‘sticks’ present in other systems for incentivising change such as a schools inspectorate or league tables. Where enthusiasm and goodwill are lacking among teachers, principals therefore do not possess the levers to force uptake of the reforms by individual staff members.
- **Pull-back by government:** The change of government in 2017 has led to some political reticence, slowing down the final phases of implementation. This applies particularly to the reporting arrangements, with politicians nervous about the scale of change and removal of ‘universal’ benchmarks. While the new government is committed to seeing implementation through, the revised Provincial Assessments may more closely resemble their predecessors than had initially been planned.

3.7.6 Evaluation

How is success or failure being characterised?

Success is characterised in different ways by different people, at different levels but notions include:

- Improved communication between schools, parents, employers and universities
- Maintaining (and being seen to maintain) educational standards and league table positions.
- Reducing the perceived need for private tuition
- Shifting priorities and pedagogical approaches in the education system.

The extent to which members of each group are aligned behind these different outcomes and the extent to which they believe the reforms are the right way of achieving the outcomes are mixed. For example, teachers want to secure the best education possible for their students and some see the reforms as a helping with this goal, while others are less convinced.

What approaches are being used to evaluate the innovation’s success?

The Ministry is currently conducting a pilot in 13 school districts. This has been underway since

2016. The pilot has been extended for another year to give schools and districts more time to embed the changes. The extent to which findings from the pilot are being effectively fed upwards to the Ministry is unclear, and some contributors suggested there was a lack of formal mechanisms for doing so.

The Ministry is conducting surveys of teachers', pupils' and parents' understanding of the assessment changes and attitudes towards these. The results were not yet available at the time of writing. Ultimately, senior policy makers within the Ministry said the reforms' success would be judged using graduation rates and the province's performance in international rankings such as PISA. Although ministry staff and politicians talked about their hope that graduation rates would rise, some contributors including administrators and teachers expressed cynicism about this, saying that the reforms make graduation easier and represent a lowering of the bar.

Success is also being judged on a more ad hoc basis, including through school and district feedback to the Ministry, the results of surveys undertaken by organisations like BCTF, and potentially through election outcomes. However some commentators working at different levels in the system expressed concerns about how well feedback

from the pilot was being captured. Some were not aware of any formal mechanisms for passing feedback back up to the Ministry.

There is resistance to more structured evaluation because of deep-set ideological views held by many teachers and superintendents when it comes to comparing and benchmarking schools and pupils. These concerns are linked to fears about excluding vulnerable pupils. Representatives from the Fraser Institute observed that while many parents and even the Ministry are in favour of monitoring system performance, the BCTF's influence makes it difficult to introduce monitoring systems.

Is the innovation believed to be achieving its intended objectives?

At the moment, many are optimistic about the reforms, believing they legitimise and embed burgeoning good practice. However there are also some concerns about implementation.

BC's performance in the 2018 OECD PISA tests placed the province slightly below the Canadian average in reading, maths and science, although the province's performance places it among the top 20 jurisdictions within the OECD, in each area.¹³ BC's levels of

13 O'Grady, K., Deussing, M.-A., Scerbina, T., Tao, Y., Fung, K., Elez, V. and Monk, J. (2019) *Measuring Up: Canadian results of the OECD PISA 2018 study*, Toronto:

educational equity were below the average for Canadian provinces in the 2018 PISA results. It is however worth noting important limitations when it comes to Canada's PISA data¹⁴.

Lessons learned

Educational and philosophical beliefs and ideologies lie at the heart of the reforms currently underway in BC. Some say that the consultation process enabled a wide variety of views to surface, and that the curriculum landed in the middle ground so as to acknowledge a range of perspectives and beliefs. A critical mass of opinion and pre-existing pedagogic practice also prepared the ground for extensive reform.

Others argue there was little meaningful opposition while the reforms were designed, claiming that groupthink during drafting and revision cycles meant that certain perspectives were not sufficiently taken into account.

Closely aligned priorities were crucial in securing buy-in at all levels, including from ministers, academics, union officials, administrators and teachers.

Council of Ministers of Education, Canada, available at: https://www.cmec.ca/Publications/Lists/Publications/Attachments/396/PISA2018_PublicReport_EN.pdf

14 Jerrim, J. 2019. Is Canada really an education superpower? The impact of exclusions and non-response on results from PISA 2015. https://johnjerrim.files.wordpress.com/2019/12/canada_papaer_october_2019_final_wp.docx

The stars therefore aligned in terms of the macro political context and the perceived educational imperative.

While there is broad agreement on the priorities that underpinned the reforms, it is increasingly clear that different people are defining success differently depending on their position in the system. For instance, students characterise success as making a successful application to university, whereas politicians see maintaining the province's international league table position as an important goal.

Paradoxically, these far-reaching reforms have in some ways had only limited impact on practice. This is because teachers have had to do a lot of interpretation about what the reforms mean, but also because there are few formal levers for changing practice.



3.8 Innovations 8 and 9: Online assessment resource banks in New Zealand and Victoria, Australia

3.8.1 Introducing the innovation

In recent years, both New Zealand and the state of Victoria in Australia have created online assessment resources. Other Australian states have also introduced similar initiatives but this study focuses specifically on the work going on in Victoria and New Zealand.

Both jurisdictions have set up one or more websites through which teachers can access a selection of resources. A limited but growing set of assessments is available in Victoria and the assessments tend to be focused on core areas of interest within the curriculum such as literacy, numeracy or cross-curricula competencies. They are intended to help gauge and map-out progression. In New Zealand, a far larger bank of resources is available in an Assessment Resource Bank (ARB), which can be used to support day-

to-day teaching in a range of subjects, through diagnostic, formative assessment.

In New Zealand, assessment resources are available through one website. In contrast, in Victoria a range of different portals exist (such as the Insights Platform), with different tools. Moves are currently underway to bring some of these together in one 'Digital Assessment Library' (DAL) and to add an online adaptive platform to this. The first elements were due to be online in mid-2019, followed by a full platform from 2020. However, this was delayed and during the second visit conducted in October 2019 it was not yet clear whether the DAL will end up including all the different resources and how it will develop.

New Zealand's ARB includes 2,800 assessments, which can be completed online, 1,200 of which are self-marking. Each tool includes an overview of the topic being assessed, guidance on how to use the resources and links to the curriculum. Benchmarking information is also available, to help teachers understand common responses and misconceptions (for example, stating '80% of pupils correctly identified that... 30% had the following misconception, you might like to...'). Pupils can access resources directly when the teacher gives them a URL and teachers can go back to previously used resources to see how pupils performed or to use previous responses as a prompt for self-reflection among pupils.

Online assessment tools and resources in Victoria include:

- adaptive online tests known as ‘assessment on demand’
- assessment rubrics
- intercultural capability
- critical and creative thinking assessment
- ABLE (ability-based learning), which helps teachers and teaching assistants make judgements about communication, literacy, interpersonal and intrapersonal skills and cognition among students with additional needs.
- a Transition Learning Development Status tool, which all government-funded early years provision must complete and which is delivered through the Insights Platform
- a guide to formative assessment designed by Melbourne University’s Assessment Research Centre (ARC), which includes examples designed in collaboration with teachers.

While many of the assessment resources that are available in Victoria may appear to be summative from an English perspective, in Australia they are viewed as formative because they provide immediate and actionable information for teachers. In both New Zealand and Victoria, the formative dimension of the assessment resources is underlined by the fact that the reporting of assessment information

(to teachers and pupils) is generally seen as equally important as, if not more important than, the assessment itself.

3.8.2 Agenda setting

What were/are the ‘pain points’ that this innovation sought to respond to?

i) Improving educational outcomes

Victoria has currently got a high-level priority around becoming ‘The Education State’ and this involves improving achievement in reading and writing (in the national ‘NAPLAN’ assessment – National Assessment Program – Literacy and Numeracy), as well as in arts and creative thinking. Some argue that this agenda is about securing ‘excellence’ rather than tackling specific problems.

In 2012, a review of Australian education by David Gonski – ‘The Gonski review’¹⁵ – highlighted a number of negative trends in the educational achievement of Australian students, namely:

- a decline in performance in comparison to international benchmarks
- a ‘concerning’ proportion of students not meeting minimum standards of achievement

¹⁵ Gonski, D. (2012) Review of Funding for Schooling: Final report, Canberra: Department of Education, Employment and Workplace Relations (Australia).

- a ‘significant gap between its highest and lowest performing students; far greater than in many OECD countries’
- a link between low levels of achievement and students from low socioeconomic and indigenous backgrounds.

While most of the recommendations focused on increasing funding, high-quality formative assessment was seen as a key way of improving teaching and learning in the context of large attainment gaps within each classroom.

The so-called ‘Gonski 2.0’ report was published in 2018¹⁶ and had a greater focus on assessment, though the Victorian Curriculum and Assessment Authority (VCAA) reports that this “confirmed they were on the right path” rather than the report driving the online assessment bank project. One recommendation of the Gonski 2.0 report specifically suggests the need for the government to ‘develop a new online and on demand student learning assessment tool based on the Australian Curriculum learning progressions’ (p xiii). It recommended that the tool should focus on measuring ‘growth’ (progress) and that it should support teachers

¹⁶ Department of Education and Training (2018) Through Growth to Achievement: Report of the Review to Achieve Educational Excellence in Australian Schools, Canberra: Australian Government, available at: <https://docs.education.gov.au/node/50516>

not just with assessment, but also with tailored next steps:

‘The tool would assist them to readily identify the stage of learning a student has reached and to provide a choice of possible appropriate interventions from which the teacher could select to provide the next challenging but achievable learning task.’ (p 61)

The focus on signposting towards next steps is also an increasingly important goal behind the ARBs in New Zealand.

The focus on assessment resources around intercultural and creative skills in Victoria ties into a widespread (and international) interest in cross-curricula skills and capabilities that has also been noted in our study of British Columbia’s reforms. Teachers’ appetite for these agendas is demonstrated by the fact that trials for new assessments in these areas have been over-subscribed.

ii) Improving the quality of assessment practice

The Gonski 2.0 review suggested that the lack of an ‘online learning assessment tool’ is a key issue in Australia, highlighting that there is ‘uneven access to validated formative assessment tools’ and that ‘teachers... rely on a mix of locally developed... idiosyncratic tools that do not provide the most reliable measure

of progress’ (p 63). This was highlighted by the teachers and leaders we interviewed who reported having to dedicate many hours to:

- designing assessment resources, duplicating work that is done in all schools but rarely shared
- attempting to analyse the results of summative tests such as Progressive Achievement Tests (PAT tests) and NAPLAN in a way that provides diagnostic information for teachers, despite the fact that they are not designed for this purpose.

On the other hand, according to some educationalists in Victoria, the DAL and Insights Platform is also intended to address shortcomings in monitoring pupil progress, as well as teachers’ difficulties planning and delivering the curriculum. This issue is linked to:

- a top-level and relatively non-prescriptive curriculum
- lack of clarity about progression among teachers
- poor-quality assessment resulting in wasted time and inconsistent scoring and reporting, which means a lot of time is wasted ‘weighing the pig’ and producing data that is not reliable, valid or comparable.

According to the VCAA, lack of monitoring is problematic because it makes it hard

to identify and learn from good practice and it is this, rather than a need for more accountability, that is driving the agenda (since accountability is being addressed through a separate reform). On the other hand, if – thanks to better assessment and monitoring – teaching and learning improve, NAPLAN results might also improve.

iii) Implementing the curriculum

The link between a relatively non-prescriptive curriculum and a perceived need for clarity about progression is common to both Victoria and New Zealand. Schools in Victoria have a particularly high degree of autonomy compared to those in some other Australian states, making it fertile ground for the innovation. Teachers in Victoria were particularly keen for support in making sense of the link between ‘learning areas’ and ‘capabilities’, which were intended to mesh together subjects and cross-curricula strands. Formative assessment resources are seen as something which could fill this gap without making the curriculum more prescriptive or increasing accountability.

Issues due to loose curricula are particularly pronounced lower down the age range, resulting in many schools just mapping back from secondary level, or having to invent their own frameworks, even if this is very time-consuming and inefficient. Saving

teachers' and principals' time is therefore an important driver behind the innovation.

There are concerns in some policy circles that autonomy has worked well for some schools but not all of them. Both the VCAA and the Ministry of Education recognise a groundswell in calls for support in understanding curriculum implementation. Assessment tools are seen as a key response to this. There also appears to be a growing interest in curriculum design, which is closely linked to the need for carefully planned assessment tools – particularly in order to map out progression.

These views were echoed by teachers we interviewed in Melbourne. Teachers viewed a non-prescriptive curriculum as having benefits and drawbacks. Benefits mainly centred around teacher autonomy and the freedom to adapt teaching to pupils' current levels. This is seen as particularly important in a system that does not set or stream pupils. However, despite this, we found that teachers still wanted their practice to be more informed by accurate assessment data and judgements:

“We want to work towards data informed teaching to pupils at their point of needs, rather than teaching the curriculum to the middle of the bell curve.”

Although teachers we spoke to were not aware of plans to develop the DAL, they recognised the potential benefit such a resource could have in supporting them to implement the curriculum.

It was clear that teachers felt that the lack of prescription in the curriculum made it difficult to make accurate assessments and judgements:

“You just don't know where your students stand against others. Compared to countries where you have that detail fleshed out in the curriculum where you can say for sure 'I know that you have to do fractions in Year 8, I know it's this kind of thing, I know what that looks like'. You haven't got that objective description to compare against.”

Additionally, teachers highlighted further downsides to having a non-prescriptive curriculum, mainly in relation to high workload and inconsistency across and within schools, leading to a situation in which no one felt sure they were delivering the correct content at the right pitch.

“We don't have a good centrally provided curriculum. We have the Australian curriculum and the Victorian curriculum but it will say for example 'you need to teach magnetic fields', then a full stop and no additional resources, so every school teaches magnetic fields and they might get other resources but there is huge variety in what that looks like.”

For some, this led to new teachers struggling with the challenge of interpreting the curriculum and the workload of creating assessments based on this interpretation, and more experienced teachers disregarding changes to the curriculum:

“It's difficult if you're a new teacher trying to find your way through. If you've been around for a little bit longer they have more idea of what level to teach that content... or they pin it to the old curriculum that they just roll out every year because there is nothing else... If they've been through a few curriculum changes [they] don't so much see the need to change things in response to curriculum changes.”

Overall, teachers felt that assessment resources – primarily rubrics, tasks and model responses to tasks – would be useful when interpreting the curriculum levels and would not have a detrimental impact on autonomy. They also did not appear concerned that extra assessment resources would be interpreted as an attack on autonomy or a push to greater accountability.

“Because we don’t have a lot of accountability tied to testing, any resources out there are going to help me curriculum planning. The problem now is that that’s not there. I don’t think it would harm the autonomy and curriculum design; it would complement it. It would be good for teachers to know if... they interpreted the little dot point of the Victorian curriculum correctly. In some cases they would see that they weren’t interpreting it right and that would be a useful conversation.”

They felt that comparative data from other schools would be particularly useful in order to provide information in advance of NAPLAN or Victorian Certificate of Education (VCE – GCSE equivalent) tests.

iv) New Zealand’s shifting ‘pain points’

The ARBs in New Zealand have sought to

address different pain points over time, with different priorities dominating for different stakeholders. For the New Zealand Council for Educational Research (NZCER) (which manages the ARBs), the primary purpose has always been pedagogical, in terms of providing support to teachers in a non-prescriptive form. However, different governments have had different goals. In particular, during the late 1990s, the government saw the ARBs as a tool in a move towards closer monitoring of national standards, particularly at key ‘transition points’. However, Cedric Croft, one of the architects of the ARBs, explains that this goal was contested (in his tellingly titled article “The Assessment Resource Banks: From National Testing to a School-Based Resource”):

‘There was no deep conviction within NZCER that intact tests [a set of tests within the ARBs that could be used as part of national testing] would necessarily be of high educational value... but they were a pragmatic response to the continuing notion of national testing, which they regarded as somewhat flawed.’¹⁷

17 Croft, C. (2002) ‘The assessment Resource Banks: From National Testing to a School-Based Resource’, New Zealand Annual Review of Education, 11, pp 229–43.

The early stages of the ARBs in New Zealand can therefore be characterised as a compromise between two competing agendas: one focused on pedagogy, the other on accountability, with some seeing ARBs as a way of holding off the ‘threat’ of national testing.

This shifted as the Labour party and teaching unions drove forward an alternative agenda around strengthening teacher professionalism and ‘authentic learning’, which they argued were under threat. This alternative pain point has since come to dominate, with the new Labour government abolishing the national standards. This has allowed the ARBs to move further towards NZCER’s preferred pedagogical agenda, for example helping teachers understand the ‘nature of science’ strand of the curriculum that had previously been neglected, perhaps due to a lack of guidance. This has close parallels with interviewees in Victoria, who also noted a tension between pedagogical and accountability-based agendas for the DAL.

Other groups continue to hold differing agendas and to point to different pain points. Some politicians and influential groups (such as the think-tank ‘The New Zealand Initiative’) see closer monitoring and better use of more standardised and reliable data as routes to tackling underachievement – particularly among Maori pupils. Meanwhile, others in

government recognise that New Zealand has an unusually well-developed data infrastructure. They argue that this needs to be used better.

Some in government also believe that the current curriculum results in an ‘anything goes’ culture and that more (quality-assured) guidance is needed. Over the course of this study, the ongoing education consultation also led to further emphasis on the notion of progression and the need to map out pupil trajectories and monitor progress rather than what thresholds they are reaching. This shift in agenda will have implications for the ARBs and it is not yet clear how this will play out, though it appeared that, in future, online assessment resources would be centred around new progression frameworks rather than the current platform.

Who and what influenced the agenda?

The overall culture around education policy in Victoria and New Zealand has played a critical role in shaping the agenda around ARBs. Both jurisdictions prioritise school and teacher autonomy and online assessment resources are seen as a way of guiding and supporting without prescribing. In New Zealand, investing in experts, including organisations like NZCER, is seen as a way of equipping teachers to act in a professionally autonomous manner;



despite the fact that, in other contexts, commissioning an external organisation to develop assessment resources might be seen as removing teachers’ independence. Meanwhile, both jurisdictions are keenly aware of the close link between curriculum, progression and assessment, and online assessments are seen as a way of tying these together.

In New Zealand, prioritising school autonomy ties into the country’s historic, geographical and cultural context. Since the country is largely rural and schools are closely tied to their communities, they are expected to tailor their provision to their communities, particularly their Maori communities given the country’s history. High degrees of autonomy can therefore be linked to an aversion to anything

that might be seen as cultural imperialism or domination. This approach is embodied in the country’s curriculum, which is described by some as ‘beloved’ and ‘world leading’ – despite others seeing it as too high level. An aversion to hard accountability also means that the ARBs give teachers (rather than schools or managers) control of the data that is produced.

As noted above, teachers have called for more guidance, particularly in terms of how they should enact the curriculum, as well as in order to reduce the workload associated with developing assessment resources. The fact that teachers and schools have been willing to pay for tools provided by the private sector has helped demonstrate their keenly felt need for a solution and the importance of this agenda. Independent Schools Victoria and Catholic schools have also helped to articulate and mobilise demand.

Meanwhile, there is a live debate in Victoria, including in the media, about how schools report to parents. Schools are free to do this how they wish, yet parents are often dissatisfied with the way they receive information. Non-prescriptive nudges towards standardisation can therefore be seen as a response to parents’ demands, but this is held in tension with many educationalists’ aversion to standardised grading. Assessment tools such as rubrics

to map progression are therefore seen by some as a way of negotiating this tension.

Teachers and parents have therefore both played a role in building the case for an online assessment tool in Victoria, though some argue they have been important supporters rather than a driving force.

3.8.3 Formulation

How was a solution arrived at?

In both jurisdictions, current work on assessment resources is part of a long-running, iterative process. As we have seen, in New Zealand, the focus of the ARBs evolved over time and in Victoria, current developments are partly a response to dissatisfaction with, or a need to evolve, the current platforms and products (with some seeing the existing platforms as a ‘sunk cost’).

In Victoria, various platforms and tools have evolved in a piecemeal and somewhat ad hoc fashion, with a range of assessment tools emerging, some from the government and some from private providers. Schools therefore have a range of solutions to choose from – for example, many use a product from the Australian Council for Educational Research (ACER). The DAL therefore evolved to draw together and upgrade existing ‘part-solutions’ and provide a state-sector, state-level solution.

In New Zealand, schools and teachers are particularly powerful players in education policy. Approaches that do not sufficiently respect their autonomy are therefore actively resisted – for example, over 200 schools (a sizeable number in the New Zealand context) refused to implement a previous set of education standards.

Research institutes have played a particularly important role in formulating a solution in both jurisdictions, whether the ARC in Victoria or the NZCER in New Zealand. The ARC has had a long-standing interest (and considerable academic expertise) in formative assessment and developing tools to assess cross-curricula skills and capabilities. The VCAA’s strong reputation strengthens the initiative. Meanwhile, the NZCER has been working on the ARBs for nearly three decades, since 1993, though its involvement with assessment in New Zealand stretches back to the 1970s.

3.8.4 Decision making

Who was involved in deciding on this approach to addressing the pain points?

The case for the DAL and a \$17 million investment in the platform was primarily made by the Head of the VCAA to the minister and minister’s office. By arguing that assessment was “the missing piece” in terms of helping

teachers understand where pupils are so that they can raise standards, David Howes (VCAA CEO) helped ensure the DAL became a “pet project” for the minister, who is said to be well-respected. Effective lobbying seems to have been a driver in moving along the ‘long and torturous road’ towards funding. It was suggested that this may have been helped by the fact that the minister’s wife is a teacher, giving the minister first-hand experience of issues such as workload that result from a lack of assessment resources.

The detail of which assessments feature in Victoria’s bank of resources has been influenced by policy makers’ interest in particular areas – for example, critical and creative dimensions of learning.

In New Zealand, although ministers develop work programmes and come with their own priorities, there is a culture of consultation, which means that if the profession and the public do not back reforms and initiatives, they may not go ahead. Sector and professional support was therefore crucial.

Why were other solutions rejected and this approach preferred?

Teachers themselves make decisions about which resources to use in both jurisdictions. Use of ARBs therefore

depends on teachers seeing them as more useful than alternative products.

In Victoria, at one point there was a grander plan, to pull together a range of tests from different providers, including PAT tests from ACER, but this grand dream was moved around from one person to another and abandoned for practical and software reasons, since it was “like bringing together a set of parts from different car manufacturers”. Another reason for this was said to be that ACER, which owns PAT tests, which are used by around half of schools in Victoria, was reluctant to hand over the management of its tool. It was concerned that the platform would be managed by tech experts, rather than assessment experts, and thus PAT tests would not be updated and maintained to a high-enough quality. ACER also believed that insufficient consideration was given to the reporting side of assessment as the plan focused on the analysis that would be possible with large sets of data rather than how teachers would use the platform on a day-to-day basis. It decided it had no motivation to join the platform and, as time went on, the idea of the creating a bank of new resources (the DAL) emerged instead.

There was also a question over whether the ministry or the VCAA should run the DAL but because the VCAA covers all schools whereas

the ministry only covers state schools, the VCAA was considered a better home. Questions remain around how the data from assessments should be collected and used. It is currently stored and this feeds fears that it could be used to monitor teachers. This potential challenge has been tackled in New Zealand by linking all data to teachers’ individual accounts (rather than schools), allowing them to delete the data if they wish to do so. Meanwhile, although the platform and the data are owned by the New Zealand Ministry of Education, the NZCER is the data custodian and would anonymise data before sharing it with the Ministry. However, the NZCER notes that the Ministry has never requested access to ARBs’ data.

In New Zealand, questions around funding also affected the formulation of the solution. It was a battle to secure funding to create online ARBs, since making an online resource is not the same as putting paper resources online; at one time the Ministry simply wanted online, markable responses. However, the NZCER wanted to continue a formative focus and felt that this required a more sophisticated (and expensive) approach. Chris Joyce at the NZCER therefore had to put in multiple requests to maintain funding and compromise was needed, for example when it came to ownership of the platform. The ARBs also shifted into the remit of the Ministry’s digital

team instead of the curriculum team. Later on, there were questions around whether the new “Progress and Consistency Tool” (PACT) should be part of the ARBs; however, this idea was rejected because of PACT’s links to the National Standards towards which there was considerable resistance.

3.8.5 Implementation

What mechanisms have been used to implement the approach?

i) Developing the initiatives

Implementation in New Zealand has proceeded in several phases:

1. **1993/94–1998:** The curriculum framework/ vision was produced as well as individual subject curricula. Some argue this was very challenging for teachers at primary level because each was very thick. It gradually became more coherent as curricula were more tied to the overall framework. The first feasibility studies for the ARBs were conducted in 1993–94 and the first ARBs were made available in 1995. During this time, the ARB remained mainly summative and multiple choice based.
2. **1998–2003:** Assessments gradually became more formative.

3. **2003:** ARBs became explicitly formative in intended purpose.
4. **2006–07:** The new curriculum led to a remapping of the resources. The Ministry wanted explicit links between resources and National Standards. This led to constant remapping. When the curriculum was released, people said “Is it detailed enough?” The response was that a second tier of supporting materials was promised, but never came in any fully developed form. ARBs filled this gap.
5. **2009–10:** National Standards led to another shift in course.
6. **2018:** National Standards were abandoned and the aim shifted to ‘student-centred formative assessment’. The focus of resources therefore began to move from ‘What do students know?’ to ‘What can they do?’. It is this phase of the innovation that we have focused on.

Over time, things shifted away from downloading resources and towards online tools. Resources also became more sophisticated, for example allowing students to draw things online or introducing auto-marking.

ii) Developing resources

In Victoria, the provision of online assessment resources began with literacy and numeracy

before moving on to other state priorities. Like in the early days of ARBs in New Zealand, simpler assessments (such as multiple-choice format) were also prioritised.

Teachers have been involved in developing assessment resources and tests in both jurisdictions, through both piloting and item creation. In both jurisdictions, resources are piloted in a mix of schools or with a range of teachers (low-, mid- and high-decile schools in New Zealand), while in Victoria, panels of teachers are selected to include teachers from both rural and urban schools. This has considerable implications in terms of needing to pay for travel but it is considered important and worthwhile.

In order to develop a replacement for the old ‘on-demand’ tests, the VCAA invited specific teachers (government school leads with a particular interest in the areas) to a two-day workshop. This began with a session in which a subject manager from the VCAA introduced the project. There was then some expert input on item development, which prepared teachers to review and develop items. The teachers then created items for the bank, which were subsequently reviewed and quality assured. Joe Pellegrino at the VCAA argues that:

“Teacher involvement genuinely improves the assessment, adding to their authenticity and credibility.”

On the other hand, he says it is important to accept that not all the material teachers submit will be up to scratch. Teachers therefore have to know they will get feedback and that this is part of professional learning.

In October 2019, the DAL was still in development, slightly behind schedule. Although the VCAA reported that the workshops had been useful, more items still needed to be developed, suggesting that this method alone had not been sufficient to fully populate the online bank.

Efforts to involve teachers in item development in New Zealand have had mixed success. Originally, teacher panels came up with ideas for the assessments and items, but this did not work at the time because they tended to come up with learning activities rather than assessments. This appeared to be because they lacked the curriculum and assessment expertise to create valid and reliable items. Part of the problem was that, as Jonathan Fisher of the NZCER explained, teachers had to think “not just how did pupil respond, but why did they respond that way”. Assessment resource developers

therefore had to become “researcher writers”. NZCER also pointed out that that resource development needs to be conceived of as a form of curriculum development because it involves mapping capabilities and concepts. This is particularly important given New Zealand’s broad, non-prescriptive curriculum.

All assessment resources were piloted and trialled in schools and this process included gaining an understanding of the reasons behind certain misconceptions and mistakes.

The ARBs in New Zealand are maintained by the NZCER team, with less input from teachers now, as new resources are not being developed. This is partly driven by funding constraints but it also highlights the challenges of depending too much on input from serving teachers. In time, the VCAA in Victoria may also have to respond to this challenge and move task creation away from teachers to expert teams.

Who are the different actors involved?

Bodies that are close to (and/or funded by) government play a critical role in both jurisdiction’s initiatives. These include non-governmental/semi-non-governmental actors such as the NZCER, Melbourne University and the VCAA. Meanwhile, alternative (and sometimes competing) solutions often come from the commercial sector.

To what extent is the reified solution consistent with the planned/intended solution?

The current innovation in Victoria looks different to the originally intended plan to bring all online assessments together under one platform. This original solution would have involved more focus on grouping different assessments and filling gaps whereas the current plan primarily involves creating new items to form the DAL.

In New Zealand where the innovation is much further along, decades from its initial inception, the ARBs have changed and evolved multiple times. The main gap between the intended plan for the ARBs and the current state of play lies in the extent to which the ARBs are updated and refreshed, according to the NZCER. Current funding levels allow maintenance of the ARBs rather than ongoing innovation and new resources.

What obstacles had to be overcome as part of implementation?

The need for constant refining and improvement is common to both jurisdictions. Interviewees in New Zealand pointed out the importance of good-quality digital infrastructure to organise the resources. This infrastructure now needs updating



and improving and in spring 2019 it was highlighted that what is possible will be limited by upcoming spending decisions. By autumn 2019, it was clear that funding would remain limited and therefore maintenance was all that was happening, despite the need for new items to be created to keep up with changes in the progression frameworks and the world at large – as well as to maintain teacher engagement in the ARBs, since it is new content that tends to drive engagement.

In Victoria, plans to develop the DAL have been in existence for a long time, but it has been hard to secure funding. Only now has this been secured - though it remains much lower than what was hoped for, and possibly what is needed. New Zealand provides a comparative or perhaps predictive lens for this issue: it is clear that funding requirements for online ARBs remain high even once the platform is up and running and if this funding is not available in the long term then it is hard to keep the resources at the desired standard.

What conditions contribute to or detract from effective implementation?

i) Prioritisation

In Victoria, the fact that the DAL is not more high profile (and is thought of as a ‘footnote’

in the budget) is considered potentially problematic as difficulties securing interest and exposure could result in low prioritisation.

Although crucial, the relationship between the Ministry and the VCAA has not always been easy and at one point became strained. This was partly because of a lack of clarity over respective remits (for example, whose role it was to say how the curriculum should be interpreted), partly because of specific individuals and partly because of a perception that the assessments were not being sufficiently prioritised by the Ministry. Changes to staffing have since resolved this problem. Changes to individuals within the Ministry have also caused challenges in New Zealand. Churn, in terms of individuals and in terms of curriculum/policy change and prioritisation, is therefore a considerable challenge in both jurisdictions.

In New Zealand, the development of a new, online digital curriculum is potentially a competing priority. In our first visit, we heard that for the last year or two, new assessments have not been developed and this is considered a potential threat to success. During our second visit, some in the Ministry suggested that if new assessments are to be developed, they would take a different form to the current ARBs and will likely be linked more closely to the new progression

frameworks and mapping tools that will replace the recently removed National Standards. However, new resources would draw on the learning from the current ARBs and maintain a focus on formative feedback for teachers.

ii) Long-term funding

As noted above, funding constraints (often driven by changing priorities) can have a huge impact on implementation. Items need near-constant updating and maintenance to ensure they evolve alongside the curriculum. This is particularly pertinent in a context where the curriculum is relatively non-prescriptive. Shifting priorities and global issues can considerably impact the interpretation of a loosely defined curriculum and what teachers teach - and thus, what assessments should cover. Resource banks can help teachers keep up with change, but only if they are sufficiently funded to do so. Without new content NZCER is concerned that engagement with the ARB will plateau or even decline.

iii) Assessment development expertise

The right combination of technical and educational expertise is considered important in Victoria, with comparisons drawn to other assessment initiatives that may have worked technically but did not work in a

classroom context. Recruiting teachers to work on creating the questions and tasks is therefore important and is currently a barrier to developing the new resources on time. The development of assessment on demand has also been delayed considerably.

Current technology and resourcing means that only the more closed assessments in New Zealand's ARB can be auto-marked online. However, the resources have moved away from 'correct/not correct' questions, making auto-marking functions more difficult.

iv) School buy-in: compatibility with other assessments

When considering what a new DAL would have to offer to secure engagement, teachers in Victoria felt strongly that new assessments needed to 'marry up' with other assessments such as NAPLAN.

For example, one Melbourne school explained it currently used teacher assessments, based on teacher-created learning tasks each lesson, as well as doing considerable analysis with NAPLAN data. It also recently began using PAT tests to help check teacher judgements and map vastly different types of assessments (namely NAPLAN and in-class assessments) onto one another. PAT tests, which are used

every six months, are intended to boost consistency across teacher judgements and check whether teacher judgements map onto Victorian curriculum levels.

However, although schools sometimes try to use NAPLAN as a diagnostic tool, it is not intended (or designed) to be a formative assessment; rather it was designed to identify 'cold spots' at a state level. Analysis takes a huge amount of time (often weeks) and relies on having staff members that can take time out to perform such analysis. Teachers explained that due to NAPLAN being the main form of school accountability, they were highly motivated to spend considerable time breaking down NAPLAN results at a pupil level, performing gap analysis and looking for trends in pupil performance:

"We would be silly if we didn't use the data in this way [as formative and diagnostic assessment] because we know it's used to judge us. It's the performance evaluation the department makes of our school and our Principal and then that our Principal makes of us as teachers and leaders."

Thus, the apparent absence of other forms of assessment that are able to identify knowledge and skills gaps that would impact NAPLAN results means that NAPLAN data ends up being used formatively. Though PAT data fills this gap to some extent, there is still considerable work to do to map results from one to another and to the curriculum (and there are questions as to whether this is pedagogically desirable). Therefore, should the ARBs allow schools to understand pupils' attainment and progress against the Victorian curriculum in a way that is predictive of NAPLAN results, engagement with the resources would be high and teacher workload considerably reduced. However, there is an obvious tension here between the original goal for the DAL of being formative and low stakes and it becoming a tool in 'teaching to the test'.

3.8.6 Evaluation

How is success or failure being characterised?

Teachers' use or misuse of the resources is seen as an important determinant and metric of how effective they are in both jurisdictions. Good uptake in New Zealand is seen as an indicator of the resources' success. Meanwhile in Victoria, usage figures

are believed to provide some information on what is most useful to teachers and where need is greatest. For example, higher uptake of existing maths assessment tools compared to English resources is believed to demonstrate low teacher confidence in the former.

The NZCER highlighted that it is interested not just in how much resources are being used, but how they are used (for example, the extent to which teachers are using supporting resources, guidance on misconceptions, the tagging system and so on). Training is therefore prioritised as a means of ensuring success.

Impact on pupils is also an important goal, but given that there is less emphasis on supporting materials in Victoria compared to New Zealand, there is some concern that this could diminish impact.

In Victoria, the VCAA's dream is that there will one day be an accurately tagged bank of high-quality items that could be combined into bespoke tests covering a wide range of subjects. However, this is a long-held dream and currently a long way off. Success is also described as "successfully deployed assessment resources used in all schools and all sectors, with high teacher satisfaction". A high ratio of usage to

helpline calls was given as a potential metric for the satisfaction element of this.

There were no plans to measure effectiveness by tracking pupil performance or using comparison groups and some emphasise that success needs to be characterised as gaining pedagogically useful information, rather than pupils achieving high marks. Meanwhile, the VCAA recognises that frequency of use should not be compared across tools too bluntly since usage will vary depending on the assessment (for example, 'on-demand' assessments are likely to be used at the start and end of the year, whereas others would be in a specific context).

In Victoria, much like in British Columbia, there is some hope that assessment innovation will drive education in a particular direction towards critical and creative skills – in this way, assessment resources might 'educate the educators' by shifting teacher practice.

On the other hand, the intention is not that online assessment tools should become the sole means of assessing pupils. Interviewees therefore emphasised that the tools should not compromise teachers' freedom to develop their own assessments.

What approaches are being used to evaluate the innovation's success?

The teachers and leaders we spoke to were not aware of plans to develop the DAL, though they could see a need for something that helped them perform accurate formative assessments linked to the curriculum and NAPLAN. This is perhaps to be expected as the development and rollout of the DAL has been delayed.

The primacy of uptake and popularity among teaching professionals as a valid measure of success in New Zealand was demonstrated by the fact that one interviewee said formal evaluation was not needed because it was clear the resources were popular. The need to see the ARBs as deeply embedded within a national culture that elevates teacher professionalism to an extremely high degree is therefore key.



4

Conclusion



In the last few months, countless remarkable responses to covid-19, from every quarter of the education sector have been a testament to the sector's endless ingenuity.

This report was written prior to the pandemic, but the innovations it catalogues show that the same ingenuity is being deployed to tackle long-standing and intractable challenges in the field of assessment. Importantly, almost all the innovations we looked at focused primarily on enhancing pupil learning, rather than the top-down accountability that many discussions about assessment focus on.

In this conclusion we summarise what we have learned about how teachers, schools, groups of schools and states can channel their professional expertise into change. In doing so, the main challenge is that the Making Waves study was not designed to measure the impact of a set of innovations - they were at too early a stage for that. Instead we followed people as they embarked on the process of innovation and our conclusions are therefore impressionistic - highlighting what we have learned from our observations and analysis.

We began with a tidy framework structured around 'the policy cycle' but soon found that the reality was far messier. Indeed, one of the most important things some of our 'wave

makers' did was adapt and change tack as they found their way through an uncertain landscape. This is not surprising. In her book *Cleverlands*, researcher Lucy Crehan notes that:

“Despite all this uncertainty, politicians and system leaders have to keep on making educational decisions. And these decisions are influenced by the beliefs and desires of the society that they represent, and mediated by the beliefs and desires of the teachers that work for them; so to a degree we all have to work with this uncertainty”¹⁸

The same holds for anyone making decisions about how to conduct assessment in schools. Professionals make highly contextual decisions based on imperfect information, refining their approach in an iterative manner. Thus, reflecting back on the last two years conducting this study we find that our conclusions are less about 'what to do', and more about 'how to do it'.

We have found that given the right conditions and support, alongside a keenly felt need, professionals are able to develop their assessment expertise and reduce

¹⁸ Crehan, L. (2018). *Cleverlands: the secrets behind the success of the world's education superpowers*. UNBOUND.

unhelpful workload linked to assessment. However, assessment innovation tends to be pragmatic and iterative. Thus, making waves in assessment is not just about identifying effective approaches to assessment, it is also about meeting teachers where they are, tapping into their motivations or pain points, and equipping them to make changes that work in their context.

4.1 Recommendations

Based on what we have observed and analysed over the last two years we make the following four recommendations to anyone seeking to make waves in assessment

4.1.1 Prepare for innovation

As we learned in *Testing the Water*, teachers and school leaders are not always confident about their assessment expertise. In this study too, we saw that innovators generally drew from a limited range of sources and options when determining their approach. Yet limited expertise constrains people's repertoire of options and their ability to make informed decisions.

In contrast, at Heathfield Community College we saw that members of the assessment

innovation team received training prior to selecting their approach. We also saw with other innovations that bringing together different partners can ensure that innovation is based on mutually reinforcing collaboration that does not begin with a blank page.

We therefore recommend that those looking to make waves in assessment grow their repertoire of potential solutions by learning more about assessment, building up their technical knowledge and forming new partnerships and collaborations.

4.1.2 Be explicit about what purpose assessment is serving in this instance

One of the reasons the innovations we studied looked so different from each other is that they focused on different functions of assessment. Rosie Osborne's approach at Eltham Hill was a great way of providing credible feedback to pupils in an informal environment, but it would not have worked as a means of identifying pupils' support needs across a multi-academy trust, or identifying misconceptions to inform lesson planning.

As Daisy Christodoulou points out in *Testing the Water*:

“At the school-level, many assessments are expected to provide both formative and summative information. That is, one assessment may be used to generate a grade that shows how a pupil is doing relative to their peer group, but it may also then be expected to provide diagnostic information about what a pupil has to do next to improve. But the kind of assessment that is ideal for these two purposes is very different”¹⁹

Without a crystal ball we cannot know which of the innovations we studied will succeed in the long run. However those we have the greatest concerns about are the ones that have tried to kill too many birds with one stone. In contrast, some wave makers designed their innovations to serve a hyper-specific purpose: not just “how shall we go about formative assessment” but “how can we ensure teachers are aware of key misconceptions at the start of a lesson,” “what impact are our (non-academic) interventions having,” or, “how can we use homework to identify what has and has not been learned”.

As Christodoulou points out, trying to find an approach that serves multiple assessment

¹⁹ Millard, W., Small, L., & Menzies, L. (2017). *Testing the Water: How assessment can underpin, not undermine, great teaching*. Retrieved from <https://cfey.org/wp-content/uploads/2017/11/Testing-the-Water-Final-Report-WEB.pdf>

functions can compromise the effectiveness of an innovation. In contrast, clarity of purpose can reduce the risk of overlooking solutions suited to hyper-specific issues.

We therefore recommend that assessment innovators, particularly at a classroom or department level, focus sharply on a specific challenge that they are seeking to address. They can then tailor their response to this, rather than seeking one approach which they hope will tackle all their pain points in one go.

4.1.3 Consider the best level to innovate at

We studied innovations originating at various levels; some were led by individual teachers, some by leadership teams, others by trusts and some by whole states.

It is important to consider how these different levels relate to each other. For example in academy trusts, how do initiatives and insights move between teachers, departments, schools and the trust? It is easy for knowledge to get stuck at different points and we have explored the benefits and challenges of various approaches to ‘cascading’ throughout this study.

We found that in secondary schools there appeared to be considerable benefits to Heads of Department leading an innovation.

Firstly, assessment looks different in different disciplines and innovating at subject level recognises this, as was clear at Heathfield where each department's approach reflected an academic field's particular characteristics.

Secondly, Heads of Department tend to have strong relationships with their team combined with a degree of authority. This allows them to take a context-informed approach when planning implementation whilst ensuring compliance with 'non negotiables' (see below).

Thirdly, where other departments could benefit or learn from an innovation, Heads of Department are well situated within school hierarchies to share insights or secure senior leadership buy-in. Annual or termly cross-department meetings could potentially help this process if middle leaders worked together to identify shared priorities and decide which approaches could be scaled up. This could provide an alternative to the proliferation of top-down, untested initiatives.

We therefore recommend that, subject-specific assessment expertise should be recognised as a critical element of professional development and competence for middle leaders. Schools could then explicitly recognise that research-informed, assessment innovation is one of middle leaders' responsibilities.

4.1.4 How tight or loose?

Few if any of the innovations we studied were particularly rigid. Even where externally provided platforms were deployed, the opportunity to make modifications was highly valued. On the other hand, this needs to be balanced with the risk of compromising impact when using a tried and tested model.

Furthermore, innovators noted that maintaining flexibility would be difficult or undesirable as an innovation scaled up. Throughout our research we heard about innovators grappling with the right balance between non-negotiables and flexibility. Explicit decision making about the parameters of what can be adapted is therefore an important consideration when planning implementation.

When implementing an external solution like Isaac Physics, Eedi or ImpactEd, innovators had to make critical decisions about how to implement their chosen approach, and they tended to adjust their approach along the way. Thus, even where a platform or tool does not afford for flexibility, schools still have to show flexibility in how they communicate or deploy their chosen approach. Identifying emerging difficulties with implementation and responding to these fast is therefore crucial.

We therefore recommend that innovators maximise opportunities to gather and act on feedback fast, rather than waiting until end-point evaluation. This involves planning opportunities for reflection and incremental steps that facilitate adaptation. However the appropriate level of flexibility will depend on the approach that is being implemented since well-evidenced, established approaches need to prioritise fidelity more than early-stage experimental initiatives.

Appendix:

Research framework



Phase	Research question	Question	Prompt
Introduction / context	(As much as possible to be covered in pre-visit communication)	<p>Check consent form etc.</p> <p>Tell me about this school/MAT/jurisdiction – what is it like?</p> <p>What is your role in the school/jurisdiction?</p> <p>How long have you been here?</p> <p>What has your role/involvement been with the innovation?</p> <p>Tell me about your/the innovation / what you know about it?</p> <ul style="list-style-type: none"> • What does this type of assessment aim to achieve? • How does information from this type of assessment feed into teachers’ practice • How does pedagogy/curriculum shape this type of assessment? • How much of the content that is covered in teaching is included in the assessment? • What is the balance between assessing skill, knowledge, technique or application, in the assessments you are conducting? 	<p>What are the demographics?</p> <p>How effective do you think it is?</p> <p>Has it always been like this/how has it changed?</p> <p>Formative pedagogical information, predictions, validated summative grading for labour market/higher education progression, accountability?</p> <p>Does it feed in to shape teachers’ lesson-to-lesson planning? Feed into reflection about effectiveness of pedagogy? Shape curriculum planning?</p> <p>Is the assessment based on the schemes of work? End of Key Stage expectations?</p> <p>How do you decide what to include and not to include?</p>
Agenda setting	<p>What are ‘pain points’ this innovation sought to respond to?</p> <hr/> <p>How did these become a priority?</p>	<p>What was/is the situation (with regard to x) like before this innovation began?</p> <p>What was/is problematic about that?</p> <p>Who recognised that problem?</p> <hr/> <p>How long had/has this been an issue for?</p> <p>How did you first become aware of the issue?</p> <p>Has anything else been tried to address the issue?</p> <p>Why was action taken at this point rather than another?</p> <p>What other priorities are there for your classroom/school/MAT/state?</p> <p>Is there any competition between these priorities and your innovation?</p>	<p>What was the impact of x?</p> <p>Was it universally recognised or were some people more aware/concerned about it than others? Why?</p>

Phase	Research question	Question	Prompt
Formulation	How was this solution developed?	<p>Who was involved in shaping your/this innovation? What role did they play?</p> <p>Did you consider any other possible solutions? Please tell me about them.</p> <p>In what ways, if any, has your idea/has this innovation changed or evolved?</p> <p>What sources of information did you use/are you using to shape your idea?</p>	
Decision making	How was this solution chosen?	<p>When was the decision made to go with this particular approach?</p> <p>Who was involved in pushing forward this particular solution?</p> <p>Who was involved in signing off or giving permission for this solution?</p> <p>Did anyone push against this approach or argue for different solutions? If so, what led to this one being the preferred option? Why were other potential solutions rejected?</p>	
Implementation	How is the innovation being implemented?	<p>What was the first step in implementing this innovation? What were the subsequent steps?</p> <p>Who has been involved in implementation? What has their role been?</p> <p>What mechanisms have you used?</p>	E.g. policies, instructions from leaders, new resources, training, communications materials.
	What has implementation looked like?	<p>Has the solution changed in any way over the course of its implementation?</p> <p>What has made it easier to implement the approach?</p> <p>What has made it harder?</p>	<p>What have been the main challenges during implementation? How were these challenges overcome?</p> <p>Have there been any key supporters?</p> <p>Have there been any expected or unexpected barriers?</p>

Phase	Research question	Question	Prompt
Evaluation	How is it being evaluated?	When considering how successful the approach has been, what do you consider? Has there been/will there be any formal evaluation? If so what?	Which factors do you look at to consider the success of the approach?
	How is success being characterised?	What would it look like for this innovation to be successful? What is needed for it to be successful? What would failure look like? What might lead that to happen? Has it/is it having any unexpected effects? What are the best things that have happened/ are happening as a result of this innovation? What are the worst things that have happened/ are happening as a result of this innovation? Overall, how successful do you think this innovation has been and why?	
	What will the medium/ long-term consequences of the innovation be?	What do you think will happen with this innovation in the next term? Next year? Longer term? What do you think will happen as a result of this innovation in the next term? Next year? Longer term?	



cfey.org
@TheCFEY
hello@cfey.org
+44 203 488 2765

©2020