EXTERNAL REVIEW INTO THE DELIVERY OF DIGITAL ASSESSMENT EVENTS: 30 OCTOBER TO 10 NOVEMBER 2023

DEBBIE FRANCIS

MARCH 2024

"For many students, their first formal external assessment is a life milestone and one of their very first engagements with the big machine of government. And at the end of the day, any exam is about the hopes and dreams of the learner, their family and their teachers. It's almost always a highly emotional thing. So it's a high importance, high stakes, high stress exercise for the all the players; from NZQA, to the School, to the learner and to their parents. It's absolutely critical it all goes right."

Review respondent

"It takes about a million things to right to stage successful large scale external assessment events each year. And each year they mostly always do. We obsess on this. We live and breathe these exams. For something like this not to work is professionally and personally devastating. "

Review respondent

Final

The Chief Executive

New Zealand Qualifications Authority

Level 13, 125 The Terrace

Wellington, 6011

19 March 2024

Dear Grant

Please find below my independent, external review report on the issues with digital assessment that occurred in late October and early November during the 2023 external examination session.

In summary, these events occurred because a number of NZQA quality safeguards failed at once. This compounded with problems with your vendor's Assessment Master platform. While this particular combination of issues is unlikely to reoccur, these events provide useful lessons learned. There are now a number of significant opportunities to improve digital examination processes and systems, both for NZQA and for the students engaged in them.

In addition, these events appear to me to be symptomatic of wider issues in NZQA organisational operating model and culture that you may wish to address. Consequently, my approach here has been to take a wide angle lens to causation and impacts.

I have been impressed by the openness and commitment to quality improvement displayed by the staff I interviewed. It is also clear to me that, while some students and stakeholders were adversely impacted by these events, schools' commitment to partnering with NZQA on the digital assessment journey is undiminished. They are keen to be involved in the work to prevent a reoccurrence of the disruption experienced by some students in the 2023 session.

Thank you for the opportunity to undertake this review.

Yours sincerely

D j francis

Debbie Francis

Reviewer

CONTENTS

Method and Approach 6
Acknowledgements 6
Limitations 6
Overview of Findings 7
Summary of Recommendations 10
Technical 10
Organisational 10
System 11
The context for this Review 12
Role of NZQA 12
NZQA Agency Strategy 12
NZQA and Innovation 13
NZQA's Role in Assessment 13
NZQA Examinations and Assessments 14
NZQA's Digital Assessment Transformation Programme 14
A snapshot of the digital assessment disruption From October 30 to November 10, 2023 18
What happened? 20
Why did it happen? 24
Technology factors 25
- Vendor relationships and contract governance 25
- Production Performance testing 27
- Website entry 28
- The Keycloak interface 29
- Assessment Master capacity issues 29

School and system factors 31
- High change tempo 31
- Schools' choice and customisation 31
- The PN model 32
- The ECM workforce model 32

Organisational factors 34
- The 'why' of digital delivery 34
- Strategic line of sight 34
- Accountability culture 36
- Process clarity 36
- Customer experience and behavioural mindsets 37

- Data capture and intelligence led design 38
- Risk management and scenario planning 38
- Communications 39

Impact on learners 41

Impact on stakeholders and partners 43

Conclusions 44

Appendix One : Terms of Reference 45

Appendix Two: Implementation map for Recommendations 47

Appendix Three: Interviews for this Review 48

METHOD AND APPROACH

The questions at the heart of this review are[1]:

- What were the root causes of the issues that disrupted delivery of secondary school and kura external assessment events held from 30 October to 10 November 2023?, and
- What should constitute an Action Plan to maintain the integrity of the delivery of the digital assessment to secondary schools?

In approaching these core questions, my intention was to identify the lessons learned from these events and the opportunities for improvement in the future delivery of digital assessments.

My method was based on interviews with a sample of both internal NZQA staff and external stakeholders, as shown in Appendix Three. The review also included a high-level desk top review of relevant documents. Specifically, I:

- Conducted one on one interviews, face to face or via online meetings, with a cross section of existing NZQA staff and managers from the relevant teams
- Conducted interviews with a selection of key external stakeholders, including principals, representatives of peak bodies, teaching staff and principals' nominees from affected schools
- Reviewed relevant internal documentation, such as strategic plans, performance information, IT vendor contracts, stakeholder complaints, internal and external communications and other materials
- Reviewed the parallel NZQA Risk and Assurance Review findings; and
- Prepared this report for review by the NZQA management team and Board. This includes themes and patterns found in interviews and practical suggestions for improving the future delivery of digital assessments.

ACKNOWLEDGEMENTS

I would like to acknowledge the openness and passion for improvement on the part of the NZQA Executive, managers and staff. My thanks also to the stakeholders who described the impact these events had on their schools and students and also had the grace to suggest many constructive ideas for improvement. I particularly want to thank the NZQA Risk and Assurance team, (who explored the underlying causes of the technology issues much more thoroughly than I) who have helpfully shared the findings of their internal review with me as external reviewer.

LIMITATIONS

This Review was conducted over a two-to-three-week period in February 2024. It is a point in time snapshot of the events in scope, based on interviews with around 30 staff and stakeholders. I am satisfied that I have heard and seen enough of how the agency operates external assessment events to be confident in the themes I describe here.

The Review is purposefully not a forensic Investigation. Rather, the approach I took was exploratory and inquisitive, based on the materials available to me and this insights of these impacted by and involved in responding to the problematic assessment events.

[1] See Appendix One below for the full terms of reference.

Each person interviewed for this Review was given an assurance that individual responses would be treated in confidence and that documentary materials shared with me would be held only by me for the purposes of this exercise and not used for any other purpose.

OVERVIEW OF FINDINGS

In the external examination session from October to November 2023, New Zealand secondary school students suffered from significant service shortfalls in several digital assessments, including the NCEA English Level 1[2] examination. These events distressed students and teachers and adversely impacted stakeholder confidence in digital assessment.

The New Zealand Qualifications Authority, (NZQA) as the exam provider and system regulator, has asked me to review the root causes of these events and to suggest a plan of action to ensure the future integrity of national digital assessments.

NZQA has had a recent history of successful innovation and piloting for digital assessment. The agency is one of a small number of qualifications providers in the world to offer a broad portfolio of online external assessment and marking.

My interviews with NZQA staff and sector stakeholders[3], supported by a review of the relevant documents, suggest that the problems in the 2023 exam session were caused by a combination of root causes and compounding issues both internal and external to NZQA, including:

1. Technology issues, such as:
    a. Problems with the vendor's Assessment Master platform and its capacity.
    b. Non optimal NZQA vendor management and contract governance
    c. Inadequate and misaligned pretesting of the Assessment Master technology platform and the end to end student experience; and
    d. A login redirect defect in NZQA's internally hosted Keycloak security application which caused significantly more login attempts between Keycloak and the vendor's Assessment Master platform than each single attempt to log in would have done. This placed significant strain on the Database. This was a defect introduced by the vendor, but not picked up in vendor testing.
2. Compounding organisational issues, including internal culture, such as
    a. Unclear internal accountability for the end to end digital assessment process
    b. Optimism bias, impacting risk management and scenario planning
    c. Lack of process clarity and SOPs[4]
    d. Patchy utilisation of available intelligence on school and student customer experience and behaviours
    e. Loss of strategic line of sight on the 'why' of digital asessment; and
    f. Institutional conservatism and a defensive communications style.

[2] National Certificate of Educational Achievement, (NCEA). The period under review here was 30 October to 10 November 2023. For context, the NCEA examination period was from 6 November to 30 November, during which 68 digital examinations were delivered, with Level 1 English the subject most affected by these events.
[3] See Appendix Three for a list on those interviewed for this Report.
[4] Standard Operating Procedures, (SOPs)

3. Wider school and systems issues, including:

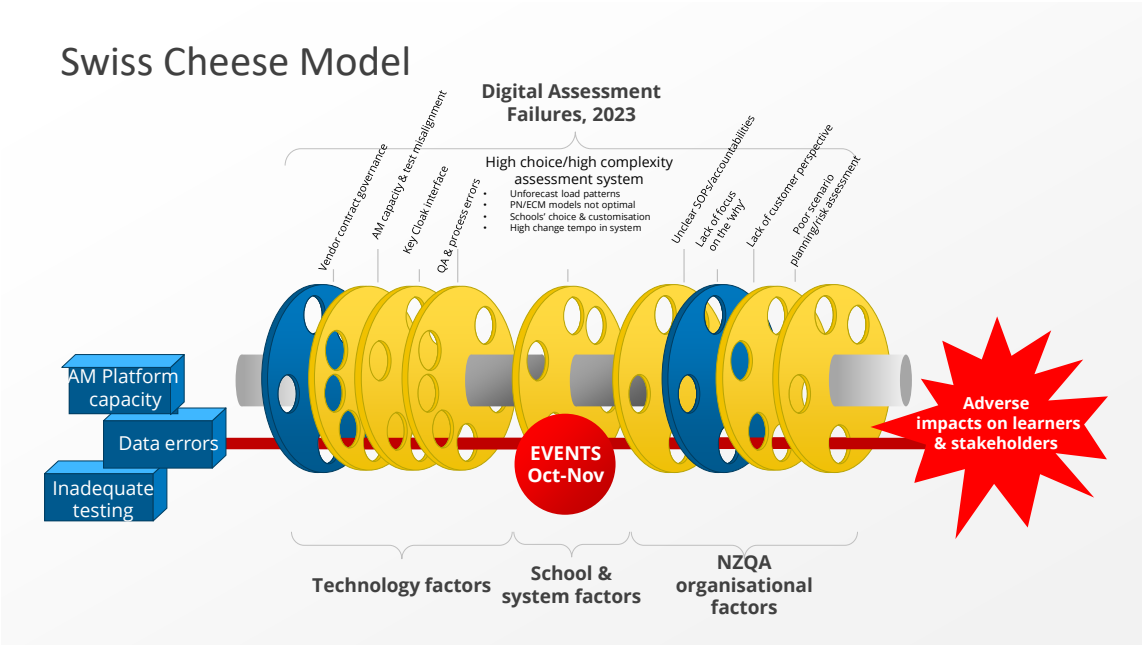
    a. High change tempo in the secondary schools sector

    b. Schools' choice and customisation

    c. The Principal's Nominee (PN) model; and

    d. The Exam Centre Manager (ECM) workforce model.

In the event, these issues compounded, as per the well known Swiss cheese model of causation[5], shown in summary form in the schematic below.



The full Report explores each of these root causes and compounding issues. It outlines their combined adverse impacts on learners, schools and the reputation of NZQA as provider.

It concludes that delivering large scale national examinations, whether in paper form or online, is always a complex and high stakes undertaking. As the respondent quoted at the head of the Report put it: *'about a million things have to go right'* to make it happen. And mostly, NZQA does get it right. Even with these digital assessment problems, the majority of the 263,000 students who completed digital assessments in more than 50 subjects in 2023 did not have a disrupted experience.

However, respondents are also correct in saying that problematic events of this kind seriously damage the confidence of those students who were affected, and that of their families, in the future of digital assessment. These issues were the last straw for tired teachers and school leaders at the end of a busy year. Their impacts were likely most harmful to disadvantaged learners. It is clear that in some schools, there is now real fear of a repetition in 2024.

From a system perspective, any repetition would not only damage NZQA's digital assessment programme but tarnish the qualifications themselves and public confidence in them.

[5] This model was first developed by James Reason, in Reason, James (1990-04-12). "The Contribution of Latent Human Failures to the Breakdown of Complex Systems". Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences.327. The model demonstrates how, generally, a failure cannot be traced back to a single root cause; accidents are often the result of a combination of factors. It suggests that most accidents are the result of latent errors, which are failures that are intrinsic to a procedure, machine, or system. It is most often used nowadays in aviation safety.

Given that it took many holes in the Swiss cheese to line up to create last year's problems, and in light of all players' strong commitment to working collaboratively on quality improvements, I believe the chances of similar problems in this year's assessment and examination sessions are low.

However, these events have highlighted a number of urgent and important opportunities for NZQA to improve both its internal operating model and its partnerships with its technology vendor and the sector to codesign workable solutions and improved safety mechanisms. The agency needs to seize these opportunities while their drivers remain fresh.

I have considerable faith that NZQA will work hard and collaboratively with the sector on the recommendations made in this Report. My discussions with managers, vendors and stakeholders suggest a powerful and shared commitment to doing so.

SUMMARY OF RECOMMENDATIONS

My Terms of Reference require me to suggest 'an Action Plan to maintain the integrity of the delivery of the digital assessment to secondary schools'. Taken together, the following recommendations should be viewed as that action plan. In Appendix Two below I also suggest implementation priorities and sequencing.

TECHNICAL

1. Review the portfolio of current technology vendor contracts with a view to ensuring that all Service Level Agreements, reporting requirements and performance metrics are current, appropriate and sufficient. Develop a plan for rolling reviews in future.
2. Following the above review, align functional and product performance testing protocols between NZQA and its technology partners and vendors.
3. Develop and promulgate a revised and fit for purpose operational incident management model, with accompanying Standard Operating Procedures (SOPs), which clearly articulates core processes, escalation points and accountabilities.
4. Consider and engage with schools on the optimal mechanisms to ensure that digital development is informed by customer (schools' and students') behaviours and experience.
5. Review and simplify the internal governance arrangements to improve the oversight of technology contracts and relationships, in order to clarify accountability, streamline decision making and take an integrated approach to this portfolio.
6. Consider the appointment of external members to a new or revamped advisory body for digital assessment and change.

ORGANISATIONAL

7. As a subset of organisational strategy, revisit, clarify and promulgate the NZQA strategy for digital assessment over the coming years, including clear outcome statements and progress metrics.
8. Ensure clear line of sight for NZQA staff on the overall strategic context within which digital assessment service delivery sits, by:
    a. Ensuring that team and individual performance measures align to strategy; and
    b. Ensuring that middle managers have the leadership skill to understand their accountability for joined up service delivery, rather than focussing only the priorities of their individual teams.
9. Review NZQA's current operating model and consider:
    a. Better integrating the CX and SRM teams into co-design, service delivery and change processes
    b. Consolidating the current assessment teams to ensure an end to end view of all assessment experiences
    c. Ensure that specialist skill sets are located in the right places in the organisational structure and are not duplicated; and

   d. Clarifying the respective accountabilities of internal teams in the end to end assessment process, including IS, operations and logistics, communications, CX, SRMs and assessment.

10. Once the operating model is reset, map, streamline and document all internal processes relating to digital service delivery and ensure these form a part of staff induction into the relevant teams.
11. Using scenario planning techniques, review and reset the risk register for digital assessment and change, including mitigation and contingency plans.
12. Review and promulgate a new success profile for exam centre managers and develop a mini workforce strategy for this group of staff. Once this is done, redevelop the training for this group to equip members for a digital world.
13. Move away from siloed data (both within NZQA and across sector partners) to joined-up analysis, intelligence, research and insights. This will require:
    a. a deeper understanding of the data that NZQA (and others) already hold
    b. investment in internal capability to turn data into insights; and
    c. capacity and capability to publish and share those insights with stakeholders more effectively.
14. Develop a templated 'hot debrief' model through which lessons learned can be captured immediately after external assessments take place, to inform ongoing service delivery improvement and risk management.
15. Develop improved processes for capturing school and learner related behavioural insights into internal service delivery and technology design processes. Consider in particular the roles of PNs and SRMs in such processes.
16. Consider opportunities for more proactive and regular system facing communications, based on in house research, about the strategy for and lessons learned from digital assessments.

SYSTEM

17. With the Ministry of Education, consider the Principal's Nominee model in schools with a view to:
    a. Developing a consistent and clear position description
    b. Understanding the support needed to those in this position
    c. Developing a plan for the professional development of this group; and
    d. As for recommendation 15 above, clarifying channels by which PN feedback and input into service delivery design can be captured to inform practice.

THE CONTEXT FOR THIS REVIEW

ROLE OF NZQA

NZQA is an independent Crown entity under the Crown Entities Act 2004. It has a range of functions, many of which are set out in the Education and Training Act 2020.

Its three main functions are described below. The third is most relevant to this review.

1. The guardianship of the New Zealand Qualifications Framework (NQF) and the international recognition of New Zealand's qualifications, to ensure that qualifications meet the needs of learners and employers and enhance the nation's social and economic outcomes
2. Quality assurance processes that provide confidence in tertiary education organisation (TEO) performance; and
3. To ensure robust and credible secondary school level NCEA assessment processes.

Specifically, NZQA is responsible for:

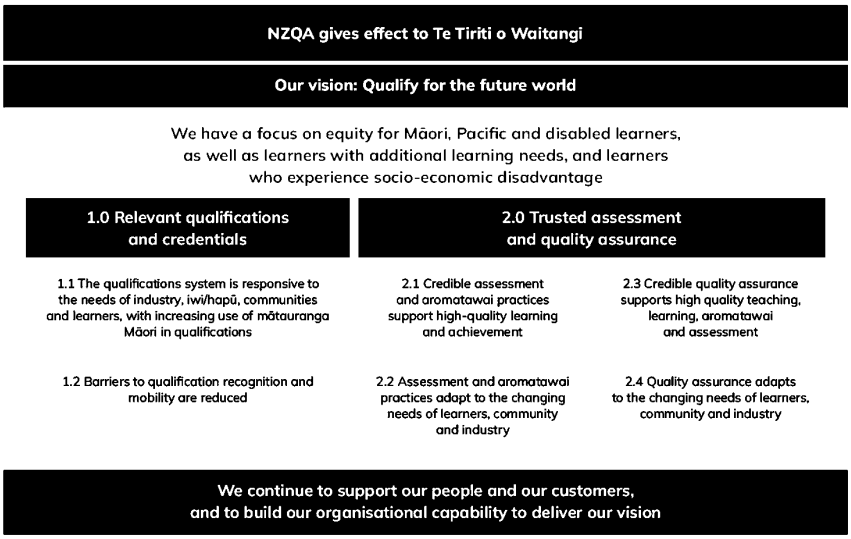
- Managing the New Zealand Qualifications Framework
- Administering the secondary school assessment system
- Independent quality assurance of non-university tertiary education providers; and
- Qualifications recognition and standard setting for some specified unit standards.

In fulfilling these roles, it operates in a manner unusual for a regulator, in that, in addition to its regulatory policy, quality assurance and monitoring work, it is also a direct provider of services in the form of secondary school external assessments and examinations.

NZQA operates within the context of a complex and fragmented education system landscape, characterised by frequent touchpoints and overlaps with other agencies, including the Ministry of Education, which, in addition to its lead policy role, has ultimate ownership of the NCEA examination.

NZQA AGENCY STRATEGY

The agency's organisational vision is to 'qualify for a future world' and its current strategic outcomes are set out in the diagram below.



NZQA AND INNOVATION

In recent years, NZQA has invested considerable internal effort to transform the qualifications and assessment landscape to meet the needs of future learners. Initiatives have included, among others:

- The application of transformative approaches to its technology projects and programmes
- The establishment of an internal customer experience (CX) team dedicated to ensuring that services meet customer needs
- Programmes targeted to improving educational outcomes for Māori and Pasifika in STEM subjects
- A range of interventions, such as Special Assessment conditions, (SAC), designed to reduce barriers for disadvantaged and high needs learners; and
- The establishment of micro credentials for tertiary students.

In addition, the agency has worked hard to ensure that a mātauranga Māori approach underpins its entire portfolio of roles and functions.

In pushing for such innovations, the agency has had to balance multiple system-facing challenges, including:

- As a small agency in a crowded education landscape, it has often had to drive change through other, bigger agencies, while also staying it its own swim lane, which has required considerable leadership nuance and skill
- The need to balance transformative approaches with prudent risk management
- Leading change at the pace that schools and other education providers can cope with, while also challenging them to operate at the limits of their comfort
- The tension between its roles as a regulator and quality assurer, with its direct delivery of complex external assessment services to large cohorts of students every year; and
- Managing a complex and demanding portfolio of change projects inside an agency with modest resources and a relatively flat baseline.

Internal challenges have also included balancing the risk aversion and conservatism typical of a regulatory agency's culture, with the need for transformative service delivery and innovation.

Overall, the agency has managed its innovation programmes and the attendant challenges expertly. In doing so, it has built strong relationship capital across the system.

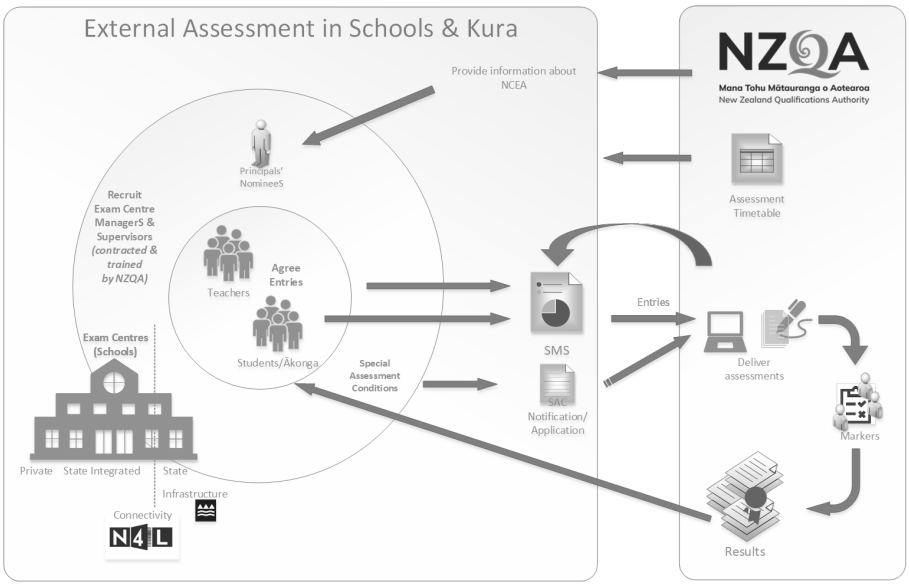
NZQA'S ROLE IN ASSESSMENT

NZQA manages quality assessment practice in the schools' sector via a regular cycle of national assessment reviews. It also directly offers training based on best practice evidence of assessment practice and administration of the NCEA.

It provides external moderation of schools' internal assessment for NCEA and publishes guidance accordingly.

The agency directly administers the national NCEA and Scholarship examinations each year and releases results. These moderation and assessment activities require the annual addition of a surge workforce of contracted staff.

The traditional approach to external assessment can be illustrated in the diagram below, which shows how NZQA interfaces with schools, kura and other agency partners in service delivery.



NZQA EXAMINATIONS AND ASSESSMENTS

The delivery of corequisite assessment events is planned and administered at schools by Principals Nominees and teachers within a period of time set by NZQA. Last year, these assessments were made available on the assessment platform over a one week period, to enable schools' flexibility in how they apply their resources (supervision, timetables, classrooms, devices, digital infrastructure) to deliver the assessments to students. Derived grades are not available to students participating in these assessments because schools have flexibility in scheduling them. In 2023, around 77,011 students participated in the LitNum digital assessment events.[6]

NCEA external assessments (examinations) are scheduled by NZQA and supervised by NZQA staff (Exam Centre Managers and supervisors). There is one national external examination opportunity each year, and students who are entered must participate in the three hour examination at the same time. In 2023, almost 40,000 students were entered by schools for Level 1 English, which comprised 16,900 digital and 20,177 paper based examination entries. However, as students can decide to complete the examination on paper rather than digitally when they commence the assessment, the number of students participating digitally is *estimated,* rather than precisely determined prior to the examination. Students can apply for derived grades if they meet approved criteria, including if they believe their performance has been significantly affected *during* an NZQA external examination session.

NZQA'S DIGITAL ASSESSMENT TRANSFORMATION PROGRAMME

By the time of the events under review, NZQA had been engaged in an ambitious transformation programme, including online assessment, since 2014[7]. Its Board and leadership team saw opportunities for innovation in the assessment system based on emerging models in other countries. They were determined to apply an innovative and collaborative approach to change.

The transformation programme had an initial vision of 'assessment: anywhere, anytime'. This was intended to allow teachers and students to tailor external assessment events to learner needs and school philosophies as opposed to assessment being driven by national set piece, paper-based examinations in the traditional way.

[6] Note: this figure includes students from both secondary schools and tertiary institutes and includes any student with N, A or V result. It does not include Absences or Z (missing results). It therefore informs the number of students that interacted with the platform in 2023.

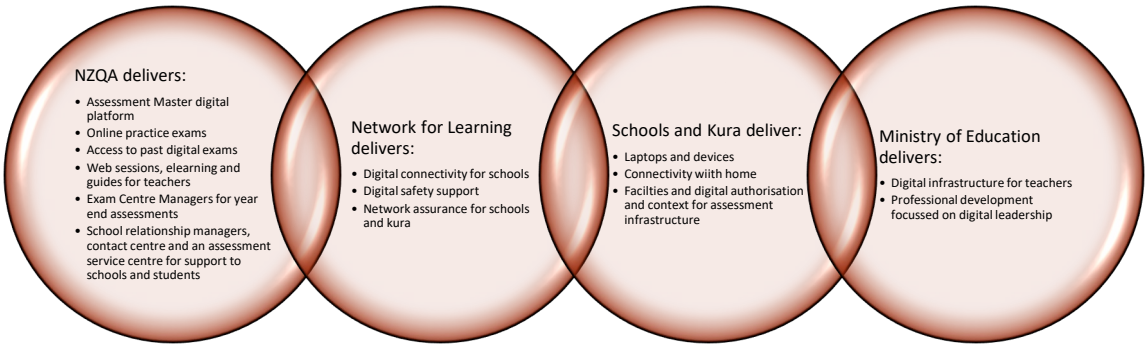
[7] For full disclosure, I was one of two external members on an advisory board to the programme from 2014-17.

By 2018, these aspirations were reflected in the NCEA Online programme, which was focussed on developing a new digital assessment platform and operating model for NCEA assessments. This project recently moved out of the agency's transformation portfolio and into business as usual management.

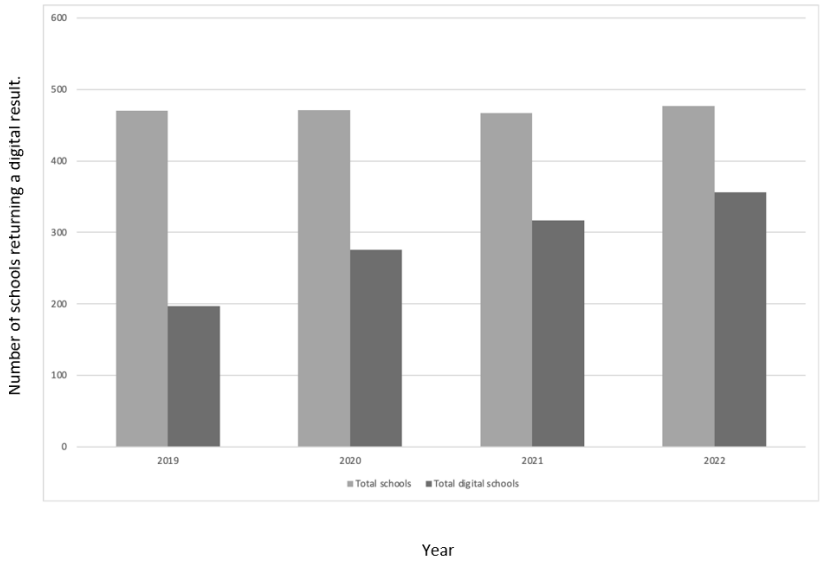
The business case for NCEA Online was based on a core set of assumptions about digital assessment, including:

- Students wanted digital assessment and that would drive schools to 'opt in' to participate in NCEA Online on NZQA's Assessment Master platform
- Digital assessment would increase digitally supported teaching & learning and this demand side pressure would resolve issues of device access equity and school infrastructure vulnerabilities
- There would be some 'hard to digitise subjects' that might be challenging to develop and deliver as online assessments, and which as a consequence, would be addressed in a later phase of development
- NZQA would partner with a coalition of willing schools to pilot digital assessments before rolling these out more broadly; and
- Engagement in digital assessment would be non-mandatory for schools and students.

Under this conceptual model, the various players in the assessment ecosystem deliver in partnership as below:



During the development phase, the programme trialled multiple pilots in schools. Since going live, it has achieved steady penetration and uptake. Many of the initial fears that those in the sector voiced in regard to digital infrastructure, access to devices and cyber safety were put to rest through the pilots. Reliance on remote learning during the Covid pandemic also pushed schools further down the digital assessment route. By 2022, 75% of New Zealand schools and kura were returning digital results, as shown in the graphic below. Much of this work was and is world leading.



Over the period from 2018, as lessons were drawn from the pilots, the foundational vision for digital assessment morphed from the original 'anytime, anywhere' to a more complex idea centred on improving equity, removing barriers for vulnerable learners and improving both quality and efficiency.

Increasingly also, some assessments, such as LitNum[8], were developed as 'digital first.'

NZQA describes the current overall intent of the changes in assessment in its 2023-6 Statement of Intent:

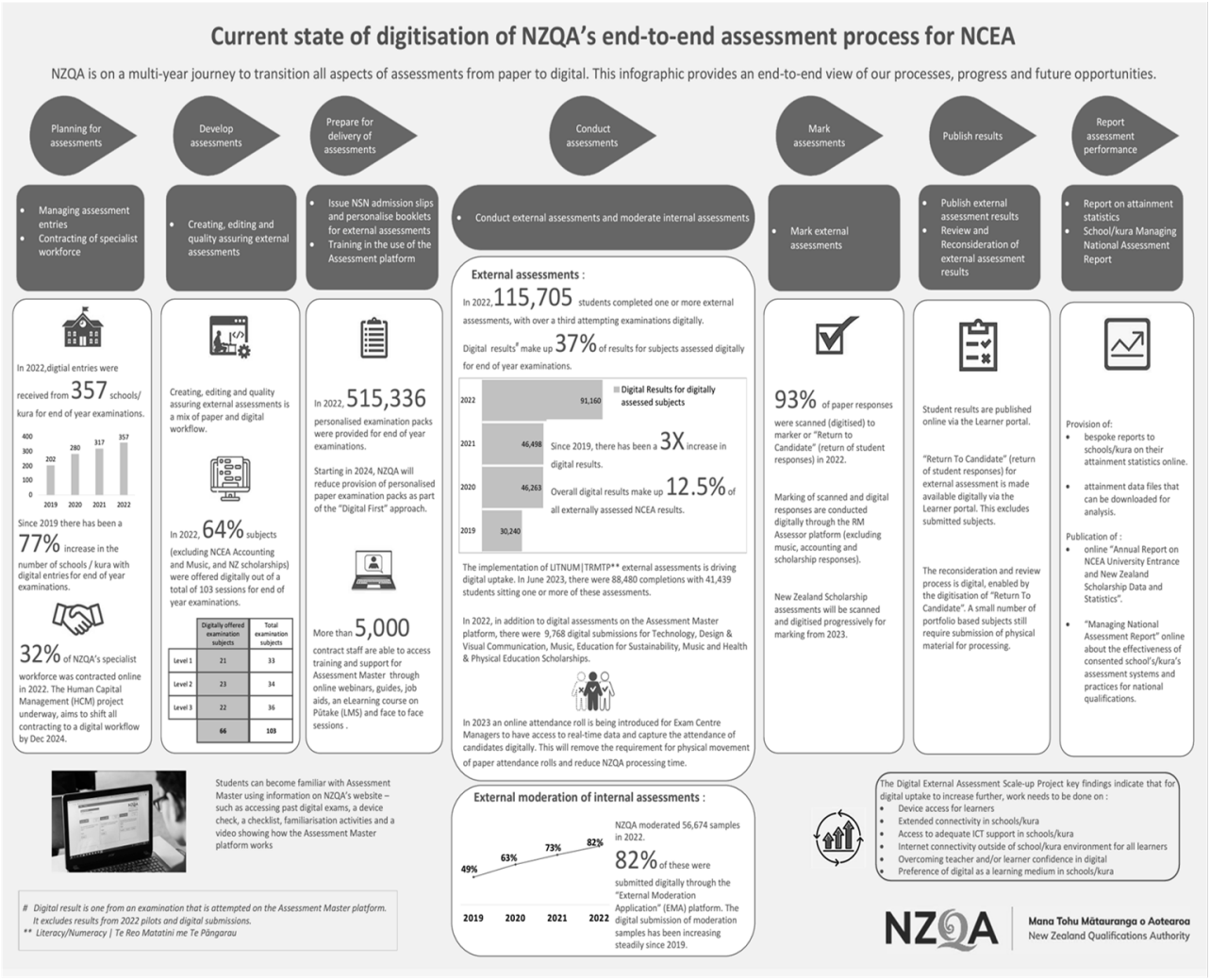
> *"NZQA is focused on setting learners up to succeed in a digital world. We are transforming assessment to respond to the different contexts and changing needs of learners, communities and employers. Digital technologies provide opportunities for new types of support and services that can significantly reduce barriers for disadvantaged learners. The benefits of digital external assessment include greater efficiency and enhanced quality assurance. It provides opportunities to include functionality such as text to speech, spell check and editing, where appropriate, for students participating in digital assessments."*

Just as the vision has been refined over time, the change leadership approach NZQA has taken to encouraging digital participation has also changed. It has become more ambitious as sector confidence has grown. The emphasis has changed from schools opting-in into digital assessment, to the notion that digital assessment is now the norm and a few schools and kura may opt out. In recent years, the agency has pushed hard to enrol more schools in digital assessment, rather than working, as it had in the initial stages of the programme, only with a coalition of the willing.

While the majority of NCEA exams remain paper based, digital assessment is now well established, with increasing numbers of schools and students participating, and rapid growth in the number of achievement standards being assessed digitally. NZQA's research data shows that overall teacher and student satisfaction with digital assessment is high.

The current state model for digital assessment and the related performance statistics (as at 2022) are summarised in the NZQA graphic below:

[8] More correctly, Literacy & Numeracy | Te Reo Matatini me Te Pāngarau (LITNUM | TRMTP) assessments.



In the 2023 school year, 263,000 students successfully completed digital assessments. To put the events under review into context, there were around 8,850 applications for derived grades, some unknown subset of which related to the disruption caused by these events.

A SNAPSHOT OF THE DIGITAL ASSESSMENT DISRUPTION FROM OCTOBER 30 TO NOVEMBER 10, 2023

On the morning of October 30, 2023, the first group of year 10 and 11 students at one of New Zealand's large secondary schools[9] started to log into their Literacy/Numeracy[10] external assessment exams online.

For the year 10 students, this would be their first experience of a formal external assessment. They were nervous. Those with special assessment conditions (SAC) were supported with preplanned, individualised arrangements reflecting their needs.

For the staff and exam supervisors, this was the first of a group of 290 students who would be tested in these subjects over the coming week. They had planned the day and time of the assessments to suit their circumstances and the needs of their learners, within the time window allowed by NZQA. It was a busy time, as the end of the school year neared, and with NCEA examinations scheduled for the following week.

For the School's leadership, this was the second year in which they had engaged in digital assessment, and they had taken care to apply the lessons learned in pilots and during the previous year's sessions. The Principal's Nominee, (PN), responsible for the interface with NZQA for assessments, had planned and practiced the arrangements carefully.

As soon as the students, (who were located in multiple classrooms) started logging in at 8.15am, problems arose with access to the assessment platform. As they, and their supervisors, became increasingly concerned, the PN started calling NZQA for support. It took 11 phone calls before he was able to reach help. While the issue was resolved by 9.15am, 42 students had their writing exams delayed by 30 minutes.

On the Thursday of that week another group of students managed to log into their assessment but a small number experienced 'digital lag', which meant they could not always edit their work. Calls from the School to NZQA on that day went unanswered. Staff and students were again distressed.

On Friday, 72 students sitting the Numeracy assessment also experienced lag effects which caused some, including six SAC students, to be logged out and unable to save their work. The PN, also logged into the platform as a supervisor, faced a blank screen when he attempted to monitor what was happening. Several of these students walked out of the exam.

The School had communicated with NZQA throughout the week and, on Thursday, had been advised to reschedule the exams on a different day to spread the load on the technology platform. This was unfeasible, given the pressure on examination facilities driven by end of year events, the need for tailored facilities for some SAC students and the need to schedule the large NCEA exams the following week.

In the meantime, the School faced multiple complaints from impacted students and their parents.

The following week, another large secondary school, also a keen adopter of digital assessments, felt sure that the lessons learned the previous week at NZQA meant that the large cohort estimated for Level 1 NCEA English would be able to engage in the examination without issue. This was the marquee examination event of the year, with predictable cohort size[11]. Instead, its students also experienced access issues and slowdowns on the platform. This School was forced to substitute this and other exams with paper based exams in the moment, which created panic on the part of the Exam Centre Manager (ECM) and logistical havoc and stress for many

[9] The events in this example are a conflation of events at several schools and indicative of events across the system.
[10] In full, Literacy & Numeracy | Te Reo Matatini me Te Pāngarau (LITNUM | TRMTP) assessments. LITNUM and Pāngarau was held during the week 30 October to 3 November; Te Reo Matatini was held from 6 -10 November 2023.
[11] Cohort size can be anticipated by NZQA based on lists of entrants provided in advance by schools and on historical patterns of entries.

students[12]. In the end, its PN applied for the entire Level 1 English cohort to receive derived grades, a less than ideal outcome for learners who had worked hard to improve their performance since sitting their test papers several weeks earlier. This School also fielded multiple complaints from students and families.

Given NZQA's solid record of balancing innovation with careful risk management, and the lessons learned from the digital assessment pilots, how did things go so wrong?

In the sections that follow, I outline what happened in more detail, my findings on the root causes of these events, a related set of compounding issues and my recommendations to prevent their recurrence.

[12] In a world of digital natives, many students are not comfortable writing rather than typing.

WHAT HAPPENED?

Twenty-nine digital assessment events were held during the period under review here, from 30 October to 10 November 2023, on NZQA's digital external assessment platform.

Exactly *what* went wrong from a technical perspective with some of these digital assessments is not difficult to establish. NZQA's own internal investigations broadly align with my independent review to produce the following set of events, shown in summary form in the table below.

Issue	Date/time detected	Date/time resolved	Technical fix	Impact on learners and stakeholders
Save Defect **Student work was not being saved during the numeracy assessment.**	30 October 11.17am	2 November 8pm	Interim close down of Assessment Master for access to the numeracy assessment at 12.06pm on 30 October. Communication to schools to extend availability of exam to 6 November to stagger load on the platform. PNs[13] instructed to provide paper alternative.	4,519 students were due to sit the exam on 30 October, with an unknown (and likely to be small) subset impacted by the saving issue. The assessment was postponed for several days with significant impacts on school operations as well as causing learner stress.
Assessment Master Platform lag and log in redirect problems **Platform lagging for LitNum/Pāngarau** **assessments.** **Further slowing was observed at 10, 000 log ins.**	2 November 9.05am	8 November 7.05pm	Access to Assessment Master closed 2 November 9.30am Communication sent to schools to advise staggering access to exam and noting options for paper responses.	Resulted in NZQA temporarily pausing student logins for LitNum and Pāngarau for at least 30 minutes to manage performance of the Assessment Master platform and minimise the risk of platform failure. This left students, exam centre managers, staff and PNs in the position of having to wait, with attendant stress. Some students left the exam.
Assessment Master Platform lag and access problems **Slowing down on platform for NCEA Level 1 English exam at 13,000 log ins.** **Issues also experienced with Level 3 Media Studies and Level 2 Le Faka Tonga exams.**	10 November 9.30am	10 November 10.18am	Platform access was locked at 9.30am. Text to Exam Centre Managers re stopping new log ins. Platform access restored at 10.18am	Resulted in problems for students in accessing digital external assessments on 10 November for up to 40 minutes to manage performance of the assessment platform and minimise the risk of platform failure Again, some students left the exam. Others reported considerable stress Many Exam Centre Managers (ECMs) and PNs saw this event in the moment as 'the last straw'. However, it should also be noted that further 52 digital

[13] PN stands for Principal's Nominee. This role is the key interface between schools and NZQA for assessment.

				examinations were subsequently delivered without issue.
Less than optimal contingency planning, with plan b approaches not conducive to a good customer experience	Throughout these events	Had in some cases to be resolved by applications for derived grades	Paper versions of exams when the digital assess failed Reducing save times to take load off the platform Asking schools to reschedule large exam events at short notice.	Learners thrown by not having prepared for a handwritten exam In some cases, teachers needed to print multiple copies at short notice and distribute to multiple exam rooms, with adverse impacts on exam protocols Timetabling and in school logistics were often such that a reschedule was not possible Reducing save time frustrated many students.
Engagement with and communications from NZQA	Throughout these events		Monday 30 October: Communication to PNs to shut down Numeracy assessment 31 October: Communications to PNs re access and extension of time for Literacy and Pāngarau to 6 November 2/3 November: comms to schools to advise staggering cohorts to reduce load and extension of time 6 November: Email providing options for paper responses 9/10 November; Emails and texts to ECMs and PNs re stopping additional logins Email to ECMs with 3 question survey Email from NZQA DCE to principals and PNs included apology and message template for students and families 16 November: Email from NZQA CE with apology and announcing independent review.	Delays in inwards call handling and 'contact with a human' frustrated some PNs PNs had to communicate with supervisors during exam and in multiple locations Different communications to PNs and ECMs created confusion in some schools Schools were receiving different advice from the contact centre, school relationship managers and other NZQA staff 10 November email came too late for some principals, who had to front families before its arrival Some stakeholders felt the tone of some NZQA Emails was defensive Many respondents felt 16 November Email came too late after the events.

As context for these issues, the graphic below shows the various system interfaces that students and administrators must use to access the Assessment Master platform and the exam.

Vendor testing currently does not incorporate Keycloak in performance testing. Hence the issue with the login redirect defect, which caused significantly more login attempts between Keycloak and Assessment Master. This placed additional strain on the vendor Database. This was a defect both introduced by the vendor, and not picked up in vendor testing.

The problems with Assessment Master pertained to lower capacity (and higher use within concentrated periods) than forecast. Platform capacity appears to have been affected not only by the number of logged in users on the platform, but by Keycloak and administrator usage, which reduced overall capacity in unforeseen ways. For example, platform capacity was adversely impacted by such factors as :

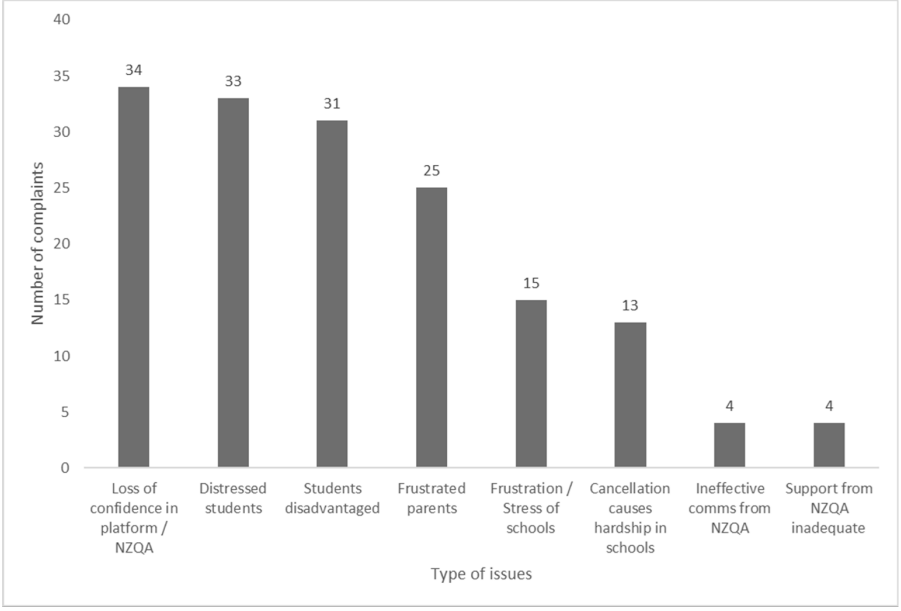
- Challenges in forecasting the traffic and spread of use for LitNum assessments. Schools' 2023 behaviours in this were not foreseen
- Internal NZQA testing and external vendor performance testing on Assessment Master gave different total capacity limits, suggesting that base assumptions and test environments differed. These differences were not fully addressed up front in part because the vendor did not, although asked, provide testing results to NZQA in advance of the examination session
- Keycloak compatibility with Assessment Master was not tested by the vendor. It was not understood than Keycloak would consume up to 30% of the platform's overall capacity; and
- There was no simple tool at the vendor end to monitor traffic on the platform in real time.

In addition, as shown in the table above, communications issues impacted the ability of schools to adapt to the issues they faced. For example:

- Some stakeholders felt that the Numeracy assessment save issues, (in response to which NZQA asked schools to defer sitting the assessment and to stagger their participation over a few days), were not communicated in a timely enough way to schools, so that they could consider both their messaging to parents and overall logistics
- Scattered ownership of the various communication channels between schools and NZQA meant that stakeholders received inconsistent information
- Information for different audiences was slightly inconsistent. For example, slightly different emailed communications on the platform capacity issues were sent to PNs and ECMs, which created confusion in some schools
- Call centre staff in NZQA struggled to find the key personnel for PNs and others to engage with in the moment
- Key facts to assist principals to reassure students and their families were not provided in a timely enough manner, leaving school leaders feeling exposed; and

- Modes of communication were sometimes inefficient, such as the banner on the NZQA website (which few PNs, ECMs and teachers appear to have noticed) and the texts and emails to busy ECMs, who were often running between multiple rooms to advise supervisors and students, and found it hard to keep up with the volume of information, particularly regarding login issues.

In the wake of these events, NZQA received a number of formal complaints from schools, families and students. Content analysis of these complaints shows pain points for the sector as per the chart below.



In the next section, I address the root causes of these problems. I also traverse several related organisational and system issues that I believe amplified their impact.

WHY DID IT HAPPEN?

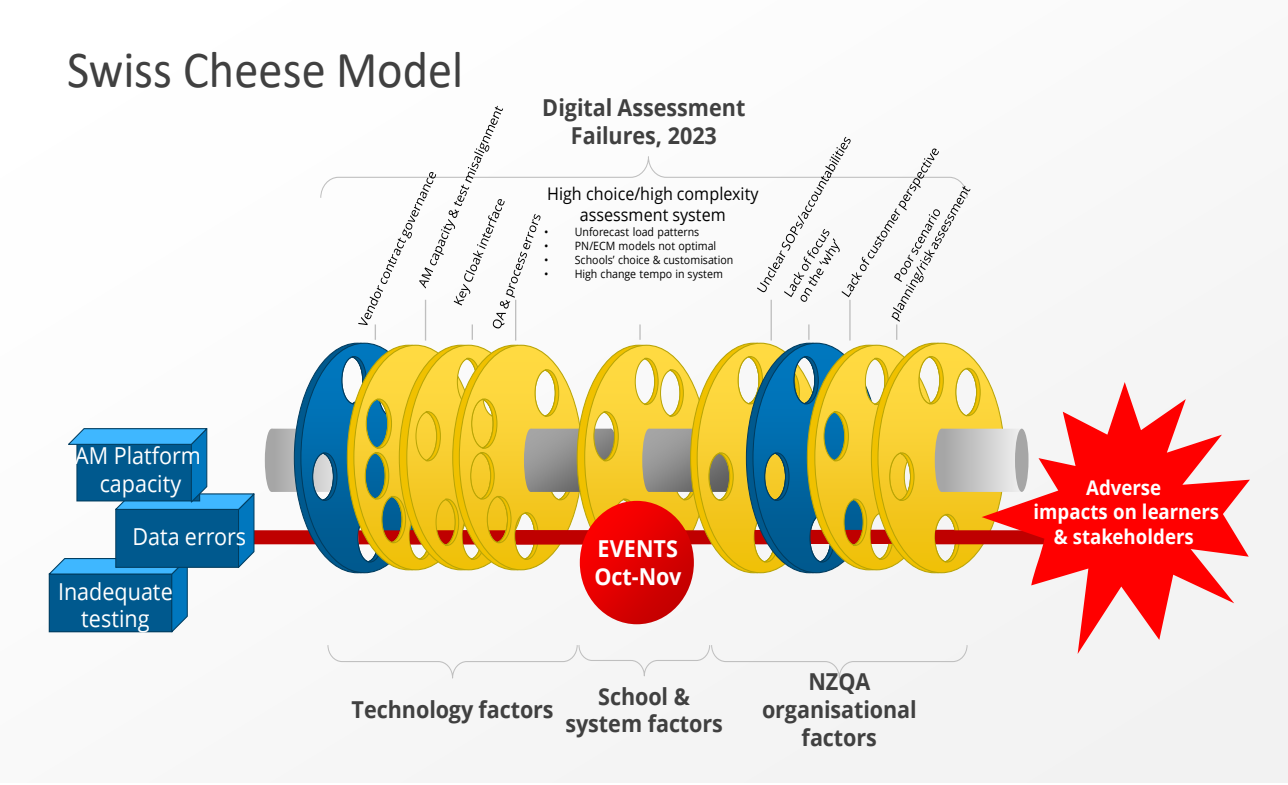
As with most service failures, these digital assessment problems resulted from a number of factors which layered up and aligned to create service disruption.

Such failures are often addressed by the well-known Swiss cheese model, in which multiple contributors (the holes in the cheese slices) must be aligned for adverse events to occur. Barriers in a system (the slices themselves) are intended to prevent errors that result in adverse events.

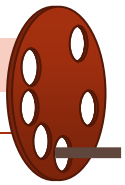
The factors which led to the events described above included a diverse combination of 'holes' at school, NZQA and system levels, as summarised in the graphic below.

Generally speaking, the reason these technical and communications issues occurred is because multiple safeguards failed at once, and problems lined up like the holes in the Swiss cheese shown below.

In this section, I categorise the 'holes' or causes in three broad categories: technological root causes, and compounding organisational (reasons internal to NZQA) and systemic reasons linked to the wider education ecosystem. The graphic below shows in summary how these factors aligned to create the digital assessment problems examined here.



TECHNOLOGY FACTORS



VENDOR RELATIONSHIPS AND CONTRACT GOVERNANCE

The vendor relationship most relevant to these events is with the provider responsible for the Assessment Master platform. This platform is NZQA's core vehicle for digital assessment. As such it is an enterprise system and mission critical. The same vendor hosts NZQA's marking platform.

When NZQA commenced its digital transformation journey, this vendor was a start-up located in Australia. There was, in the early days, a keen sense of partnership between agency and vendor, given that both were developing an application that was extremely innovative by world standards. With the vendor located locally, fast fixes were possible and the vendor had a good understanding of NZQA's wider strategic context for digital assessment.

In time however, this vendor was acquired by a larger company, with headquarters in the United Kingdom. Some staff respondents told me that since then, they believe the relationship has diluted. From the vendor's point of view however, they continue to regard NZQA as a beacon client, a world leader in digital assessment and a long term partner.

Over this same period, the contractual conditions and appended service level agreements changed little.

Following the handover of its NCEA Online Programme to business as usual, NZQA now governs and manages this vendor contract and relationship at three levels:

- At the operating level, there is regular contact between the vendor relationship manager, the NZQA IT manager and the assessment cycle production team
- At management level, the relationship is operated via a steering group compromising NZQA middle managers, along with a more senior vendor representative. The steering group meets on a quarterly basis; and
- In addition there is a strategy board which meets twice yearly. This involves the NZQA Deputy Chief Executive Assessment and the vendor's Chief Operating Officer. This body resolves any disputes, agrees user number projections and so on.

With regard to the business as usual management of digital assessment, NZQA manages this via two groups:

- The external assessment technical committee which includes IT and corporate staff; and
- The external assessment cycle advisory board, which is chaired by the Director Assessment Production and Delivery. This group meets monthly. It monitors test reports on platform performance, approves change requests and so on. It also regularly discusses risks and mitigations in regard to platform performance.

I should also note that in an agency as small as NZQA, there are challenges in managing complex contracts given the light staffing in the procurement team.

Internal respondents raised the following issues in regard to these arrangements and the relationship with the vendor more generally:

- The existing multi-level governance and management processes, while complex, are appropriate, but still require maturation to ensure all members of these bodies take an end to end and enterprise view of the digital assessment process, rather than members seeing themselves as representatives of an internal team

- However, given the many internal groups touching this vendor contract and relationship, it is hard to see who has ultimate accountability for the contract. Contractual issues also sometimes get lost in translation between these groups
- The vendor's monthly SLA reporting is light on detail and should be updated to reflect the new priorities and direction of the digital assessment programme
- The interface between vendor and client IT teams on the ground is not always as positive as that at more senior levels
- While these internal groups are adept at future thinking and regularly address risks and consider scenarios around the Assessment Master platform, they can also suffer from optimism bias. They struggle, within their available resources, to plan for a wide range of black swan events and worst case scenarios
- There is no consolidated joint planning in advance with NZQA staff *and* the vendor, in order that multiple scenarios are considered from different perspectives and appropriate contingencies are jointly developed
- Likewise, the vendor has had limited input into NZQA's business continuity planning in the event of a service problem
- Information from the schools' relationship managers (SRMs) who are best placed to provide intelligence about ground level behaviours and logistics in schools, is not systematically fed into all levels of NZQA governance and management or to the vendor. Certainly SRMs themselves are unclear about whether their insights on digital assessment are being utilised and if so, where
- As the vendor has expanded and become more remote, it has a less granular understanding of NZQA's wider strategy and desired outcomes. Internal respondents suggested that NZQA could be more proactive in sharing enterprise strategy with the vendor; and
- NZQA may have been a little unassertive in ensuring that vendor performance information in regard to Assessment Master is shared across the agency and is of appropriate rigour and depth.

It can be easy, in long term vendor relationships, for all parties to lose sight of the bigger picture within which they are operating. Until recently, NZQA seems to have been overly relaxed in its management of this relationship. It has certainly been insufficiently rigorous in its reporting requirements from the vendor and in keeping contractual terms and SLAs relevant and up to date. For example, I find it odd that there are few relevant penalty clauses in the master contract, given the criticality of the assessment platform to the delivery of large scale national examinations.

Once these events happened, vendor and client quickly activated an incident management team and worked together to develop fixes. However, it is not clear to me that a 'hot debrief' was done post event to capture the lessons learned as soon as possible after the event. This would be best practice going forward. It also appears that the operational incident management framework that was followed was developed somewhat on the fly rather than being a tried and tested approach.

It is important that NZQA communicates its long term strategy for digital assessment and its wider organisational strategy to the vendor on a regular basis. The vendor would also value more information on the customer experience and the issues schools face to assist with their contribution to the partnership.

The vendor believes, and I agree, that it will be important, going forward, to undertake joint and consolidated scenario and contingency planning, well in advance of exam sessions.

The vendor is also now signalling a readiness to review SLAs, and this will be an important opportunity for NZQA to bring to bear the lessons learned from these events. The agency is currently taking steps to address the issues in relation to platform performance and capacity with the vendor.

NZQA should also consider the vendor relationship from a partnership perspective first, as opposed to focussing only on the contract. Given the criticality of this platform to NZQA's core business, the partnership needs to be deeper, more open and more anticipatory from both sides. More joint activity and due diligence should be engaged in in peacetime, to reduce the need to respond reactively once a crisis happens.

As a part of this more rigorous approach, NZQA should consider the optimal management and governance arrangements for this critical vendor relationship. Accountabilities need to be clearer, and there may be benefits from introducing external scrutiny into the governance structure.

While this is not within my Terms of Reference, it may be worth reviewing other key vendor contracts to ensure they do not reflect similar issues and that best practice contract lifecycle management systems are in place.

RECOMMENDATIONS

Review the portfolio of current technology vendor contracts with a view to ensuring that all Service Level Agreements, reporting requirements and performance metrics are current, appropriate and sufficient. Develop a plan for rolling reviews in future.

Develop a templated 'hot debrief' model through which lessons learned can be captured immediately after external assessments take place, to inform ongoing service delivery improvement and risk management.

Review and simplify the internal governance arrangements to improve the oversight of technology contracts and relationships, in order to clarify accountability, streamline decision making and take an integrated approach to this portfolio.

Consider the appointment of external members to a new or revamped advisory body for digital assessment and change.

PRODUCTION PERFORMANCE TESTING

Respondents raised a number of issues with me with regard to product performance testing, and this issue is also addressed in depth in the findings of NZQA's internal review. Problems in this slice of cheese/layer of protection included:

- The base assumptions and test environments in play from the vendor and NZQA were different. For example, the vendor's summaries of performance test results, (shared with NZQA on 6 April 2023 and 21 September 2023), showed that the platform could support up to 38,500 concurrent users for any one of the 2023 external assessments, based on agreed AutoSave functionality. NZQA's own performance tests produced significantly different results where the maximum capacity of concurrent users was only 25,000 users. While NZQA used this lower figure for prudence in its own planning, the inconsistencies and the differing assumptions which clearly lay behind them should, in hindsight, have been explored more fully, well in advance of the examination period
- NZQA performed user acceptance testing on Assessment Master but did not regularly undertake additional performance testing on the platform. In 2023, once the load capacity issues were identified, diagnostic performance testing showed potential looping and capacity problems. These were then applied to contingency plans which NZQA did in the event, use

- Testing came too late in the year and too close to the examination period. Post testing, there were a number of late patches put into Assessment Master by the vendor, without informing NZQA, that contributed to the problems with the LitNum assessments. NZQA testing did not identify any defect. This went to the save issue experienced during the LitNum assessments
- Vendor performance testing of the Assessment Master platform was undertaken in relative isolation from the other technology components of the student user's end to end assessment process. In other words, it was not tested holistically, in terms of its interactions with the web portal and NZQA's Keycloak security application. Nor was the potential impact on Assessment Master on internal monitoring systems and dashboards considered in vendor testing
- The vendor was not initially transparent about the assumption they had applied to their testing about the autosave time for the platform. It was not until last October, (very close to the exam session) that the vendor mentioned that the 38,500 projection could not be achieved within the expected autosave timeframe
- NZQA did test for Keycloak impacts, which contributed to its lower capacity projection. However, the login redirect defect, which caused students to make repeated login attempts between Keycloak and the assessment platform, placed unanticipated strain on the Database; and
- Testing was largely uninformed by understanding of potential customer behaviours with regard to how schools might schedule and manage cohort size and timing of the examinations. It took place within an IT and corporate bubble and seems not have been informed by the internal CX team's insights or those of SRMs. It appears not to have considered a wide range of worst case scenarios. As one internal respondent put it: *'Our testing was narrow and technical rather than based on performance and user experience'.*

These events provide an opportunity for NZQA to review its testing processes and approach. It will need to balance testing for a wider range of scenarios within its available resources.

In future, NZQA and the vendor also need to undertake testing on the same datasets and assumptions, and ensure that this is done well in advance of the exam session. Work has already commenced in this regard.

On the positive side, the agency did identify the potential for what later occurred and had developed contingencies which could then be applied in real time. The need for such contingencies was, however, not well enough communicated in advance, both internally and externally to schools.

RECOMMENDATION

Following the above review, align functional and product performance testing protocols between NZQA and its technology partners and vendors.

WEBSITE ENTRY

Some students were unable to login to the platform via the website entry portal (shown as a banner on the site) for around half an hour on October 30.

In addition, several stakeholders remarked on the fact that in 2022, entry via the website was done by direct click on a large banner on NZQA's homepage. In 2023, the banner was smaller and not live as a portal. The portal for login was high on the screen in a non-intuitive spot that required the student to scroll to it. As one PN put it: *"absolutely anyone looking at that website prior to the exams would have seen that the banner was going to be a problem.'*

School relationship managers told me that the absence of the log in tile from the NZQA website in advance of exams makes it hard for them to train PNs and others in the log in process, outside exam season. It also drives calls to the contact centre from students when they cannot locate the login while preparing for their exams.

Respondents internal to NZQA identified the following related matters:

- Maintenance of the URL inside NZQA is a manual process which can introduce human error. There is no advance UAT testing to check the accuracy of the URL
- Some schools had requested PN and/or student access to the platform the Friday prior to the examination period to allow students to set up logins. This is not something that NZQA supports. The banner portal tile was displayed only on the day of the assessment. Earlier access however, even if only for PNs and administrators, may have identified the problem in advance
- Various teams within NZQA engage with the website and thus the URL process. The channels team for example, (part of the CX function) expected the operations and logistics group to have pretested the link; and
- When the login redirect problems occurred, these should have been escalated to an internal operational incident management process, which does not appear to have happened.

These issues are administrative rather than technical and suggest the need for greater process clarity in NZQA about what team is responsible or what. The agency appears to already have these matters in hand.

THE KEYCLOAK INTERFACE

As noted above, Keycloak is NZQA's internally hosted application to provide security on the platform. Its interface with Assessment Master was not pretested by the vendor. It was not understood that Keycloak logins would degrade the platform's overall capacity. This also goes to my point above about the need for a more holistic and end to end approach to *joint* scenario and contingency planning.

In part this also seems to have been the result of poor handover management when Keycloak moved from project status to BAU within NZQA. While Keycloak has expert NZQA staff attached to it, in house understanding of the relationship between the application and the overall platform seems to have been patchy.

In hindsight, this issue could have been avoided had more comprehensive end to end testing been undertaken from a customer perspective, by both NZQA and the vendor.

ASSESSMENT MASTER CAPACITY ISSUES

NZQA has been aware, in previous exam sessions, of capacity issues on the vendor hosted assessment platform. In 2023, the vendor's pre session testing showed that maximum capacity of the Assessment Master platform in the New Zealand setting was 38,500 users, as noted earlier. NZQA's own testing showed 25,000. For prudence, the lower figure was used by NZQA in its planning for the large cohort in level 1 English.

NZQA planners then applied an attrition assumption, based on prior year attrition rates, of some 20% of exam entrants, (i.e. students who were entered but would not actually present for the exam on the day), and used this lower figure for their capacity sufficiency decision making. A further assumption was made for the LitNum assessments, that schools would want to undertake the assessment at various times of day (and in regular classes) across the available exam window of one week. As one respondent described it:

> *"38,500 users was the preliminary target agreed with the vendor. This is because early projections in July were up to 35,000 entries for Level I English, plus 10% for media studies. Then we took an 80% of max figure for testing purposes. In the past actual uptake has been more like 50-75%. So*

all this maths ended up with a 24,000 projection, which we also used in the 2022 performance testing."

Based on these calculations, it was assumed that there would be sufficient redundancy in the system to manage demand. However, these projections failed to account for:

- The fact that most users for the corequisite exam appear to have logged in at around the same time and day of the week, which was inconsistent with historic demand patterns
- The degradation of capacity created by Keycloak; and
- Additional degradation of capacity from other platform monitoring activities.

NZQA did ask the vendor for more information to explain the apparent misalignment of capacity. The vendor declined to provide this at the time, given the proximity of the request to the exam session. They have since undertaken to provide more details about their test assumptions and parameters.

This particular set of issues may be unlikely to arise in future, provided the vendor can deliver on its current internal (and very ambitious) project to scale up its infrastructure and increase overall capacity to 150,000 users this year. Vendor progress towards this target appears partial at the time of writing.

This also illustrates a need for NZQA and the vendor to:

- Better document and map the end to end system interfaces, from the web portal to Keycloak to Assessment Master
- Be more open with each other about assumptions and adjunct applications
- Ensure that fixes and changes are finished and tested well in advance of the exam window
- NZQA should revisit its attrition assumptions, in partnership with schools, so that it has a clear ex ante picture of the future pattern of consumption; and
- If necessary (and practically feasible), require schools to stagger participation in the corequisite to reduce interim loads until platform capacity is lifted.

As noted above, in addition to these capacity issues, the vendor appears to have introduced an unknown bug into Assessment Master with some changes made close to the commencement of exams. This contributed the slowdown in the Numeracy exams. The vendor told me the change freeze period would be extended in future. NZQA would do well to capture this undertaking in the revised SLAs I suggest.

NZQA should also thoroughly challenge the vendor as to the feasibility of its current capacity enhancement project. Their aspirational scale up to around four times current platform capacity may not be a realistic target in advance of the 2024 sessions.

RECOMMENDATIONS

Develop and promulgate a fit for purpose operational incident management model, with accompanying Standard Operating Procedures (SOPs), which clearly articulates core processes, escalation points and accountabilities.

Consider and engage with schools on the optimal mechanisms to ensure that digital development and assessment planning is informed by customer (schools' and students') behaviours and experience.

SCHOOL AND SYSTEM FACTORS

HIGH CHANGE TEMPO

The backdrop to these events is one of rapid and disruptive change in education, both in New Zealand and elsewhere. This is driven by increasing diversity of learner needs, modern technologies (such as AI) , new awareness of the need to address educational inequality, demand for education to address indigenous learning, a hollowed out teacher workforce and many other trends.

In New Zealand, secondary schools are experiencing a high change tempo with changes in both national curriculum and the NCEA. This can make it difficult for schools to engage in co design with government agencies, properly service pilots and reviews and, in this context, to arrange and deliver external assessments with the perfect configuration of facilities and staff. It also tends to create significant workforce exhaustion at year end.

For example, many schools told me of students sitting in multiple rooms for a single exam due to the needs of SAC students for quiet environments and the pressure on facilities at year end. When problems occurred, this made moving to a paper contingency or getting messages out to supervisors a logistical challenge.

Teachers and PNs also told me that the problems with digital assessment seemed like the final straw at the end of a frantic year when teachers and students were tired. This no doubt contributed to the emotionalism of many respondents, as outlined in the later section on stakeholder impact.

SCHOOLS' CHOICE AND CUSTOMISATION

At system level, New Zealand's highly devolved and 'self-governing schools' based model of governance, the complexity of its secondary school curriculum and the heavy emphasis on assessment, drive inconsistencies, inequities and risks into the system even without the added issue of increased digitisation of learning and assessment. [14] System incentives are for schools to differentiate themselves competitively and they do this at least in part by developing differentiated approaches to the digitisation of learning and their approaches to assessment.

The self-governing schools model also means that every school or kura has to have one of everything, which creates at best inconsistency and at worst, inequity. A PN at one school might be vastly experienced and a member of the School's leadership team. The PN at the school next door may be neither. Remote or small schools can find it hard to resource a PN position at all or to hire exam centre managers with the right skills.

The size of some schools lends itself to mass assessment events. In others, facilities must be cribbed from business as usual. These variables, and there are many others, drive diverse behaviours and approaches to the management of assessment events. This in turn implies that NZQA must be responsive to an enormous range of service delivery needs and aspirations across New Zealand's 500 secondary schools.

As one respondent put it:

> *'At its root, this is about schools' choice. It's really hard to deliver fail safe service when the 'customer' can consume the service in so many different ways'.*

It may now be timely to work with schools to co design a restricted set of choices with regard to digital assessment delivery in the interests of consistency and service quality. For example, co requisite exams may in

[14] See https://conversation.education.govt.nz/assets/TSR/Tomorrows-Schools-Review-Report-13Dec2018.PDF

future need to be delivered on a regionally staggered basis, at least for 2024, until platform capacity can be lifted. Large cohort NCEA exams may need to have two national sittings. Technology experts assure me that this can be done without damaging security and hence exam authenticity.

THE PN MODEL

A number of respondents drew a line between these events and what they saw as overloading of the PN model in schools. PN are the primary in-school contact for NZQA. They are administrators for examination entries and have access to the platform to monitor delivery during exams. While in the past the PN role was largely administrative, it has expanded over time and now constitutes a considerable workload.

Although NZQA typically meets with PNs on a regular basis, several PNs told me, in the words of one: *'PNs can provide insight into what works best on the ground. But is anyone listening?'* This cadre should be a critical source of intelligence for NZQA and if well engaged, a mechanism for driving greater consistency across the system.

PNs also felt under informed about NZQA's overall digital assessment strategy and outcomes. *'What will 'done' look like?'* said one. PNs told me they would value more two way engagement and better training experiences. *'We're keen to be part of the solution for NZQA'* said one respondent.

NZQA staff respondents also felt there was opportunity here, '*If we can invest in them they are high value partners for us'*, said one.

THE ECM WORKFORCE MODEL

The recruitment of appropriately skilled Exam Centre managers, who are recruited by each School or Kura but employed by NZQA, is a challenge for many schools. The role is poorly paid. As many ECMs are retired teachers, it can be difficult to find the technology savvy managers (and the supervisors who work to the ECM) needed for a digital assessment. When these events occurred, especially with log in redirects, some schools told me that their ECM panicked and did not know what to do. Teachers familiar with the platform had to be brought into the rooms to assist the students, which was in breach of external exam protocols.

With regard to the paper contingency, some ECMs were well prepared to execute this plan and others were not. Some schools described running frantically around with pencils and printer paper to print off paper copies of the exams in a hurry.

As fixed term employees of NZQA, ECMs in the past received a one day training package prior to the exam session. This year their training was reduced to a half day due to budget pressures within NZQA. Webinars were used to supplement in person training (perhaps not the best channel, given the need to upskill the IT skills of this audience). Respondents were consistent in their view about the need to, as one put it bluntly:

> *'Sort the ECM model. We've known about the workforce issues for ages and they'll only get worse. Either train and pay them properly or get teachers to do it.'*

RECOMMENDATIONS

With the Ministry of Education, examine the Principal's Nominee model in schools with a view to:

Developing a consistent and clear position description

Understanding the support needed to those in this position

Developing a plan for the professional development of this group; and

As for recommendation 15, clarifying channels by which PN feedback and input into service delivery design can be captured to inform practice.

Review and promulgate a new success profile for exam centre managers and develop a mini workforce strategy for this group of contract staff. Once this is done, redevelop the training for these fixed term staff to equip them for a digital world.

ORGANISATIONAL FACTORS

Within NZQA, I see these assessment events as a microcosm illustrative of a range of organisational issues, some of which will require considerable leadership effort to address as they go directly to internal culture. In my view, these matters compounded the technology issues discussed above.

For some readers it may seem a long bow to draw inferences about organisational culture from this series of events. Having undertaken an earlier review of the organisation's performance in 2021, however, it appears that some of the issues identified then, such as a very siloed operating model, did contribute to these events.

THE 'WHY' OF DIGITAL DELIVERY

NZQA has been on its trail blazing journey towards digital assessment for a number of years now. New Zealand is among the most advanced jurisdictions in the world in this regard.

Over time, as noted earlier, the core vision for digital assessment has changed from a focus on 'anytime, anywhere' access, to a deeper mission centred on learner equity, support for disadvantaged learners and greater assurance and efficiency, all underpinned by the agency's strong commitment to supporting mātauranga Māori provision.

In the meantime too, NZQA's overall organisational strategy has evolved and changed. Almost all respondents expressed a desire to better understand both the enterprise strategy and how digital assessment fits into that strategy. Schools, NZQA staff and vendor representatives all told me that they lacked clarity about what the 'end game' for digital assessment was and how success should best be measured.

As a result, while NZQA's people are passionate about their work and feel a strong personal sense of mission, they do not appear to see the agency's current strategy as critical context for their own work.

Some internal respondents were unclear about both what success means with regard to digital assessment and the contribution they could make to that success. This seems to be making it difficult for the teams involved to prioritise their work.

The 'why' (vision, plan and KPIs) for digital assessment is not presently sufficiently galvanising to unify an agency characterised by deeply embedded organisational silos.

This risks adversely impacting the development of a customer-led, systems-based and holistic view of the project, consistently applied across the agency.

A refreshed digital assessment 'why', along with progress and summative measures, will be critical if NZQA is to create a shared focus for the next phase of the digital assessment journey. It will need to be clearly and succinctly communicated through the organisation and outward to partners and stakeholders.

RECOMMENDATION

As a subset of organisational strategy, revisit, clarify and promulgate the NZQA strategy for assessment over the coming years, including clear outcome statements and progress metrics.

STRATEGIC LINE OF SIGHT

On a related point, while both the NZQA Board and executive team are clear about both enterprise strategy and the place of digital assessment in that strategy, there seems to be poor line of sight on the strategy and outcomes, *through* the organisation to its frontline.

This is in part an operating model issue. NZQA has a siloed organisational structure characterised by multiple teams touching a process. For example, for digital assessment there are two assessment teams (internal and external), a digital assessment and change team, the operations and logistics team, the CX team, IT teams, SRMs and quality assurance personnel. Some activities are duplicated across these groups.

Organisational structure is also quite steep given the size of the agency.

I suggest there is also a 'frozen middle' problem. I am not convinced that all third tier managers in the agency have their eyes firmly fixed on the vision for digital assessment and can communicate it well to staff and stakeholders. Nor am I convinced they are consistently looking *across* the enterprise as well as down into their silo responsibilities. NZQA activities sometimes appear to be driven by reactive developments and personalities, as opposed to being rooted in a clear vision, strategy and operating model.

When these events took place, much of the response and recovery effort, while urgent and committed, was based on interpersonal interaction and ad hoc fixes, as opposed to clear frameworks and processes which had been designed in advance to reflect overall strategy.

NZQA should consider the best mechanisms to improve line of sight and thus a shared focus on outcomes. These might include:

- A reset of the operating model to support the next stage in the organisation's development, including consideration of where critical functions best sit and how existing silos can be consolidated
- Ensuring that the CX team works across all functions, so that a customer mindset is imbued throughout the agency and not just in this team
- Ensuring the induction and professional development of third tier managers supports them to take an enterprise view. Their performance management arrangements should consider their contribution to the whole of NZQA rather than being focussed only on their intra team leadership; and
- Ensuring that strategic outcomes are cascaded into the performance measures of all staff.

RECOMMENDATION

Ensure clear line of sight for NZQA staff on the overall strategic context within which digital assessment service delivery sits, by:

ensuring that team and individual performance measures align to strategy; and

Ensuring that middle managers have the leadership skill to understand their accountability for joined up service delivery, rather than focussing only the priorities of their individual teams.

Review NZQA's current operating model and consider:

Better integrating the CX and SRM teams into co-design, service delivery and change processes

Consolidating the current assessment teams to ensure an end to end view of all assessment experiences

Ensure that specialist skill sets are located in the right places in the organisational structure and are not duplicated; and

Clarifying the respective accountabilities of internal teams in the end to end assessment process, including IS, operations and logistics, communications, CX, SRMs and assessment.

ACCOUNTABILITY CULTURE

These line of sight issues contribute to a low accountability culture within NZQA, below the senior team. This was a strong theme amongst staff respondents, who felt that the complex organisational structure and siloisation allowed everyone to duck ultimate accountability for the successful operation of digital assessments. It will be critical that as part of the work on operating model design, the agency maps accountabilities, decision paths and escalation protocols.

In this it needs to move from an informal, personality based culture where things often get done though relationships, to a more systematic and mature culture where things get done according to clear accountabilities, formal frameworks and well understood targets and measures.

A responsibility assignment matrix (commonly known as a RACI[15]) does not appear to have been clearly articulated for the end to end digital assessment process.

NZQA should consider developing, and then publishing, accountability statements and maps for all staff involved in this core process from first line leadership upwards. The process of developing the statements and maps will provide the agency with the opportunity to carefully reflect on who should be accountable for what, and why.

PROCESS CLARITY

NZQA also appears not to have fully mapped its core processes as they relate to digital assessment. Process maps and standard operating procedures (SOPs) lack currency, are used only in pockets or exist in an overly complex form that makes it hard for all staff to understand them.

This was a common concern for internal respondents. They suggested that there was little clarity in regard to process steps, and around key handoff and sign off points. Decision pathways were also unclear, in that staff were not sure where definitive decisions would be made on particular matters.

These events also highlighted gaps in the operational incident management processes for digital assessments. While NZQA's Assessment Division has a documented CIMs[16] process for use in major events, the process was problematic in that:

- There are definitional issues about what constitutes a major incident
- Thresholds for escalation are poorly defined; and
- The roles and responsibilities of the various players are not clear.

In the wake of these events, and as part of its operating model design work, NZQA should work across teams to create both current and desired future state process maps of the core processes entailed in digital assessment. It should create simple and visual SOPs which can be used for incident management.

The same rigour should be applied to the processes by which stakeholders can engage in the event of a crisis. During these events, stakeholders reported considerable confusion as to who to contact for what. Staff in the

[15] A RACI sets out responsibilities ('who does'), accountabilities ('who approves'), who is consulted ('who has input') and who is informed ('who sees the output') in relation to activities in cross-functional projects and processes.
[16] CIMs is the acronym for 'Coordinated Incident Management System'.

contact centre felt similarly confused. This exercise could also consider the need for additional surge capacity in the call centre during the examination period.

RECOMMENDATION

Once the operating model is reset, map, streamline and document all internal processes relating to digital service delivery and ensure these form a part of staff induction into the relevant teams.

CUSTOMER EXPERIENCE AND BEHAVIOURAL MINDSETS

NZQA manages its interfaces with schools via school relationship managers, (known internally as 'SRMs' and with roles similar to those of account managers). These staff tend to be held in high regard by the schools they service.

In an ideal world, these staff would be the eyes and ears of NZQA. They would be transmitting insights and intelligence to NZQA about schools' intentions for the exams. They would also be describing the behaviours and pain points they are seeing in schools around digital assessment.

SRMs with whom I spoke felt that they often had, as one put it: *'24 carat gold intel'* but that it was hard to know where to take it within NZQA and to see what action was taken as a result. Some felt that corporate staff distrusted them and believed them to have been 'captured' by schools' perspectives.

Similarly, system stakeholders felt that many of the problems that occurred in 2023 could have been avoided if NZQA had a better sense of the operational realities in school settings. One said, in a comment typical of many:

> *'The kernel of all this is a lack of understanding about how schools operate, and the end of year pressures on our teachers, facilities and timetable. Our SRM is awesome, but do they have a voice to inform practice inside the mothership?'*

I suggest that NZQA should see SRMs as the face of NZQA into schools and a key portal for issues requiring troubleshooting. They should shield the customer from having to learn how to navigate the regulator themselves. This would imply that SRMs have a higher internal profile and deeper understanding of internal processes, capabilities and accountabilities than they presently do.

In addition to the SRMs, NZQA has a pool of specialist talent in its CX team, which undertakes research into school and learner experiences and also manages the website. This team sits in the corporate part of the organisation and appears somewhat disconnected from core service delivery, including digital assessment.

If NZQA is in future to design and deliver digital assessment with a strong customer focus, any future processes must provide for inputs from both the on ground 'account managers' and the in house customer experts to analyse school and student behaviours, stressors and drivers. If this had been done systemically and prior to the 2023 exam period, some of the LitNum issues could have been prevented or at least mitigated.

Once NZQA can harness its internal capability to gain insights into how schools and learners make decisions, it will help identify problems and enable it to tailor solutions that are potentially more efficient and effective. In the complex world of education, NZQA, from the perspective of both its regulatory and service delivery functions, must incorporate behavioural research into its work.

RECOMMENDATION

Develop improved processes for capturing school and learner related behavioural insights into internal service delivery and technology design processes. Consider in particular the roles of PNs and SRMs in such processes.

DATA CAPTURE AND INTELLIGENCE LED DESIGN

Related to this, NZQA needs explicit mechanisms for capturing data related to assessment and turning it into useful intelligence. At the moment, relevant information and data analytics sit in multiple organisational functions and it can be challenging to create a joined up picture. This also means that internal teams are sometimes operating on different and non-explicit assumptions.

As part of its operating model rethink, NZQA should consider the data and intelligence related to digital assessment and other critical services that it needs to collect, hold or access.

In the absence of sound analytics there is a risk that decisions about technology may dictate business requirements, rather than business and customer requirements driving technology decisions.

NZQA should also make sure it has an appropriate level of understanding and access to analysis of customer wants, needs and behaviours, especially regarding vulnerable learners. Internal expertise on customer analytics exists in the CX team but is not yet sufficiently broad or deep. An understanding of behavioural economics is also needed.

Finally, while NZQA rightly prides itself on its innovative and partnership based approach to technology solutions, I am not sure that it is yet sufficiently nimble to apply design tools to service delivery improvements. There is an opportunity now to use some design tools in conjunction with the customer, and particularly with PNs and SRMs, to prototype, iterate and refine a failsafe approach to the 2024 exam sessions.

RECOMMENDATION

Move away from siloed data (both within NZQA and across sector partners) to joined-up analysis, intelligence, research and insights. This will require:

a deeper understanding of the data that NZQA (and others) already hold

investment in internal capability to turn data into insights; and

capacity and capability to publish and share those insights with stakeholders more effectively.

RISK MANAGEMENT AND SCENARIO PLANNING

NZQA has, in recent years, strengthened its risk management and assurance activity. Clear risk appetite statements exist, with the statement relevant to digital assessment calling for ‘zero appetite’ for external exam delivery related risks.

The risk registers pertaining to digital assessment appear to be broadly sound, but I am not convinced that a sufficiently mature and sophisticated approach to managing risk is yet being applied at all stages in the digital assessment process. My specific concerns are:

- Risk registers appear to be viewed as a static compliance exercise, as opposed to teams constantly asking themselves 'what if?' and dynamically reviewing and managing risks
- Identified risks are more tactical than strategic
- There appears to be optimism bias, whereby NZQA staff tend to plan for the best, rather than game out worst case scenarios
- Scenario thinking appears underdeveloped, with the possible exception of the cyber security area, in that teams do not typically have the resources and time to engage in wide ranging scenario planning. When they do so, it tends to be internal and not to include the relevant vendor or customer facing staff; and
- The link between lessons learned processes, data analytics and risk identification is not clear.

There is opportunity now for NZQA to hold some reflective sessions, involving diverse (and ideally some external) participants, which convert the lessons learned from these events into a refreshed risk register and business continuity plan. This should involve all of the management and governance bodies for digital assessment.

There is something deeper here too, in that while NZQA has always tended to be an extremely risk averse organisation with a conservative culture, some staff seem to struggle to speak out about difficult issues and probe relentlessly for the weak spots in service delivery. Managers need to model and the organisation needs to be comfortable with lively debate about risks and how and where to manage them.

RECOMMENDATION

Using scenario planning techniques, review and reset the risk register for digital assessment and change, including mitigation and contingency plans.

COMMUNICATIONS

Not only does NZQA have an internally conservative culture, but it also tends to be modest in regard to external communications such as those relating to strategy and research outcomes. As one respondent said:

> *'Get out there are talk more about digital assessment to the public. NZQA does great research and is leading this in the world. But they tend be defensive rather than proactive.'*

Proactive communications, to partners, the sector and to parents, which address both the success of and challenges entailed in digital assessment, would have the effect of expanding NZQA's license to operate and improving stakeholder understanding of the risks and challenges.

Most schools told me they felt the communication from NZQA during these events, while copious, was lacking in timeliness, had a somewhat defensive tone and was confusing in that there were often subtle differences in the communications sent to different audiences. Principals, in particular, felt that NZQA's communications did not equip them with sufficient facts to inform and support students and their families.

Schools would like to see more timely, urgent and open communications about any challenges that may impact examination events. As one PN put it: '*we know Assessment Master is a work in progress. We want to be invited to be part of the solution.'*

From an internal perspective, NZQA will need to reflect on how best to ensure proactive communications strategy, collateral development, channel management and approvals processes to support timely and joined up communications with stakeholders. This should be a specific focus of the work suggested here on operating model refinement.

RECOMMENDATION

Consider opportunities for more proactive and regular system facing communications, based on in house research, about the strategy for and lessons learned from digital assessments.

IMPACT ON LEARNERS

There is no question that these events disrupted, distracted and stressed some students. The impact appears to have been most pronounced for those students sitting an external assessment for the first time and students with special assessment conditions, some of whom are vulnerable learners.

Impacts varied however, from school to school and from student to student. Some schools engaged closely with NZQA around contingency planning after the LitNum assessment events and were well prepared with a paper contingency for NCEA English. Some had technically skilled ECMs or supervisors who could help with the URL problems by logging in through a known back door. In other schools there was panic, with the ECM and teachers running from room to room carrying messages or papers and sizeable number of students walking out of the exam. The complaints received from parents show that even those students who got into the exam with no trouble and completed their entry felt that they did so under poor conditions, given the considerable distraction.

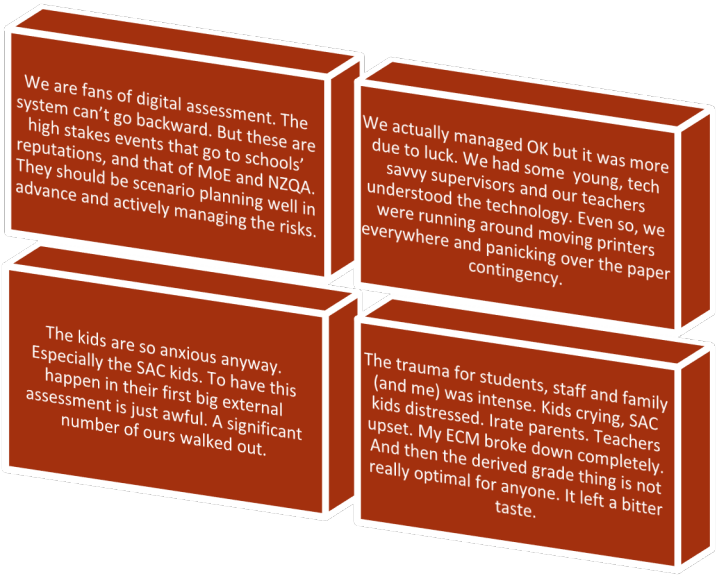
Some students reported that they were horrified to have to produce a handwritten exam, as they had no experience of handwriting for lengthy periods. They worried that their work was illegible or incomplete. Others became frustrated by their experiences in regard to the slow save fix that NZQA instituted after the platform load became critical. Still others were deeply distressed by being closed out of the system.

As a consequence of these quite different experiences it is difficult to put a number on the students who were disrupted. Some schools took the blanket view, after the NCEA English event and disruption to other NCEA exams such as media studies, that their entire cohort should apply for a derived grade.

Overall we know that in the 2023 school year, 263,000 students successfully completed digital assessments. During these events, 16,884 students successfully completed the digital exam for NCEA English Level `1.There were subsequently around 8,850 applications for derived grades, some subset of which related to the disruption caused by these events.

Interestingly, of those who applied for derived grades and also completed the exam, 59% got a better result from the digital examination than from their prior test scores. (For a derived grade, NZQA takes whichever of the two grades is higher.)

In the graphic below, I share some indicative verbatim responses from teachers and PNs when asked about the impact on their students.

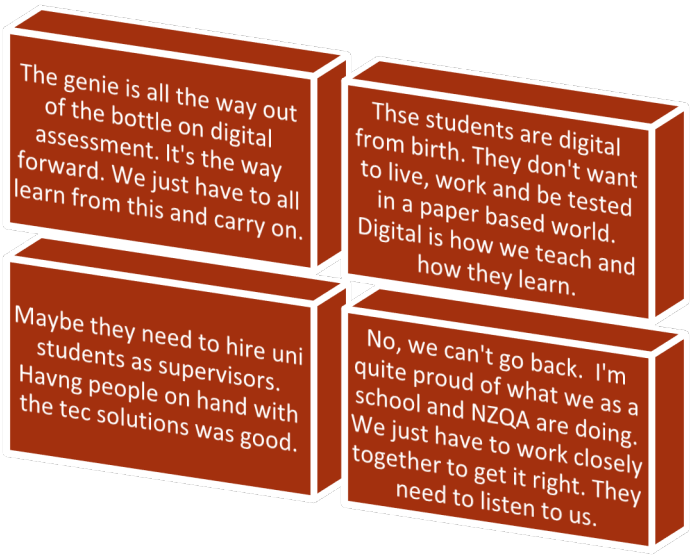


A number of respondents also felt these events had a knock on impact to the reputation and credibility of the qualification itself. As one put it: '*Yes, it damages progress with digital assessment, but beyond that, it damages trust in the actual qualification'*.

Others also suggested that the experience of these events was already creating exam hesitancy amongst students for the 2024 exam period. One teacher said:

> *'Those kids who were most impacted, particularly the SAC kids, are having a tough time thinking about the next external assessment. We are trying to build their confidence and allay their fears but we are also worried and they can sense our uncertainty. 2024 is high stakes for everyone, us, the kids, the sector and NZQA'.*

However, when asked about whether these experiences had completely turned schools or learners off the digital assessment experience, respondents were universally positive about continuing. A verbatim sample representative of stakeholder responses is shown below.



IMPACT ON STAKEHOLDERS AND PARTNERS

Although some stakeholders, and particularly school principals, were very vocal about the adverse impacts on students caused by these events, they were also, without exception, committed to continuing to partner with NZQA on the digital assessment journey.

When asked about what they saw as the root causes of these events, most respondents volunteered comments similar to this one:

> *'Even though we were keen adopters, all the way through we've felt that they don't understand the modern realities of school operations and students' needs at the coalface. Like, telling us to do the exam again 'after morning tea'. Haven't they ever met the realities of a school timetable?'*

However, also without exception, all stakeholders were keen to be involved in planning and testing activities that would help prevent a recurrence of the 2023 exam events. As one said:

> *'If you are really being agile then you have to partner with us to do the design work. I'm not sure NZQA is really using the insights of PNs and others to drive the implementation. That's a real opportunity.'*

When asked about solutions, stakeholders also stressed the need for:

- Improved clarity about the overall strategy and key milestones for digital assessment
- Stringent internal quality controls over all aspects of the examination, from entries to testing
- More fully developed contingency plans, communicated to schools and ECMs well in advance
- Improved training for ECMs, or the development of an alternative model
- SRMs to be well equipped with the information needed to support schools with examination session planning and communications; and
- Clear SOPs, incident management protocols and communication channels for escalation in the event of problems.

For some stakeholders, there was an elevated level of emotion in these events, in that they were the 'last straw' at the end of a busy year. For others there was also a feeling of moral hazard, as here from a school respondent:

> *'We were given assurances from NZQA and we relied on them to allay the fears of students and families. When it all went wrong we felt personally and professionally devasted. We felt we had acted in bad faith to students.'*

Some NZQA staff shared this feeling:

> *'It's like we lost sight of the why of digital assessment and focussed on 'sales' to schools. I feel so compromised that I sold a faulty product.'*

In some schools however, there was a sense that these events, while undesirable, had been well prepared for on the basis of NZQA's advance communications to PNs:

> *'We were all ready to go with the paper option, we had prepared the kids for that and my ECM was organised. The initial log in thing was painful but we just stayed calm and worked through it.'*

There was an almost universal sense amongst respondents that this review was a positive signal that NZQA wanted to learn from these events and would move heaven and earth to prevent a recurrence in the 2024 exam session. Balancing this, however, was a strong view that, in the words of one; *'this won't be fixed by tweaks. They have to go all in and get this right. Plus, they have to do it with us, not to us.'*

CONCLUSIONS

Delivering large scale national examinations, whether in paper form or online, is always a complex and high stakes undertaking. As the respondent quoted at the head of this report put it: *'about a million things have to go right'* to make it happen. And mostly, NZQA does get it right. Even with these digital assessment failures, the majority of students did not have a disrupted experience.

However, respondents are also correct in saying that problematic events of this kind seriously damage the confidence of those students who were affected, and that of their families and teachers, in the future of digital assessment. It is clear that in some schools, there is real fear of a repetition in 2024.

From a system perspective, any repetition would not only damage the digital assessment programme but tarnish the qualifications themselves and public confidence in them.

Given that it took a lot of holes in the Swiss cheese to line up to create last year's problems, and in light of all players' strong commitment to quality improvement, I believe the chances of similar problems in this year's exam session are low.

However, these events have highlighted important opportunities for NZQA to improve both its internal operating model and its partnership with the sector to codesign workable solutions. The agency needs to come out of its current somewhat shy positioning and seize these opportunities while their drivers remain fresh.

In closing, I have considerable faith that NZQA will work hard and collaboratively with the sector on the recommendations to enhance digital examination integrity made in this report. My discussions with managers, vendors and others suggest both a powerful commitment to doing so and constructive first steps.

In closing, I have two cautionary notes about the implementation of the recommendations made here.

My first concern is that the agency will adopt incremental or partial solutions to the matters identified as organisational root causes. Some of the internal cultural issues I traverse here are of long standing. They should be addressed by a thorough reset of the operating model in the context of current strategy. The matters surfaced by these events are almost certainly symptomatic of wider organisational issues.

Second, I worry about the tension between increased expectations on NZQA to meet the challenges identified here and its current financial position and modest baseline. It will be important, as part of the reset of the operating model, that NZQA's leadership considers how organisational efficiencies can best be secured, to allow them to reinvest resources in reliable and failsafe mission critical services such as digital assessment.

APPENDIX ONE : TERMS OF REFERENCE

EXTERNAL REVIEW INTO THE DELIVERY OF DIGITAL ASSESSMENT EVENTS:

PURPOSE

The Chief Executive and Deputy Chief Executive, Assessment of NZQA have initiated an external review into the issues that arose with the delivery of digital assessment events between 30 October to 10 November 2023. This document sets out the Terms of Reference for the review.

BACKGROUND

A core function of New Zealand Qualifications Authority (NZQA) is administering the national assessment system for secondary school learners. This includes the New Zealand Scholarship examinations and external assessments for NCEA.

While the majority of NCEA exams are still paper based, digital assessment is now well established with increasing numbers of schools and students participating, and with an increasing number of achievement standards being assessed digitally, where appropriate.

Twenty-nine digital assessment events were held during the period 30 October to 10 November 2023 on the digital external assessment platform.

This included Literacy & Numeracy | Te Reo Matatini me Te Pāngarau (LitNum | TRMTP) assessments and the English Level 1 exam, which were characterised by having large cohorts of candidates entered:

- LitNum | TRMTP (5 standards) - 129,806 digital entries across 59,463 students, and
- English Level 1 (RAS + BAU) - 27,320 students with digital entries.

Students sitting these digital assessments experienced various issues on the assessment platform on the day of the assessment. These issues also impacted other cohorts of students that sat their digital assessments in other subjects during this time period. The issue with access to the assessment platform for students doing the assessments on 2, 3 and 10 November has damaged the confidence that teachers, other school staff, students and their whānau have in NZQA and digital assessment. This may impede our ability to successfully drive the uptake of digital external assessment across NCEA and NZ Scholarship.

REQUIREMENT FOR AN EXTERNAL REVIEW

In response to the issues that affected digital assessment delivery, NZQA's Chief Executive and Deputy Chief Executive Assessment propose to address the performance and stakeholder/public confidence issues with a planned series of steps.

The first of these, is through an internal review commissioned immediately following the last external assessment on the digital assessment platform (refer to Appendix 1 for the Terms of Reference of the internal review).

The internal review aims to:

- define the root causes to the issues that disrupted delivery of external assessment events held from 30 October to 10 November 2023, and
- investigate and provide recommendations resulting in an action plan to maintain the integrity of the delivery of the digital assessment to secondary schools.

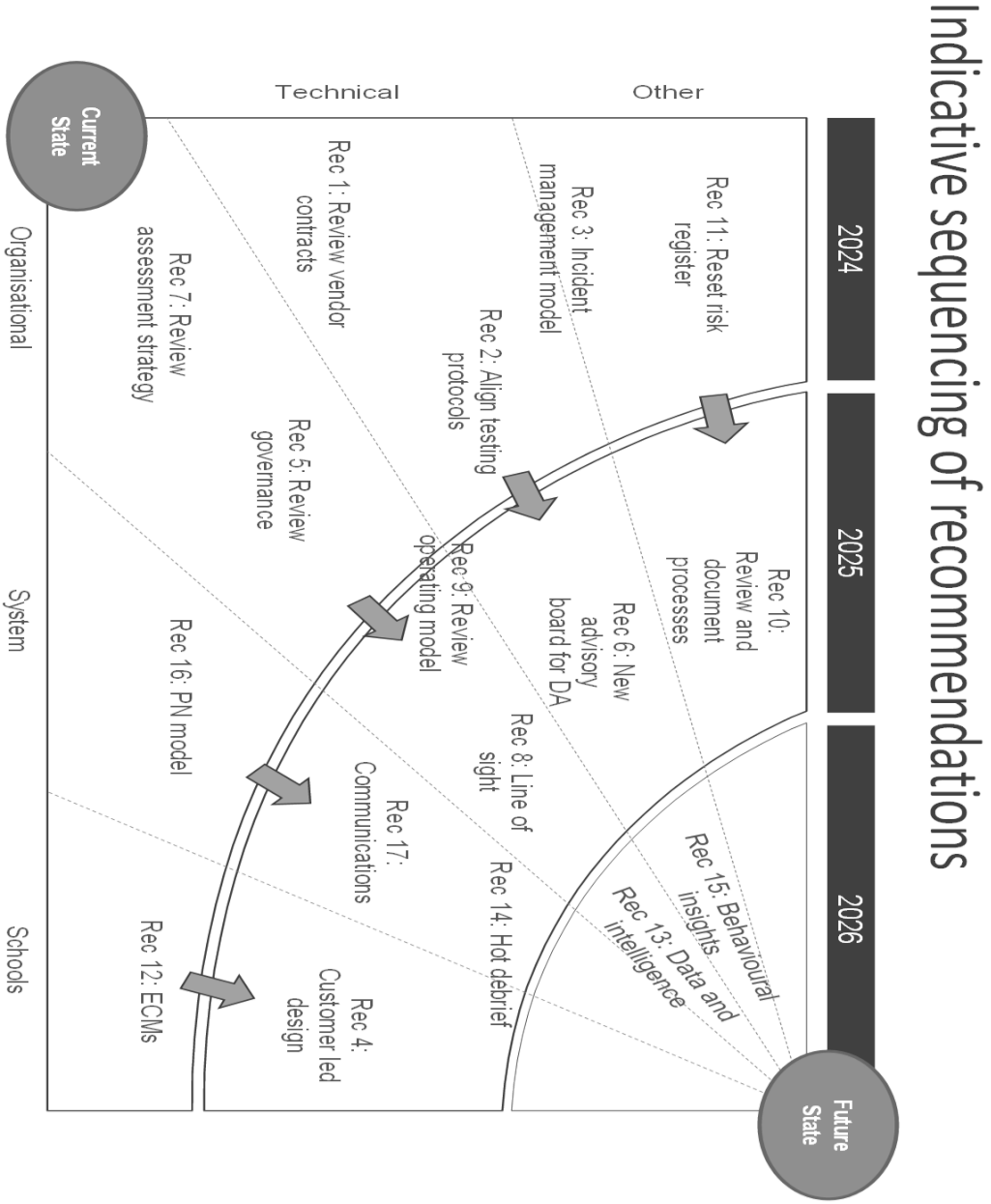
The second review will be performed by an independent consultant based on information provided by NZQA and primary results of the internal review. The consultant will report their findings to the NZQA Board, and the summary of the report will be made public.

The external review aims to:

- provide an overview of the development, successes and challenges for NZQA's digital assessment transformation journey since its introduction in 2015.
- outline the issues and root causes that occurred with the delivery of digital assessment events between 30 October to 10 November 2023, focusing on;
    - the completeness/adequacy of end-to-end functional testing and load testing
    - the alignment of the vendor and NZQA's testing environments, and of NZQA's testing and production environment; and
    - key programme decision points,
- assess the impact the issues had on various stakeholders,
- provide key recommendations on issues to be resolved for the delivery of digital assessment in the future which will enable NZQA to rebuild stakeholder confidence.

The report will be shared with a sample of the external stakeholders for their feedback prior to finalisation of the external report. The external consultant will report their overall findings to the NZQA Board, and the report summary will be made public.

APPENDIX TWO: POSSIBLE IMPLEMENTATION MAP FOR RECOMMENDATIONS



APPENDIX THREE: INTERVIEWS FOR THIS REVIEW

Confidential interviews were held with samples of stakeholder groups as shown in the graphic below: