



NEW ZEALAND QUALIFICATIONS AUTHORITY
MANA TOHU MĀTAURANGA O AOTEAROA

QUALIFY FOR THE FUTURE WORLD
KIA NOHO TAKATŪ KI TŌ ĀMUA AO!

National Certificate of Educational Assessment (NCEA) External Digital Assessment 2016 Trials and Pilots

User Experience Evaluation Report

May 2017

CONTENTS

GLOSSARY	3
Executive Summary	4
2016 Digital Trials and Pilots	6
Purpose of this document	6
Description	6
Assessments Trialled and Piloted	7
Scope reductions and challenges during 2016	8
Music Trial	8
French Pilot	8
Tahi	8
English Pilot	8
Special Assessment Conditions	11
Key demographics	12
Participation in Trials	12
Participation in Pilots	13
Summary of Findings	15
Student experience	15
English Pilot	15
Media Studies and Classical Studies Pilots	16
Digital Trials	17
Teacher experience (Managing Trials)	19
Marker experience (Pilots)	19
Examination Centre Manager / supervisor experience (Pilots)	20
Psychometric analysis (English Pilot)	22
Discussion	22
Project investment objectives	22
Trial and evaluate critical elements	22
How to deliver	23
Identify operating model changes	23
Understand support needed	23
Understand how learners interface	24
Identify management and logistics benefits	24
Understand how to transition and embed change	24
Reflection on NZQA approach to working with school leaders, teachers and students	24

GLOSSARY

<i>API</i>	Application Programme Interface
<i>DAT</i>	Digital Assessment Transformation
<i>eQA</i>	NZQA qualifications information management system
<i>IAM</i>	Identity and Access Management
<i>NCEA</i>	National Certificate of Educational Achievement
<i>NCEA Online</i>	NCEA Online: Transforming Assessment for Learners
<i>N4L</i>	Network for Learning
<i>NZQA</i>	New Zealand Qualifications Authority
<i>2016 Trials & Pilots</i>	NCEA Digital Assessment: External Assessment Trials and Pilots Project 2016
<i>BYOD</i>	Bring Your Own Device – school policy for student access to digital devices (e.g. laptops, Chromebooks, iPads etc)

Executive Summary

Students in New Zealand are increasingly exposed to digital learning environments, which assessment needs to match. As part of NZQA's digital transformation, we are investigating technology-based methods of assessment, with the goal of having all examinations available online by 2020, where appropriate, and ensuring the necessary systems and processes are in place to support them. The multi-year NCEA Digital Trials and Pilots project is a key element of this work.

NZQA's intention has been to keep building our own understanding and work with the sector to understand what it will mean to be delivering digital assessments at scale and, potentially, anytime. In particular, NZQA is seeking to understand some of the critical success factors that will affect school uptake of digital assessment and successful outcomes for learners so they "Qualify for the Future World: Kia Noho Takatū Ki Tō Āmua Ao".

Each year we are increasing our knowledge base to facilitate the transition. Compared with the position after the 2015 Trials and Pilots, NZQA and the sector now have a richer understanding of how to successfully implement digital examinations. Evaluating and reporting on each year's Trials and Pilots enables NZQA to show how the lessons from the previous year have been factored in. For example, in 2016 we provided student access to familiarisation activities for the assessment software as a direct result of 2015 feedback.

The 2016 report also helps schools understand how NZQA worked with them to meet the challenges presented during the year. This enables schools to confidently make decisions about their involvement in the 2017 Trials and Pilots.

The 2016 findings are also directly influencing the 2017 Trials and Pilots. Examples include implementing new ways to understand in more detail the student and school by school experience of digital examinations. This confirms the value in supporting schools to participate in Trials and Pilots according to their individual circumstances. The work also confirms the value of undertaking psychometric analysis of results and the next steps to be taken to enhance that evaluation framework.

As part of the 2016 Digital Trials and Pilots project, feedback was sought from the students, teachers, markers and examination centre managers who participated in those Trials and Pilots. Their feedback, captured in this report, has been valuable in providing insight into their experience, and is already informing the next steps NZQA takes with digital assessment in collaboration with schools and the wider education sector.

The 2016 NCEA Level 1 Digital Trials and Pilots involved 100 unique schools and 4386¹ unique students. A total of 3039² unique students and 72 unique schools took part in the Trials, and a total of 1563 unique students and 46 unique schools took part in the Pilots.

The ethnicity data show that students undertaking Trials were New Zealand European (59.4%), Asian (16.6%), New Zealand Māori (13%) and Pasifika (6.3%). The ethnicity data show that students undertaking Pilots were New Zealand European (63.9%), Asian (14.7%), New Zealand Māori (14.8%) and Pasifika (5.2%).

¹ Because students doing the Samoan and Spanish Trials were using a different system we are unable to identify uniqueness in relation to student participants in the other Trials and Pilots.

² As per the footnote above, because students doing the Samoan and Spanish Trials were using a different system we are unable to identify uniqueness in relation to student participants in the other Trials.

Ten Trial subjects (where the results did not count towards NCEA) and three Pilot subjects (where the marks did count towards NCEA) were offered. The Trial subjects were Business Studies, Economics, Geography, History, Health, Physics, Samoan, Science, Spanish and Te Reo Rangatira. The Pilot subjects were English, Classical Studies and Media Studies.

A two-fold approach was taken for the user experience evaluation approach. First, NZQA surveyed participants – learners, teachers, Principals' Nominees, markers and Examination Centre Managers/supervisors. Second, NZQA engaged external experts to psychometrically analyse the student results from the Level 1 English Pilot.

The survey response rates were low, in some cases too low to draw valid conclusions. However, survey data has provided sufficient insight, corroborated by other interactions with the surveyed participants, for NZQA to enhance its approach in 2017 and the future.

The psychometric and statistical analysis was undertaken only for English in the 35 participating schools. The format of the digital and paper-based examinations was identical in all respects apart from the medium. The performance of the students who took the digital English examination was compared to the performance of the students who took the paper-based examination in the 35 participating schools.

There was no evidence of disadvantage to those students who undertook the examination in a digital medium. There were some differences in the results distributions for the group of students who took the digital English examination compared to the group who took the paper-based English examination, in favour of the digital group of students. This is most likely due to more able students opting for the digital examination.

The relationship between performance in internal and external assessment for English was not affected by whether the assessment was in digital or paper format.

Some students undertaking the English Pilot experienced issues with the authenticity software implemented to prevent cheating. This was because the software was more sensitive than expected to the multiplicity of devices being used and the background processing on these devices, resulting in some students being locked out. A significant proportion of these effects were experienced between the time the student logged on and when they were issued the code to commence the examination, rather than during the examination itself. Approximately one third of students did not receive a warning and were not locked out. Well over half the students received a warning but were not locked out. NZQA made limited changes to the software to improve the experience of students undertaking the subsequent Classical Studies and Media Studies Pilots. NZQA also visited the remaining Pilot schools to provide support in advance of those Pilots for examination set-up and management.

The *Trial – Pilot – Review, co-design with the sector, opt in model* for digital assessment transformation, and the ongoing work to understand the variety of user experiences, will continue to be a major influence on the next steps.

2016 Digital Trials and Pilots

Purpose of this document

This document outlines the major user experience evaluation findings arising from surveys of participants in the 2016 NCEA Level 1 Digital Trials and Pilots. The participants are learners, teachers, Principals' Nominees, markers and Examination Centre Managers. The document also outlines the findings from the psychometric evaluation of the results from the Level 1 English Pilot. Psychometric analysis for other Pilot subjects is not included because of the limited sample size.

Description

The 2016 Trials & Pilots project was built upon the learning from the 2014 and 2015³ Digital External Assessment Prototypes Projects. It aimed to test the end-to-end examination cycle, to introduce capabilities into Trial assessments that were available only in a digital mode, and to offer all externally examined standards for the Trialed or Piloted subjects, rather than individual standards.

NZQA's intention has been to keep building our own understanding and work with the sector to understand what it will mean to be delivering digital assessments at scale and, potentially, anytime. NZQA's goal is to have all NCEA examinations available online, where appropriate, by 2020. In working towards achieving that goal, NZQA is seeking to understand some of the critical success factors that will affect school uptake of digital assessment and successful outcomes for learners.

The development and robustness needed for external assessments requires that testing must go through different levels to ensure the digital assessments offered meet strict criteria before progressing to the next stage. These testing stages, as they were conceived of in 2015 and early 2016, are outlined below:

Testing stage	Description
Trial	Designed for schools to use as practice assessments, and will be marked by teachers. Student results will not count towards NCEA, but they can be used as evidence towards a derived grade if needed.
Pilot	Designed to be used in place of the paper-based assessments. All Pilots will have sample assessments developed to provide learners and teachers with an idea as to the "look and feel" of a digital assessment. Schools/students will opt into the Pilot and some subjects may be used with a selected number of schools ⁴ . Assessments will be marked by NZQA markers and results will count towards NCEA. The rigour of paper-based assessments will be consistent in these digital examinations.
Actual⁵	Designed to be used in place of assessments for <u>all</u> candidates, and the results will count towards NCEA. Assessments will be marked by NZQA markers. The rigour of paper-based assessments will be consistent in these digital examinations.

³ For the evaluation of the 2015 project see <http://www.nzqa.govt.nz/assets/About-us/Future-State/DEAP-Summary-Report.pdf>

⁴ At this stage there has been no focus on having some subjects being undertaken by some schools.

⁵ So far, this definition has not been actively used, and may be unnecessary, depending on the transition approach

Assessments Trialled and Piloted

In 2016 Level 1 Trials and Pilots were run with subsets of the respective cohorts. The Pilot subjects were selected for two reasons:

- English, Media Studies and Classical Studies could make a simple transition from paper to digital
- French was successfully trialled as part of the 2015 Digital External Assessment Project.

The trial subjects were selected to try and provide a range of assessment types – short answer, paragraph, some equations (Physics and Science) languages, the use of stimulus material, and two subjects (Geography and Economics) that utilise other forms of presenting student evidence that is not text based. The objective was to see how these assessment types work in a digital context and what the issues may be.

The initial scope of the NCEA subjects applied to the 2016 Trials and Pilots is outlined below.

Testing stage	Subject		
Trial	<ul style="list-style-type: none"> • Business Studies • Economics • Geography • Health 	<ul style="list-style-type: none"> • Music • Physics • Science • History 	<ul style="list-style-type: none"> • Samoan • Spanish • Te Reo Rangatira
Pilot	<ul style="list-style-type: none"> • Classical Studies • English 	<ul style="list-style-type: none"> • French • Media Studies 	
Actual	<ul style="list-style-type: none"> • None in 2016 		

The approach to the Trials and Pilots included that:

- all Trials and Pilots would be at Level 1
- the Trial subjects were to provide NZQA with an opportunity to introduce some innovative digital assessment methods (“Augmentation”⁶, such as “drag and drop”, “dynamic graphing”) without disadvantaging students
- the Pilot subjects were to be digitised versions of the paper-based examination (“Substitution”)
- all assessments were to meet the requirements of the standard, and therefore be usable (e.g. for derived grades in the case of the Trials) and comparable to the paper-based examination, in the case of the Pilots
- psychometric analysis of the learner results from the digital Pilots was to be undertaken to understand their validity and level of comparability with the learner results from the paper-based examinations
- the Trials were to be available for two two-week periods in late Term 3 and early Term 4, whereas the Pilots were to be undertaken at exactly the same time as the paper-based examinations
- familiarisation exercises (for Trials) and sample assessments (for Pilots) were to be made available for learners and teachers to experience key elements of the digital experience prior to undertaking Trials or Pilots
- technical readiness assessment tools and minimum recommended specifications for digital devices were also to be made available for schools to better understand the technical environment that would best support their learners

⁶ As per the SAMR model by Dr Reuben Puentedura, e.g. <http://hippasus.com/rrpweblog/archives/2014/01/15/SAMRABriefContextualizedIntroduction.pdf>

- deeper insight and understanding from learners and teachers about the proposed digital assessments was to be sought through closer relationships with four kura and six focus group schools
- students would have online access to their marked scripts.

All elements of the approach above were put in place as intended.

A programme of relationship building and co-design with students and key teachers from six English-medium focus group schools and four Māori medium kura provided NZQA with early insights into the student experience that was likely to be required of the digital Trials and Pilots. Overall, students were excited in advance of the digital assessments. NZQA guaranteeing the authenticity of the examination (the right student undertaking the right examination under examination conditions, i.e. no cheating) was among the first few issues raised consistently by students, along with needing assurance about network and device reliability so that the digital experience did not disadvantage them. The students also provided insights into various usability features, though diverse views were expressed about the same design elements.

Scope reductions and challenges during 2016

Music Trial

After some consideration of the match between the assessment software solutions chosen for the Trials and Pilots, the externally examined Music standards were taken out of scope for the Trials. This was because the types of assessment that would have been useful to include in the Trial were not possible to achieve in the time available to configure the software. The decision not to include a Music Trial was taken in March 2016.

French Pilot

In October 2016, after the Spanish Trial was observed in one school, and after other Pilot preparation testing was carried out in other schools, NZQA decided to withdraw the French Pilot. This was because we could not guarantee a consistent user experience with the wide range of browsers and devices expected to be in use by learners. The issues were specific to language examinations and options for managing them for each student device were logistically untenable at the time and potentially unreliable.

Tahi

The intention was to use Network for Learning's *Tahi* identity and access management solution to demonstrate that a digital assessment solution could interface successfully with a sector identity and access management solution. This intention was met, to the extent that it was used by one school successfully for the Media Studies Pilot. It was feasible to use it for the Pilot examination because the Pilot was turned on once for a known cohort of students and then turned off. In contrast, to give schools flexibility to use the Trial examinations at any time within the two two-week windows in September / October, it was not easy to implement *Tahi* in this context. This was not a reflection of the *Tahi* tool, but of the way in which NZQA chose to set up the Trials administration to better support schools' own management of the Trials.

English Pilot

The English Pilot worked well for approximately half the candidates, but the authenticity (anti-breach) software, which was designed to ensure students stayed in the examination, was more sensitive than expected due to the multiplicity of devices and background processing on these devices. The authenticity software had the ability to detect if students moved away from the examination and alerted them to return with a pop up warning. If students did not respond to the pop up warning within a specific timeframe they were locked out and needed a supervisor to unlock them. The sensitivity of the software was such that students may not have explicitly lost focus on the examination before the software was activated. The figures below show the level of instigations of the authenticity software and the associated lockouts (where

they occurred) for the English Pilot. Figures 1 and 2 show the instigation of the authenticity software during the whole examination period (from the time the students entered the assessment software) and the timed period of the actual examination respectively. If a student noticed the warning received they could immediately return to the examination software without getting locked out.

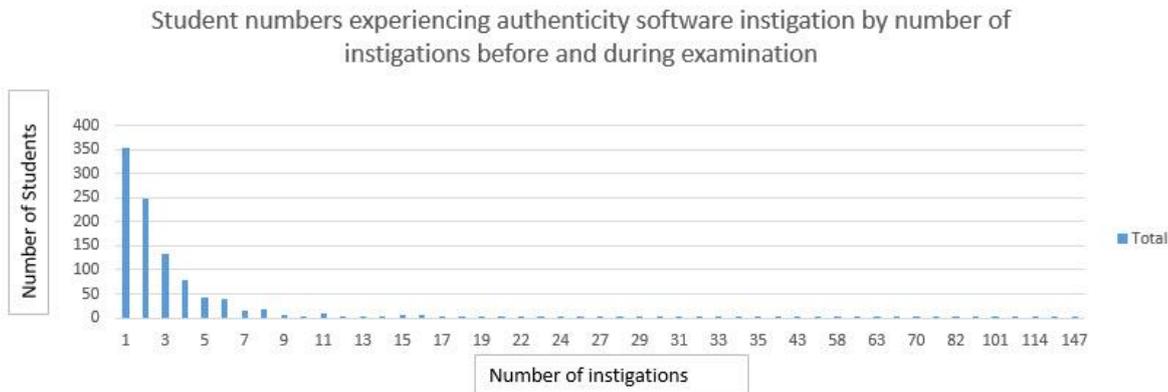


Figure 1: Instigations of the authenticity software *before and during* the examination

Figure 1 shows that approximately 350 students had one instigation of the authenticity software; approximately 240 had two instigations; just under 150 students had three instigations, and so on, from the time they were set up to commence the digital examination. The numbers of students are not to be added together; rather, a student who experienced three instigations also experienced one and two.

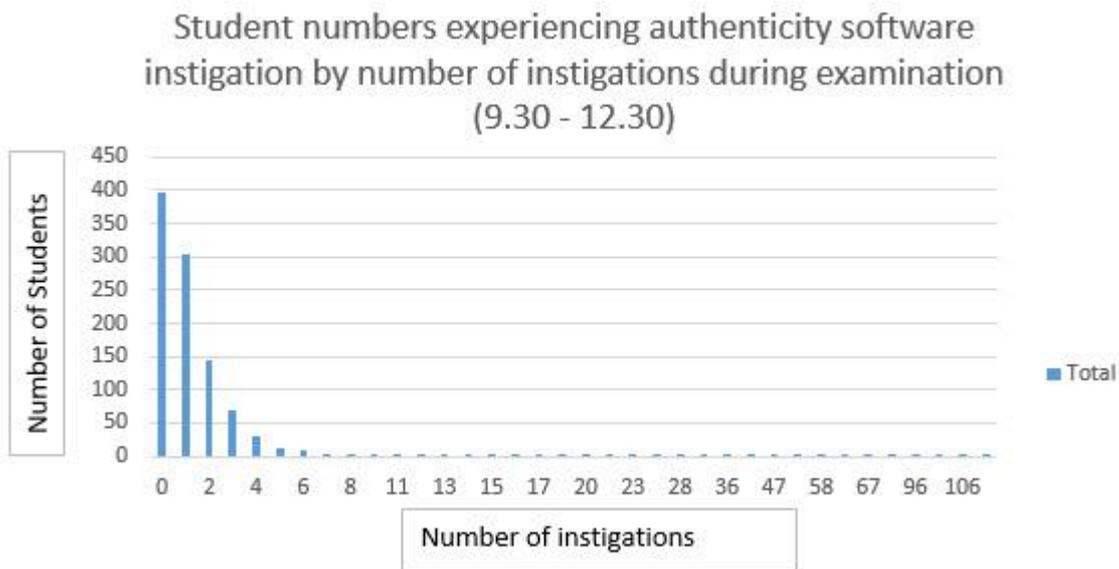


Figure 2: Instigations of the authenticity software only *during* the examination

Figure 2 shows that 400 students experienced no instigations of the authenticity software during the examination; approximately 300 experienced one instigation, and so on. Those students who were set up and ready to enter the examination code were more likely to experience the authenticity software early.

Around half the students experienced one or more lock outs. A lock out occurred if the student who had received a warning from the authenticity software did not confirm within 5 seconds that they wished to return to the examination software. A software error prevented Examination Centre Managers and supervisors from being able to immediately unlock the students and the (planned for) NZQA telephone support was activated to assist with unlocking the students' examinations.

Figures 3 and 4 below show how the lockouts were experienced before and during the examination respectively.

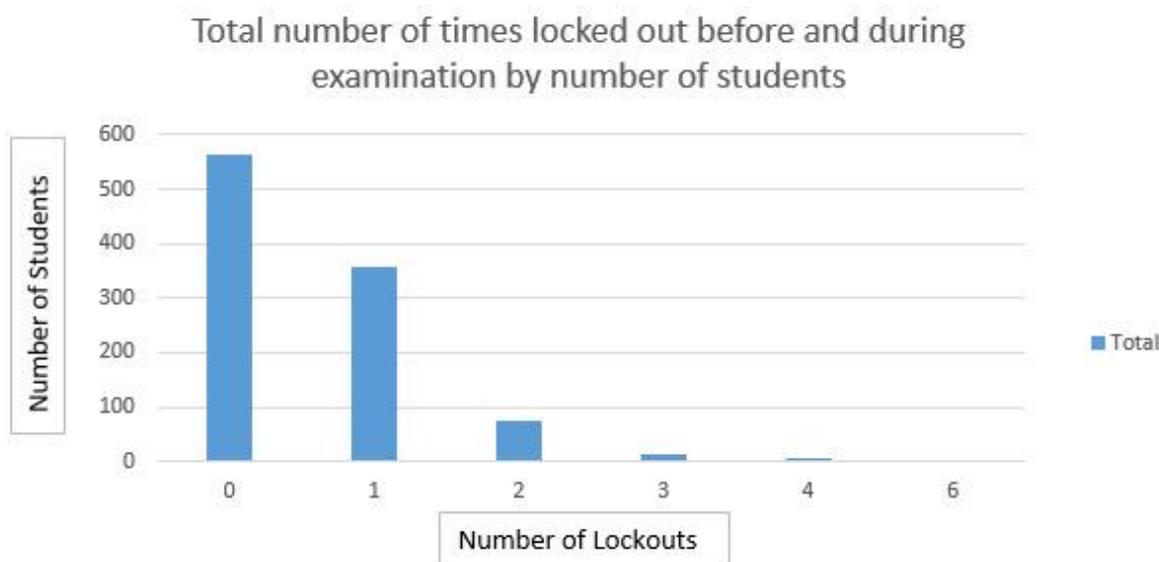


Figure 3: Lockouts by the authenticity software *before and during* the examination

Figure 3 shows that just over 550 students were not locked out at all, either before or during the examination. It also shows that approximately 350 were locked out once, approximately 75 twice, and so on.

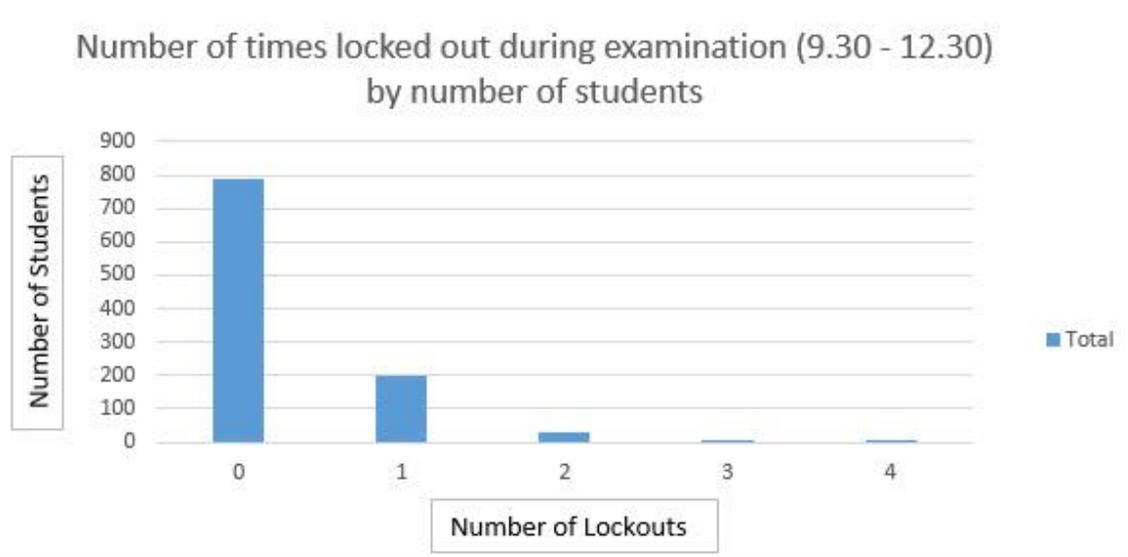


Figure 4: Lockouts by the authenticity software only *during* the examination

Figure 4 shows that almost 800 students were not locked out at all during the examination. It also shows that approximately 200 were locked out once, and around 30 students twice, and so on.

Of the approximately 1400 students who commenced the examination digitally, almost 400 reverted to the paper version which had been planned for as a contingency. Examination Centre Managers applied their usual procedures for managing disruptions to the students' examination.

After the examination, NZQA appreciated the opportunity to discuss the issue of the English Pilot with Pilot school principals or Principals' Nominees. Before the Media Studies and Classical Studies Pilots were held NZQA staff visited the relevant schools and, with updated software for unlocking the students and a walk-through of the logistics, their experience was significantly better.

Although the number of instigations of the authenticity software was in excess of NZQA's expectations, our analysis of students' online activity before or during the digital examination does not indicate any actual breaches of examination conditions by students.

Special Assessment Conditions

When Expressions of Interest to participate in the 2016 Trials and Pilots were sought there was a high degree of interest from schools with Special Assessment Conditions students. This reduced significantly when we clarified that a spellcheck facility would not be made available, as this had the effect of reducing the conditions that some Special Assessment Conditions students already had when using computers for examinations. Spellcheck will be offered for some Trials and all Pilots in 2017. Text to speech is under consideration and we are considering trialling strategies for visually impaired and/or blind students.

Key demographics

Number of unique schools in Trials and Pilots: **100**

Number of unique students in Trials and Pilots: **4386** ⁷

Participation in Trials

Candidate entries: **4086**

Number of unique students: **3039** ⁸

School entries: **137** schools

Number of unique schools: **72**

Table 1: Participation in Trials by Subject

	Number of Candidates	Number of Schools
Trials		
Economics	576	22
Business Studies	203	6
History	550	21
Health	207	7
Geography	687	20
Science	1373	33
Physics	71	4
Te Reo Rangatira	43	5
Spanish	309	16
Samoan	67	3
Total	4086	137

⁷ Because students doing the Samoan and Spanish Trials were using a different system we are unable to identify uniqueness in relation to student participants in the other Trials and Pilots.

⁸ As per the footnote above, because students doing the Samoan and Spanish Trials were using a different system we are unable to identify uniqueness in relation to student participants in the other Trials.

Table 2: Participation in **Trials** by **Decile**

Deciles	School	School	School	School	School	School	School	School	School	School
	Economics	Business Studies	History	Health	Geography	Science	Physics	Te Reo Rangatira	Spanish	Samoan
Low Decile 1-3	1		4	2	1	6		5	2	2
Mid Decile 4-7	8		10	1	7	15	1		2	1
High Decile 8-10,99	13	6	7	4	12	12	3		12	

Table 3: Participation in **Trials** by **Ethnicity** (unique students, ethnicity data from National Student Index)

Ethnicity	Students	
NZ Maori	395	13.0%
Asian	503	16.6%
NZ European	1805	59.4%
Pasifika	191	6.3%
Other/Unspecified Ethnicity	107	3.5%
NOT FOUND	37	1%
Total	3038	100%

Participation in Pilots

Candidate entries: **1698**

Number of unique students: **1563**

School entries: **53**

Number of unique schools: **46**

Table 4: Participation in **Pilots** by **Subject**

	Number of Candidates	Number of Schools
Pilots		
English	1393	35
Media Studies	234	13
Classical Studies	71	5
Total	1698	53

Table 5: Participation in **Pilots** by **Decile**

Deciles	School	School	School
	English	Media Studies	Classical Studies
Low Decile 1-3	6	2	1
Mid Decile 4-7	14	3	
High Decile 8-10,99	15	8	4

Table 6: Participation by in **Pilots** by **Ethnicity** (unique students, ethnicity data from National Student Index)

NZ Maori	232	14.8%
Pasifika	82	5.2%
Asian	229	14.7%
NZ European	999	63.9%
Other/Unspecified Ethnicity	21	1.3%
	1563	100.0%

Summary of Findings

The separately published appendices accompanying this report provide the detail of the analysis that has informed the Summary of Findings. Each appendix has its own summary level information, as well a wider variety of verbatim comments, survey questions and the data tables. For that reason, this summary focuses only on the key points from the surveys. The discussion on achievement of the project investment objectives outlines where more learning is required, where changes have been made for 2017 or planned for future consideration.

Student experience

English Pilot

The student survey was designed to measure students' self-reported satisfaction and experience of the digital examinations, including establishing whether the students regularly use electronic devices at home and at school.

As noted above, there were technical issues experienced during the English Level 1 digital examination that were resolved by the time the Media and Classical Studies Level 1 digital examinations took place. Students who participated in the English Level 1 Pilot examination were emailed the survey to complete as it was not made available to them at the end of the examination. Only 34 out of a possible 1393 responses (0.29%) were received. This not considered a representative number of respondents to draw any conclusions of students' experience of the English digital examination. Nonetheless, the results for the closed ended questions have been compiled and are presented here.

Note that not all 34 respondents answered all the questions in the survey.

Overall satisfaction

24 of 27 respondents agreed or strongly agreed with the statement 'Overall I found doing this exam digitally was a positive experience'.

Exam preparation

18 of 32 respondents found the sample assessments very useful in preparing for the digital exam. 10 respondents reported not being aware of the existence of the sample assessments.

23 of 26 respondents disagreed or strongly disagreed with the statement 'I spent more time preparing for this digital exam than I would have if I had been doing it on paper'.

On the day

30 of 31 respondents reported they used their own device to complete their examination, and one respondent reported that they used a school provided device. 29 of 31 respondents reported they used a laptop. One respondent used a desktop and one respondent used a tablet.

21 of 34 respondents experienced network or device issues accessing or completing the digital examination.

27 of 29 respondents reported that they found it easy or very easy to navigate through the digital examination.

24 of 29 respondents reported that they found it easy or very easy to enter their responses to the examination questions.

Digital technology at home

29 of the 30 respondents reported having more than one device at home. 20 of the 30 respondents reported having three different types of devices available at home.

Digital technology at school

24 of 30 respondents reported that digital technology is used in class to support their learning very often or quite often. No respondents reported that digital technology was never used in class.

28 of 30 respondents reported that digital technology is used for homework to support their learning very often or quite often. No respondents reported that digital technology was never used for homework.

20 of 30 respondents reported that digital technology is used in internal assessments to support their learning very often or quite often, nine respondents reported it was used occasionally, and one respondent reported it was never used.

Media Studies and Classical Studies Pilots

The student survey was designed to measure students' self-reported satisfaction and experience of the digital examinations, including establishing whether the students regularly use electronic devices at home and at school.

Of the 305 students who participated in the Media Studies and Classical Studies Level 1 Digital Pilot examinations, 100 (33%) completed the online survey. Not all the students answered all the questions in the survey. Of the 100 responses received, 83 out of a total of 234 Media Studies students participated in the survey, with 17 out of a total of 71 Classical Studies student participants.

The survey was made available to students within the assessment solution, directly after they submitted their exam. The survey consisted of 13 questions. Some of them were open ended questions. The survey was designed to take approximately five minutes to complete.

The students who participated in the survey were generally satisfied with the digital examination, with most respondents agreeing or strongly agreeing that they found undertaking the examination digitally to be a positive experience.

“Overall I really enjoyed being able to type out my exam, it was a lot faster and easier than doing it all by hand in a booklet”.

“This is a good way to do assessments like this, media studies. That’s just because there is a lot of writing involved so it just makes things quick and easy.”

“It was very good, but it was sometimes confusing and to log on we started later than the time we should have because of trying to find our way into it.”

Most responses to the question ‘What did you like most about completing the exam digitally’ included that it was easier and faster to type than to handwrite, and that essay answers could be easily reviewed and edited and looked much neater than handwritten answers. Some students also mentioned that they were more confident using a keyboard, that they liked the word count and timer, and that the instructions were easy to understand.

“Not having to write with a pen because it hurts my hand after 3 hours. That I could delete writing without it taking up space on the page. That it was new and a trial, which made me feel excited.”

“The digital examination was easy to navigate through and it was relatively faster to complete than by using pen and paper. The user interface was relatively friendly and easy to use.”

*“No need to hand write
Simple instructions
Easier process of exam”*

The most common response to the question ‘What did you dislike most about completing the exam digitally?’ was that the sound of other students typing was distracting and noisy. Other dislikes included:

- using the school rather than their own keyboard
- time to set up and log on
- eye strain
- no spelling check
- that the font and text boxes were too small and required constant scrolling
- worrying about doing something wrong, and
- warning boxes appearing too quickly.

Digital Trials

The student survey was designed to measure students’ self-reported satisfaction and experience of the digital Trial examinations, administered by schools during September and October 2016. The survey was also used to trial the questions in anticipation of the surveys to be included with the Pilots.

The survey included establishing whether the students regularly use electronic devices at home and at school. 10 Level 1 digital Trial examinations were offered (Economics, Business Studies, History, Health, Geography, Science, Physics, Te Reo Rangatira, Spanish, and Samoan). 4,086 students participated in the digital Trial examinations.

Of the 4,086 students who participated in the digital Trial examinations, 663 (16%) completed one of two almost identical online surveys. Not all the students answered all the questions in the surveys.

The responses to identical questions from each survey have been combined in this report. The survey was made available to students within the assessment system, directly after they submitted their trial exam. The survey consisted of 12 questions. Some of the questions were open ended questions. The survey was designed to take approximately five minutes to complete.

There were some technical issues with some of the digital Trial examinations, which impacted negatively on the experience of some students sitting those examinations. Comments made by respondents to the survey reflect that.

Limitations

Only 663 of the possible 4,086 responses were received. This is not considered a representative number of respondents to draw any conclusions of students’ experience of the Level 1 digital Trial examinations. Nonetheless, the results and comments have been compiled and are presented here.

Due to some technical issues experienced during the trials, responses to two of the survey questions were not captured. These were:

Question 8: How easy was it to navigate through the digital assessment?

Question 9: How easy did you find entering your responses to the digital assessment?

Responses to a further two of the survey questions were captured for only 267 of the 663 respondents. These were:

Question 10: Do you agree or disagree with the following statement "I needed more preparation time for this digital assessment than for a paper-based assessment"?

Question 11: Do you agree or disagree with the following statement "Overall, I found doing this assessment digitally was a positive experience"?

Responses to a further two of the survey questions were captured for only 397 of the 663 respondents. These were:

Question 4a: How is digital technology used to support your learning at school? (In Class)?

Questions 4b: How is digital technology used to support your learning at school? (Homework)?

Overall satisfaction

Respondents were divided about whether their experience of undertaking the digital Trial examination was positive, with more than half agreeing or strongly agreeing it was a positive experience, and approximately a third disagreeing or strongly disagreeing that it was a positive experience.

The aspect of the digital Trial examination that respondents said they liked most was the ability to type rather than handwrite.

"I liked the digital assessment because I think I have the ability to type faster than I can write. It's probably a bad thing in most exam cases but I really do like the fact that I can type in especially my History Exam. This means I have more time to think and if I forget to write something in a paragraph I can just as easily go back and fix it. I had no distractions because I turned off my notifications and this helped me as I feel this was a good way to do my exam. I think the only difficulty I had was the fact that I didn't have auto correct or spell check which would tell me when I had a word spelt wrong but overall I could see when I made a mistake."

"I liked not having to write things out by hand as my writing tends to get messy and i become flustered. My hand also becomes sore from writing things by hand. I loved doing it digitally."

"I liked how it was easy to change our answers if we did answer them wrong at first, it was a lot easier than the paper version in this aspect."

Common dislikes expressed by students included navigation issues, the ability to cheat, issues with the graphing tool, the font size being too small, issues with typing and the noise of keyboards, and issues with assessing if they had written enough.

"It was hard to navigate the Punnett squares and spelling errors were more likely due to typing speeds and the small size of the font."

"I did not like how time consuming some parts of the digital assessment were. Logging in was time consuming, and lots of people had technical difficulties throughout the exam that wasted time from their 3 hours. I liked how the exams were set out in a simple way to read and look through."

"I felt it was hard to focus when typing and I was not able to organise my ideas as easily compared to the paper assessment. I also found it harder to read over work afterwards, and while typing I felt my thoughts weren't as processed compared to when writing in paper examinations. Also, I believe that the exams are not secure enough cheating wise, and this insecurity could possibly be exploited"

if these exams replace paper examinations. On a positive note, typing up assessments is sometimes faster than writing them down.”

“I didn't like that we haven't been taught how to use computers and laptops to type properly. It is also a hassle for exams where you need to draw diagrams to explain something (where you couldn't for this exam).”

“... because I am used to judging if I have written enough on paper, it is confusing digitally because typed words take up far less space than written, and the box does not indicate how much you are supposed to write. This could mean people spend way too long on questions that are not supposed to use all the time.”

“The tapping of the keyboards is so loud and distracting! “

Teacher experience (Managing Trials)

The teacher survey was designed to measure teachers' self-reported experience of managing digital Trial examinations in their schools.

Of the 322 teachers who managed digital Trial examinations, 63 (20%) completed the online survey. Just under half of the respondents were from either Auckland or Wellington schools.

The respondents did not all answer all the questions in the survey.

The survey consisted of 24 questions. Some of them were open ended questions.

Limitations

The response rate was 20%, which is low. Those who chose not to take the survey may have greater or lower levels of satisfaction than those who responded.

Below is a summary of the main findings. Appendix 1 contains a count of responses for each closed ended question. Appendix 2 is a full list of the survey questions.

Overall assessment of participating in the digital Trials

30 of 48 respondents agreed or strongly agreed that the Trials helped them identify gaps in their students' knowledge.

31 of 47 respondents agreed or strongly agreed that the Trials are a good preparation for NCEA examinations.

31 of 56 respondents agreed or strongly agreed that the Trials should be extended to Pilots.

28 of 56 respondents agreed or strongly agreed that they would encourage their students to sit the digital Trial or digital Pilot examinations in 2017.

8 of 55 respondents agreed or strongly agreed that the 2016 digital examinations encouraged them to use more digital tools in their teaching.

Marker experience (Pilots)

The marker survey was designed to measure markers' self-reported experience of marking a digital examination, including their experience of the marking tool.

Of the 52 markers who marked digital examinations, 21 (40%) completed the online survey. The respondents did not all answer all the questions in the survey.

The markers who answered the survey had an average of 19 (lower quartile 14, upper quartile 22, median 18) years' teaching experience and an average of nine (lower quartile six, upper quartile 10, median seven) years' marking for NZQA.

The survey consisted of 20 questions, seven of which were open ended questions.

Limitations

The response rate was 40%, which is low. Those who chose not to take the survey may have greater or lower levels of satisfaction than those who responded.

Overall, markers were positive about the digital marking process. 15 of the 21 respondents either agreed or strongly agreed the digital marking process was satisfactory.

Examination Centre Manager / supervisor experience (Pilots)

The Exam Centre Manager / supervisor survey was designed to measure Exam Centre Manager (ECM) and/or supervisors' self-reported experience of invigilating a digital Pilot examination, including their experience of the dashboard.

Of the 250 ECMs/supervisors who supervised the digital examinations, 120 (49%) completed the online survey. The respondents did not answer all the questions in the survey. For the purposes of the analysis of the responses, we have excluded 20 respondents who answered only one or two questions. This brings the response rate down to 40%.

Of the 120 survey participants, 41 were ECMs and 79 were supervisors. The response patterns for ECMs and supervisors do not differ in a statistically significant way. Therefore, the responses are presented together.

Of the 120 survey participants:

- 83 supervised only the English Level 1 digital examination
- 13 supervised the English Level 1 and either the Media Studies Level 1 digital examination or Classical Studies Level 1 digital examination, and
- 24 supervised either the Media Studies digital examination or Classical Studies digital examination or both.

There were some technical issues experienced during the English digital examination that were resolved by the time the Media and Classical Studies digital examinations took place. For those questions where the response patterns for ECMs/supervisors who supervised the English digital examination differed in a statistically significant way from those for ECMs/supervisors who supervised the Media and Classical Studies digital examinations, these are reported on separately.

80% of survey participants had been an ECM/supervisor for previous examination rounds, and 89% indicated they were involved in invigilating both digital and paper-based examinations.

The survey consisted of 23 questions, nine of which were open ended questions. The first three questions were for ECMs only.

Limitations

The response rate is 49%, and lowers to 40% once respondents who answered only one or two questions were excluded. Those who chose not to take the survey may have greater or lower levels of satisfaction than those that participated in the surveys.

The respondents did not all answer all the questions in the surveys. Some are easily explained, for example those for Questions two, three and four were the ECM only questions (so could only be answered by some of the respondents) and questions 16 and 17, which were about whether switching to paper during the examination made supervision harder or was disturbing to some candidates. Nonetheless, a sizeable proportion of respondents chose not to answer each of the close ended questions.

Overall experience compared to paper-based invigilation

Respondents rated their overall experience as worse if they supervised the English digital examination than if they supervised the Media Studies and Classical Studies digital examinations. This was expected given the technical difficulties experienced during the English digital examination.

Many of the respondents who supervised the English digital examination agreed or strongly agreed that digital examinations were more difficult to manage than paper-based examinations. Less than half of respondents who supervised the Media Studies and Classical Studies digital examinations agreed or strongly agreed that digital examinations were more difficult to manage than paper-based examinations.

Comments substantiating respondents' rating of why managing a digital examination was more difficult varied, some were more positive and some negative.

Positive comments

"I agree only in the sense that being a new environment and the need to deal with some of the problems that arose, there was some extra effort and time required. There was also the learning curve of dealing with the real live environment. However, once things had settled, I found very little difference. Not having to collect completed papers was a bonus."

"Because it was a first, it was more stressful and supervisors, [Principals' Nominees], and I were nervous. For me, there was more work (by choice) to ensure that every candidate/supervisor had a copy of the URL and candidate log in details by their computer. 'Firsts' are always more difficult."

Negative comments

"The technical issues that were unavoidable such as computers freezing, etc. added excess stress for students. I found that we had to constantly be hanging around the students more closely, which made them nervous and there was constant disruption to others during the exams when a student was having difficulty with their computer."

"Starting the exam is a tricky process as it's hard to enforce exam conditions while students still need to set up after entering the room. There are also more random errors that cannot be controlled."

Working in a digital environment did not appear to affect the respondents' confidence in managing possible breaches of the examination rules, and there was no statistically significant difference in the responses of those supervising the English digital examination and those supervising the Media Studies and Classical Studies digital examinations.

Most respondents agreed or strongly agreed that they felt confident managing possible breaches of the rules in the digital examination room. Reasons provided by respondents for their high level of confidence included good training, backup from NZQA staff, the number of supervisors, and the small number of participating

students. Many respondents commented that because there were many accidental lockouts it would have been hard to detect an actual breach.

Psychometric analysis (English Pilot)

The psychometric and statistical analysis was undertaken only for English in the 35 participating schools. The format of the digital and paper-based examinations was identical in all respects apart from the medium. The performance of the students who took the digital English examination was compared to the performance of the students who took the paper-based examination in the 35 participating schools.

There was no evidence of disadvantage to those students who undertook the examination in a digital medium. There were some differences in the results distributions for the group of students who took the digital English examination compared to the group who took the paper-based English examination, in favour of the digital group of students. This is most likely due to more able students opting for the digital examination.

The relationship between performance in internal and external assessment for English was not affected by whether the assessment was in digital or paper format.

Discussion

Project investment objectives

Based on internal review processes, feedback from stakeholders and the survey results, NZQA is confident that we made progress towards realising all investment objectives. The end to end process was substantially tested to the extent that the examinations were designed and developed online, had digital supervision, and were marked online. Processes to register digital assessment candidates and transfer their results to NZQA's qualifications management systems were a mix of electronic and manual processes.

While we learned a great deal from the 2016 Digital Trials and Pilots, we are mindful of the ongoing development process. For example, where we found an element of the Trials or Pilots worked well we will have to make a judgement about whether or how that is scalable; where we learned something did not work well, we will need to work on establishing credible alternatives.

Trial and evaluate critical elements

Key elements such as security and authenticity of the examination were tested. For the purposes of this discussion, security covers all stages of the examination development, delivery and results management process and authenticity covers the integrity of the student response (is the response their own work?). Several important lessons were reinforced or learned, including:

- The need for an ongoing security risk assessment and management of authenticity was reinforced. Security risks generally emerge slightly ahead of the known and affordable options to mitigate them and process mitigations (such as supervisors watching candidates or monitoring activity from a dashboard) are likely to have a role in the immediate future.
- Managing authenticity with software embedded in the examination is one solution which does not require adding new software to school or student devices, or pre-cleaning the devices. The 2016 experience suggests that the variables encountered with the multiplicity of devices used by students mean that alternatives will need to be explored.

- The profile of security and authenticity risks may present differently in the future if the nature of assessment evolves to take advantage of digital assessment opportunities and to reflect the changing needs of users of assessment information (e.g. employers, tertiary providers) who are looking for deeper evidence of skills and capabilities other than content knowledge. For example, partial or full access to the resources outside the digital examination may feature in the future.

How to deliver

The key learning, borne out by the psychometric analysis of the 2016 English Pilot, is that substitution examinations (digitising the paper examination), are likely to continue to deliver the robust digital assessments that remain fit for purpose and maintain the integrity and fairness of the external assessment process. The 2017 Trials and Pilots are expected to furnish NZQA with a deeper understanding of this through enhanced data collection and analysis, and greater sample sizes.

NZQA is acutely aware of the need to ensure ongoing confidence in the NCEA qualification as it steps with the sector in a planned, staged and managed way into digital assessment. This will be particularly important as we respond to calls from some schools for greater exploitation of the digital assessment opportunity to deliver examinations that better reflect their digitally enabled teaching and learning. At the same time, while NZQA is focused on ensuring students undertaking digital assessment are not disadvantaged, in the future the focus may shift to ensuring that students *not* able to undertake digital assessment will not be disadvantaged. In summary, NZQA needs to ensure that the examinations, both digital and paper-based, are of comparable difficulty and offer no candidate a particular advantage.

Identify operating model changes

The engagement with schools during 2016, whether or not they were directly participating in the Trials and Pilots, has provided valuable insights into future operating model changes. Educators in a number of forums have seen the advantages of “anytime” assessment, or at least the opportunity for students to access digital external assessments more than once a year. This is a key component of implementing personalised learning pathways for all students and enables a greater focus on assessment for learning, with faster feedback loops becoming inherent in the external examination process.

NZQA will be looking to co-create our medium to longer term operating model with the sector to support a phased delivery of “anytime” assessment. We recognise the importance of maintaining confidence in the NCEA qualification and its credibility, ensuring that student learning is reliably reflected in the “anytime” assessment approach and that it is logistically feasible.

Conversely, as more schools take up digital assessment and the cohorts accessing the digital assessment solution at one time grow, this will place significant operational priority on “anytime” opportunities.

The 2016 Trials and Pilots have helped crystallise some of the dynamics of the current operating model that will be stretched and need to be adapted as digital assessment is taken up at scale.

Understand support needed

The delivery of a fair assessment experience naturally requires different support mechanisms. For example, some non-Bring Your Own Device (BYOD) schools provided back-up devices for students in the case of device failure (which did occur for a small number of students); some schools used the readiness assessment capabilities, familiarisation activities and sample assessments provided by NZQA to help prepare themselves and their students for the examinations; managing a digital examination requires more developed help desk support from NZQA than a paper-based examination, where the longstanding and well-understood expertise in managing examinations under diverse conditions rests with Examination Centre Managers. In-school IT support also needs to be factored in.

NZQA has learned that we need to back up written communications about support requirements with more opportunities for schools to walk through their own specific scenarios with NZQA providing feedback and prompts for further discussion about how best to provide a conducive digital examination environment. NZQA will need to develop sustainable ways of supporting schools as participation scales up.

Understand how learners interface

The majority of this learning has been derived from the student survey information and the psychometric evaluation. As noted above, the focus groups and relationship building with kura has shown that students will engage deeply, frankly and imaginatively in the task of informing the design of digital examinations. This has reinforced NZQA's commitment to a co-construction model where this is practicable.

Identify management and logistics benefits

The 2016 Trials and Pilots identified that there are likely to be significant logistics benefits in the future, such as the reduced movement of paper and the potential to add new business continuity options into the mix with multiple sittings of digital examinations.

The project also confirmed new challenges in staging digital examinations in the current school environments. These include access to sufficient devices that are suitable for undertaking digital examinations and that are familiar to the students (school or student provided). Access to appropriate spaces for students to be working on a device under examination conditions (e.g. with reduced opportunities for screen peeping, managing power supplies, managing temperature) is also important. Schools that are examination centres are responsible for providing the appropriate environment.

NZQA has confirmed with this project the ongoing need to continue to work closely with the Ministry of Education and Network for Learning to map out the implementation approach to digital assessment so that their work with schools' physical environments and technical infrastructure, including device availability, is fully informed by the projected rate of schools opting into digital assessment.

Understand how to transition and embed change

The key reinforced learning about transitioning and embedding the change is that the opt in, Trial, Pilot and review model is working well to lay the ground work for wide uptake of digital assessment in the future. Finding more sustainable ways to co-create the solution(s) for digital assessment with students, teachers, examiners, markers, examination centre managers and supervisors will be a starting point for understanding the transition and embedding the change. This is important to retain the underlying principle that assessment is driving the transition, rather than technology.

NZQA recognises that the change will happen largely in schools and at the pace at which schools are ready. The Trials and Pilots of digital assessment provide a safe way for schools to take advantage of the digital opportunity now and to better understand the opportunity in their own context for the future.

Reflection on NZQA approach to working with school leaders, teachers and students

Some elements of the approach taken in 2016 worked well in themselves, but need further development to be effective in supporting schools to make the transition to digital assessment. Schools and kura will want to position themselves to take advantage of the opportunity that it creates for their students without introducing a sustained increase in workload.

The 2016 work with focus group schools and relationship building with kura has evolved in 2017 to the concept of Co-Managed Trials, which we hope will provide deeper understanding of critical factors in schools and ways of responding to them. The expectation is that with the cross section of schools and kura

expected to participate, we can package the relevant aspects of their experience – the challenges, potential solutions, good ideas – and communicate that to others with similar characteristics.

We know that schools must process large amounts of information from NZQA, and recognise that in 2016 the timeliness, relevance, targeting and delivery medium of the information about the Trials and Pilots meant that schools were not always provided with what they needed, when they needed it. We have started to address this with:

- notification in late 2016 of the 2017 Trials and Pilots scope
- use of the Principal Nominee Seminars early in the year to provide more information and context
- streamlining of the digital assessment communications into the familiar, efficient channels used by NZQA for all other communications to schools
- working in depth with a number of schools and kura
- working on our internal communications to ensure School Relationship Managers are better-informed in advance of key milestones, or breaking issues.

Schools have been constructive and realistic partners during the 2016 Digital Trials and Pilots project. As part of our co-construction approach for 2017 and beyond we will continue to value that as it evolves.

In addition, NZQA has analysed the assessment software enhancements required for the 2017 Trials and Pilots. This was informed by user feedback. Not all enhancements can be made in 2017, but we have sought to strike a balance between what is of value to give students a further opportunity to experience a wide range of digital assessments in 2017, and what is best implemented in the future.

Although some high-level insights were gained into what could work well for Māori students, and Te Reo Rangatira was included as a Trial subject, we have only limited data analysis capability deriving from the 14.8% of the Pilot participants who identified as Māori. A similar situation applies for Pasifka students (5.2% of the Pilot participants) where the insights that can be gained are limited. This is something NZQA is seeking to address with schools this year, linking it to our goal of working with other sector agencies to lift by 50% the Level 3 achievement rate of Māori and Pasifka students in STEM⁹-related subjects by 2020.

⁹ Science, Technology, Engineering and Mathematics