Digital Statistics 2020 Innovation Trial Evaluation and Closure Report

NCEA Online Programme

NEW ZEALAND QUALIFICATIONS AUTHORITY
MANA TOHU MĀTAURANGA O AOTEAROA

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

Digital Statistics 2020 Innovation Trial Evaluation and Closure Report

NZQA would like to acknowledge the participating schools, teachers and students for their support of this innovation trial and contribution to its success.

While we have done our best to accurately represent the findings of this innovation trial and the views expressed by participants, NZQA accepts responsibility for any omissions or inaccuracies in this report.

Table of Contents

NEW ZEALAND QUALIFICATIONS AUTHORITY
MANA TOHU MĀTAURANGA O AOTEAROA

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

NEW ZEALAND QUALIFICATIONS AUTHORITY
MANA TOHU MĀTAURANGA O AOTEAROA

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

1. Innovation Trials and their place in the NCEA Online Programme

The work of innovation trials is informed by NZQA's strategic thinking as expressed in the Digital Assessment Vision[1], Te Kōkiritanga strategy and its overarching vision: 'Ākonga Māori qualify for the future world'.

In line with that vision, the NCEA Online Programme prioritises work that aims to optimise the examination experience for Māori and Pacific learners.

All data collected in innovation trials must be able to be disaggregated to enable NZQA to identify and learn about the needs or preferences of Māori, Pacific and all learners. To this end, innovation trials will collect data that enables identification of student ethnicity. Inclusion and publication of ethnicity data will be suppressed for privacy reasons where student numbers are below a level where this is appropriate.

2. Trial overview – Digital Statistics 2020 Innovation Trial

The evaluation of the Digital Mathematics Innovation Trial 2019 recommended that a trial be undertaken for an externally assessed statistics Achievement Standard.

This 2020 trial enabled NZQA to trial and evaluate options for the use of digital activities in assessing students against the externally assessed Achievement Standard (AS) 91586 Apply probability distributions in solving problems. This aligns with the NZQA's long term goal for digital assessment and dovetails with changes that are expected to occur as a result of the current Review of Achievement Standards (RAS).

The choice of the Level 3 AS91586 was driven by the time scale of the RAS, and the opportunities and challenges for modifying and redefining assessment activities for this Achievement Standard.

The trial helped NZQA to continue the conversation with the mathematics and statistics sector about digital mathematics and statistics examinations. This was achieved by including a greater number of schools in the development and trialling of the activities and by reporting on the trial at meetings of the regional mathematics associations.

The Digital Statistics 2020 Innovation Trial focussed on:

1. building on the lessons learned and recommendations from the Mathematics 2019 Innovation Trial – see Appendix One - including improving the literacy of the questions in the activities and making familiarity with the tools used a requirement of participating in the trial.
2. evaluating digital statistics tools to determine the most suitable tool(s) for the trial
3. using the selected digital statistics tools to develop prototype digital assessment activities that would provide an opportunity to digitally assess students for AS91586 Apply probability distributions in solving problems
4. devising and trialling activities which modify and redefine traditional questions and problems for this AS.

Following the investigation and evaluation of digital tools, four teachers and the National Kaiarahi for the Mathematics and Statistics Networks of Expertise attended a workshop in Wellington on 3

[1] https://www.nzqa.govt.nz/about-us/future-state/digital-assessment-vision/

March. This group were the developers of activities to be trialled. The purpose of the workshop was to initiate the development of digital assessments for trialling with students, incorporating lessons learned from the 2019 Digital Mathematics innovation trial. The digital tools that were selected are tools developed by Anna Fergusson and other employees in statistics at the University of Auckland (UoA). This was the only tool that enabled the reimagining of activities for this achievement standard. Even so it needed, and received, further development from UoA for the purposes of the innovation trial. In all, three activities were then developed by the teachers via remote collaboration. They were then trialled with students at 18 schools during term 3 2020.

Integration of selected tools with the NCEA Online platform (Assessment Master) was not considered as it was out of scope for the trial.

Participating students and teachers were asked to complete an online survey once the activities had been completed. NZQA's National Assessment Leader (Mathematics) also conducted focus interviews with two teachers.

3. Trial objectives

The objectives of the Digital Statistics 2020 Innovation Trial were:

1. To identify and select appropriate digital environment(s).
2. To develop prototype digital assessment activities for the assessment of AS91586 apply probability distributions in solving problems.
3. To understand how the use of the chosen digital environment(s) enables the modification and redefinition of assessment activities.
4. To utilise the opportunities within a digital environment to increase the engagement and resilience of NCEA students when taking part in external assessments. In the context of the trial resilience is defined as when a student perseveres with the activity rather than giving up if challenged.
5. To use lessons learned from the Digital Mathematics Innovation Trial 2019, applicable for this innovation trial, to improve the design of activities.
6. To investigate whether the chosen digital environment(s) would either hinder or provide support for equity of access for Māori and Pasifika students.

4. Summary of the trial of Digital Statistics 2020

4.1 Trial phases

The trial milestones are described in the table below. A number of milestones were delayed due to the COVID-19 level 4 lockdown in term 1.

Milestone	Original data for completion	Date completed
Workshop with development group	March 2020	March 2020
Identification of, and engagement with participating schools	February 2020	n/a

Milestone	Original data for completion	Date completed
Teacher survey about commonly used digital teaching tools for statistics	March 2020	March 2020
Request schools register interest in trial	Not in original proposal	End of term 2 (July 2020)
Development group collaborate remotely to build activities and develop marking schedules.	May 2020	August 2020
Webinar	Not in original proposal	July 2020
Critique by members of examination panel and subsequent revision of activities and marking schedules	May 2020	August 2020
Showcases	No dates set	September – December 2020
Trial the activities with teachers and students	July 2020	October 2020
Evaluate trial	August 2020	November 2020

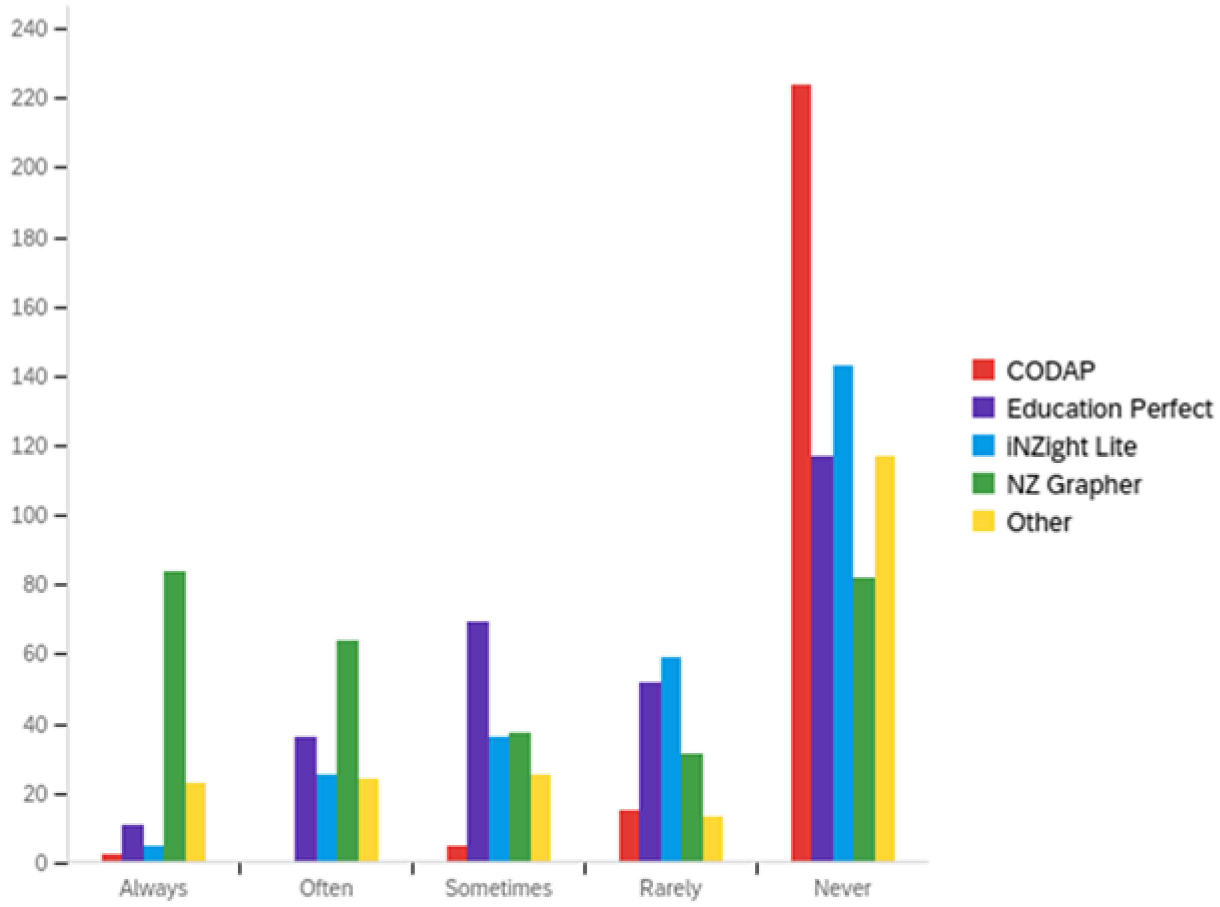
4.2 Evaluation of digital platforms and tools

Prior to the initial workshop in Wellington, a range of digital statistics tools were evaluated to determine the most suitable tool(s) for the trial. The tools were evaluated by a representative from the New Zealand Association of Maths Teachers (NZAMT) and the National Assessment Facilitator (Statistics) from NZQA. The evaluation included the following activities:

1. Surveying Heads of Departments (HODs) / teachers in charge (TIC) of Mathematics and Statistics in New Zealand Secondary Schools about the use of digital tools used in the teaching and learning of level 2 calculus and level 3 probability distributions The survey questions were designed, and the survey created, by NZQA and distributed to HODs and TIC via NZAMT. See Appendix One for a list of survey questions.
2. Evaluating a number of digital statistics tools against a set of evaluation criteria. See Appendix Two

The survey about the use of digital environments in the teaching and learning of statistics was a voluntary survey advertised via the email lists of the National Kaiarahi for the Mathematics and Statistics Networks of Expertise, who also advertised the survey on three Facebook pages: New Zealand Mathematics Teachers, New Zealand Statistics Teachers and New Zealand Calculus Teachers. The survey also asked for information about the use of digital environments in the teaching and learning of mathematics, for a sister project.

There were 358 survey responses, 199 from curriculum leaders and 159 from classroom teachers.



The responses to question 4 (Do you use any of the following digital statistics applications for teaching and learning at your school for Level 3 Achievement Standard 91586 Apply probability distributions in solving problems?) demonstrated a preference in the sector for NZ Grapher.

The evaluation criteria were used to evaluate the digital platforms and tools. The Education Perfect platform is heavily used but is not free to use for teachers or students. None of the tools that were asked about in the survey, provided the opportunity to reimagine problems for this achievement standard. Further, comments from survey respondents indicated that some schools were already using tools developed by the UoA.

For this reason, it was decided to use Microsoft Word as the platform for the activities and to use tools developed by lecturers in statistics at UoA. These tools were developed further to enable the investigation of probability distributions as models for real data. These tools are free to use for teachers and students. Support materials were developed by UoA so that teachers and students could become competent and familiar with using it.

Question 5 (Do you have any comments or feedback on digital statistics applications you have used for teaching and learning?) was an open-ended question. It is common practice to assess students in a digital environment for the internally assessed standards. These are often completed in Office 365 or Google Drive using browser-based applications to perform the statistical analysis and draw graphs.

These applications do not have the capabilities needed for the externally assessed achievement standards, but Office 365 and Google drive could be used. Two respondents had used the tool chosen for the trial and were supportive of the use of this tool.

For example [as received]:

- *I find NZGrapher easy to use and flexible. Students at our school have MAC laptops and are not able to download iNZight. I like the flexibility to change the labels on the graphs and the appearance of the graphs, such as, the size of the dots on a plot.*
- *NZ Grapher is very useful for stats.*
- *We need simple tools that are easily accessible and easy to use to enable a better way to teach probability and stats concepts.*
- *Tend to use NZ grapher for all statistics work so students are familiar with it but have not used for distributions*
- *Anna's tools are awesome, my faculty use them to teach concepts*
- *Are more concerned with teaching understanding of concepts than the use of statistical software. There is a cost associated with learning how to interact with each different platform. We almost exclusively use NZGrapher for the production of displays at Level 3. Many other online sources of information are used, particularly animations, for the purpose of learning concepts.*
- *Anna Ferguson (University of Auckland) has created some powerful, visual and really solid modelling that I am a huge fan of.*
- *All have their strengths and weaknesses. For example, CODAP is stunning to look and operate and play with (with no fear!) but has limited functionality. Similar with NZGrapher. iNZight is very powerful but is awkward to load on macs.*
- *NZ Grapher is an exceptionally good tool!*
- *There are very few digital statistics applications to use for teaching and learning Probability. The CODAP application is clunky and difficult to use - and is the only one I know of.*
- *Digital statistics applications are widespread in the teaching and assessment of internal standards. For external standards, they are mainly used to enhance understanding of concepts. Exam preparation requires thorough understanding of concepts and proficient use of graphics calculator. [sic]*

4.3 School participation

The original intention was to select a small number of schools (approximately 10) for trialling, based on the recommendations of the NZQA School Relationship Managers (SRMs) and the National Kaiarahi for the Mathematics and Statistics Networks of Expertise. It was anticipated that the resulting selection would provide diversity of decile, region, rural vs urban location and of student ethnicity (to enable comparison of Māori, Pasifika and other students' data).

The onset of the COVID-19 Pandemic and the consequent shift to alert level 4 (a lockdown with schools closing) forced a change in direction for the proposed innovation trial. The activities in development were clearly going to be a valuable resource for online teaching and learning as well as a potential assessment activity if derived grades were needed. In the light of this it was decided to invite all secondary schools that wished to participate in the trials to do so.

NZQA used the following channels to invited teachers to participate in the trial.

- Email lists from national and regional Kaiarahi for the Mathematics and Statistics Networks of Expertise.
- Posts on the NZ Mathematics and NZ Statistics Teacher Facebook pages.
- Posts on the NZAMT website and those of the regional associations.

These resulted in 57 registrations from teachers, from 52 schools, to participate in the trials.

The planning for this project took place before COVID-19 had been heard of. The changing alert levels in response to the COVID-19 pandemic had a significant impact on the ability of registered schools to engage with the trials or prepare students for the activities.

Schools were asked to complete the activities and provide feedback via separate student and teacher surveys by the end of term 3. Teachers were provided with a link to the tools that could be used to familiarise themselves and their students with the tools. Teachers were advised that students could access and complete the activities at any time and activities did not all need to be done at once. Feedback was received via the student and teacher surveys from 18 schools. These schools gave an adequate representation of spread across the regions and rural vs urban location. There were however no decile 1-3 schools, and survey responses revealed there were an inadequate number of Māori or Pasifika students to draw meaningful conclusions [from survey feedback]. 199 students completed the survey about the trial activities. Of these 8 (4%) identified their ethnicity as Māori and 4 (2%) with Pacific Peoples. 10 (5%) students were year 12, 189 (95%) year 13.

4.4 Sector Engagement

All teachers who had expressed an interest in the activities and trials, as well as representatives from MOE, NZQA, NZAMT and the NZ Statistics Teacher Association were invited to attend a webinar which showcased the type of activity and offered support for teachers to prepare students for trialling. 82 expressions of interest were received, and 36 participants attended this webinar. The webinar was recorded and made available online for parties who could not attend the live webinar.

Whilst the trials were in progress further webinars and presentations were made available to the sector in order to create awareness of the trial. Audiences included:

- Southland Heads of Department
- University of Canterbury mathematics staff
- New Zealand Statistics Association Education Committee

Invitations have been received to present, post-trial, at:

- University of Auckland Statistics Teachers Day
- Victoria University Mathematics Day

4.5 Summary of feedback and survey responses

Survey responses were received from 199 students and 16 teachers (with 2 more teacher responses by email) who had participated in the trial. Sections 4.5.1 to 4.5.5 include data from these responses. Teacher survey responses are presented as a count of responses by question, student survey responses are presented as percentages of responses received by question.

4.5.1 Evidence of all levels of achievement

Three activities were collaboratively produced by developers (selected New Zealand statistics teachers).

The process used was a continuous cycle of creation, review and improvement. A key part of the review process was the input of subject experts, teachers, the National Assessment Leader, a representative from the New Zealand Association of Mathematics Teachers and a representative of the Education Committee of the New Zealand Statistical Association to ensure that the activities offered students the opportunity to demonstrate all three levels of achievement. A schedule was also developed for each activity.

Each activity was similar in design to one question of a traditional, paper, external examination. See Appendix Seven. Students could be awarded a similar number of u (Achieved), r (Merit) and t (Excellence) marks for answering sections of the activity. The activity itself and the assessment schedule were reviewed by two materials critiquers, including a member of the NZQA examinations panel for AS91586. This feedback was used to refine the activities and schedule to ensure that they met the national standard, before they were released to the sector. Note: Quality assurance procedures were not as robust as for external examination papers, and the sector was made aware of this.

Teacher survey responses

All teacher-developers and the critiquers agreed that each of the three activities offered students the opportunity to demonstrate all three levels of achievement.

4.5.2 Equitable access on BYO device

The varying alert levels for COVID-19 during the trial period forced a shift to allow schools to self-select to participate in the trial and then whether to complete the trial. One consequence of this was that there was only a very small number of students of Māori and Pacific ethnicity who participated in the trial (based on survey responses where students could optionally select their ethnicity or ethnicities). It was not possible to draw any meaningful conclusions about whether the digital format of the activities had enabled equitable access for Māori and Pacific students.

See [Answering the trial's evaluation questions] for discussion re the intention of this question and the trial objectives.

Student survey responses

The vast majority of students used a laptop to access the activities.

Device

Q19 - What sort of device did you use for these activities? Select all that apply...	Count
Chromebook	21
Desktop	6
iPad	8
Laptop	147
Other	4
Tablet	5

<table>
<tr>
<td>The vast majority of students used a Chrome browser</td>
<td>Browser
<table>
<tr><th>Q20 - What browsers did you use for these activities? Select all that apply.</th><th>Count</th></tr>
<tr><td>Chrome</td><td>153</td></tr>
<tr><td>Firefox</td><td>5</td></tr>
<tr><td>Microsoft Edge</td><td>10</td></tr>
<tr><td>Microsoft Internet Explorer</td><td>4</td></tr>
<tr><td>Other</td><td>2</td></tr>
<tr><td>Safari</td><td>19</td></tr>
</table>
</td>
</tr>
<tr>
<td>The vast majority of students completed the activities in school.</td>
<td>Where did you do the activities?
<table>
<tr><th>Q22 - Where were you when you did the activities? Select all that apply - you may...</th><th>Count</th></tr>
<tr><td>At home</td><td>6</td></tr>
<tr><td>At school</td><td>180</td></tr>
</table>
</td>
</tr>
</table>

17 of the students reported problems with devices or internet access. 10 of these were related to faulty laptops which had to be restarted or exchanged for a new device. The remaining seven students reported difficulties with the school network.

Two students who used an iPad reported that it was not easy to complete the activities on this device. They reported problems with switching between browser windows and cutting and pasting graphics.

A common thread from the students was the difficulty of typing Greek letters (common for parameters) and inserting images.

Teacher survey responses

One teacher reported that students using Apple computers found it harder to complete the activities.

Teacher comments supported the students' concerns about Greek letters and images.

4.5.3 Modification and redefinition of assessment questions

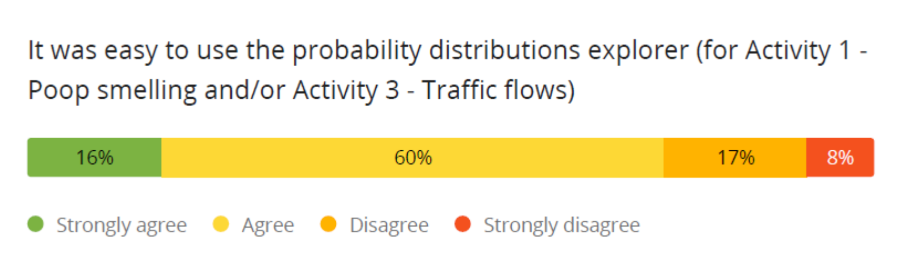
The activities that were developed retained the format of the traditional written assessments whilst utilising the benefits of technology to overcome the literacy demands in the introduction of the contexts and to pose investigative problems that are more in line with the teaching, learning and assessment that are expected by the New Zealand Curriculum.

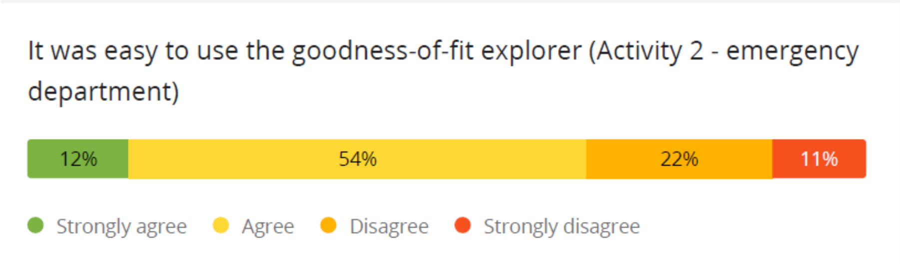
The activities were presented to students in a Word document. Students were able to write text and paste snips from browser windows to support their reasoning.

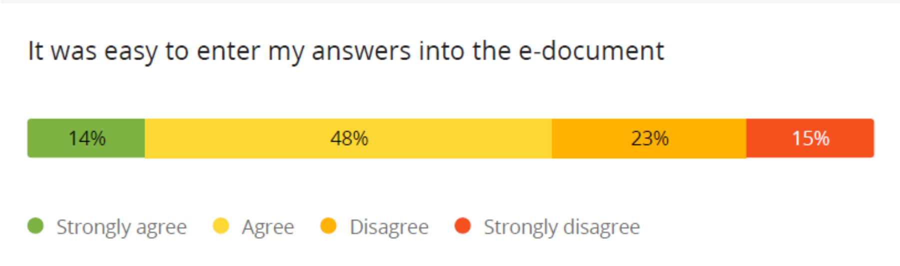
1. The context of the question was introduced in a variety of media.

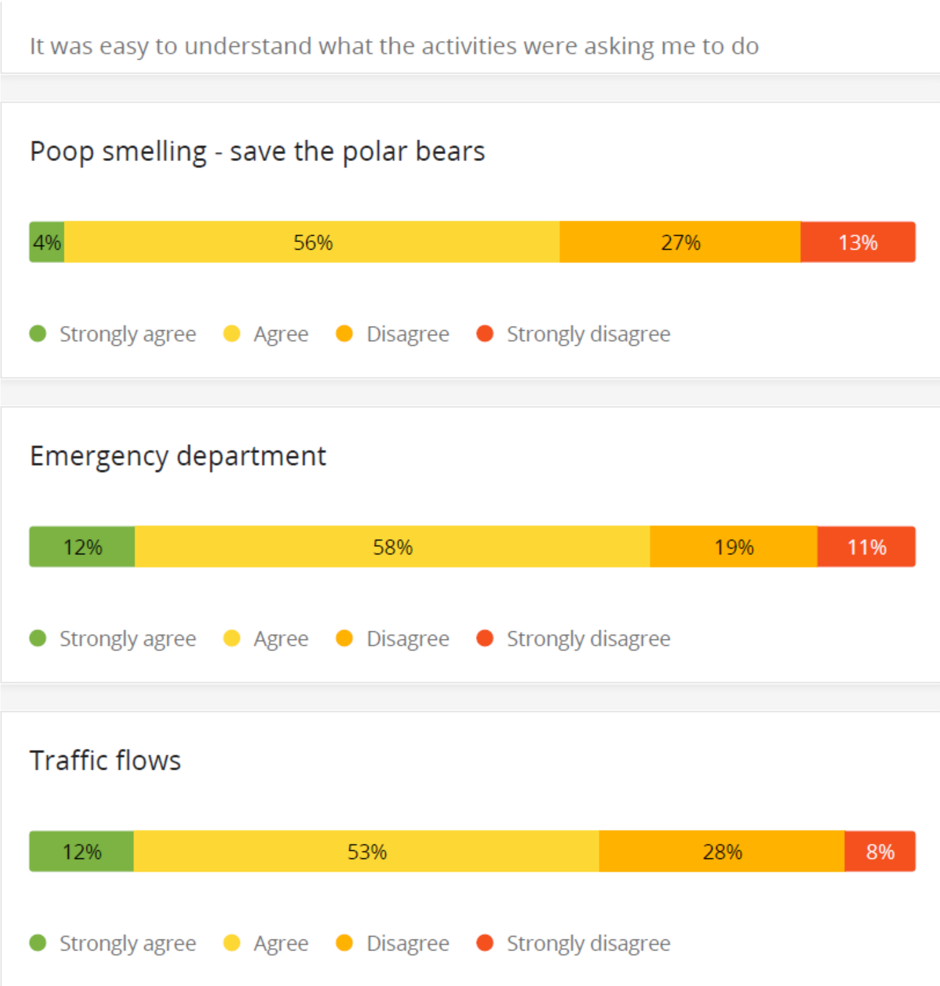
2. Comic strips, narration and videos were integrated within the Word document to overcome the traditional literacy demands of context-based assessment problems. These were provided in addition to written text, allowing students to use whichever medium as best for them to understand the context of the problem.
3. An easy to use online tool for probability distributions was made available.
   i. This could be used instead of, or alongside, the more traditional graphical calculator. The online calculator was simple to use and provided a graphical and numerical answer to problems
   ii. The same online calculator also enabled students to investigate in real time the impact of changing distribution parameters.
   iii. This simple to use feature enabled student to vary the parameter(s) of a distribution and see the ensuing changes in the graph of the distribution and the answers to calculations. It invited students to explore how changing the parameter(s) might improve the model for the context. This is a key idea of the curriculum which cannot be easily addressed in a traditional paper assessment.

Student survey responses









Teacher survey responses

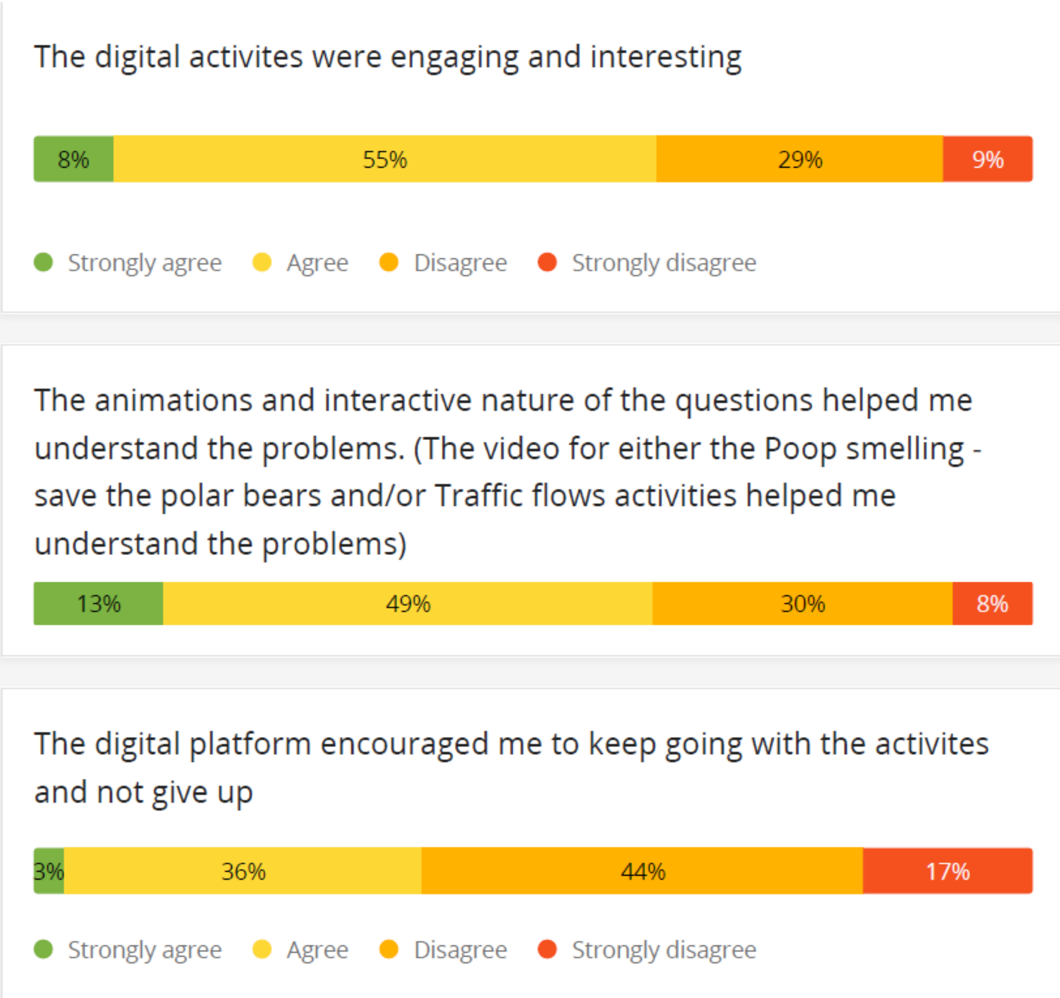
Most teacher comments were favourable for these criteria [Comments as received]:

- *It's another way to engage with the content and have a bit of fun exploring what it can look like in an easier way than number crunching.*
- *I like that students could visually see what the distributions looked like and the probability they were calculating.*
  Students had to focus on thinking about the best model distribution and thinking about context.
  For Q2b and Q3d students could explore adjusting models which is not able to be easily done in a paper exam.
- *Using the tools allowed students to think more about the distributions and the context rather then the mechanics.*
- *Questions became less mathematical and more conceptual and based around modelling real data.*
- *Should have been quicker to use the tool to compare theoretical and experimental data*
- *Focus on thinking about what the data means rather than the calculations to get to an 'answer'.*
  Contexts were accessible/equitable in that all students would know or not know in equal measure
- *Nice initial move forward.*
- *Easy for the students to pick up and explore with the tech available.*

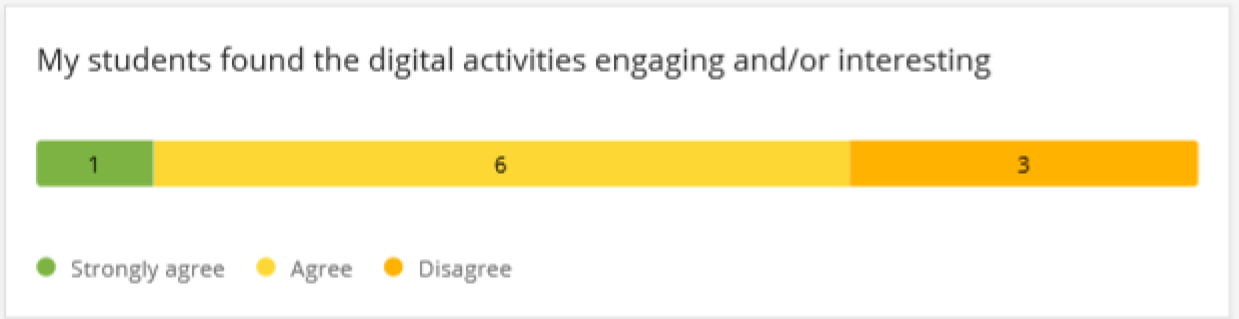
- *My students especially enjoyed using the tools for exploring the data. I liked the style of question and that it asked students to THINK about what they were seeing in terms of fit/modelling/simulation.*
- *The questions demanded a better understanding of sampling/distribution modelling/fit than what they experience in the paper exam. this threw them.*
- *Once they got the hang of it, they enjoyed using the tools.*

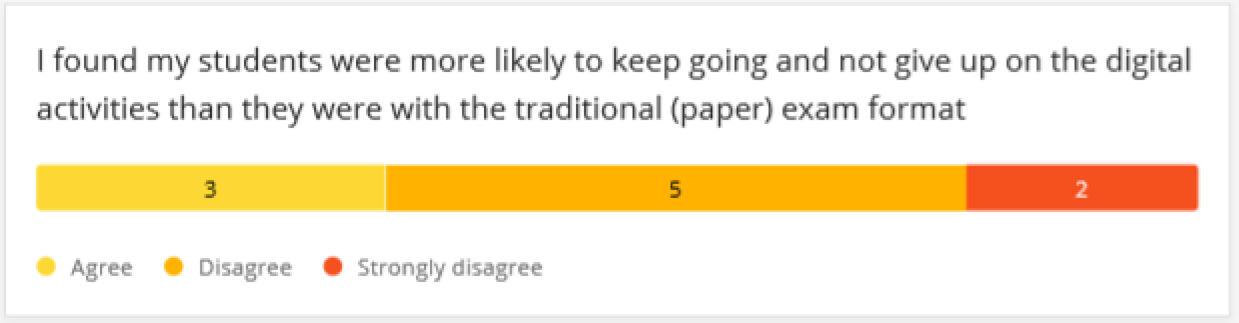
4.5.4 Encouraging engagement and resilience

Student survey responses



Teacher survey responses

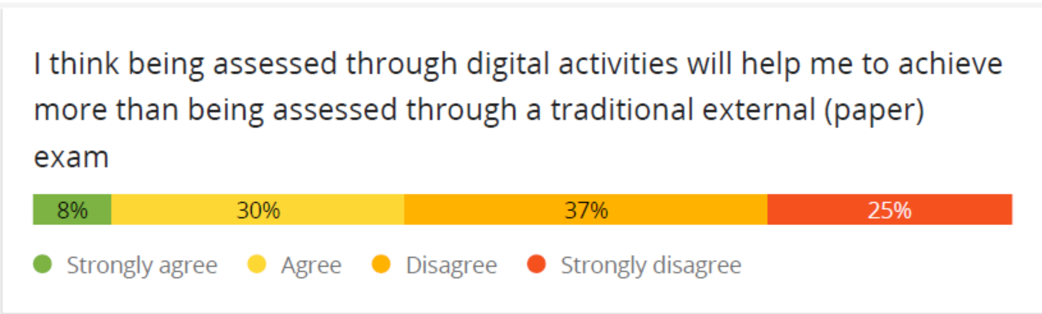




I thought the questions were interesting and the content engaging.

4.5.5 Maximising achievement

Student survey responses



22 students indicated that they much preferred paper assessments and were opposed to a move towards digital assessments [Comments as received]:

> *Don't do it online it's crap. It makes no sense we learned through books and calculators and get given an online exam during a year of Covid when the exams actually count. It's dumb. stick to paper its just better.*

> *I find it easier to look at paper and write my ideas down as it helps me understand the question. I also struggle to stay focused and concentrate looking at a screen for long periods of time.*

A further 11 said that they found entering mathematical text and symbols difficult and that this need to be resolved.

> *Please just stick to traditional paper exams!!! It's way easier to draw mathematical symbols and numbers than on a laptop.*

By contrast six students indicated that they were comfortable this style of digital assessment.

> *The tool is really interesting to use. I like that you can flick through the screens and see the changes to the distributions quickly and easily.*

> *Much easier experience compared to written exams*.

Five students commented that the problems were difficult to understand.

> *It seemed a bit ambiguous.*

> *The questions were written so confusingly.*

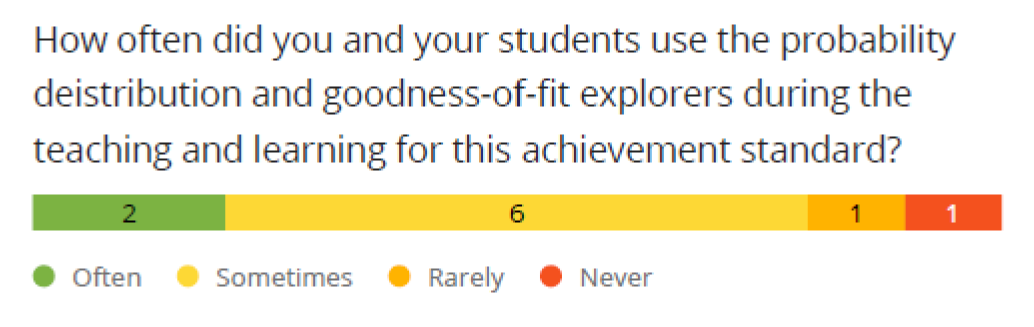
Four students reported that they found they had insufficient time for the activity

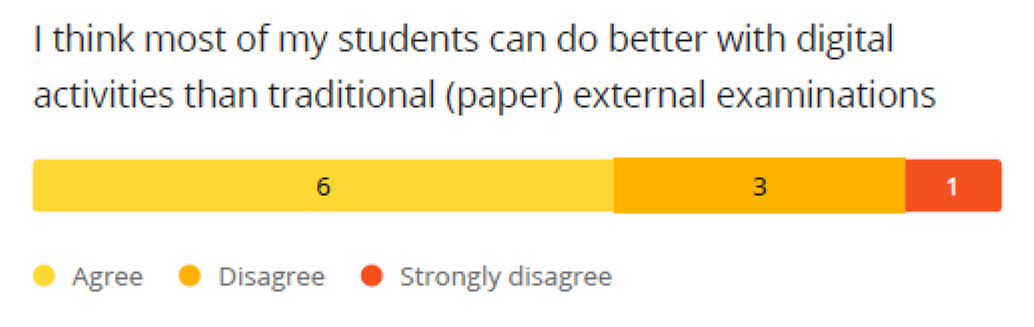
> *There was not enough time to complete the activities. I think that 45 minutes would have been better than 30 minutes.*

> *Not enough time to complete the tasks given.*

Teacher survey responses

The trial information stressed the need to familiarise students with the tools and to utilise the training materials. Time pressures prevented teachers from fulfilling this requirement as well as they might.





The new nature of the questions and the literacy levels required were concerns for some teachers [Comments as received]:

- *Quite different to the current questions in NCEA assessment.*
- *They jumped in far too soon looking at modelling. The pupils need to actually do some problems mechanically to achieve understanding. This is the first year they have encountered distributions and they need to be learning the skills, not the applications. The Achieved questions were fine as they could use their calculators, but the Merit and Excellence ones were confusing and poorly written.*
- *Students found the tasks confusing - more so that NCEA past papers. They preferred to use their GC.*
- *The questions set at the higher levels of achievement were quite different in the thinking required compared to traditional paper exam - more interpretation and comparison between sample and model, rather than extended problems requiring 2/3 processes or combination of models to acquire solution.*
- *My students actually said they would prefer to do the exam on paper and not digitally.*
- *My class were clear that they liked the tools for learning but would much rather write a paper exam.*
- *I thought they were great.*

4.5.5.1 Feedback from teacher interviews

School 1

This school does not enter students for any digital externals. Pen and paper were not offered to students whilst accessing the activities.

The students had about 90 minutes practice with the tools before using them in the activity. Found them very easy to use. Teacher had looked at the support resources and thinks she should have spent more time on them to help her students. The tools are wonderful she says, as were the activities – a step in the right direction to align with the investigative nature of the internals.

Her teachers much prefer teaching skills than problem solving, and her students give up too easily. This was no different on these activities to paper and pen versions. The teacher is not sure that external assessment should go down this route.

School 2

This school enters students for digital assessments in two subjects. There is a considerable overlap in the students entered for these standards and those trialling the activities. Students had paper and pencil to hand when doing the activities, but rarely used them. The students are fluent with word processing etc.

The teachers spent two lessons helping students to get familiar with the tool. They found this intuitive and easy to use. The teacher looked at the support materials available. On reflection she should have used the ideas in these videos with the students to support them before trialling the activities.

The students had no preparation on how to tackle the investigative questions at the end of the activities.

The teacher liked the activities and the students were engaged. She felt they would work well if students were better prepared for them, but that her students would probably prefer a traditional examination.

4.6 Outcome of the trial

The planning for this project took place before COVID-19 had been heard of. The changing alert levels in response to the COVID-19 pandemic had a significant impact on the ability of participating schools to engage with the trials or prepare students for the activities.

One consequence of this is that the volume of data collected at the end of the trialling period was not as comprehensive as it might have been. Although 52 schools registered, only 18 provided feedback about their participation in the trial.

A second consequence of this was an inadequate number of Māori and Pasifika participants, preventing any meaningful comparison of the experience of these groups with other groups.

The trial has demonstrated that this externally assessed Statistics achievement standard can be assessed in a digital environment. It has also suggested that this style of assessment is unlikely to be successful unless it is integrated within a teaching, learning and assessment programme so that learners are immersed in a cohesive experience through all three phases.

4.6.1 Answering the trial's evaluation questions

Do the activities provide evidence for the standard at all levels of achievement?

All three digital activities offered students opportunities to demonstrate all three levels of achievement.

Do the activities provide equitable access to the activities on BYO devices?

The intent of this criterion was twofold:

- to establish whether any BYO devices were more challenging to use than others
- to establish whether Māori and Pasifika students met greater challenges with BYO devices than other ethnic groups.

There were very few issues reported to do with accessing the activities on any devices. Those problems that were reported were related to faulty devices or isolated problems logging onto school networks. The vast majority of participants accessed the activities at school, on a laptop, using a Chrome browser. There were a few comments that it was difficult to complete the activities on an iPad.

Do the activities provide opportunities for modification and redefinition of assessment questions?

The activities developed were significantly different to traditional written assessments. Using a digital environment enabled the writers to introduce and explain the context in innovative ways, and to calculate probabilities using a tool which was easier to use than a graphical calculator. The style of question was similar to questions at Achieved and Merit level, but the digital environment allowed the writers to pose Excellence questions that were more investigative in nature. This is in line with the expectations of the New Zealand Curriculum.

These activities could not have been developed without the support and expertise of Anna Fergusson and the staff from UoA. The curriculum knowledge and the programming skills required to develop the tools are significant and are unlikely to be available within NZQA.

Do the activities encourage the engagement and resilience of students in completing the activities?

Feedback from both students and teachers reported that:

- the activities did encourage engagement by the participants. In particular, the way that the context was introduced with the aid of audio and visual experiences was found to be very helpful in this respect.
- the activities were less successful in encouraging resilience. Some participants reported that they found the literacy levels of the questions challenging and were unable to decode what they were expected to do.
- the shift of focus to a more investigative nature of question at Excellence was not welcomed by most participants.

Both of the latter issues could be alleviated by more support for students in using the technology and the type of question before completing the activities.

Do the activities support students to maximise their level of achievement?

The innovative ideas used in the design of the activities encouraged students to engage with the activities, but this was countered by the difficulties some students had with the technology and the style of Excellence question. Most participants responded that they preferred the traditional pen and paper assessments, although a respectable proportion were positive about the experience. Teachers were, in general, more positive about this criterion of the trial than the students.

4.7 Evaluation of Digital Statistics 2020 Innovation Trial

4.7.1 Indicators and success

Evaluation question	Criterion or Indicator	Data collection Method	Source of data	Standards (What constitutes success)	Response
1. Do the activities provide evidence for the standard at all levels of achievement?	Data reveals whether all levels of achievement are accessible to all students.	Via developer emails and online meetings and materials critique.	Teachers (including one current panel members for AS91586) and National Assessment Facilitator (NAF)	Evidence shows clearly that all levels of achievement are accessible.	Evidence from developers, materials critiquers and subject experts showed clearly that all levels of achievement are accessible within all three activities.
2. Do the activities provide equitable access to the activities on BYO devices?	Data reveals the benefits and challenges of using BYO devices.	Via end of trial surveys	Teachers and students from the groups Māori, Pasifika, all other	Evidence gathered identifies benefits of the use of BYO devices for students in the groups Māori, Pasifika, all other and suggests how challenges could be met.	Clarification of question: The intent of this criterion was twofold: • to establish whether any BYO devices were more challenging to use than others Response: some BYOD devices e.g. iPads were more difficult to use • to establish whether Māori and Pasifika students met greater challenges with BYO devices than other ethnic groups.

Evaluation question	Criterion or Indicator	Data collection Method	Source of data	Standards (What constitutes success)	Response
					Response: not enough data for meaningful analysis/conclusion
3. Do the activities provide opportunities for modification and redefinition of assessment questions?	Data reveals that the activities developed can modify and redefine existing patterns of questions.	Via review of final items	Workshop participants and subject matter experts	Evidence highlights how questions have been modified and redefined and suggests how further progress can be made.	The activities modified and redefined questions by: • introducing the context of the question via comic strips and videos • making available an easy to use online calculator for probability distributions • enabling students to investigate in real time the impact of changing distribution parameters.
4. Do the activities encourage the engagement and resilience of students in completing the activities?	Data reveals that students were engaged whilst undertaking the assessment.	Via end of trial surveys	Teachers and students from the groups Māori, Pasifika, all other	Evidence highlights what secured and maintained engagement from students in the groups Māori, Pasifika, all other and suggests further opportunities.	The activities did encourage engagement of the participants but were less successful in encouraging resilience.
5. Do the activities support students to maximise their level of achievement?	Data reveals how the performance of students in the groups Māori, Pasifika, all other compared to	Via end of trial survey and interview	Teachers	Evidence highlights how student achievement was maximised.	Teachers were, in general, more positive about this criterion of the trial than the students. Note: there are far fewer teacher responses in comparison to student responses.

Evaluation question	Criterion or Indicator	Data collection Method	Source of data	Standards (What constitutes success)	Response
	that under traditional (paper based) assessment format.				

4.7.2 Management of the evaluation

NZQA Title or role	Responsibilities
Project lead	Drafting of questions Analysis of data
NAF	Drafting of questions Evaluation of activities
Senior Advisor	Review of evaluation questions
Senior Project Manager	Evaluation management

5. Evaluation of the trial process

- **Was the testing adequate and capable of telling us what we need to know?**

The advent of a global pandemic and the impact on schools and students made it difficult to meet all objectives. The rise and fall of the alert levels in New Zealand posed challenges for the teaching and learning in schools. In particular schools chose to reduce the amount of teaching and learning and had to provide online teaching and learning. The decision to widen the opportunity to all schools was made in the light of these challenges. Whilst the initial interest from schools was significant the final number of schools and students involved in the trial eventually reduced to the number expected from the initial selection without the diversity of schools that were planned to reflect the overall schools' profile. No schools of decile levels 1 -3 participated in the trial and the number of Maori/Pasifika students was very low.

- **Did the trial test what we set out to test for Māori/Pasifika/all learners?**

No: the reduced number of Māori and Pacifika students participating in the trial meant that the disaggregated survey data was insufficient to elicit any conclusions about the ease of access to the activities for Māori and Pacific students.

- **Could we have better identified aspects worth testing for Māori/Pasifika/all learners?**

Yes: it would, in hindsight, have been better when planning the trial to work with students and teachers of these students to incorporate Māori and Pasifika student voice when planning the trial.

- **Is the evidence robust enough to support reliable decision-making?**

Yes, except as noted for Māori and Pasifika student cohorts. This was due to both COVID-19 changes to school engagement (see School participation) and trial design (see Lessons learned).

5.1 Performance against plan

5.1.1 Review of trial objectives

Proposal objective	Performance against objective
To identify and select appropriate digital environment(s)[2].	This objective was met by working with an expert teacher development group and NZAMT representative.
To develop prototype digital assessment activities for the assessment of AS91586 – Apply probability distributions in solving problems.	This objective was met.
To understand how the use of the chosen digital environment(s) enables the modification and redefinition of assessment activities.	This objective was met.
To utilise the opportunities within a digital environment to increase the engagement and resilience of NCEA students when taking part in external assessments.	This objective was partially met via feedback received via survey responses. The engagement of NCEA students when taking part in external assessment was increased. The resilience of NCEA students when taking part in external assessment was not increased.
To use lessons learned from the Digital Mathematics Innovation Trial 2019, applicable for this innovation trial, to improve the design of activities. See Appendix One	This objective was met. Note: Some students and teachers still reported difficulty with the literacy demands of the activities.
To investigate whether the chosen digital environment(s) would either hinder or provide support for equity of access for Māori and Pasifika students.	This objective was not met. This was because of • the impact of COVID-19 on schools and students' ability to participate in the trials • the trial design was not explicit about engaging Māori and Pasifika students (noted in Evaluation of the trial process).

[2] Environment: platform activities delivered by (MS Word) and tools (Anna Fergusson's Probability Distribution Explorer and Goodness-of-fit explorer)

5.1.2 Review of benefits

5.1.2.1 NCEA Online benefits

NCEA Online benefits	This trial's contribution to benefit(s)
Students' experience of equitable access and outcomes and relevant credentialed learning sustains their ability to succeed in higher education and the workplace.	The trial resulted in digital statistics activities which are accessible to students and provide opportunities to meet the national standard. However, teachers and students need support to ensure that students have the necessary technological skills. Once the latter issue is overcome then digital statistics assessments are likely to be accessible to a broad range of candidates. It is expected that this will improve equity of access to NCEA external assessment for Māori and Pasifika students and provide wider opportunities for students to succeed.
Sustained confidence in the assessment system as demands of it change.	This trial demonstrated that there is the potential to capitalise on the use of digital activities to transform assessment and enable students to demonstrate higher-level thinking skills in a greater variety of ways. This will help sustain long-term confidence in the assessment system.
Greater efficiency and enhanced quality assurance.	The trial demonstrates that statistics can be assessed digitally. If digital assessment of statistics can be made easily accessible and widely available, then this will contribute to the overall quality assurance and efficiency benefits of NCEA Online.

5.1.2.2 Other benefits identified for (potential) realisation

It was identified that there is a need for professional learning and development for teachers as they embark on a digital assessment journey. This offers a unique opportunity to support teachers to develop appropriate pedagogy to support Māori and Pasifika students. This could be realised by working with students and teachers to develop a coherent teaching, learning and assessment package that uses digital activities.

The digital activities that were developed for the trial were recognised by teachers, as excellent vehicles for teaching and learning.

6. Related work, planned or underway

Related work	Impact
Where to next for Mathematics and Statistics	• PLD to support new, cohesive environments • Closer alignment of learning and assessment
Digital tools and applications investigation (including SoNET capability).	Work is currently underway to identify any NCEA subjects where external assessment has not been digitised due to the technicalities and challenges in digitisation. Subjects that require graphing or formula are being considered included statistics. The outcomes of this project will help inform the next steps for digital statistics. This trial supports the need for further investigation – student comments included "It's too hard to enter numerical symbols and equations. We thus cannot use the same level of reinforcement with equations and proofs for our argument as on paper."
Assessment Master UI improvements	This trial supports the importance of a good/easy/intuitive user experience. Students commented on UI issues of entering symbols and also "It was much harder to use a device than paper for me. Switching between tabs and having to type maths into a doc was much worse."
Review of Achievement Standards (RAS)	National Assessment Leader (Mathematics) will maintain close contact with the RAS process currently underway with the MOE so that any plans are fully aligned with the direction of the new achievement standards.

7. Lessons learned

For this trial specifically:

- Teachers need professional learning and a comprehensive set of support materials for teaching and learning so that the assessment mirrors the teaching and learning experience.
- Students need comprehensive support and time to become familiar with the digital environment(s) used.
- Clear guidelines, expectations and where necessary agreements or e.g. memorandums of understanding with suppliers, need to be agreed on as early as possible when working with external parties, even if there is no payment involved.
- The tools used within the activities are sophisticated. It would be time consuming and expensive to recreate them. The expectation is that third-party support would be required (and that we do not have the development expertise within NZQA).
- Relating to equity:

  - It would, in hindsight, have been better when planning the trial to work with students and teachers of these students to incorporate Māori and Pasifika student voice when planning the trial (see Evaluation of the trial process).
  - Extending the invitation to all schools to participate in the trial increased the pool of schools that wanted to participate but resulted in the cross section of students involved being narrower than needed to provide adequate data for measuring equitable outcomes for Māori and Pacific students.

For implementing innovation trials/research:

- NZQA policies prevent the sharing of documents via Google format. For many schools the preferred environment is Google. Transfer from Office 365 to Google can cause problems. This is a significant issue when working with schools and needs further consideration in terms of NZQA policy.
- Documents needs stripping of personal information before sharing outside NZQA.
- Zoom and Teams are both good vehicles for remote meetings. For this trial Zoom had the greater flexibility and was accessible for more people outside NZQA. Some external people had difficulty accessing meetings via Teams.
- The teachers we worked with preferred to meet remotely from 5 p.m. to 6 p.m.
- Qualtrics is a powerful tool for creating surveys, analysing data and presenting graphical displays. Improved confidence and competency by team members could increase its usage and better utilisation.

Appendix One: Lessons Learned from the 2019 Digital Mathematics trial

i. The tool (Desmos) does not solve the issue of manipulating algebraic equations without the use of paper or a touch screen.
ii. The teacher participants were improving their knowledge of how to use the tool whilst developing assessments which met the requirements of the Achievement Standard. Some questions contained higher levels of literacy requirements than those normally found in external assessments. The formatting of the questions could have been improved, so all aspects required to answer the question are easily accessible to students.
iii. Teacher participants focussed on developing assessments which mirror current paper assessments to ensure that they would assess against the Achievement Standards. This restricted their work. There is the need to re-imagine questions. Digital assessment offers a limited advantage for Substitution and Augmentation, it offers a greater advantage for Modification and Redefinition in the SAMR continuum. Non-timebound assessments may also provide a better opportunity for Modification and Redefinition.

Appendix Two: Initial survey of statistics teachers re use of digital platforms and tools in the teaching of statistics.

Statistics teachers were asked to complete a short survey to inform the choice of digital environment for the trial.

1. Please select your school/kura from the list below
2. Your role in the statistics department
3. Are you completing this survey for mathematics, statistics or both?
4. Do you use any of the following digital statistics applications for teaching and learning at your school for Level 3 Achievement Standard 91586 - Apply probability distributions in solving problems?
    a. CODAP
    b. Education Perfect
    c. iNZight Lite
    d. NZ Grapher
    e. Other
5. Do you have any comments or feedback on digital statistics applications you have used for teaching and learning?

Appendix Three: Evaluation criteria for digital platforms and tools

Criteria	Description
Functionality	The tool(s) must provide functionality to create problems and reimagine statistics and probability problems.
Cost	The environment(s) must be free to teachers and students.
Accessibility	The tool must be easily accessed by students, especially at BYOD schools where it will be accessed from a number of different devices and platforms.

Appendix Four: Student Survey

Q1 What is the name of your school or kura? If your school or kura is not on the list please select 'Other'

Q2 Which school year are you in?

Q3 Which ethnicity do you identify with? Select all that apply.

Q4 I think being assessed through digital activities will help me to achieve more than being assessed

Q5 The digital activities were engaging and interesting. (Strongly agree to strongly disagree.)

Q6 The video for either the Poop smelling - save the polar bears and/or Traffic flows activities helped me understand the problems. (Strongly agree to strongly disagree.)

Q7 The digital platform encouraged me to keep going and not to give up with the activities. (Strongly agree to strongly disagree.)

Q8 It was easy to enter my answers into the e-document. (Strongly agree to strongly disagree.)

Q9 It was easy to use the probability distributions explorer (for Activity 1 - Poop smelling and/or Activity 3 - Traffic flows) (Strongly agree to strongly disagree.)

Q10 It was easy to use the goodness-of-fit explorer (Activity 2 – Waiting Times) (Strongly agree to strongly disagree.) (Strongly agree to strongly disagree.)

Q11 It was easy to understand what the activities were asking me to do. (Strongly agree to strongly disagree.)

Q12 What sort of device did you use for these activities? Select all that apply.

Q13 What browsers did you use for these activities? Select all that apply.

Q14 Did any devices, browsers or applications work better than others while completing the activities? If so which ones?

Q15 Where were you when you did the activities? Select all that apply - you may have done one at home and one at school for example.

Q16 Did you experience any network or device problems while doing these activities? Select all that apply.

Q17 Finally, is there anything else you would like to tell NZQA about your experience with these activities?

Appendix Five: Teacher Survey

Q1 What is the name of your school or kura? If your school or kura is not on the list please select 'Other'

Q2 What role do you have in your school?

Q3 How often did you and your students use the probability distributions and goodness-of-fit explorers during the teaching and learning for this achievement standard?

Q4 I think most of my students can do better with digital activities than traditional (paper) external examinations. (Strongly agree to strongly disagree.)

Q5 My students found the digital activities engaging and/or interesting. (Strongly agree to strongly disagree.)

Q6 I found my students were more likely to keep going and not give up on the digital activities than they were with the traditional (paper) exam format. (Strongly agree to strongly disagree.)

Q7 I found the marking schedules easy to use. (Strongly agree to strongly disagree.)

Q8 The activities provided a good opportunity for students to be awarded these grades. Select all that apply.

Q9 How likely is it that you would recommend these activities to a colleague?

Q10 What did you like about these activities?

Q11 What didn't you like about these activities?

Q12 Please give examples of any questions related to the technology that your students asked while completing the activities. For example, questions about devices or the probability explorer tools.

Q13 What advice would you give to another teacher who was considering trying these activities?

Q14 Finally is there anything else you would like to tell NZQA about these activities?

Appendix Six: Teacher interview questions

These questions were used with two of the teachers who completed the survey. The teachers were selected at random and the interview was conducted by telephone.

1. Did you ensure that your students were familiar with the apps [platform and tools] and style of Excellence questions before asking them to complete the digital activities?
2. Did you suggest or allow them to use pen and pencil to sketch out solutions before entering them into the tools?
3. Does your school enter students for existing digital NCEA external assessments? If so, had the students who participated in the trial completed one of these digital assessments?

Appendix Seven: What were the activities like?

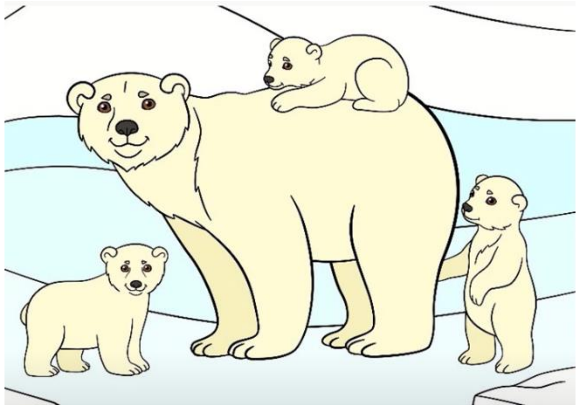
Introducing the context via different media:

Assessment Question: Activity 1 – Poop smelling – save the polar bears

Adapted from a National Geographic article "Elvis the dog sniffs out pregnant polar bears.

Watch this video (make sure you have headphones on if you are close to other people) before answering the questions below.

Polar Bears video

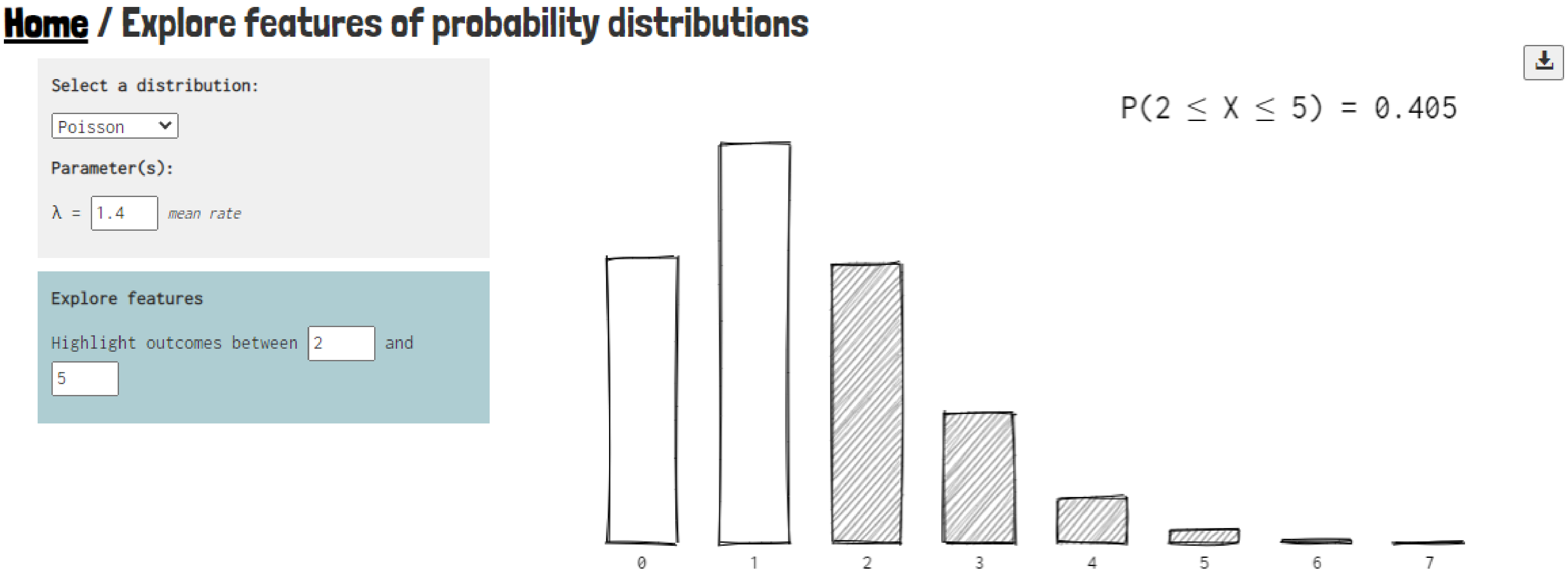


Assessment Question: Activity 3 – Traffic Flows

This link (Wellington traffic camera) shows traffic flowing on Oriental Parade in Wellington during one morning in August 2020. Watch and count how many cars travel towards you in a one minute interval.

The tools:

Link to tool: [Explore probability distributions tool]



Link to tool: [Explore goodness-of-fit for probability distributions tool.]