

Squirm Germ Innovation Trial 2020 Summary

NCEA Online Programme



NEW ZEALAND QUALIFICATIONS AUTHORITY
MANA TOHU MĀTAURANGA O AOTEAROA

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

Table of Contents

1. Trial overview	3
2. Trial objectives	3
3. Trial description	3
4. Objectives evaluation	4
5. Evaluation description	5
6. Gathering credible evidence	5
6.1 Qualitative evidence	5
6.2 Quantitative evidence	6
7. Evaluation results	6
8. Evaluation review	7
9. Trial outcome	8
10. Recommendations	8
Appendix A:	9
Appendix B:	9
Appendix C:	10



NEW ZEALAND QUALIFICATIONS AUTHORITY
MANA TOHU MĀTAURANGA O AOTEAROA

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

1. Trial overview

This innovation trial set out to use an engaging gamified assessment¹ as the basis for designing and developing alternative forms of digital student assessment. Biology standard AS90927 was selected because this externally assessed standard has real life applications that are engaging and relevant for NCEA Level 1 students. The assessment allowed a range of students to answer questions and submit responses online in an engaging gamified environment.

Evaluation included feedback on student experience and feedback from subject matter experts, including teachers, in relation to credibility aspects including:

- the value of a gamified method of assessment as another option for students
- whether this prototype meets the requirements of the standard, AS90927
- how engaging this type of assessment is for students, relative to a paper-based exam.

2. Trial objectives

1. To develop a prototype of an engaging gamified external assessment as one possible alternative to paper-based examinations.
2. To assess student experience of gamified external assessment and the extent to which it provides Māori, Pacific and all students with the opportunity to demonstrate their skills and knowledge other than via a paper-based response.
3. To have subject/assessment experts evaluate the credibility of this gamified external assessment.
4. To compare spread of performance from gamified assessment with that of paper-based profile of expected performance.

3. Trial description

“Squirm Germ”, the game, was created in 2019 during a previous innovation trial which recommended further developing the prototype game into an engaging gamified external assessment for trialling in 2020. The first innovation trial was developed from a prototype designed by NZQA’s Pacific STEM Ambassadors. They also provided feedback on various iterations of the game as it developed, which informed the direction of this second innovation trial.

Secondary Examinations and an externally contracted developer, Gamefroot, worked together to further develop the prototype game, Squirm Germ. This included creating explicit assessment questions, that students attempted, and capturing the student responses.

Once the assessment questions were added to Squirm Germ it was to be trialled with students from a range of different schools and subject experts or teachers, however participation was impacted by COVID-19. There was initially high interest from approximately 50 schools to participate in the trial during Term 2, but the number of schools who committed to completing the trial after level 4 lockdown was considerably lower. Four schools participated in term 3, involving a total sample size of 86 students.

¹ Gamified assessments apply virtual game-design elements, principles, activities and processes to solve problems in non-game contexts, e.g. in this case, assessment of NCEA standards.

There were no controls set around the way the game was used by students (e.g. no instructions to complete the assessment, or undertake the assessment in or outside of classroom), as teachers were responsible for managing participation in the trial and game access for students.

Students involved were selected through ongoing involvement with NZQA and Gamefroot's established network. Student responses were captured digitally and transferred to the current format of paper-based external exams to retain the "look and feel" of an exam, so that markers were able to mark the responses easily. This was deliberately selected to provide markers with a normal marking experience.

We selected external markers, who are subject matter experts, using our existing contacts through the National Assessment Facilitator for Biology. Individual student results were sent to schools through Principal's nominees.

4. Objectives evaluation

1. To develop a prototype of an engaging gamified external assessment as one possible alternative to paper-based examinations.

This trial successfully made further developments to the Squirm Germ 2019 prototype to develop an engaging gamified external assessment. Gamefroot removed the learning prompts from the original game prototype and forced students to complete the assessment before they could progress fully through the game. Student responses were securely captured and marked by an experienced Biology marker. Results were securely shared with schools.

2. To assess student experience of gamified external assessment and the extent to which it provides Māori, Pacific and all students with the opportunity to demonstrate their skills and knowledge other than via a paper-based response.

Trial participation was impacted by COVID-19. There was large participation interest, but the number of schools who committed to completing the trial after level 4 lockdown was considerably lower. NZQA engaged schools who were able to participate in the trial.

Student responses were captured in Google Drive files which were accessed by the markers. Student and teacher survey responses were captured by Gamefroot and shared with NZQA. The length of time it took students to navigate through and complete the assessment was captured and reported on. Separate Māori and Pacific students' analyses were excluded due to the low participation numbers, four Māori and seven Pacific students.

The prototype only assessed a portion of the AS90927 and was anticipated to take less than 60 minutes based on the current timing of achievement standards. Analysis of student completion time showed that on average students spent just under 30 minutes engaging with the assessment. See data in [Appendix A](#).

The survey responses provided the most valuable method for assessing student experience with Squirm Germ. There was a good mix of feedback received, both in support of gamified external assessment and describing why they would prefer to maintain more traditional written methods of assessment. See full survey feedback in [Appendix B](#).

3. To have subject/assessment experts evaluate the credibility of this gamified external assessment.

To evaluate the credibility of this gamified external assessment, the marker engaged for the trial and the Biology teachers from the participating schools were asked to evaluate the credibility of this gamified external assessment. The majority of the feedback was positive, and the subject experts felt that the assessment met the requirements of the standard whilst being an engaging and interesting format for students. See teacher comments in [Appendix B](#).

4. To compare spread of performance from gamified assessment with that of paper-based profile of expected performance.

We reevaluated this objective due to low participation because of COVID-19. The low participation numbers made completing any spread of performance comparison impossible, therefore it was not completed.

5. Evaluation description

We gathered both qualitative and quantitative evidence to evaluate the trial. For qualitative information, we surveyed students, teachers and the Biology marker on aspects such as engagement, comparison with paper-based exams, and feedback was sought on what could be done differently. Quantitative data was gathered to measure how long it took the student to complete this assessment. Initially we planned to make student achievement comparisons between this trial and historic assessment results, but this wasn't possible due to low participation numbers due to COVID-19 and the differing conditions that the assessments were completed under.

6. Gathering credible evidence

6.1 Qualitative evidence

- **How this meets the requirements of standard 90927.**

Participating teachers felt that the assessment did meet the requirements of the standard, but some feedback suggested they would have preferred the game to incorporate the whole standard, rather than a portion of the standard as Squirm Germ did.

The marker also agreed that the assessment met the requirements of the standard. They suggested that students would benefit from some scaffolding and the questions were potentially too open ended which could make the assessment harder for lower ability students.

- **Is this more engaging than paper based? Why/why not?**

Most of the students who completed the assessment and survey found it engaging, with one student commenting “...*this experience was much better than paper-based assessments*”. Students who preferred paper assessments found that traditional methods of assessment felt “... *more formal and serious*...”.

Teachers really enjoyed the look and feel of the assessment and the Biology marker commented that “...*students are used to digital presentations and I would think moving animations and game play would be more inviting/engaging than a paper test*”.

- **Do student responses to this assessment show their understanding?**

Students worked through the game and answered the assessment questions. Schools had administered the assessment under a variety of conditions, meaning that the responses were not necessarily a reflection of how they would respond under exam conditions. The Biology marker commented that some students had evidently worked together on their responses and other students had given very brief responses. Despite the differences in responses submitted, the marker felt that students were able to demonstrate their understanding through the assessment.

- **Do student responses to this assessment show that they can apply their learning? How?**

The marker acknowledged that student responses did demonstrate they could apply their learning. This was more evident in students who met the assessment criteria of the standard. The marker suggested incorporating a combination of question levels to cater to a larger range of student abilities, to support all students applying their learning. They also suggested the assessment could include other ways for students to respond, for example verbally or by drawing.

6.2 Quantitative evidence

Data was gathered to measure how long it took students to complete this assessment. Analysis of student completion time showed that on average students spent just under 30 minutes to complete the assessment. Full data can be seen in [Appendix A](#).

Data was captured from the following three schools:

- School A. Average time of engagement: 21.85 minutes.
- School B. Average time of engagement: 36.33 minutes.
- School C. Average time of engagement: 27.75 minutes.

7. Evaluation results

As a result of COVID-19, school engagement was a struggle as they focussed on completing teaching content, organising internal assessments and gathering evidence for internally and externally assessed standards. Taking part in this trial was not a priority for schools. From almost 50 schools that indicated their interest to participate initially, only four schools recorded entries. The outcomes are summarised below:

- 86 grades from four schools were used in the analysis.
- Participant schools were School A (20 responses), School B (9 responses), School C (38 responses), and School D (19 responses),
- Separate Māori and Pacific student analyses were excluded due to the low number of responses - four for Māori and seven for Pacific students.

Student and teacher feedback were very positive, and they enjoyed the game. Gamified assessment is a very different format to paper-based assessments and the digital NCEA Online assessments and some students felt it was so different it didn't feel like a serious or formal assessment.

We set out to look at student views of gamified assessment and to explore whether there were any trends from Maori and Pacific students. Several attempts were made to reach out to kura and schools with a high proportion of Pacific students, but resources were stretched in schools. Students had lost too much of their learning time to take on anything else. The second lockdown in Auckland also impacted on participation from schools that have high proportions of Pacific students. Our internal networks were utilised, but the fact remained that time was a limitation for schools. Ethnicity breakdown showed very little participation from Māori and Pacific students. See [Appendix C](#).

The data that we planned to analyse to compare spread of performance from gamified assessment with that of paper-based profile of expected performance was not possible as the engagement was too low.

8. Evaluation review

1. Was the testing adequate and capable of telling us what we need to know?

COVID-19 significantly impacted this innovation trial. We planned to have students engage with Squirm Germ at the 2020 ASB Polyfest, but unfortunately due to COVID-19 Polyfest was cancelled. There was initially a high level of interest from schools to participate in the trial, but actual engagement was heavily reduced and there were only four schools who participated. The timing of the trial could have also contributed to the lack of participation, as we had to delay communications with schools and participation took place in term 3, instead of the planned term 2. The further COVID-19 lockdown in Auckland saw Auckland schools withdrawing from the trial.

Despite the low levels of participation, there were good responses to the student and teacher surveys and the Biology marker provided valuable insights from a subject matter experts perspective. Both this trial and the previous Squirm Germ trial have allowed NZQA to gain beneficial insights into gamified assessment.

2. Did the trial test what we set out to test for Māori/Pacific/all students?

The cancellation of the 2020 ASB Polyfest and low participation impacted on the numbers of participating Māori and Pacific students. In collaboration with the NZQA's Data team, it was decided that the breakdown into ethnic groups was not viable.

The low participation numbers also meant that our fourth objective (*to compare spread of performance from gamified assessment with that of paper-based profile of expected performance*) couldn't be achieved.

3. Could we have better identified aspects worth testing for Māori/Pacific/all students?

We considered further development to the assessment, like incorporating speech to text and translating the game, but these were large changes to the initial scope of the work and the cost, time and resourcing involved in making these changes was too great. Gamifying an assessment was a quantum leap for both NZQA, schools and students and it was deemed worthwhile to trial it without further changes.

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

There is enough evidence to support further work as only a narrow aspect of this achievement standard was "gamified". This standard deals with fungi, virus and bacteria. The gamified assessment only had a prototype for bacteria. The assessment was developed on top of a National Assessment Facilitators existing workload and at times capacity was quite heavily stretched. Dedicated time on this project has the potential to develop an exciting prototype that fully assesses a standard.

5. Were those involved in the trial well enough prepared to be able to carry out their trial roles effectively?

Definitely. The tasks were developed as a prototype for learning and then as an assessment. The game developers came up with their ideas and the National Assessment Facilitator worked with other subject experts to confirm questions and expected evidence, based on the current way of assessing external assessments. An assessment schedule was also developed together with the assessment questions and quality assured to ensure that it met the current assessment requirements in terms of validity and reliability.

The schools were communicated with and supported throughout the trial by the National Assessment Facilitator.

The maker was experienced and worked with the National Assessment Facilitator throughout their marking.

6. Was the trial as set up capable of determining whether the trialed product/process meets NZQA's digital assessment needs?

Processes were set up to align with NZQA's digital assessment needs. For instance, assessment attempts by students were linked to their NSN so these could be traced as individuals' authentic attempts. Marked results were sent directly to the school contact person, so results could be released in a secure way and they were kept confidential to each school. Timing was traced so that we were able to see how long each student spent on each question.

9. Trial outcome

The project started out as a way of looking to assess differently from paper-based externally assessed standards. Gamified assessment was proposed by STEM Ambassadors, who were Pacific students in their 2nd year of studying medical science courses at University of Auckland in 2018.

The trial was adversely impacted by COVID-19 as schools struggled to use up remaining time for learning and assessing, leaving no time for trials such as this. However, schools are keen to participate in future trials and such initiatives by NZQA. Some of the schools participated without students having the prior knowledge of the content of the achievement standard. The quantitative results were inconclusive as a result.

Qualitative results were positive and showed a high level of interest and engagement from both teachers and students.

Despite the impacts of COVID-19, the trial was still successfully delivered to a range of users and results were assessed and feedback received. Quantitative results were impacted, but the trial managed to provide enough information to achieve most outcomes and close the trial.

10. Recommendations

The recommendation is to inform the Ministry of Education of the work that has been undertaken which may stimulate thinking for the Review of Achievement Standards.

Appendix A:

Data on submission times in minutes:

School A

Submission times in minutes: 16, 5, 21, 20, 17, 16, 25, 21, 28, 21, 16, 37, 22, 25, 41, 24, 24, 23, 13, 22.

Mean average: 21.85 minutes.

School B

Submission times in minutes: 50, 51, 47, 29, 52, 47, 27, 10, 14.

Mean average: 36.33 minutes.

School C

Submission times in minutes: 14, 30, 41, 32, 44, 48, 42, 36, 39, 40, 39, 26, 39, 35, 39, 24, 34, 21, 11, 15, 24, 22, 17, 18, 11, 20, 23, 19, 21, 35, 17, 7, 18, 9, 35, 49, 33.

Mean average: 27.75 minutes.

Appendix B:

Student survey comments:

- Easy to use, interactive aspects, engaging, not so monotonous, more appealing than paper-based, not so stressful, feel good environment, should be used as end of year exam, the ability to have fun and recall what you have learnt at the same time, easy to visualise the life process of bacteria and the effects of antibiotics on it, I can actually relate to it better in this way, etc.
- I feel like it is very easy for people to be engaged by this game. It was a good test and fun to play.
- Compared to paper-based assessments, it is a lot more fun (obviously) and a lot less stressful, and it doesn't feel like a chore.
- I think that this experience was much better than paper-based assessments. This is because it was really fun to do (whereas paper-based assessments are as fun), and it was easy for me to engage in the assessment.
- it makes me actually want to do an assessment :/
- This is easier to do because it doesn't require us to write our answers. We could just type it. Typing allows us to put our answers down quicker compared to when we have to write it by hand which is a pain and most of the time our hands get tired.
- I prefer paper-based assessments as they are more formal and serious, so it is easier to prepare for them.

Teacher survey comments:

- This looks cool, how cool is this,
- This looks amazing. Is there a possibility of trialling this?
- Love it
- I had a play with it; looks really cool
- Wow! This is how all assessment should be!
- This looks incredible! Is it available to all schools?
- This looks was better than the research report we made them write
- Loving the solutions; amasing, awesome etc
- The game is only associated with nutrition (inaccurate) and reproduction as for life processes. No other life processes are addressed. Also, the other microorganisms are locked. Is there any way to unlock these?
- A few students prefer to do exams on paper; they are not used to digital platform. Some students claimed they preferred the formal nature of the paper-based exams.

Kura survey comments:

- Students were focussed and accepted this form of learning and assessment.
- Translated version in te reo would be great, together with audios.
- Students mostly use ipads rather than keyboards.

Appendix C

Ethnic breakdown of students

New Zealand European	37
Asian	28
MELAA	7
Pacific Peoples	7
Māori	4
European	3
Total	86