evaluation associates
Te Huinga Kākākura Mātauranga

2022 NCEA Te Reo Matatini me te Pāngarau | Literacy and Numeracy Pilot Evaluation

Evaluation Report One

October 2022

Tēnei rā te whakairi ake i te kete o te wānanga,
Tōna mauri nō runga, nō Rangi, nō raro, nō Papa,
Tēnei te mauri o te mātauranga ka whakatakina ake,
Kia wātea ai ēnei pūkenga,
Hui e, tāiki e!

Bestow the life force of learning,
Upon these repositories,
So that aspiration of quality learning for our children is paramount,
And remains at the forefront of all of our works,
So that they may flourish and thrive,
For all eternity!

Table of contents

Executive summary 5

Summary of recommendations 10

SECTION A: Introduction 13

1 Pilot of Te Reo Matatini me te Pāngarau | Literacy and Numeracy standards 13

1.1 2021 Pilot 13

1.2 2022 Pilot 13

2 Evaluation Approach 15

2.1 Scope 15

2.2 Evaluation methods 16

2.3 Sources of data 16

2.4 Limitations of the evaluation 18

2.5 Report structure 19

SECTION B: Findings for objective one 20

3 How well are the assessments performing? 20

3.1 Overall results 20

3.2 Achievement by year level 23

3.3 Attainment rates compared with e-asTTle scores 25

3.4 Psychometric analysis for literacy and numeracy assessments 27

3.5 Teacher perceptions of how well the Literacy and Numeracy standards are performing 27

3.6 Te Reo Matatini me te Pāngarau Common Assessment Activity 30

4 How equitable are the assessments? 31

4.1 Assessment results by ethnicity, gender, school decile, ELL status and tertiary/alternative education status 32

4.2 Student perspectives on the assessments 43

4.3 Teacher perspectives on the equitability of the literacy and numeracy standards 49

Summary of Objective One 58

Observations and recommendations for objective one 58

SECTION C: Findings for objective two 60

5 Understanding student readiness 60

6 Tools and resources 63

6.1 Uptake of literacy and numeracy tools 64

7 Reported impact of the implementation of the Te Reo Matatini me te Pāngarau | Literacy and Numeracy standards and assessments on teaching and learning 68

7.1 Changes to systems and programmes 68

7.2 Impact of changes on pedagogy and students 70
7.3 Impact of changes on students who access special assessment conditions (SAC) 72
8 How have kura, schools/organisations experienced the pilot and what opportunities are there for strengthening implementation? 74
8.1 Perspectives on piloting the literacy and numeracy standards 74
8.2 Strengthening Implementation 76
8.3 Perspectives on the implementation of Te Reo Matatini and Pāngarau standards 78
Summary for Objective Two 81
Appendices 82
Appendix 1: List of Figures and Tables 82
List of Figures 82
List of Tables 83
Appendix 2 84
Psychometric analyses Summary Report: 84
Appendix 3 90
Tai Huki - Initial readiness report 90
References 91

Executive summary

Introduction

Evaluation Associates | Te Huinga Kākākura Mātauranga was contracted in 2022 to evaluate the NCEA Te Reo Matatini me te Pāngarau | Literacy and Numeracy standards, as a critical part of the quality assurance. This evaluation of the standards and assessments and will inform a transitional year in 2023, which will lead to the full implementation of Te Reo Matatini me te Pāngarau | Literacy and Numeracy standards nationally in 2024. The pilot of the NCEA Te Reo Matatini me te Pāngarau | Literacy and Numeracy standards began in 2021 with the purpose of understanding whether the standards, and methods of assessing these are appropriate, and whether the participants were ready for change.

The 2022 pilot of the Literacy and Numeracy standards involves 162 New Zealand schools, seven tertiary/alternative education organisations, and seven schools from Realm countries. Ten Māori-medium kura and three secondary schools providing te reo Māori education are involved in piloting the Te Reo Matatini me te Pāngarau standards. Ākonga/students from these schools, kura, and organisations participated in two assessment events.

The Evaluation

This evaluation of the 2022 pilot of the new standards and assessments focuses on whether the Common Assessment Activities (CAAs) are "fit for purpose" and are equitable for all students and explores how schools and organisations are building their capability for implementing the standards and assessments and what is needed to further support going forward. The structure of the report follows these two objectives, and each section includes an analysis of data, followed by key observations and recommendations for consideration.

This evaluation report is the first of two for the 2022 pilot year and follows the first assessment event held from 26 June to 1 July. The second report will follow the September assessment event and is due to be completed December 2022. The evaluation uses a mixed-methods approach, which includes an analysis of the achievement results, survey data from teachers in the pilot schools and organisations, and survey data from students who had completed their CAA. Feedback about the Te Reo Matatini me te Pāngarau standards and assessments was gathered through survey responses gathered by Tai Huki Consult Ltd[1] and contribute to the findings of this report.

This report is in three sections.

1. A description of the pilot and details of the evaluation approach.
2. Findings relating to evaluation Objective One: How the design of the assessments can be further refined and improved?
3. Findings relating to evaluation Objective Two: Understand how schools and organisations are building their capabilities and what is needed going forward.

[1] The link to the full report by Tai Huki Consult Ltd is in Appendix 3.

Objective one: understand how the design of the assessments can be further refined and improved

The achievement rates in each assessment area varied, with Literacy (reading) having the highest rate of achievement (64%), followed by Numeracy (56%), and then Literacy (writing) (34%). The patterns of achievement across the 2021 and 2022 pilots of the Literacy and Numeracy standards are similar.

The analysis of available e-asTTle data indicates that a proportion of the students who undertook the assessments were not achieving at the recommended curriculum level (4A). For Literacy (reading) this was 16.6% of students, for Literacy (writing) 26.6%, and for Numeracy 21.1%. So, it is important to note that when the results for students below the recommended curriculum level are excluded, the achievement rates are considerably higher.

The achievement rates for Pāngarau (18%) and Te Reo Matatini (24%) were lower, and there is less similarity between the two pilot years for Te Reo Matatini me te Pāngarau standards. Given the small number of ākonga participating in the first assessment event of 2022, it is difficult to draw any substantive conclusions. Feedback from pilot kura and schools notes a number of competing and overlapping activities which impacted on the level of ākonga participation in the first assessment event. There is expected to be more data available following the second assessment event in September, allowing for more substantive commentary about the performance of the Te Reo Matatini and Pāngarau assessments.

In looking at how the standards are performing in terms of levelling and alignment with other measures of attainment, the findings from the 2022 results data shows that the Literacy and Numeracy assessments continue to generally perform well against the standards. Comparison with e-asTTle results suggests that the e-asTTle assessment could appropriately be used as a readiness indicator for sitting the standards. What the readiness level should be set at requires further clarity regarding the expected achievement rate.

The relationship between the e-asTTle and the CAAs is particularly strong for Literacy (reading) and Numeracy, but there is less alignment with the Literacy (writing) assessment. It is noted by NZQA that for writing there are greater differences between the assessment tasks in the e-asTTle and the CAA than in the reading and numeracy assessment areas.

The overall indication from the statistical/psychometric analysis (Item Response Theory), which New Zealand Council for Education Research (NZCER) carried out, is that the items in the three Literacy and Numeracy assessments performed well. It is important to note that analysis or critique of the assessment tasks is not within the scope of this evaluation.

Examining the equitability of the new achievement standards involved comparing achievement results for various sub-groups of interest (ethnicity, gender, school decile, ELL, tertiary/alternative education, and Realm countries) and collecting the views of teachers and students. The patterns of achievement reveal differences that are likely related to factors such as ethnicity and, particularly, decile. Where possible, e-asTTle data was used to support comparisons with these sub-groups and the overall cohort. The achievement results for the small number of ELL students who were at the recommended readiness level for the assessments are not lower than the overall cohort.

Tertiary/alternative education and Realm country students also had lower rates of achievement, but less is known about this group of students participating in the pilot due to the absence of comparative e-asTTle data. Until information about students' previous curriculum level scores, which is an indication of readiness, supports the achievement data, limited conclusions can be drawn from the sub-group comparisons.

There were mixed views from teachers' views on the equitability of the Literacy and Numeracy standards, but for each assessment area there were more respondents who disagreed than agreed that the CAA was equitable for all learners. The Numeracy standard was seen as the least equitable, (with 55% disagreeing/strongly disagreeing) largely due to the amount of reading and writing within the assessment.

Objective two: understand how kura, schools, and organisations are building their capability and what is needed going forward

A range of experiences and perspectives on piloting the new standards and assessments were gathered through the teacher survey. The pilot schools and organisations reported different understandings of readiness and different approaches for determining student readiness for participating in the Common Assessment Activity. This is reflected in the numbers of students who participated in the assessments and who had e-asTTle scores below the minimum recommended level. Many schools and organisations were using the mid-year assessments as an opportunity to test their own processes and accuracy in identifying student readiness. The diagnostic tools teachers most commonly used to assist with determining readiness were teacher judgement and e-asTTle.

Pilot participants were asked about the tools and resources available to prepare students for the new Literacy and Numeracy standards and assessments, and views about the resources on the NCEA and TKI website were largely negative due to lacking relevance to actual teaching and learning. Participants' most-requested resources were CAA-related resources such as practical CAA tests, marking schedules, exemplars, and task activities that teachers could use directly in classrooms. The responses provided some insight into the expectations that "ready-made" resources would be provided for them, not only to assist student learning but also to help teachers better understand the Common Assessment Activities and the requirements of the standards.

The impacts on teaching and learning due to implementing the new standards and assessments varied. A small number of schools and organisations described relatively complex whole-school systems change focused on teaching and learning programmes and wider systems and processes. However, smaller scale changes to teaching programmes were more commonly reported. Some schools and organisations were waiting for the results of the CAAs before starting on any internal changes. Impacts on students were tentatively and minimally described, which is not unexpected given that schools and organisations are in the early stages of implementing this large change process.

Teachers also provided a useful and practical perspective on their experiences of the pilot and the opportunities were for strengthening implementation. While the majority agreed that administering the CAAs was straightforward, a range of mainly logistical and resourcing issues were raised as problems to be addressed going forward.

For kura and secondary schools using the standards and assessments for Te Reo Matatini me te Pāngarau, most kaiako felt that their ākonga were mostly ready to complete the assessments and had implemented changes to their learning programmes to support ākonga readiness and success. As with schools using the Literacy and Numeracy assessments, similar challenges were identified with the administration of the assessments. The challenges of particular note include reliable internet connection and ready access to devices, staffing capacity to assist at the assessment time, and access to critical and timely information.

Furthermore, there are two assessment options available for Te Reo Matatini me te Pāngarau, that is, ākonga can choose to demonstrate their achievement of the standards using a CAA or submitting a portfolio. In the first assessment event of 2022, all ākonga used the CAA. There has been a positive response from kura and secondary schools about the opportunity for ākonga to use portfolio, and while no portfolios were submitted in the first assessment event, the Ministry and NZQA remain committed to developing resources to support the use of portfolios.

General Observations

- Feedback from those involved in the pilot focused on what could be improved to make the new standards work better. While issues, concerns, and problems have been identified and while the MOE and NZQA continue to refine and simplify pilot processes and systems to make them work better, it is important that the changes recommended in the evaluation reports are used to review and, if necessary, to refine processes, systems, and supports.
- Further understanding is needed around "readiness". The analysis of achievement results and e-asTTle data suggests that e-asTTle results provide a good prediction about the likelihood of students achieving the Literacy and Numeracy standards. However, greater clarity about the expected achievement rates is required to better determine the appropriate curriculum level for readiness.
- A small number of teachers felt there are opportunities to improve the assessments' contexts and questions to be more inclusive and relevant to a wide range of students. There are also opportunities to address some of the technical and logistical issues that may be impacting on equity of access. However, it is likely that the factors that most significantly impact on patterns of achievement are deeply entrenched in an education system and the wider social and economic systems that perpetuate inequality.
- There appears to be a dichotomy between schools and organisations who have already been through or are part way through a "systems shift" in literacy and numeracy practice in their schools, and those pilot participants who appear to have adopted a "wait and see" approach before implementing any changes. How extensive any change or challenge to existing practice in this second group of schools will remain to be seen. More detailed questions in the second literacy and numeracy survey and in the focus group interviews may garner more understanding about this issue.

- Teacher apprehension about whether students are "ready" for the Common Assessment Activities was a common theme in the teacher survey. Much of this anxiety can be explained by teacher reactions to a very new and different way of assessing the literacy and numeracy levels of their students and to uncertainty about the levels required for students to pass or fail the Common Assessment Activities. Subsequently, they found the readiness of students to sit the assessments difficult to gauge. With focused professional development and support, teacher knowledge and practice will improve, and the readiness of students will be more readily identifiable.
- The co-requisite nature of the Te Reo Matatini me te Pāngarau | Literacy and Numeracy standards poses a barrier to priority learners being able to achieve an NCEA qualification. It is, therefore, important to consider what processes, programmes, and interventions will be required to remove this roadblock. This could include timely and adequate learning support, academic mentoring, and careful monitoring of student progress towards achieving the co-requisite NCEA standards. Such approaches will require schools to allocate resource and put in place effective processes that involve students, teachers, and family/whānau.

Summary of recommendations

Objective one: understand how the design of the assessments can be further refined and improved

1. The rates of those achieving the standard in Literacy (writing) is lower compared to Literacy (reading) and Numeracy, even if they have achieved similar levels in e-asTTle. For example, those students who had scored 5A in e-asTTle had a 98% achievement rate in Literacy (reading), 96% achievement rate in Numeracy, and only a 77% achievement rate in Literacy (writing). This warrants further analysis if similar rates of achievement are wanted across all three standards.
2. Further clarity is required regarding expected achievement rates for each of the assessments in order to determine the appropriateness of the readiness indicators used in the pilot.
3. Many survey respondents criticised the Numeracy CAA because of the amount of reading and writing involved. It will be important to ensure that the teaching and learning resources developed appropriately prepare students for these types of questions.
4. The breakdown of achievement by ethnicity again highlights the wider inequity of the education system and the importance of maintaining, and potentially increasing, the wider capability building and support to remedy this. NZQA will continue to make further improvement to the assessments, however there is a limit to what the assessments can do without wider change occurring.
5. Understanding more about the significant differences in achievement for low decile schools and tertiary/alternative education organisations is important. The second phase of this evaluation will seek the available e-asTTle data for these sub-groups, and further analysis will better determine whether the variation in results are reflective of the already known differences in achievement or if the CAA design is inequitable.

Objective two: understand how kura, schools, and organisations are building their capability and what is needed going forward

Student readiness

6. Further understanding of student readiness for the CAAs is required. Many teachers stated they did not understand the student level required, and some schools indicated they used the pilot to understand the level of readiness needed rather than use selection methods such as data from the recommended readiness tools.
7. School, kura, and organisations could benefit from more focused PLD about the curriculum levels of the readiness indicators and the tools that can help them indicate student and ākonga readiness. Webinars and in-school PLD could support teachers' understanding and practice in this area.
8. Some pilot teachers asked for specific support in unpacking the Big Ideas and Significant Learning areas that support the Literacy and Numeracy standards.

9. More availability of resources and activities set at the CAAs levels of difficulty may also support teacher confidence and understanding of the levels required. To prevent teaching to the test, it would be necessary to identify activities and tasks that could broaden units of work and teaching and learning programmes (versus one-off activities aimed to "teach to the test").

10. Many teachers indicated they would like the return of student papers as it would give them the opportunity to understand the marking of the papers and the standard that students need to reach to pass the exam.

Tools and Resources

11. A range of case studies showcasing effective literacy and numeracy practice for students and tips for managing the CAA processes effectively would provide the opportunities for schools and other organisations to see and understand how the Literacy and Numeracy Standards and Common Assessment Activities are best applied in a variety of settings. These could be in the form of webinars, written examples, and materials.

12. The development and sharing of different literacy and numeracy approaches should be encouraged. The setting up of spaces in which to collaborate and discuss resources and approaches, plus brainstorm solutions, may increase confidence levels and feelings of support. The use of online hubs and professional learning groups are good ways to encourage this and might improve students' levels of preparedness for the assessments. Likewise, the use of online hubs and professional learning groups are good ways to create effective professional practice.

13. Illustrating different ways of incorporating activities into units of work would prevent "teaching to the test." A cross-curricula approach to incorporating literacy and numeracy skills will promote rich curriculum experiences and also possibly preventing the use of ability groupings and the narrowing of the curriculum.

14. Further resources to support the preparation and understanding of the levels required would make teachers feel more supported. This could include marking schedules, exemplars and task activities, and question banks. This will require careful design and messaging so that these are used in ways that support the overall purpose of the standards.

15. Regarding the option for portfolio assessment for ākonga in kura and secondary schools, NZQA has already begun developing more specific resources to assist kaiako and ākonga to understand how learning can best be captured and presented using portfolios. Feedback from assessment experts notes the additional complexity to undertaking portfolios and the time this requires to grow and build capability, so this will be an area of ongoing development to monitor closely.

16. For kura and organisations delivering Te Reo Matatini and/or Pāngarau programmes, further support, tools, and guidance are required to set these assessments within the context of their marau ā-kura, the redesign of Te Marautanga o Aotearoa (Te Tīrewa Marautanga), and the development and implementation of Hei Raukura Mō Te Mokopuna.

Schools/organisations teaching and learning

17. Teachers play a vital role in supporting student success, and it is important to reinforce this. This role includes students being engaged in effective teaching and learning and being well prepared for the CAAs.

18. Positive experiences should be highlighted within resources such as case studies. Despite this being early in the process for most, some teachers have commented on increased confidence and student engagement. Engaging students only when they have a good chance of achieving the standard should be a key message.

19. Encourage schools to consider what specific actions are needed to support student learning and to prepare them in ways that gives them a good chance to achieve the standards.

20. Webinars and case studies could provide a useful framework for schools/organisations implementing the standards in 2023 and 2024; these should highlight the need for schools to prepare students for the assessment events.

21. Deeper insight into the experiences and challenges of certain groups – such as Realm country schools, tertiary and alternative education institutions, English language learners and those requiring special assessment conditions – is recommended. This insight will reveal any specific issues or supports that they may require.

Strengthening Implementation

22. For further assessment events, schools will need to consider logistics carefully. Aspects such as technical issues, resourcing, workload, budget, and staffing were identified as issues in their schools. Increased support for school administration and examples of effective administrative practice would be beneficial for all schools.

23. Leaders and teachers must understand the purposes of the pilot and their role in preparing students for the Common Assessment Activities. Several participants did little to prepare their students or were using the assessments as a way of identifying students' strengths and weaknesses before changing teaching and learning programmes.

24. Some participants indicated there was a requirement to administer the assessment at the same time for all students, and that it made logistics difficult. Some consideration could be given to allow the sitting of Common Assessment Activities in sessions, so adequate classrooms and staffing are available.

25. Embedding the NCEA Te Reo Matatini and Pāngarau standards more firmly within this vision of Hei Raukura Mō Te Mokopuna will create greater cohesion with the overarching focus of Te Tamaiti Hei Raukura and Te Tīrewa Marautanga. This would support a necessary shift in thinking to consider more broadly the definition of what ākonga achievement and progress looks like, sounds like, and feels like, particularly within the local context of each kura and school.

SECTION A: Introduction

Evaluation Associates | Te Huinga Kākākura Mātauranga conducted this evaluation for the 2022 pilot of the NCEA Te Reo Matatini me te Pāngarau | Literacy and Numeracy standards. These new standards, part of the wider NCEA Change Programme, aim to ensure that ākonga/students[2] with an NCEA qualification will have a good level of foundational literacy and numeracy and te reo matatini and pāngarau. The achievement of these standards will be a co-requisite for ākonga/students to be awarded an NCEA qualification.

The pilot of the NCEA Te Reo Matatini me te Pāngarau | Literacy and Numeracy standards began in 2021 with the purpose of understanding whether the standards – and methods of assessing these – are appropriate and whether the participants were ready for change. 2022, the second phase of the pilot, is on a much larger scale. The findings from this evaluation will inform a transitional year in 2023, which will lead to the full implementation nationally of Te Reo Matatini me te Pāngarau | Literacy and Numeracy standards in 2024.

This evaluation report, the first of two for the 2022 pilot-year, follows the first assessment event held in June/July. The second report will follow the September assessment event and is due for completion in December 2022.

1 Pilot of Te Reo Matatini me te Pāngarau | Literacy and Numeracy standards

1.1 2021 Pilot

2021 was the first year of a two-phase pilot seeking to understand whether the five new standards and methods of assessing these new standards were appropriate and whether participants were ready for change. NZQA and the Ministry of Education, with involvement from Massey University, conducted the 2021 pilot, which was small scale. The pilot included 13 secondary schools, six kura and two tertiary providers. COVID-19 affected the number of entries received from these participants, and only 2,313 ākonga/students from the 5,810 entries submitted or completed assessments. The 2021 pilot involved a qualitative and quantitative analysis and made several recommendations that have informed the 2022 pilot.

1.2 2022 Pilot

Phase two (2022) of the pilot is on a much larger scale and provides greater detail, critical for quality assurance purposes ahead of a transition year in 2023 and the full implementation of Te Reo Matatini me te Pāngarau | Literacy and Numeracy standards in 2024. Evaluation Associates | Te Huinga Kākākura Mātauranga will carry out the evaluation of the 2022 pilot.

[2] Throughout this report, students from English-medium school/organisations are called learners or students and those from Māori-medium schools/organisations are called ākonga. When referring to both groups of students, they are called ākonga/students.

Piloting the Literacy and Numeracy standards involves 198 New Zealand schools, seven tertiary/alternative education organisations, and seven schools from Realm countries. Ten Māori-medium kura and three secondary schools providing te reo Māori education are involved in piloting the Te Reo Matatini me te Pāngarau standards. Ākonga/students from these schools, kura, and organisations participated in two assessment events.

Ākonga/Students were selected to participate in two assessment events, and these were held in June/July and September. This first report presents the findings following the first assessment event held in June/July.

Pilot participants

The pilot schools, kura, and organisations are diverse, provide national coverage, and represent different timetable structures and teaching and learning approaches. The selection of schools, kura, and organisations entered reflect the diversity of NCEA learning environments throughout Aotearoa and support and understanding of how the new requirements work in practice in a variety of contexts. Not all the schools, kura, or organisations who participated in the pilot entered ākonga/students in the common assessment event in July, as shown in Table 1.

Table 1: Numbers of pilot schools, kura, and organisations who entered students in the June/July 2022 Common Assessment Activity

Type of school/organisation	Number participating in pilot	Number who entered students in one (or more) CAA for event one
English-medium schools	162	140
Māori-medium kura and schools providing te reo Māori education	10 (kura) 3 (secondary schools)	4
Tertiary/Alternative Education[3]	7	4
Realm country schools[4]	7	6
TOTAL	162	154

Table 2 shows the numbers of ākonga/students who participated in the June/ July 2022 Common Assessment Activities for Te Reo Matatini me te Pāngarau | Literacy and Numeracy.

[3] The organisations categorised as "tertiary/alternative education" include five tertiary providers, and two alternative education providers. One of the alternative education providers works with adult students as well as rangatahi.

[4] The Realm countries are Niue (1 school) and the Cook Islands (6 schools).

Table 2: Number of participating ākonga/students in CAA by assessment standard (June/July 2022)

Subject	Provider Type	Students from Realm countries	All participating students
Reading	Secondary	234	9,346
Writing	Secondary	235	8,822
Numeracy	Secondary	240	13,402
Pāngarau	Secondary	-	95
Te Reo Matatini	Secondary	-	38
Reading	Tertiary	-	40
Writing	Tertiary	-	33
Numeracy	Tertiary	-	39

Table 3 details the combinations of standards that participating ākonga/students were engaged in. Just over a third of the participating students were assessed in Literacy (reading), Literacy (writing), and Numeracy standards. Nearly 40% of participating ākonga were assessed in both of the Te Reo Matatini me te Pāngarau standards.

Table 3: Proportion of students participating in CAA by assessment standard combination type (July 2022)

Combination of Standards	Number	Proportion of Participants by Medium
Reading and Writing	1,791	11.0%
Reading and Numeracy	730	4.5%
Writing and Numeracy	429	2.6%
Reading, Writing, and Numeracy	6,235	38.3%
Reading only	630	3.9%
Writing only	400	2.5%
Numeracy only	6,047	37.2%
Pāngarau and Te Reo Matatini	37	38.5%
Pāngarau only	58	60.4%
Te Reo Matatini only	1	1.0%

2 Evaluation Approach

2.1 Scope

This evaluation measures the extent to which the objectives of the 2022 pilot of NCEA Te Reo Matatini me te Pāngarau | Literacy and Numeracy standards and assessments have been achieved. The evaluation aims to do the following.

1. Understand how the design of the assessments can be further refined and improved.
2. Understand how schools, kura, and organisations are building their capability and what is needed going forward.

The two phases of the evaluation are structured around the two common assessment events the pilot schools participated in. The first phase of data collection and analysis followed the common assessment event held in July, and the evaluation of this data is presented here, in this report – Report one: 2022 NCEA Literacy and Numeracy | Te Reo Matatini me te Pāngarau Pilot Evaluation. It is important to note that analysis or critique of the assessment tasks is not within the scope of this evaluation.

2.2 Evaluation methods

This evaluation uses a mixed-methods approach, which involves collecting, analysing, and mixing both quantitative and qualitative data. This approach provides a comprehensive account of what is happening.

Quantitative methods:

- Analysis of the NCEA assessment results by year level, gender, decile, ethnicity, ELL status, and type of organisation.
- Psychometric analysis (completed by NZCER)
- Analysis of relationship between e-asTTle and NCEA results.
- Analysis of quantitative survey data (teacher survey and learner survey).

Qualitative methods:

- Thematic analysis of the qualitative survey data (teacher survey and learner survey).
- Thematic analysis of the information gathered from focus groups[5] and key informant interviews (for Report 2).
- Thematic analysis of data and information gathered about NCEA Te Reo Matatini me te Pāngarau by Tai Huki Consult Ltd (via a survey and direct feedback, guidance from NZQA markers about the assessments conducted in July, and summary discussions from hui about Te Reo Matatini me te Pāngarau.

2.3 Sources of data

Literacy and numeracy teacher survey

All English-medium schools/organisations who participated in the July assessment event for the 2022 pilot received the Literacy and Numeracy Teacher Survey. The survey gathers their perceptions and experiences from either a lead teacher, numeracy/literacy co-ordinator, or principal's nominee in each of the pilot schools/organisations. The table below details the survey respondents.

Table 4: Type and numbers of respondents to the mid-year Literacy and Numeracy teacher survey

Type of school / organisation	Numbers that participated in assessment event one	Numbers of survey responses
English-medium schools	140	141
Tertiary/Alternative Ed	4	5
Realm country schools	6	5
TOTAL	150	153

[5] Focus groups and interviews will be carried out after the second assessment event, and this data will be discussed in the second evaluation report.

Although the overall number of responses to the New Zealand Schools survey was sufficient to draw conclusions, the small number of responses from the other organisations (tertiary/alternative education and Realm countries) means any conclusions drawn will have limited generalisability.

The survey, customised and collated using a Teams survey form, contained both quantitative and qualitative questions. The Ministry of Education reviewed these questions. Of the 153 schools/organisations who responded to the survey, 76 (49%) completed the numeracy survey questions, 44 (29%) completed the reading survey questions, and 33 (22%) completed the writing questions.

The survey questions focused these aspects of the pilot:

- the selection process used for students who participated in the Common Assessment Activities (CAAs)
- teacher perceptions about the CAAs
- teacher perceptions about the equitability of the CAAs
- views on supporting resources
- impact of the new standards on systems or programmes in the school/organisation
- how the implementation of the Literacy and Numeracy standards could be strengthened.

Data from quantitative questions (agree/disagree Likert scales) was analysed and presented in bar graphs. Thematic analysis was used to analyse the qualitative data.

Limitations of the survey

- Because most teachers had not been able to view the CAA, their responses were limited.
- The survey responses sit within the context of pilot participants not having seen the assessment results.
- The names of the schools/organisations were not recorded in the survey.
- Small numbers of responses from tertiary institutions, alternative education providers, and Realm countries meant that their data cannot be considered representative. However, it does provide insights into their unique context and perspective.

Learner survey

The NZQA developed the learner survey and administered it following each of the July common assessment events. The survey, optional for students to complete, involved giving an agreement rating (strongly agree to strongly disagree) for four statements about the assessment, and also, if relevant, about the text-to-speech function. There was also an opportunity to submit general comments.

The response rate for the learner survey can be seen as statistically representative of the wider group of students who participated in the mid-year CAA. This was Literacy (reading) 32% (3050), Literacy (writing) 31% (2756), and Numeracy 23% (3087).

Results data from the Common Assessment Activities (Literacy, Numeracy)

NZQA provided the results data to Evaluation Associates | Te Huinga Kākākura Mātauranga in tables that reported the assessment results for each of the five pilot standards by year level, gender, decile, ethnicity, ELL status, and type of organisation.

NZQA also provided data that allowed for analysis of the relationship between e-asTTle and NCEA results.

New Zealand Council for Educational Research (NZCER) conducted a psychometric analysis for each CAA, and this was also provided to Evaluation Associates.

Focus groups and interviews

In phase two of the evaluation, focus groups and interviews will give evaluators deeper insight about:

- the experiences of piloting the standards
- to what extent the sector is ready for full implementation in 2024
- what future support the sector requires.

This information will be analysed in the second evaluation report.

Data approach for Te Reo Matatini me te Pāngarau

Given the smaller number of pilot Māori-medium kura and secondary schools providing te reo Māori education, it was agreed that a different approach was required to gather data and information about the Te Reo Matatini me te Pāngarau standards. This would give recognition to kaupapa Māori philosophies, as outlined by Tai Huki Consult Ltd[6], as well as the general feedback about ākonga and kaiako wellbeing.

Consequently, data has been drawn from an online survey and direct discussions with some pilot participants Tai Huki Consult Ltd conducted. Data also arose from feedback gathered at recent hui about Te Reo Matatini me te Pāngarau, involving multiple stakeholders in the development of PLD and other supports for kura and secondary schools using these standards. NZQA also provided a report which summarised the feedback from the group of markers for the Te Reo Matatini and Pāngarau assessments. A further, expanded online survey will be conducted following the second assessment event in September.

2.4 Limitations of the evaluation

- The completion rate for the learner survey (Reading 32%, Writing 31%, and Numeracy 23%) suggests a representative sample; however, because the completion of the learner survey was optional, it is unclear if the individuals who chose to complete the survey adequately represent the full cohort of pilot participants.
- The response rate for the learner survey was slightly higher for those students who achieved the standard than for those who did not.
- This pilot of the te reo matatini and pāngarau standards involves only a small number of kura and secondary schools, which makes the size of the sample statistically unrepresentative. While this is the case from a statistical perspective, the findings at this point do reflect and support the recommendations for the Māori-medium and kaupapa-Māori sector outlined in ERO's evaluation of Te Reo Matatini me te Pāngarau from 2021.

[6] The link to the full report by Tai Huki Consult Ltd is in Appendix 3.

- There are small numbers of tertiary institutions, alternative education providers, and Realm countries involved in the pilot, and even fewer who took part in the survey. Careful consideration of the generalisability of findings in the report is needed.
- Some of the questions in the teacher survey were based on an assumption that the teachers in the pilot would have access to the content of the Common Assessment Activity (CAA). While some teachers "looked over students' shoulders", the level of teacher knowledge about the CAA questions asked is minimal.

2.5 Report structure

This report has the following three sections.

1. Section A describes the background and the pilot for the Te Reo Matatini me te Pāngarau | Literacy and Numeracy standards. It also details the evaluation purpose and approach.
2. Section B focuses on objective one for the evaluation, which is to understand how the design of the assessments can be further refined and improved. The first part of this section reports how well the assessments are performing. The second part compares the results for different groups of students and reports teacher perceptions about the how equitable the assessments are.
3. Section C focuses on objective two, which is to understand how schools are building their capability and what is needed going forward. The first part of this section looks at the use of literacy and numeracy tools and resources to support teaching and learning. The second part describes the impacts from the implementation of Te Reo Matatini me te Pāngarau | Literacy and Numeracy standards that various kura and schools/organisations reported. The final part of this section discusses the pilot from an implementation perspective and identifies further opportunities for NZQA and the Ministry of Education to effectively implement the standards nationally in 2024.

Observations and recommendations are included at the end of each section. A full list of these appears after the executive summary. A list of figures and tables in the report is included in the appendices.

SECTION B: Findings for objective one

Objective One: Understand how the design of the assessments can be further refined and improved.

The 2022 pilot provides further information and insights about the Common Assessment Activities and to what extent they are fit for purpose. Two aspects of being "fit for purpose" are discussed. First, a consideration of how well the assessments are performing in terms of levelling and alignment with other measures of attainment. Second, the equitability of the design of the assessments and results.

3 How well are the assessments performing?

This section, which builds on the analysis that was carried out on the results data from the 2021 pilot, aims to further identify whether the assessments are set at the right level, whether the assessments are assessing the standards accurately, and what changes might be recommended for 2023.

The 2021 evaluation concluded that the assessments (CAA) generally performed well against the standards. This was demonstrated through initial testing of the assessment tasks, comparing assessment results with expected attainment rates (as determined by e-asTTle, teacher judgement, national benchmarks), and psychometric analysis from NZCER. The 2021 evaluation report also noted that the patterns around student results in the CAA broadly match those seen in wider studies, for example, the National Monitoring Study of Student Achievement[7] (NMSSA).

In this section of the report, the 2022 results data is analysed and reported in a similar way to the 2021 evaluation[8]. Firstly, the overall results for each standard are presented; results are looked at by year level and are also compared with students' e-asTTle scores. Then, a summary of the results and recommendations from the psychometric analysis is provided. Finally, teacher perceptions of how well the standards are performing are discussed using the mid-year survey data.

3.1 Overall results

The results of the July 2022 assessments for Te Reo Matatini me te Pāngarau | Literacy and Numeracy pilot are reported in Table 5 with the numbers participating and achievement for each of the five standards assessed in the pilot. For comparative purposes, the 2021 student numbers and percentage of achieved students are also included. The 2022 results represent a total 16,368 ākonga/students who participated in one or more of the July assessments.

[7] https://nmssa.otago.ac.nz/

[8] As the 2021 evaluation validated the content of the standards, this has not been examined in the 2022 evaluation.

Table 5: Achievement by standard (July 2021 and July 2022 results)

Standard	2022			2021	
	Participating Students	Achieved Students	Achieved Students	Participating Students	Achieved Students
	(n)	(n)	(%)	(n)	(%)
Reading	9,386	6,016	64%	609	67%
Writing	8,855	3,029	34%	575	35%
Numeracy	13,441	7,512	56%	1070	65%
Pāngarau	95	17	18%	29	28%
Te Reo Matatini	38	9	24%	30	83%

Overall results for the Common Assessment Activities

- The achievement of the CAA for each assessment area varied. Literacy (reading) had the highest rate of achievement, followed by Numeracy, and then Literacy (writing). With only 34% of students achieving the Literacy (writing) assessment, nearly a third more students in the pilot achieved the Literacy (reading) standard than the numbers who achieved the Literacy (writing) standard.
- In understanding these results, it is important to note that the analysis of available e-asTTle data for participating students indicates that a sizeable proportion of students who had e-asTTle scores below the minimum recommended curriculum level (NZC 4A) were entered for one or more of the CAAs. For Literacy (reading) the proportion was 16.6%, for writing 26.6%, for Numeracy 21.1%. In section 3.3, Figure 2 shows the significant increase in achievement when the data is cut to include only those students who had a recent e-asTTle score and met the recommended readiness criteria.
- Recommendations made in the evaluation of the 2021 pilot included to monitor results of all students participating in writing assessments and to provide appropriate support for the teaching and learning of writing. With performance again being low for the Literacy (writing) standard, this recommendation remains relevant. Due to the small sample size of Māori-medium pilot participants, the results for Te Reo Matatini and Pāngarau cannot be seen as statistically representative[9]. The achievement of Te Reo Matatini and Pāngarau showed lower rates for Pāngarau.
- Due to the range of events and disruptions that kura reported around the time of the first assessment event, some kaiako had considered waiting until the second assessment event in September 2022. An increase in ākonga participating in Te Reo Matatini and Pāngarau assessment is expected for the next assessment event.
- There are two assessment options available for Te Reo Matatini me te Pāngarau, that is, ākonga can choose to demonstrate their achievement of the standards using a CAA or submitting a portfolio. In the first assessment event of 2022, all ākonga used the CAA. There has been a positive response from kura and secondary schools about the opportunity for ākonga to use portfolio, and while no portfolios were submitted in the first assessment event, the Ministry and NZQA remain committed to developing resources to support the use of portfolios.

[9] With approximately 3,300 secondary-aged ākonga in Māori-medium education, a representative sample (margin of error of plus or minus 5 percent at a 95% level of confidence) would be 345 ākonga.

Comparison of overall results between 2021 and 2022 pilot data

- The patterns of achievement for the 2021 and 2022 pilots of the Literacy and Numeracy standards are very similar, with Literacy (reading) being the standard most achieved, followed by Numeracy, and then Literacy (writing).
- The rates of achievement for Writing and Reading were very similar in both pilot years. There was greater variation in the Numeracy standard, with the 2022 achievement rates being nearly 10 percentage points less than the 2021 rate.
- While there was a small increase in the number of ākonga (30 to 38) who participated in Te Reo Matatini between 2021 and 2022, the results show a significant decrease in achievement from 83% to 24%. Furthermore, the number of ākonga who participated in the Pāngarau assessment increased from 29 to 95; however, the achievement rates decreased from 28% to 18%.
- As noted in the 2021 pilot evaluation report, the small sample size and the absence of national measures for these two standards make it difficult to analyse suitably the results data. While the numbers of results for Te Reo Matatini and Pāngarau are low and not representative, investigating and understanding the reasons for this level of achievement in the pilot is important.
- The decrease in achievement in Te Reo Matatini and Pāngarau could be attributed to different factors, which will be tested further in the second 2022 evaluation report. These factors raised by kaimahi from Māori immersion kura include, but are not limited to, the following.
    - The overlapping of other significant events for ākonga in kura and secondary schools with the timeframe for the first assessment event.
    - The load and pressure on the smaller teaching workforce in Māori-medium settings and the impact on capacity and capability for this change process.
    - Responding to the impacts of the pandemic, including the need to place greater emphasis at this time on the wellbeing of ākonga, their whānau, and kaiako.
    - Variability within and for kura in terms of internal capacity and capability to fully engage in preparing and administering the assessments, access to external support and guidance, the provision of quality tools and resources to support ākonga and kaiako.
    - Differences in the cohort sizes, or in the decision-making processes about which ākonga would participate, or in how best to prepare ākonga for these assessment events.
- Since the approach used to determine which students would participate in Common Assessment Activities was different between the 2021 and 2022 pilots, some differences were expected in the results. In the 2021 pilot, full cohorts of students were entered, whether or not teachers thought they were ready. In the 2022 pilot, the intent was for student selection to instead be based on readiness (this is discussed further in section 5). The similarity in the 2021/2022 results may indicate that the two cohorts were not that different and that, potentially, the guidance around readiness was neither well understood nor applied by schools/organisations.

3.2 Achievement by year level

Table 6 below shows achievement by standard and year level. Year 10 students are the largest group of participating ākonga/students (81.9%). Only small numbers of year 12, 13, and 14 students participated in the assessments.

Table 6: Achievement by standard and year level for secondary students

Year Level	Standard Name	Participating Students (n)	Achieved Students (n)	Achieved Students (%)
9	Reading	472	294	62.3%
10	Reading	8,094	5,208	64.3%
11	Reading	705	469	66.5%
12	Reading	44	13	29.5%
13	Reading	25	10	40.0%
14	Reading	1	0	0.0%
9	Writing	369	128	34.7%
10	Writing	7,723	2,692	34.9%
11	Writing	665	197	29.6%
12	Writing	37	4	10.8%
13	Writing	25	3	12.0%
14	Writing	1	0	0.0%
9	Numeracy	737	413	56.0%
10	Numeracy	11,088	6,278	56.6%
11	Numeracy	1,489	789	53.0%
12	Numeracy	59	11	18.6%
13	Numeracy	29	8	27.6%
14	Numeracy	1	0	0.0%
9	Pāngarau	21	1	4.8%
10	Pāngarau	40	11	27.5%
11	Pāngarau	34	5	14.7%

Year Level	Standard Name	Participating Students (n)	Achieved Students (n)	Achieved Students (%)
9	Te Reo Matatini	25	10	40.0%
10	Te Reo Matatini	25	3	12.0%
11	Te Reo Matatini	29	8	27.6%

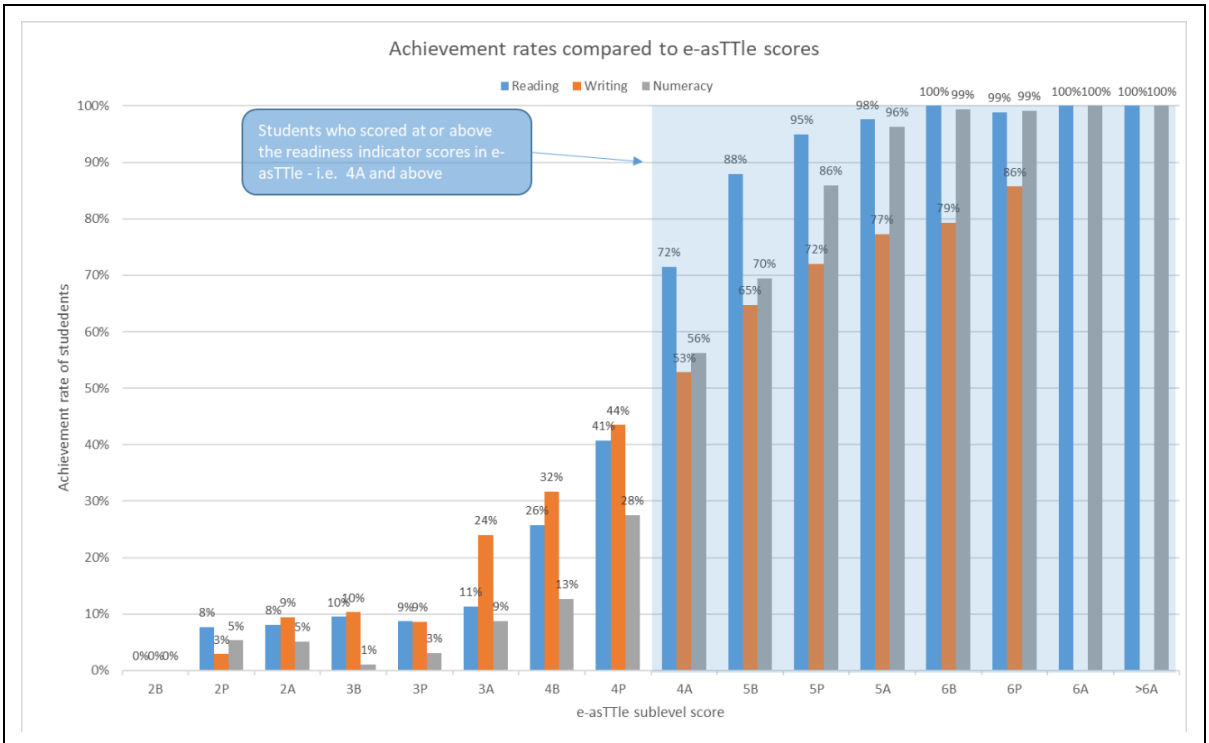
- The Literacy (reading) achievement was similar for Years 9, 10, and 11 but lower for Year 12, 13, and 14. The Writing achievement was similar for Years 9 and 10 but lower for Year 11 and much lower for Years 12, 13, and 14. The Numeracy achievement was similar for Years 9 and 10, lower for Year 11, and much lower for Years 12,13, and 14.
- Information from schools about how students were selected for the assessment activity could explain the relatively high performance of Year 9 students. For example, some schools entered their Year 9 extension classes, and some schools used the student's e-asTTle level rather than year level.
- The relatively low performance of Year 11, 12, and 13 students is also likely related to the reasons these students were selected. Many of the Year 11, 12, and 13 students entered may not have engaged in other ways of gaining their literacy or numeracy credits, so it was seen as an opportunity for them to achieve these standards.
- Pāngarau achievement was lowest in Year 9 (1 out of 21 ākonga achieved) and highest in Year 10 (11 out of 40). Conversely, Te Reo Matatini achievement was lowest in Year 10 (3 out of 25) and highest in Year 9 (10 out of 25).

3.3 Attainment rates compared with e-asTTle scores

NZQA provided comparative data for those students who had a recent[10] e-asTTle result in either Reading, Writing, or Mathematics. Over half of the students participating in the pilot had recent e-asTTle scores that could be used for this analysis. e-asTTle[11] has been identified as one of the assessment tools that teachers can use to help identify those students who are likely ready to attempt an assessment for the Literacy and Numeracy standards. The minimum e-asTTle level NZQA has recommended to indicate readiness for the CAA is 4A, meaning high or late level 4 of the New Zealand Curriculum.

Due to small numbers and lack of available e-asTTle data, an e-asTTle analysis was not undertaken for ākonga participating in Te Reo Matatini me te Pāngarau.

Figure 1: Literacy and numeracy standards achievement rate compared to e-asTTle scores

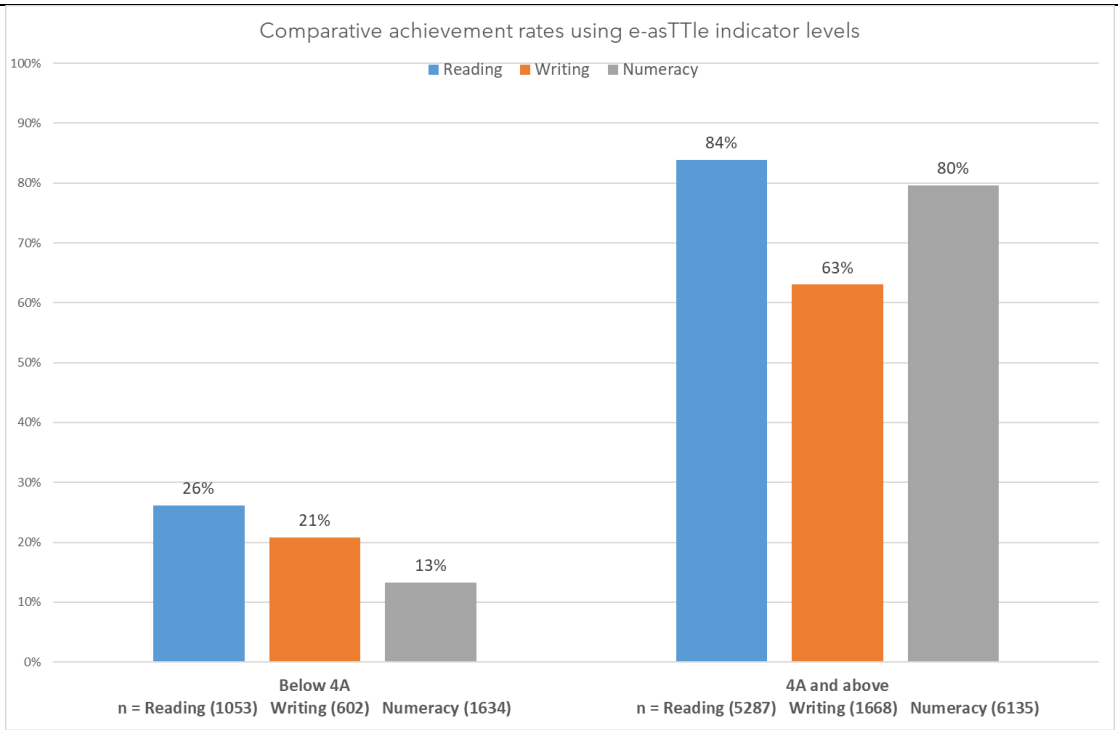


This chart shows the relationship between the achievement rates for the Literacy and Numeracy standards and student e-asTTle scores in Reading, Writing, and Mathematics. For each e-asTTle sublevel score, the bar shows the percentage of students who achieved the CAA standard. 72% of students who scored 4A in Reading achieved the Literacy (reading) standard, 53% of students at 4A for Writing achieved the Literacy (writing) standard, and 56% of students at 4A for mathematics achieved the Numeracy standard.

[10] Recent was defined as January 2022 or later, but for those not recent, they needed to be above NZC 4A.

[11] The analysis of e-asTTle data was carried out because it is data held by NZQA. Other assessment tools were also used by schools to determine readiness, but this data is not accessible for the required analysis. e-asTTle provides an indication of the level of the curriculum students are working at for reading, writing, mathematics, and in pānui, te pāngarau, and tuhituhi.

Figure 2: Literacy and numeracy achievement rates for students below and above readiness indicators



This chart shows the achievement rates for those students either below or at and above the e-asTTle sublevel (4A or above), which is being used as a readiness indicator for these standards.

- Comparing the e-asTTle results with achievement rates reveals a strong relationship between the two assessments for all three assessment areas, but particularly for Literacy (reading) and numeracy.
- Students scoring 4A or above had an 84% chance of achieving the standard for Literacy (reading), an 80% chance for Numeracy and a 63% chance for Literacy (writing) being over 70%. Once students scored 5A or above, it appears there is little chance of not achieving the assessment standard with achievement rates of over 95%. Literacy (writing) shows a weaker relationship to e-asTTle scores. While scoring 4A or above still produced over 50% achievement rates, higher e-asTTle scores did not ensure achievement to the same levels that Reading and Maths showed. In fact, 6P scores were the only e-asTTle scores that produced greater than 80% achievement rates. This indicates that even our best writers are not certain of achieving the standard in Literacy (writing).
- Possible reasons for this difference could be that the Literacy (writing) assessment activity places greater emphasis on meeting specific criteria in a piece of writing, rather than just on the quality of writing. As teachers become more familiar with the types of writing tasks being set and teach the skills needed to meet the requirements of the standard, it is likely that students' performance will improve.
- Overall the achievement results show that e-asTTle can be used as an appropriate indicator of readiness for sitting the Literacy and Numeracy standards.

- The recommended readiness indicator of high level 4 and above is dependent on the expectation of the national achievement rates when sitting the CAAs. It seems to be appropriate if seeking approximately 60% to 80% of students achieving the standard. If a higher pass rate is intended, where more than 80% of students achieve the standards on their first attempt, then a higher readiness level should be recommended.
- While the results provide good insights, they should also be read with caution as it cannot be assumed that in all cases the e-asTTle score is a full and accurate measure of the curriculum level, particularly in writing which is teacher marked.

3.4 Psychometric analysis for literacy and numeracy assessments

NZCER conducted statistical analyses on the July assessments for the three English-medium standards: (Literacy (reading), Literacy (writing), and Numeracy. This analysis included applying Item Response Theory (IRT), which provides an understanding of how each assessment is performing from a statistical/psychometric viewpoint.

Two questions are investigated:

- Was the difficulty of the items appropriate for the intent of the standard?
- Did each item contribute information to the measurement of student achievement?

Psychometric analyses have not yet been completed for the Te Reo Matatini and Pāngarau assessments. The small numbers participating limit the usefulness of a statistical methodology; however, NZQA is anticipating that these will be produced for the NCEA external assessments in November. Consideration is being given to the type of analysis that will be useful to inform the development of these assessments going forward.

A summary of the psychometric report from NZCER is in Appendix 2; however, the indication is that the items in the Literacy (reading), and Literacy (writing), and Numeracy assessments performed well This finding is consistent with the psychometric analysis on the assessments used in the 2021 pilot, which also pointed to the appropriateness of these assessments.

The specific recommendations for each assessment are in the full report, but an overarching recommendation is to consider the impact of measurement error when determining what score students are expected to score, overall, and on each separate outcome. Measurement error captures how precise the assessment instrument is in locating achievement on a scale.

3.5 Teacher perceptions of how well the Literacy and Numeracy standards are performing

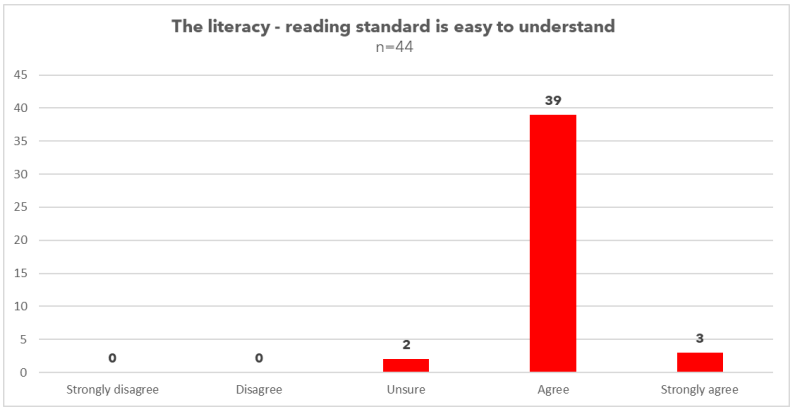
As part of the mid-year Literacy and Numeracy Survey, teachers were asked to rate the extent they agreed/disagreed with two statements.

- That the standard is easy to understand.
- That the Common Assessment Activity (CAA) allowed for valid and reliable assessment.

The rating used a five-point scale, from strongly disagree, disagree, unsure, agree, to strongly agree. For each of the three assessment areas, the responses to these statements are shown in the charts that follow.

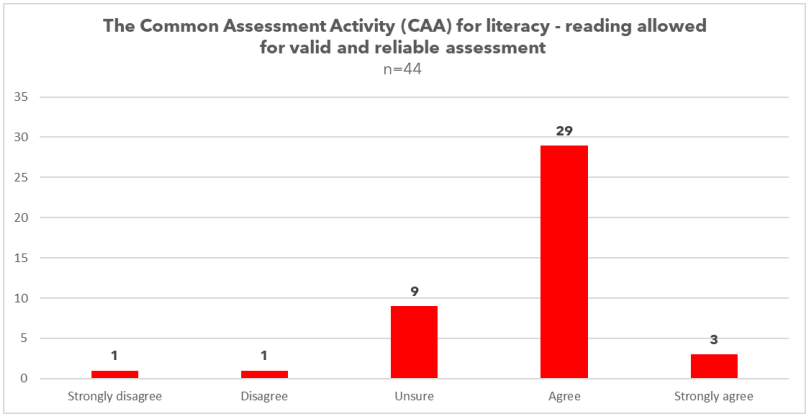
Literacy (reading) standard

Figure 3: Agreement ratings for "the literacy (reading) standard is easy to understand"



An overwhelming 96% (42) out of the 44 pilot participants surveyed agreed or strongly agreed that the Literacy (reading) standard was easy to understand. While none disagreed or strongly disagreed with the statement about the Literacy (reading) standard, two pilot participants were unsure whether they agreed or not with the statement about the Literacy (reading) standard.

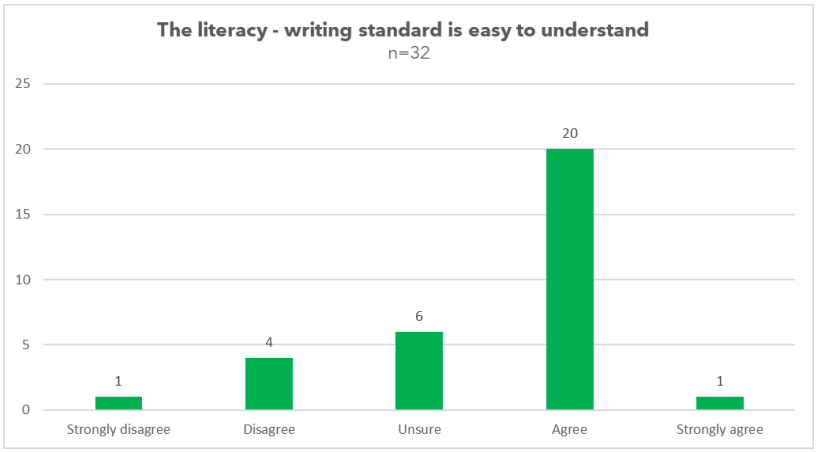
Figure 4: Agreement ratings for "the CAA for literacy (reading) allowed for valid and reliable assessment"



77% (36) of the 44 participants who responded to the statement that the Literacy (reading) CAA was valid and reliable assessment strongly agreed or agreed with this statement, while 23% (9) were unsure about this statement. Only two schools/organisations disagreed/strongly disagreed with the statement.

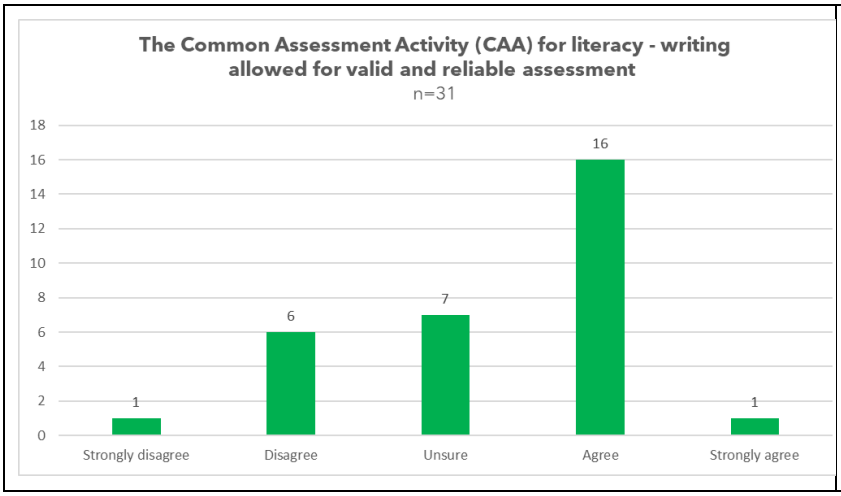
Literacy (writing) standard

Figure 5: Agreement ratings for "the literacy (writing) standard is easy to understand"



66% (21) of the 32 pilot participants interviewed in the survey agreed or strongly agreed that the Literacy (writing) standard is easy to understand, while 18% (6) were unsure about the statement. 16% (5) disagreed or strongly disagreed with the statement.

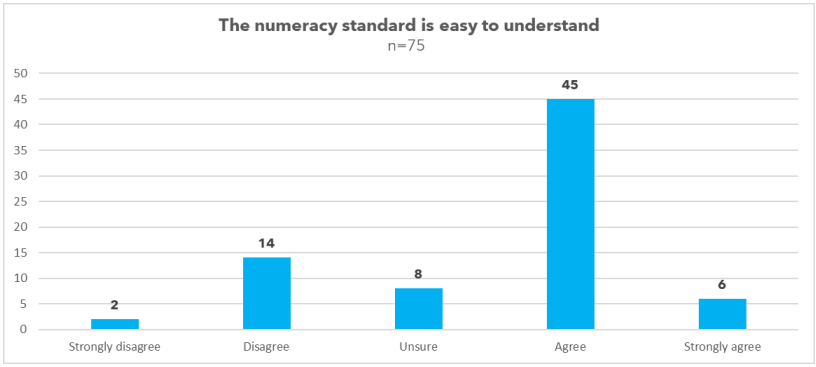
Figure 6: Agreement ratings for "the CAA for literacy (writing) allowed for valid and reliable assessment"



55% (17) of the 31 participants who responded to the statement that the Literacy (reading) CAA was valid and reliable assessment strongly agreed or agreed with this statement, while 22.5% (7) were unsure about this statement. 22.5% (7) participants disagreed or strongly disagreed with the statement.

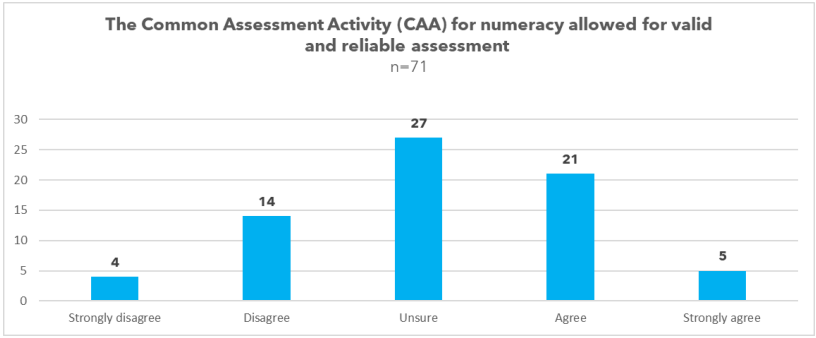
Numeracy standard

Figure 7: Agreement ratings for "the numeracy standard is easy to understand"



68% (51) of the 75 pilot participants who responded to the survey agreed or strongly agreed that the Numeracy standard is easy to understand compared to 21% (16) who disagreed or strongly disagreed agreed with this statement. 11 percent (8) of participants were unsure whether the Numeracy CAA standard was easy to understand.

Figure 8: Agreement ratings for "the CAA for numeracy allowed for valid and reliable assessment"



39% (27) were unsure about this statement and were the largest group of pilot participants who responded to this survey statement. 36% (26) agreed or strongly agreed with the statement, while 25% (18) participants disagreed or strongly disagreed that the CAA for Numeracy allowed for valid and reliable assessment. Four schools did not complete this section of the survey.

- The agreement rating teachers gave about the assessment being valid and reliable needs to be understood in the context of many participants not having seen the questions in the CAA and that teachers completed the survey prior to results being published.
- Of the two statements about the standards (easy to understand and allowing for a valid and reliable assessment), the second statement was not agreed with as strongly as the first statement, and this was true for each assessment area.
- Most Literacy (reading) (96%), Literacy (writing) (66%), and Numeracy (68%) survey participants agreed that the Literacy and Numeracy standards were easy to understand.
- Responses to the statement that the Common Assessment Activities (CAAs) allowed for valid and reliable assessment was agreed with by most of the Literacy (reading) (77%) and Literacy (writing) (66%) participants. Fewer participants agreed with this statement for Numeracy (36%).
- For Numeracy, the most common response was that respondents were unsure (39%) whether they agreed the CAA allowed for a valid and reliable assessment. Numeracy also had the highest rate of participants disagreeing (25%) with the statement.
- These statements are a snapshot of the perceptions of pilot participants. It may be valuable to investigate these responses further, especially given the differences across the different assessment areas and also to understand if receiving the achievement results has impacted teacher views about the standards and CAA.

3.6 Te Reo Matatini me te Pāngarau Common Assessment Activity

As part of the ongoing improvement to the assessments for Te Reo Matatini me te Pāngarau, NZQA has provided a summary of the guidance gathered from their team of markers.

For Te Reo Matatini, the assessment task includes the following components:

- pānui (reading)
- tuhituhi (writing)
- reo-ā-waha (listening)
- kōrero (speaking).

The team of markers identified a potential issue with the clarity of the instructions provided for ākonga and felt this may have contributed to the low achievement overall in Te Reo Matatini in the first assessment event. The markers wondered whether ākonga have not yet developed the level of fluency and language knowledge required for the assessment, particularly if the ākonga had not gained or been learning consistently through Te Marautanga o Aotearoa up to curriculum level four.

For Pāngarau, the assessment task focussed on foundational knowledge including mathematical language and terms, understanding and solving mathematical problems and operations, interpreting data representations and graphs, and applying these to real life contexts. Guidance from markers indicates that ākonga require further learning extension and experience in context to successfully meet the outcomes of the assessment tasks.

As part of work alongside kura and secondary schools, Tai Huki Consult Ltd sought feedback from kaiako regarding the assessment tasks for Te Reo Matatini me te Pāngarau. When considering whether these assessment tasks are fit for purpose, kaiako responses varied, which may be a result of the variable opportunities and space to engage deeply with the content and scope of learning required for ākonga success.

Summary: performance of the standards and assessments

This section has examined the performance of the Literacy and Numeracy | Te Reo Matatini me te Pāngarau common assessments. Building on the findings from the 2021 evaluation, the 2022 results data has been similarly analysed.

For the Literacy and Numeracy standards, this involved making comparisons with the 2021 data, analysing the relationship between e-asTTle scores, and drawing on the psychometric analysis. The findings for the 2022 results further reinforce the view that the assessments (CAA) perform well against the standards. This is particularly true for the Literacy (reading) and Numeracy common assessment. However, the e-asTTle Writing data indicates that a greater proportion of students should be achieving the standard if it were to perform similarly to the Literacy (reading) and Numeracy standards, although NZQA noted that there are greater differences between the e-asTTle writing assessment tasks and the CAA writing tasks.

Feedback from the teachers participating in the Literacy and Numeracy pilot provides strong agreement that all three standards were easy to understand, with agreement for the Literacy (reading) standard being particularly high. The view that the assessments allowed for valid and reliable assessment was less strong but was still agreed with by the majority of respondents for the Literacy assessments (both reading and writing). Participants were less confident in the Numeracy assessment being valid and reliable, and this appeared to be related to their concerns about the amount of reading and writing that is required as part of the assessment.

The small sample size and lack of national measures for the Te Reo Matatini me te Pāngarau assessment make it difficult to draw meaningful conclusions about the performance of the standards. However, the report from the NZQA markers for the two standards identifies there are opportunities to improve the clarity of the questions. In the second phase of the 2022 evaluation, it will be important to seek further insights from the kaiako involved in the pilot. As recommended in the 2021 evaluation report, work has begun to further explore ways to psychometrically analyse the performance of the Te Reo Matatini me te Pāngarau standards.

4 How equitable are the assessments?

An area of particular interest for this evaluation is how equitable the assessment tasks are. Quality assessment design should ensure that assessment tasks are culturally inclusive and support all NCEA user-groups to equally access the assessment. There is a particular focus on providing assessment items and contexts that will empower Māori and Pacific students.

To understand how well the Literacy and Numeracy standards and assessments meet these goals around equitability, the results data has been compared by gender, ethnicity, school decile, English language learner status, and tertiary/alternative education status.

This is followed by a discussion of the perspectives of teachers who responded to the mid-year survey questions about how equitable the assessments are perceived. Finally, the findings from the "Learner's Survey" are discussed, and this gives an understanding of student's views.

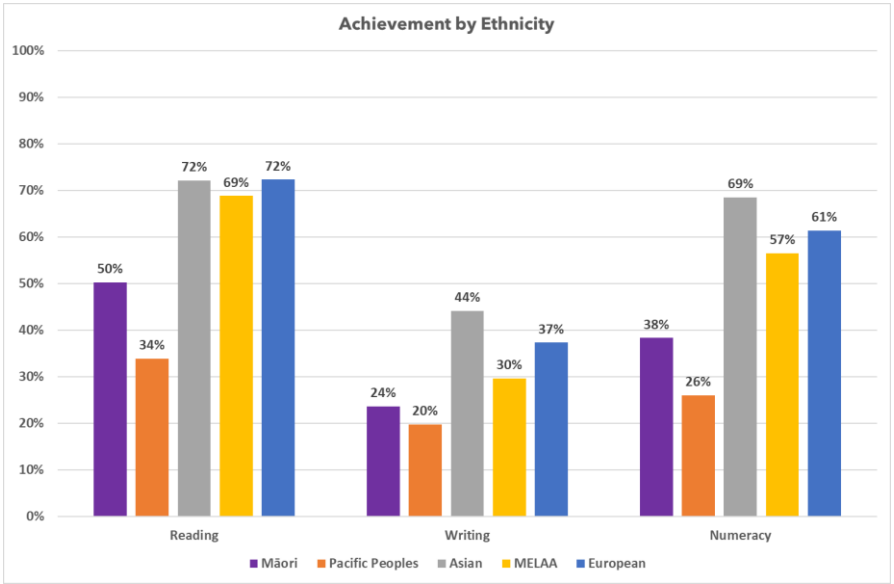
4.1 Assessment results by ethnicity, gender, school decile, ELL status and tertiary/alternative education status

Looking for any differences in the assessment results by ethnicity, gender, school decile, ELL status, Realm nations, and tertiary/alternative education status is part of understanding any potential issues regarding equitability. The results for the various pilot sub-groups are presented and discussed below.

The achievement rates are mostly reported as a percentage to best show the differences between groups and differences between standards. However, the group sizes vary considerably so, where needed, the numbers of students are also given to provide a clearer understanding of the data.

Results by ethnicity

Figure 9: Achievement by ethnicity for literacy and numeracy assessments



Across the three areas, Asian, MELAA (Middle Eastern, Latin American and African), and European students performed similarly with Asian students who have the highest achievement rates overall. Māori and Pacific students had lower achievement rates with Pacific students attaining the lowest achievement rates across all three tests.

- Looking at Literacy and Numeracy achievement by ethnicity, the patterns for 2021 and 2022 are the very similar, with Asian students achieving the highest pass rate in all subject areas, followed by European, MELAA, Māori, and Pacific students. Within this pattern, some variation in achievement between the two pilot years exists, but this is likely to be explained by the relatively small cohorts participating in the 2021 pilot. The numbers of students from these ethnic groups that participated in the 2022 pilot are detailed in Table 7.

Table 7: Achievement by ethnicity for literacy and numeracy assessments

Ethnicity	Standard Name	Participating Students	Achieved Students (n)	Achieved Students (%)
Māori	Reading	1,852	929	50.2%
Pacific Peoples	Reading	1,352	459	33.9%
Asian	Reading	1,489	1,074	72.1%
MELAA	Reading	221	152	68.8%
European	Reading	5,880	4,249	72.3%
Māori	Writing	1,666	393	23.6%
Pacific Peoples	Writing	1,319	260	19.7%
Asian	Writing	1,430	631	44.1%
MELAA	Writing	203	60	29.6%
European	Writing	5,562	2,072	37.3%
Māori	Numeracy	2,598	996	38.3%
Pacific Peoples	Numeracy	1,643	428	26.0%
Asian	Numeracy	2,087	1,429	68.5%
MELAA	Numeracy	285	161	56.5%
European	Numeracy	9,049	5,555	61.4%

Table 8: Achievement by ethnicity for te reo matatini me te pāngarau assessments

Ethnicity	Standard Name	Participating Students	Achieved Students (n)	Achieved Students (%)
Māori	Te Pāngarau	93	17	18.3%
Pacific Peoples	Te Pāngarau	16	3	18.8%
Asian	Te Pāngarau	2	1	50.0%
MELAA	Te Pāngarau	1	0	0.0%
European	Te Pāngarau	2	1	50.0%
Māori	Te Reo Matatini	37	9	24.3%
Pacific Peoples	Te Reo Matatini	7	0	0.0%

In the 2021 pilot it was noted that the unequal results for Māori and Pacific students are likely to be reflective of the wider inequity within the education system (that is, are expected). While the 2022 results reinforce the need for an ongoing focus on these groups and the systemic inequities being experienced, it does also suggest that the Common Assessment Activities are producing similar results to other assessment tools that have been used within the sector for some time.

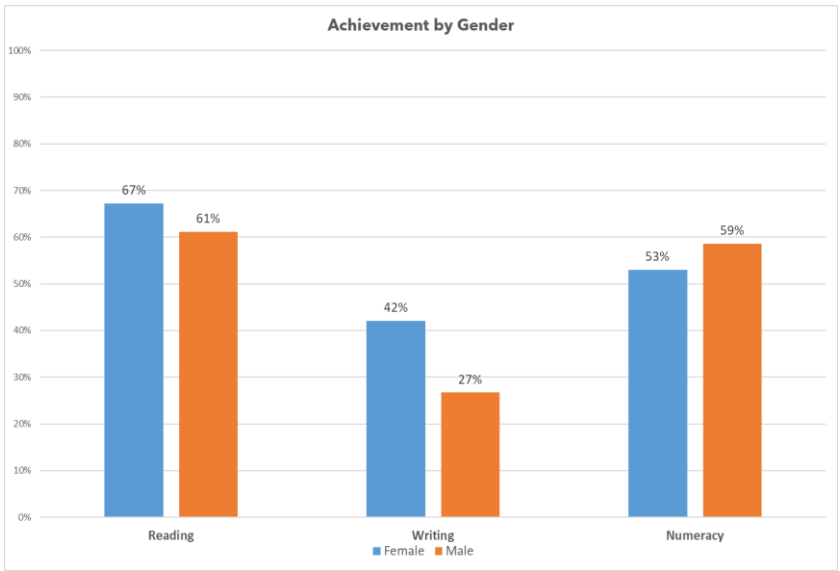
Table 9: Achievement by ethnicity for Te Reo Matatini me te Pāngarau

	Māori	Pacific peoples	Asian	MELAA	European
Pāngarau	18% (n=93)	19% (n=16)	50% (n=2)	0% (n=1)	50% (n=2)
Te Reo Matatini	24% (n=37)	0% (n=7)	N/A	N/A	N/A

- A limitation of this data is the low numbers of Pacific Peoples, Asian, MELAA, and European students have taken the assessments. This is indicated by the numbers in brackets for each ethnicity.
- All groups with more than six ākonga taking the assessment had less than 25% achievement rates with none of the seven Pacific Peoples students achieving Te Reo Matatini. Māori students performed slightly better in Te Reo Matatini than Pāngarau.

Results by gender

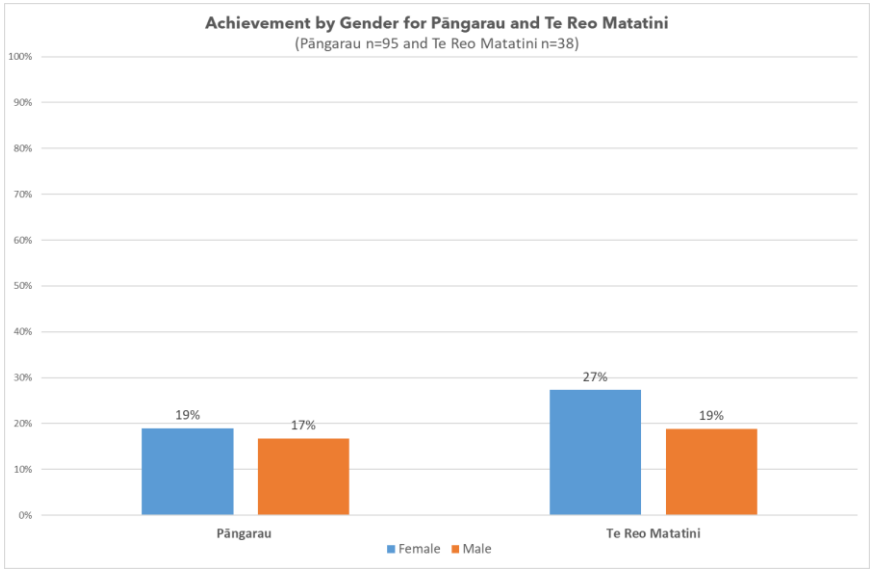
Figure 10: Achievement by gender for literacy and numeracy standards



Males and females performed similarly in Literacy (reading) and Numeracy with females performing slightly better in Literacy (reading) and males slightly better in Numeracy. Females had a significantly higher achievement rate in Literacy (writing than Males. 42% (1820) of females achieved in Literacy (writing) compared to 27% (1209) of Males.

- The 2021 pilot also examined patterns of achievement by gender and ethnicity. The 2022 results data follow trends similar to the 2021 data. In both years female students had higher pass rates in Literacy (reading) and Literacy (writing). Male students had higher pass rates in Numeracy. In 2021 the biggest difference in achievement between male and female students was in the Literacy (writing) assessment where the female pass rate was 14.5 percentage points higher. In 2022 the difference was 15 percentage points.
- Looking at the data across all three assessments, for the 2022 pilot the average rate of achievement for males was 48.8% and for females 54.1%. This is a somewhat similar pattern to the 2021 data where, on average, 56.8% of males and 59.1% of females achieved the standard.

Figure 11: Achievement by gender for Pāngarau and Te Reo Matatini standards

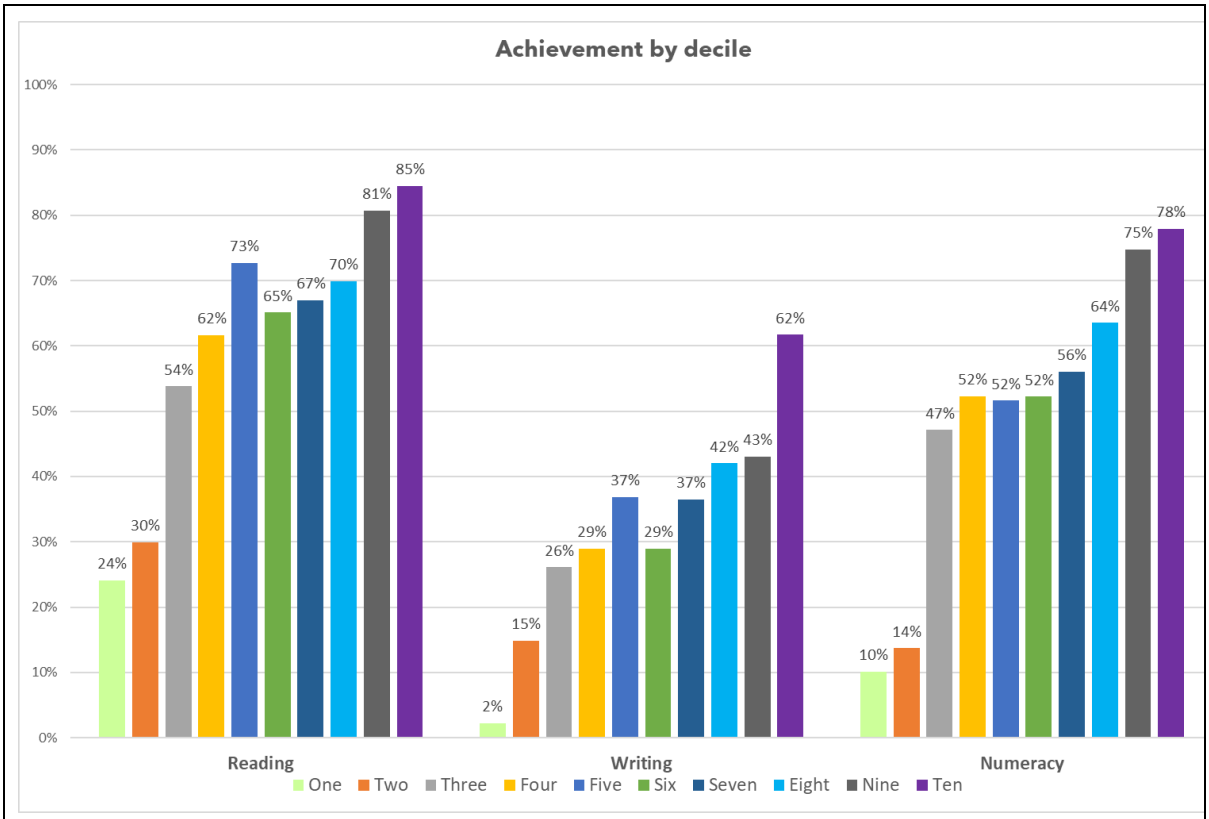


Males and females performed similarly in Pāngarau with a 19% (10) achieved rate for females and a 17% (7) achieved rate for males. Females had a higher achievement rate in Te Reo Matatini with a 27% (6) achieved rate compared to a 19% (3) achieved rate for males.

- Given the small numbers of ākonga, no substantive conclusions can be made for gender achievement rates for Pāngarau and Te Reo Matatini.

Results by school decile[12]

Figure 12: Achievement by decile for literacy and numeracy standards



There is a strong relationship between decile and achievement rates with higher decile schools generally showing higher achievement rates. Decile 1 and 2 schools performed significantly lower than other schools/organisations, with only 2% (3) of 135 decile 1 students receiving an achieved grade in Literacy (writing). Decile 5 schools were the only group that outperformed higher decile schools overall, but this did not diminish the overall trend. Decile 10 schools performed significantly higher than all other schools in Literacy (writing) was also higher in Literacy (reading) and Numeracy.

Table 10: Student achievement of literacy and numeracy standards by decile (count and percentage)

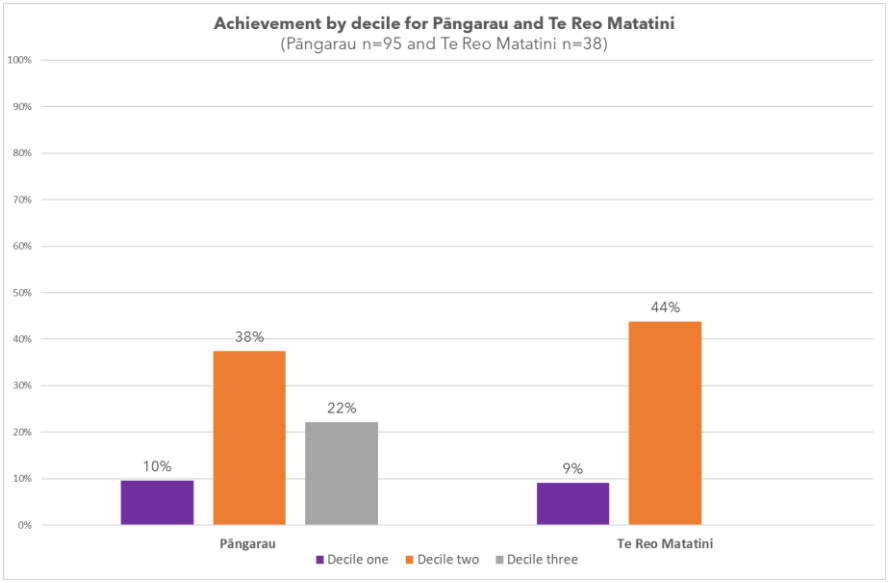
Decile	Standard Name	Participating students (n)	Achieved students (n)	Achieved students (%)
1	Reading	195	47	24.10%
2	Reading	421	126	29.90%
3	Reading	424	228	53.80%
4	Reading	1,570	967	61.60%
5	Reading	462	336	72.70%
6	Reading	1,495	973	65.10%

[12] The school decile describes the extent to which a school draws its students from low socio-economic communities. Five factors contribute to the decile: household income, occupation, household crowding, educational qualifications, and the income support levels of houses in the geographical areas a school draws its students from. Schools are ranked and then divided into 10 groups, called deciles.

7	Reading	1,711	1,147	67.00%
8	Reading	1,643	1,149	69.90%
9	Reading	545	440	80.70%
10	Reading	624	527	84.50%
1	Writing	135	3	2.20%
2	Writing	396	59	14.90%
3	Writing	567	148	26.10%
4	Writing	1,406	407	28.90%
5	Writing	345	127	36.80%
6	Writing	1,414	408	28.90%
7	Writing	1,659	605	36.50%
8	Writing	1,572	660	42.00%
9	Writing	551	237	43.00%
10	Writing	525	324	61.70%
1	Numeracy	178	18	10.10%
2	Numeracy	371	51	13.70%
3	Numeracy	683	322	47.10%
4	Numeracy	1,879	983	52.30%
5	Numeracy	1,418	731	51.60%
6	Numeracy	1,784	931	52.20%
7	Numeracy	2,600	1,455	56.00%
8	Numeracy	2,148	1,364	63.50%
9	Numeracy	1,033	772	74.70%
10	Numeracy	1,049	817	77.90%

- Analysing the rate of achievement for the three common assessments in relation to the decile rating of the student's school revealed a clear relationship, with the general trend indicating the higher the decile, the higher the achievement rate.
- This analysis of the Literacy and Numeracy achievement results for students in low decile schools draws attention to the additional challenges these students face. Other research looking at school demographics and student achievement also report that students enrolled in larger schools and high decile schools (often large schools are high decile) performed higher on average than other students (Caygill & Sok, 2008; Yuan, Turner & Irving, 2010).

Figure 13: Achievement by decile for Pāngarau and Te Reo Matatini standards

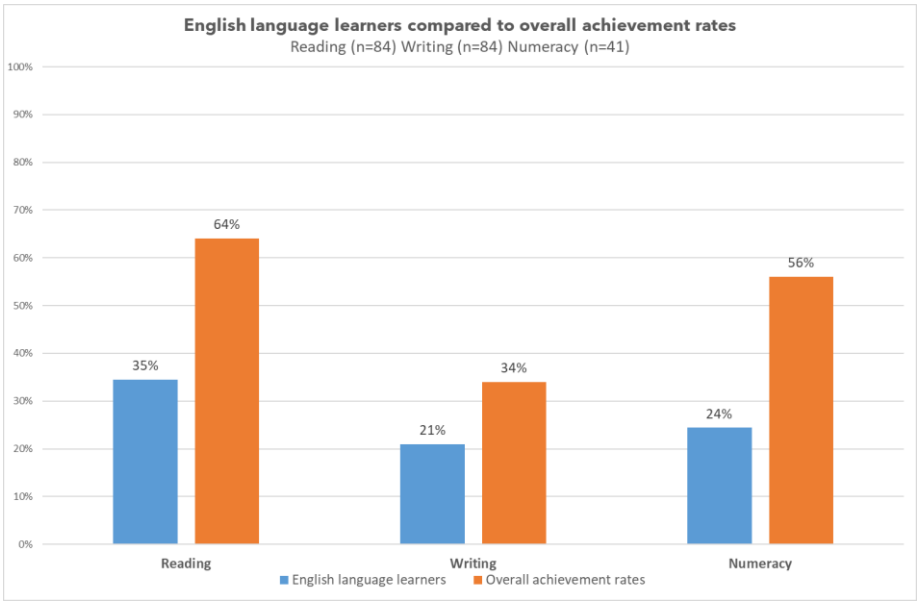


For the 95 ākonga who participated in the Pāngarau assessment, those from a decile 2 kura had a higher rate of achievement 38% (6 out of 16) than those from the decile 1 kura 10% (5 out of 52). With 22% (6 out of 27) achievement students from decile 3 kura also had a lower rate than the decile 2. Achievement rates in the Te Reo Matatini standard also had lower levels of achievement; decile 1 kura 9% (2 out of 22) comparted to the decile 3 kura 44% (7 out of 16).

- A limitation of the decile analysis for Pāngarau and Te Reo Matatini was that the kura participating in the pilot are either decile 1, 2, or 3. For both Māori-medium standards, the decile 2 kura had higher achievement rates than decile 1 kura.
- Conclusions cannot be drawn about the lower achievement rates for Pāngarau in the decile 3 kura, as overall the Māori-medium data is likely to be less reliable due to the small number of kura, possible differences between kura in how ākonga were selected to participate, and/or how readiness was determined.

Results for English language learners

Figure 14: Achievement for English language learners compared to overall achievement rates



For each assessment area, ELL achievement rates are lower than the overall achievement rate. The most significant difference for ELL students was in Numeracy where the achievement rate 24% (10) is less than half that of the overall achievement rate of 56% (7499).

- The initial analysis of the data indicates significantly lower achievement rates for English language learners, suggesting an inequity in the assessment. However, further analysis gives further understanding of differences in the readiness levels of the ELL students who participated in the pilot versus the overall cohort (indicated by current achievement using e-asTTle).
- The e-asTTle data indicates that the proportion of ELL students entered in the CAA who were not at the recommended curriculum level (NZC 4A) is higher than in the overall cohort[13] of students. The percentage of the overall cohort with e-asTTle scores below 4A were Literacy (reading) 16.6%, Literacy (writing) 26.6%, and Numeracy 21.1%. For the ELL sub-group, the percentage of students entered with scores below 4A were Literacy (reading) 34%, Literacy (writing) 71%, and Numeracy 64%.
- While this is a very small sample, a further explanation for ELL students having lower rates of achievement in the CAA is due to this cohort including a greater proportion of students who had not met the readiness standard.
- Using the Numeracy standard as an example, Table 11 shows a comparison between e-asTTle scores and achievement rates for ELL students helps to understand this further. The Literacy standards show similar patterns.

[13] The cohort of students with e-asTTle scores that were entered in the CAA.

Table 11: Relationship between e-asTTle scores and achievement of the numeracy CAA for ELL

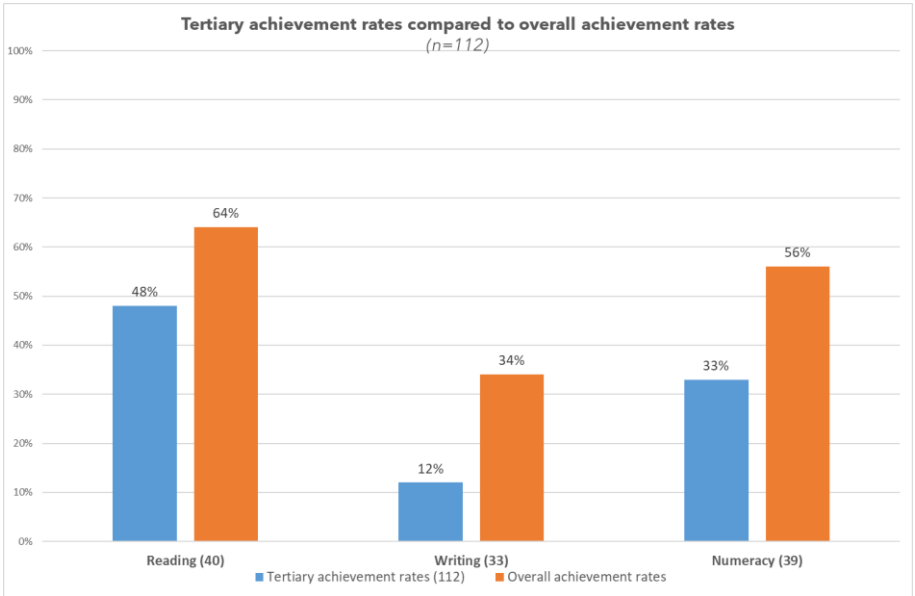
e-asTTle Maths score	No. of ELL students participating	No. of Achieved	Achievement rate
3B	1	0	0%
3P	6	0	0%
3A	4	0	0%
4B	2	0	0%
4P	1	0	0%
4A	2	1	50%
5B	1	0	0%
5P	1	1	100%
5A	3	3	100%
6B	1	1	100%

Although the number of ELL students with valid e-asTTle results (22) who participated in the Numeracy CAA is a small sample size, it does give an indication that ELL students at or above minimum "readiness" levels are as likely as non-ELL students to achieve the Numeracy CAA. Of the eight students who had an e-asTTle score 4A or above (as shown in the darker shaded portion of the table), six received an achieved grade, a 75% success rate. None of the 14 students who scored below 4A in e-asTTle received an achieved grade (0%).

- The relationship shown between the standard achievement and the e-asTTle level in Table 11 above could indicate that the e-asTTle is also an appropriate readiness indicator for ELL students

Results for tertiary/alternative education

Figure 15: Achievement for tertiary/alternative education compared to overall achievement rates



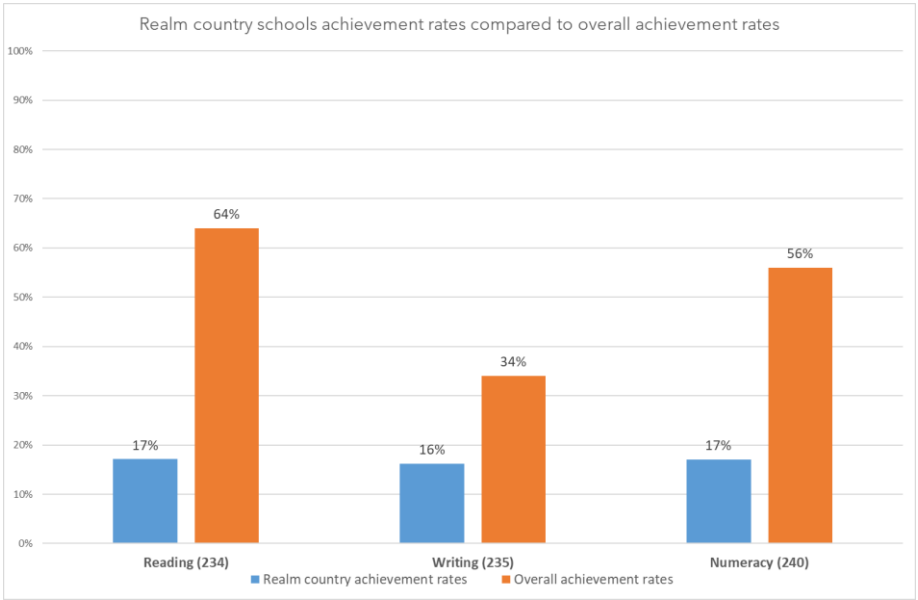
112 tertiary/alternative education students participated in one or more of the three CAA. The graph shows lower rates of achievement from this group. The highest levels were seen in Literacy (reading) 48% (19). This compares to 64% for the pilot overall. Literacy (writing) had the lowest achievement level for tertiary participants 12% (4); this compares with 34% overall. Numeracy achievement was 33% (13) for tertiary compared with 56% overall.

- Two tertiary providers and two alternative education providers entered students into the CAA. Although, it is possible that the students from the alternative education organisations were also adults due to the range of programmes being offered by these organisations.
- Across all three assessment areas, the tertiary/alternative education rates were lower than overall achievement rates, with rates of achievement in Literacy (writing) being almost three times less than the overall rate. Literacy (reading) is the highest performing area for tertiary students with nearly half receiving achieved grades.
- Analysis of the very limited e-asTTle data provides some insights into the achievement results for this particular sub-group. There were only eight tertiary students who had a relevant e-asTTle score across Reading, Writing, and Maths, and these were either at or above the readiness mark of 4A. Of these eight results, seven achieved the standard.
- No ākonga at a tertiary or alternative education organisation attempted the Te Reo Matatini or Pāngarau assessments in the July assessment event.

Realm countries

Realm country schools were invited to participate in this pilot. In total, seven schools were entered into the pilot with five of those schools entering students in assessment event one. In total, 234 students entered Literacy (reading), 235 Literacy (writing), and 240 Numeracy.

Figure 16: Achievement by Realm countries compared to overall achievement rates



Achievement rates for Realm country schools were significantly lower than the overall achievement rates in all three standards. 17% (40) achieved Literacy (reading) and 16% (38) achieved Literacy (writing). In numeracy the rate), was 17% (41). This means approximately only one in six students assessed achieved the standard in Realm country schools.

- The numbers of students who participated from schools in Realm countries are greater than the two other sub-groups of participants that have been looked at in this section (ELL, tertiary, and Realm countries).

- Achievement rates for Realm country schools were very similar across the three standards and were significantly lower than the overall achievement rates.
- There was either no or limited readiness screening for realm country students; thus, it is likely that this will have impacted on the results. The absence of e-asTTle data for this sub-group of participants means understanding more about the lower achievement results is difficult.

Summary of achievement results for pilot sub-groups

This section has looked at the achievement results for sub-groups of students who are of particular interest. To understand how equitable the assessments are, results data is presented and discussed by ethnicity, gender, school decile, ELL, tertiary/alternative education, and Realm countries. Because of the small numbers of ākonga who piloted the standards, the sub-group data has not been looked at for Te Reo Matatini me te Pāngarau standards.

Looking at differences in the achievement levels in the Literacy and Numeracy standards for the pilot sub-groups is useful for understanding equitability; however, the very small numbers in some of the sub-groups makes it difficult to draw meaningful conclusions. Potential differences in the sub-group cohorts could be due to schools and organisations making different decisions about who sat the assessments, so it cannot be assumed that we are comparing "apples with apples".

Further understandings around the equitability of the CAAs could also be gained by examining and comparing the achievement patterns in the e-asTTle curriculum levels for these sub-groups.

The second phase of this evaluation will seek further information to understand some of the potential differences with respect to ethnicity, school decile and also tertiary/alternative education settings. This will be analysed and discussed in Report Two.

Table 12 summarises the achievement ratings for each of the pilot sub-groups looked at in this section of the evaluation.

Table 12: Achievement results for pilot sub-groups

Sub-group	Achieved Literacy (reading)		Achieved Literacy (writing)		Achieved Numeracy		Average achieved across all CAA
	%	n	%	n	%	n	%
All students in pilot	64%	6,016	34%	3,029	56%	7,512	51.3%
Māori	50%	929	24%	393	38%	996	37.3%
Pacific Peoples	34%	459	20%	260	26%	428	32%
Decile 1and 2	27%	173	8.5%	62	12%	69	15.8%
ELL	35%	29	21%	18	24%	10	26.6%
Tertiary/AE	48%	19	12%	4	33%	13	31%
Realm	17%	40	16%	38	17%	41	16.6%

Ethnicity: Achievement of the Literacy and Numeracy standards by ethnicity is similar across the three assessment areas, with the highest pass rate being achieved by Asian students, followed by European, MELAA, Māori, and Pacific peoples. This pattern follows the 2021 findings and the achievement trends across the wider education sector.

Gender: The gender differences are less pronounced but do reveal higher pass rates for male students in Numeracy and for female students in Literacy (reading) and Literacy (writing).

School decile: The analysis of achievement by school decile illustrates a general trend of the higher the decile, the higher the achievement rate. The combined decile 1 and 2 results were the lowest overall for the sub-groups looked at. Similar trends are reported in other research examining school demographics and student achievement.

English language learners: The initial analysis of ELL data showed that assessment results were lower for this sub-group of students. Availability of e-asTTle data for this cohort allowed for further analysis, which indicates a possible explanation for the lower pass rate is that a greater proportion of ELL students had not met the readiness standard (NZC 4A) for the CAA than the general cohort.

Tertiary/alternative education: Across all three assessment areas the achievement results for tertiary/alternative action students were significantly lower than the overall rate. The small number of students in this sub-group means the understanding of trends for this sector is limited.

Realm: Achievement rates for Realm country schools were significantly lower than the overall achievement rates. However, the absence of e-asTTle data for this cohort means understanding these results is difficult.

4.2 Student perspectives on the assessments

At the end of each Common Assessment Activity, students were asked to rate the following statements in relation to the assessment, using a five-point scale from "Strongly agree" to "Strongly disagree".

1. I felt ready for today's assessment.
2. I had everything I needed to complete the assessment today. For example, Log on, quiet classroom.
3. The skills I used today will be meaningful in the rest of my life.
4. The questions in today's assessment reflect students like me.
5. Did you use text-to-speech to listen to today's assessment text being read aloud? (Writing and Numeracy)
6. I found the text-to-speech functionality was useful. (Writing and Numeracy)
7. Do you have any further comments about today's assessment? (Reading)

The survey, which was optional, yielded a 28% response rate. The table below details the numbers for each assessment.

Table 13: Response rate for the mid-year learner survey

	No. of survey respondents	No. participating in assessments	Response rate
Overall	**8893**	**31,462**	**28%**
Reading	305	9392	32%
Writing	2756	8862	31%
Numeracy	3087	13,208	23%

There was a slightly higher response rate from students who achieved the standard (31%) compared to those who did not achieve the standard (25%). However, the differences are not significant enough to skew the survey data.

For each of the seven statements, the agreement ratings students gave are discussed below, with consideration being given to any differences in views between achieving/non-achieving participants, gender, ethnicity, and school decile.

Discussion of students' views

I felt ready for today's assessment.

This statement asked for student perceptions about how ready they felt for the assessment. This potentially covers a multitude of factors including having practice questions, teacher discussions, self-perception of being a mathematician, or simply how prepared they were on the day.

- Overall, between a half and two-thirds of students felt ready for the assessments - (reading (47%), writing (61%) numeracy (52%).
- Overall, students who achieved the standard felt more ready than those that did not achieve it.
- Decile rating was a strong factor in ratings for this statement. In general, the lower the decile, the less ready they felt.
- While there was little difference across ethnicities in Literacy (writing), Māori had significantly lower agreement ratings compared to other ethnicities in Numeracy and Reading.
- Pacific students had lower agreement ratings in Numeracy.
- Gender was a factor in Numeracy with significantly more males (58%) agreeing with the statement than females (44%).

I had everything I needed to complete the assessment today. For example, Log on, quiet classroom.

This statement focussed on preparedness for the assessment and the quality of the administration of the assessment (for example, quiet space).

- There was strong agreement from students across all three assessments – 83% to 86% either agreed or strongly agreed.

- Achieved students were slightly more in agreement than those students who did not achieve.
- The agreement ratings across all ethnicities are similar, but Māori and Pacific students' ratings are slightly lower across all assessments.
- Decile 1 and 2 students rated this statement significantly lower than other students across all three assessments.
- No significant differences between genders arose.

The skills I used today will be meaningful in the rest of my life.

The statement focuses on the future relevance and usefulness of the skills and behaviours used to address the assessment.

- Approximately a third of students neither agreed nor disagreed with this statement.
- Achieved students generally had higher agreement ratings in Literacy (writing) and Numeracy than non-achieving students. The ratings for both groups were quite similar for Literacy (reading).
- No significant differences across ethnic groups, although Māori had the lowest percentage in agreement ratings and highest in disagreement ratings in all three assessment areas.
- No specific trends with decile ratings in Literacy (reading).
- Decile 1 students had a very low agreement percentage compared to other groups in Literacy (writing). Only 33% of decile 1 students gave an agreement rating. The next lowest percentage was 58% from decile 8 students. Decile 5 students gave the highest agreement percentage at 78%.
- Decile 1 and 2 agreement percentages were lower than all other decile groups in Numeracy.
- No significant difference between gender although boys were slightly more in agreement in Numeracy.

The questions in today's assessment reflect students like me.

This statement focuses on the relevance of the assessment to the student and whether they feel they can relate to the questions.

- A relatively large percentage of students in all three assessments neither agreed nor disagreed with this statement (42% Reading, 33% Writing, 43% Numeracy).
- Similar ratings arose for those who achieved and did not achieve in Literacy (reading) and Literacy (writing). Those who achieved in Numeracy had higher agreement percentages than those who did not achieve.
- No significant patterns emerged in ethnicity data with groups providing similar ratings.

- Decile 1 students had a much lower agreement rating than other decile groups in Literacy (reading) and Numeracy. In Numeracy decile 1 and 2 students had the lowest agreement percentages and the highest disagreement percentages. In comparison, decile 10 students had the highest agreement percentages and the lowest disagreement percentages.
- Females had slightly higher agreement percentages in Literacy (reading) and Literacy (writing), but it was the males who had a higher agreement percentage in Numeracy (46% M, 39% F).

Did you use text-to-speech to listen to today's assessment text being read aloud? (Writing and Numeracy)

Students were asked to indicate if they had used the text-to-speech function during the assessment in either Writing or Numeracy with either a Yes or No option.

Only 4% (138) in Numeracy and 7% (184) in Literacy (writing) indicated they had used the text-to-speech function.

I found the text-to-speech functionality was useful (Writing and Numeracy)

This was a follow up statement after having asked whether students had used the text-to-speech function (Polly).

It appears a limitation of this survey is that students who did not use the text-to-speech functionality could continue to rate its usefulness. While only 138 had indicated they had used it for Numeracy in the previous question, 569 students gave ratings. It was similar in Literacy (writing) with 184 indicating they had used the text-to-speech function with 485 giving ratings.

- A third of those who gave ratings in Numeracy agreed it was useful; a third felt they neither agreed or disagreed and a third disagreed it was useful.
- It was different in Literacy (writing) with 50% agreeing it was useful, with only 22% in disagreement.
- A bigger percentage of students who did not achieve the standard felt the text-to-speech function was useful, compared to those students who received an achieved grade.
- There were similar ratings across ethnicity groups in the Numeracy assessment, but Pacific students found it more useful than other ethnicities in Literacy (writing).

Comments in relation to the text-to-speech function

Some students were frustrated that the text-to-speech functionality was available, but they were not allowed to use it because they either did not have headphones available or were not allowed to use headphones during the assessment.

> *"We were not allowed to use the text to speech, I tried and it was good but they said no."*
>
> *"The text to speech seems pointless because we aren't allowed headphones and we aren't allowed to make noise."*

Some students found the function annoying as it would pop up without being asked.

> *"The text-to-speech function was useless and annoying. Every time I went to delete something, it came up. Ruined the test for me."*

Some students suggested that words should be highlighted to follow the speech.

> *"The text to speech should highlight the words as it reads them to help students follow along."*

Do you have any further comments about today's assessment? (Reading)

This was an opportunity for students to provide some feedback on their Literacy (reading) Assessment experience.

- 294 pieces of written feedback were received.
- Nearly half of all comments were positive (47% n=137).
- 32% (n=93) of all comments felt improvement was needed.
- The most common comment in relation to the experience was that the progress bar had not performed correctly.
- Some students felt the layout of the onscreen assessment needed improvement.
- Other issues raised, listed in the table that follows, and include:
  - the progress bar not working
  - being locked out
  - crashing/loading issues
  - poor image quality
  - colour schemes not good for the colour blind
  - not enough preparation and support.

Table 14: List and count of further student comments regarding the reading assessment

Categorised response from students	Count
Progress bar not working/it's saying they haven't completed all the questions	17
Layout needs improvement	12
Wording hard to understand	9
Locked out	9
Crashing/Loading slow/Answers not saving	7
Poor image quality/not good for colour blind people	6
Doesn't work on their computer/browser (examples of Apple, Microsoft Edge, battery life)	4
Not enough preparation support	4
Issue with teacher/location	3
Too much scrolling	3
Other	21

Summary of student feedback

The learner survey produced some interesting results and feedback. Some of the most interesting points include the following.

- Most students agreed they had everything they needed to complete the assessment today. Each CAA had between 80% and 90% agreement.
- Students least agreed with the statement "The questions in today's assessment reflect students like me."
- In general, males were slightly more in agreement with the statements in Numeracy and females in Literacy (reading) and Literacy (writing).
- There was a strong relationship between feeling ready for the assessment and decile rating. Students in lower decile schools felt less ready than those in higher decile schools.
- Decile 1 and 2 students consistently rated lower or significantly lower across all statements.
- There were no significant differences across the statements for ethnic groups, but Māori and Pacific students generally rated the statements lower than other ethnic groups.
- Students who achieved the standard were generally more in agreement than for those students who did not achieve the standard.
- Only a small number of students used the text-to-speech feature in Numeracy and Literacy (writing). Of those who did, Pacific students found it more useful than other ethnicities, and those students who did not achieve the standard found it more useful than those that did achieve the standard. Males appeared to find the text-to-speech much less useful than females.

Recommendations based on student feedback:

- Make both teachers and students aware that the text-to-speech function is available in Literacy (writing) and Numeracy and to bring/allow headphones if it is being used.
- Make it clear to teachers that the text-to-speech function is an authentic part of the assessment and students should be able to use it if they wish.
- Highlight words as they are spoken if this functionality exists.
- Minimise the auto pop-up feature of text to speech so it does not annoy those students not employing it.

4.3 Teacher perspectives on the equitability of the literacy and numeracy standards

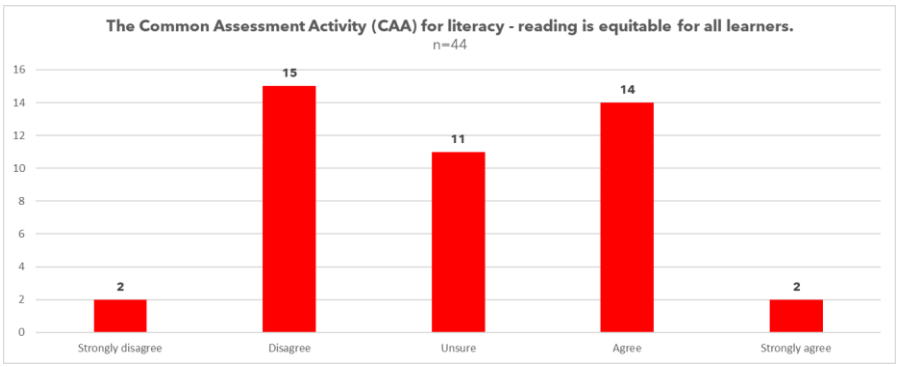
The teachers who participated in the mid-year literacy and numeracy survey were asked to rate their level of agreement or disagreement with two statements: that the CAA was equitable for all students and that it was equitable for English language learners (ELL). They were then asked to explain their ratings and share any suggestions for what they thought could be done to improve equitability. It is important to note that the survey respondents had not seen the content of the assessment (although some may have viewed parts of the assessment while students were completing it). This suggests that for many, the responses may have been an appraisal of their views on CAAs generally.

Equity for all students is discussed in the first section. The second section looks at participants perceptions of equity with a focus on English language learners (ELL). The third section describes the views of tertiary/alternative education organisations participating in the pilot. These organisations work with students who have not experienced success in mainstream education, and they frequently face a different set of challenges to support these students to participate in, and achieve success in, NCEA.

Equitability for all students

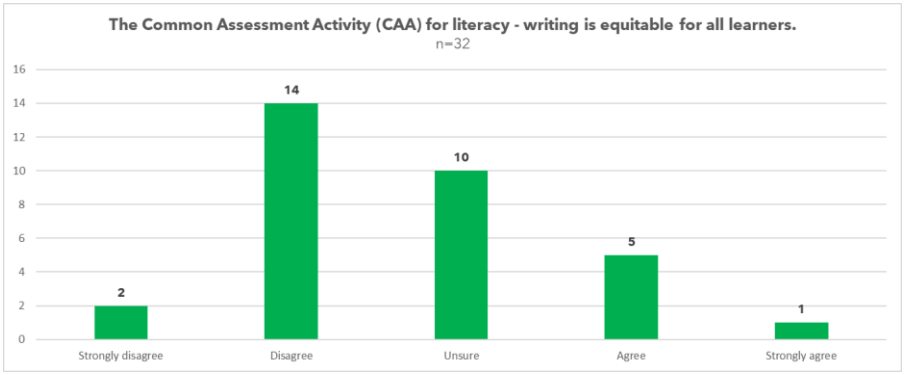
The three graphs below show the respondents' agreement ratings for the statement "the CAA is equitable for all learners" that respondents provided from each of the three assessment areas.

Figure 17: Agreement ratings for "the CAA literacy (reading) is equitable for all learners"



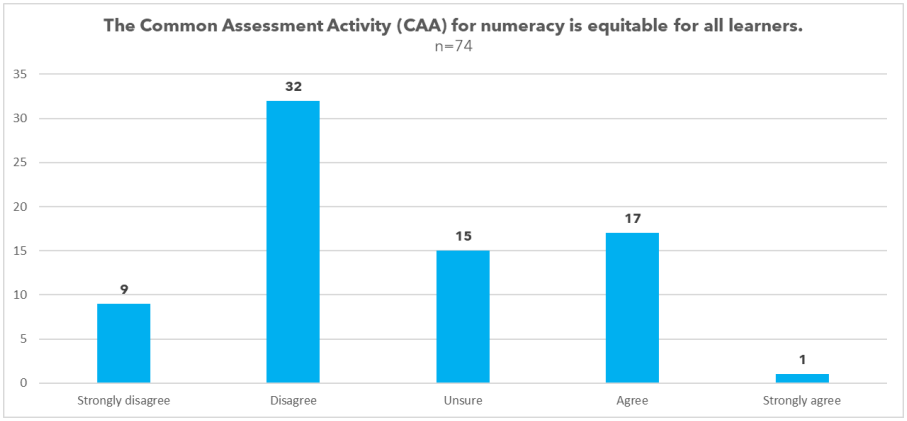
Views were divided on how equitable participants felt the Literacy (reading) assessment activity was. Nearly an equal split exists between those that disagreed/strongly disagreed 39% (17) and those who agreed/strongly agreed 36% (16). 25% (11) participants were unsure.

Figure 18: Agreement ratings for "the CAA for literacy (writing) is equitable for all learners"



50% (16) of participants disagreed/strongly disagreed that the Literacy (writing) assessment is equitable for all learners and 31% (10) were unsure. Only 19% (6) agreed or strongly agreed.

Figure 19: Agreement ratings for "the CAA for numeracy is equitable for all learners"



55% (39) of the participants were not in agreement with this statement, with 43% (32) disagreeing and 12% (9) strongly disagreeing. 20% (15) were unsure and 25% expressed agreement (17) or strong agreement (1).

- Across each of the three assessment areas, the view held by the largest group of respondents was "disagreeing/strongly disagreeing" with the statement that the CAA was equitable for all learners (Reading 39%, Writing 50%, Numeracy 55%).
- The Literacy (reading) CAA was the assessment area that was viewed to be most equitable, and the CAA for Literacy (writing) was seen as the least equitable.
- The teacher ratings for the equitability statements suggest that priority user groups would benefit from further work being carried out to address equitability, particularly for the Numeracy CAA.

Reasons for agreeing or disagreeing that the CAA was equitable for all learners

The pilot participants were asked to explain their ratings, and these are outlined below.

View of those who disagreed/strongly disagreed (39% reading, 50% writing, 55% numeracy)

Assessments were deemed less equitable based on the level of inclusiveness or diversity reflected in the various contexts or topics. While not everyone had seen the questions, there were some who identified a need for improved cultural inclusivity.

> *"The content of the reading was very Euro-centric."*

Also relevant to the assessment design of the Numeracy CAA was the view that the amount of reading particularly, but also writing, was too much. This was the reason a third of the respondents, who disagreed that the Numeracy CAA was equitable, gave. Some participants identified that this would impact ELL students particularly.

> *"One has to have a high literacy knowledge to get into the question to access the actual numeracy."*

> *"Our student feedback is that 38% of the students found the questions confusing and hard to understand, before they even attempted the maths."*

Other respondents raised concerns about the assessment method. Many identified the "exam" style as a barrier for some students, particularly those students who are anxious or not digitally confident. Others suggested that numeracy was not well suited to an online assessment.

> *"It's an on-line assessment, which does not suit all learners. Also, it is a one-off assessment where all competencies are assessed at once and gives a pass or fail result."*

Concerns were identified about how equitable the CAA was for learners who were achieving at the recommended curriculum level but who had specific learning needs, for example ADHD.

> *"It does not work for students who are ADD or ADHD as it requires long periods of concentration."*

Another query raised was whether the formatting of the CAA could be designed to better cater for students with dyslexia.

Logistical issues were raised as a contributing factor for the assessment activities not being equitable, for example the timeliness (late) and clarity of the SAC instructions, and a view that only some schools would be able to provide students with "unlimited time" to complete the assessment.

Finally, survey respondents who were from the tertiary or alternative education sector expressed views that the CAA would create significant barriers for their learners. One reason for this was if the assessment event occurred only twice a year, this would mean that their students, who are often engaged in education for shorter periods or time or are more transient, may miss the opportunity.

Views of those who were unsure (25% reading, 31% literacy, 20% numeracy)

Nearly two thirds of the "unsure" respondents said this was because they had not actually seen the CAA (questions or results). The other explanations given were similar to the respondents who disagreed/strongly disagreed, with a number identifying that the topics or contexts may not be understood or relatable to all learners.

Views of those who agreed/strongly agreed (reading 36%, writing 19%, numeracy 25%)

The most common reason provided for why they believed the CAA was equitable for all students was that the questions and/or contexts were appropriate for all students. In response to the Numeracy assessment, one participant said:

> *"The questions covered a range of situations, names, and concepts. I would be surprised if there were any students or groups of students who could not relate to any of the contexts."*

Another explanation was in support of the method of assessment, with the national consistency helping to reduce variability across schools.

> *"It is done under uniform conditions and will be marked independently."*

Other participants felt that the availability of special assessment conditions (SACs) and the text to speech tool (Polly) helped to support equitability. A participant reflecting on the Literacy (reading) CAA stated that *"allowing existing SACs to stand certainly made the CAA accessible to our students."*

Positive recognition of their school's/organisation's preparedness to administer the assessment, including having good access to devices or feeling that their students' familiarity with digital assessment and devices, was also given, and this was seen to contribute to the CAA being equitable.

There were also some who, like many of the "unsure" respondents, felt they could not give a reason for their rating due to not having seen the assessment, nor the assessment results.

Suggestions for making the CAA more equitable for all learners

The suggestions participants made for improving the equitability of the CAA related to the following main ideas.

- Standardise the upper time limit so all learners experience the same conditions.
- Tailor the question contexts and settings to be more relevant to the range of learners sitting the assessment.
- Provide a paper-based alternative to the online assessment.
- Have more than two assessment events per year.
- There was also an idea that a way of reducing the impact of learners sitting the test on a "bad day" was to divide the assessment up into a few sessions over the year with the results being collated at the end.

Some of the suggestions specifically focused on the numeracy assessment area. These included:

- reduce the amount and level of reading and writing required
- reduce the length of the assessment
- address the difficulties created by the online functionality, for example, not require scrolling
- students being able to show their working.

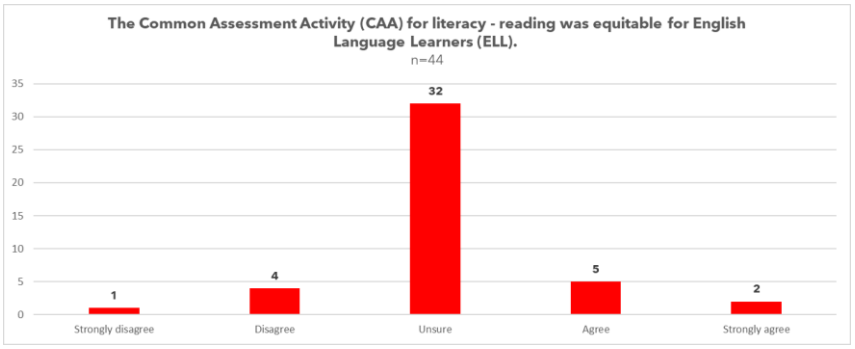
Some participants suggested that different types of assessments be used altogether, for example, portfolios and open-book assessments.

Other suggestions participants made were more focused on supporting learners to achieve the assessment standard. This included, returning the marked papers to students and providing more information and resources for teachers to better prepare learners.

Equitability for English language learners (ELL)

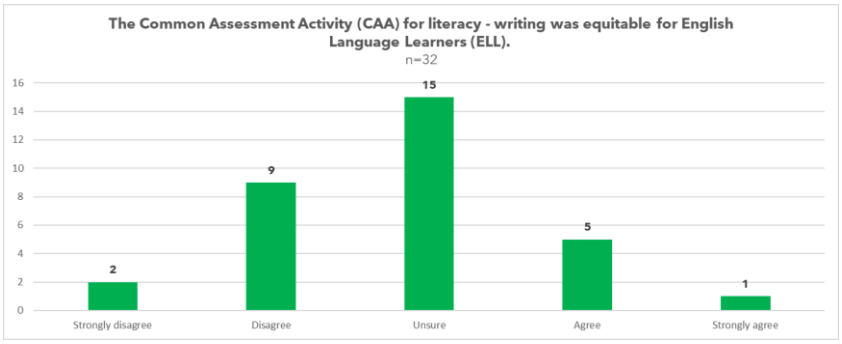
Pilot participants were asked to rate their level of agreement or disagreement with the statement that the Common Assessment Activity was equitable for English language learners. The responses for each of the three assessment areas are detailed in the charts below.

Figure 20: Agreement ratings for "the CAA for literacy (reading) was equitable for (ELL)"



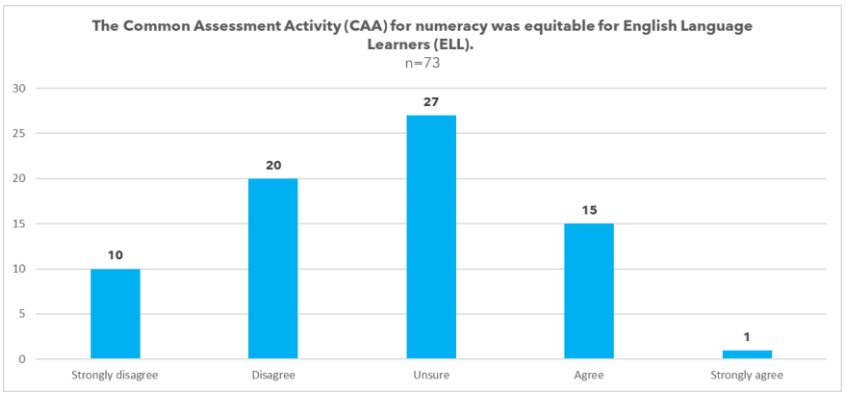
The majority 73% (32) of participants rated themselves as unsure if the Literacy (reading) assessment activity was equitable for ELL. For the remaining participants, there were slightly more, 16% (7), who agreed/strongly agreed with the statement than those that disagreed/strongly disagreed 11% (5).

Figure 21: Agreement ratings for "the CAA for literacy (writing) was equitable for English language learners (ELL)"



47% (15) of participants were unsure if the Literacy (writing) assessment was equitable for ELL. A smaller proportion disagreed/strongly or disagreed 22% (7) and only 18% (6) agreed or strongly agreed.

Figure 22: Agreement ratings for "the CAA for numeracy was equitable for English language learners (ELL)"



37% (27) of participants were unsure if the Numeracy assessment is equitable for ELL. A larger proportion, 41% (30), disagreed or strongly disagreed that it is equitable. The least number of participants, 22% (16), reported agreeing/strongly agreeing with the statement.

- Of the three assessment areas, the Numeracy CAA was viewed to be the least equitable for English language learners with 41% disagreeing/strongly disagreeing with the statement.
- For all assessment areas, there were higher levels of uncertainty expressed about how equitable the CAA was for ELL than for all students.

Reasons for agreeing or disagreeing that the CAA was equitable for English language learners (ELL)

The pilot participants were asked to explain their rating for the statement that the CAA is equitable for ELL. This section describes the main views and ideas that were expressed, firstly for those who disagreed/strongly disagreed, secondly for those who were unsure, and thirdly for those who agreed/strongly agreed that the CAA was equitable for ELL.

Views of those who disagreed/strongly disagreed (11% reading, 22% writing, 41% numeracy)

A number of reasons were given for disagreeing with the statement. The amount of reading and writing required is the main concern raised about the equitability of the Numeracy assessment for ELL. Another explanation for disagreeing is that ELL may not have been familiar with some of the question contexts in the Numeracy CAA. The issue of contexts was also an explanation for the CAA (reading) and CAA (writing) being inequitable for ELL.

Views of those who were unsure (reading 73%, writing 47%, numeracy 37%)

For each of the three assessment areas, there were a high proportion of participants who were unsure about how equitable the CAA was for ELL. The majority explained that this was because they had not seen the assessment. There were also some "unsure" respondents who reported that they did not have any ELL students sitting the CAA.

Views of those who agreed/strongly agreed (reading 16%, writing18%, numeracy 22%)

A number of those who agreed with the statement the CAA was equitable for ELL still identified some possible barriers. This included acknowledging that some of the question contexts could have been less understood by ELL, and, in the case of the Numeracy CAA, that the level of literacy required was too high.

Suggestions for making the CAA more equitable for ELL

The pilot participants made some suggestions for how to improve equitability for English language learners (ELL). Many suggestions were similar to those made for improving the equitability more generally, but the list below is more specific to ELL.

- Take care in selecting resources.
- Use the ELLP instead.
- Provide students with any Māori words that they will be expected to know for the CAA.
- Provide teachers with material about the contexts a week in advance so students can gain familiarity prior to the assessment.
- Simplify the language used.

The following support options were suggested for specifically for ELL sitting the Numeracy CAA.

- Questions to be available in other languages.
- Provide interpreters to explain unfamiliar words.
- Use of a translation app/dictionary.
- Use more symbols and iconography.
- Use of text to speech tool (Polly).

Equitability for students in tertiary/ alternative education

Seven organisations who participated in this pilot were categorised as tertiary or alternative education providers. These included:

- five tertiary providers
- two alternative education providers.

Four of the seven organisations had students participate in the Common Assessment Activity, and five responded to the mid-year survey. In total, 112 students participated across the three CAA assessments. This is a relatively small sample of this sector[14]; however, phase two of the evaluation will involve interviews with some of these organisations involved in the pilot to provide more information for the second 2022 evaluation report.

It is also worth noting that some of the secondary schools involved in the pilot also align with this sub-group. This includes two teen parent units, Te Aho o Te Kura Pounamu[15] and a school that provides education for young people in the care of Oranga Tamariki, including those in youth justice/care and protection residential facilities and community-based campuses.

In all these schools/organisations, students who face significant barriers to succeeding in education receive support and specialised learning environments that holistically address their social, behavioural, and academic needs. Feedback about the equitability of the new Literacy and Numeracy standards expressed concern that the co-requisite aspect could

[14] Approximately 3500 secondary aged students participate in Alternative Education each year; nearly 6000 are enrolled at Te Kura.

[15] Te Aho o Te Kura Pounamu plays a significant role in the education of at-risk and disengaged learners. Education agencies increasingly use Te Kura as a place of enrolment for non-enrolled and high needs learners.

impact their students' engagement and achievement. The new assessment approach, seen as unsuitable for many of their students, could likely reduce the numbers who would be able to gain an NCEA qualification. This is an area for further investigation in phase two of the evaluation.

The feedback previously gathered from this part of the education sector, as part of the NCEA review (2018), appears to be at odds with the assessment approach being piloted. In the 2018 report, perspectives of students and staff from alternative education, foundation-level tertiary courses, teen parent units, health schools, and secondary–tertiary vocational education indicated students appreciated the current flexibility of NCEA and valued internal assessments, particularly the opportunity to resubmit.

A strong theme in their feedback was dislike of external assessments such as exams. Students wanted to see less reliance on traditional written assessments and more choice of assessment format, including online or paper-based formats. Their suggestions for alternative, more relevant, and authentic methods of assessment include portfolios, conversations about learning, and video recordings of their practical work.

These themes are echoed again in this pilot evaluation with concerns expressed in teacher feedback about exams and assessments that required long periods of concentration. Other equitability issues were raised such as the significant differences in resourcing and professional development for teachers and tutors working in these schools/organisations. No, or limited, access to literacy advisors, learning support, and supports for SAC were all deemed barriers to ensuring equity in this assessment approach.

Some were concerned that the CAA was only going to being available for students to enter on a twice-yearly basis. This was seen as another factor that would affect equity of access. Some of these organisations provide only a 20-week course (Gateway courses), and others noted that often their students were transient or had short-term engagement in their education programmes.

While the voice of these organisations is relatively small in this pilot, the students who are represented by these schools/organisations are priority learners, are predominantly Māori, with most being at risk of low educational, social, and vocational outcomes. Understanding their experiences and views of the pilot and the anticipated impact of the new Literacy and Numeracy standards is important.

Summary: equitability of the standards and assessments

This evaluation has examined how equitable the standards are by comparing achievement results and drawing on the perceptions of teachers. The patterns of achievement reveal some differences that are related to factors such as ethnicity and decile but understanding more about these differences is needed. ELL students appear to achieve at similar levels as the overall cohort.

Participants rated the extent they believed the CAA was equitable for all learners and also specifically for English language learners (ELL). Feedback suggests a mixed view, but more commonly participants disagreed that it was equitable. The CAA for Numeracy was least favourably rated for how equitable it was.

The ratings for ELL had greater numbers saying they were "unsure" how equitable the assessments were for ELL, particularly for Literacy (reading) and Literacy (writing). The Numeracy CAA was again least favourably rated for equitability.

The more negative ratings of the Numeracy CAA for ELL and for "all" students, was most often linked to the view that there was an excessive amount of reading and writing within the assessment. Other reasons given were related to the assessment approach (an exam and online). There were mixed views on how appropriate or relatable the question contexts were, and there were also mixed experiences of applying SACs.

The feedback from staff in tertiary and alternative education schools/organisations represents a cohort of priority learners who face significant barriers to succeeding in education. There were concerns raised about the assessment approach, the barrier created by only two assessment events a year, and wider equitability issues within the sector that affect staff being able to adequately prepare learners for these assessments. The implications of the co-requisite aspect of these standards were also seen to compound the already concerning educational, social, and vocational outcomes for many of these students.

The findings from the learner survey suggest that students who had participated in the common assessment events generally had positive views about their readiness to participate and that the questions were relevant to them.

Summary of Objective One

Objective one of this evaluation aimed to understand how the design of the assessments can be further refined and improved. Looking at how the standards are performing in terms of levelling and alignment with other measures of attainment, the findings from the 2022 results data shows that the Literacy and Numeracy assessments continue (from 2021) to generally perform well against the standards. This is particularly true for Literacy (reading) and Numeracy, but the lower rates for the Literacy (writing) CAA indicate a need for further monitoring, and a consideration of what achievement levels should be expected. The results data for ākonga who participated in the Te Reo Matatini me te Pāngarau assessment is inconclusive and requires further piloting.

Examining the equitability of the new standards involved comparing achievement results for various sub-groups of interest and collecting the views of teachers and students. The patterns of achievement reveal inequities that are related to factors such as ethnicity and, particularly, decile. Regarding the Numeracy standard, teachers deemed this to be least equitable, largely due to the amount of literacy within the assessment.

There are potentially opportunities to improve the contexts and questions in the assessments to ensure they are more inclusive and relevant to a wide range of students. Or, to improve the technical and logistical issues that may be impacting on equity of access. However, it is likely that the factors that most significantly impact on patterns of achievement are deeply entrenched in an education system and the wider social and economic systems that perpetuate inequality.

The co-requisite nature of the Te Reo Matatini me te Pāngarau | Literacy and Numeracy standards pose a risk that less priority learners will achieve an NCEA qualification. It is, therefore, important to consider what processes, programmes, and interventions will be required to remove this roadblock. This could include timely and adequate learning support, academic mentoring, and careful monitoring of student progress towards achieving the co-requisite NCEA standards. Such approaches will require schools to allocate resource and put in place effective processes that involve students, teachers, and whānau/family.

Observations and recommendations for objective one

- The rates of those achieving the standard in Literacy (writing) is lower compared to Literacy (reading) and Numeracy even if they have achieved similar levels in e-asTTle. For example, those students who had scored 5A in e-asTTle had a 98% achievement rate in Literacy (reading), 96% achievement rate in Numeracy and only a 77% achievement rate in Literacy (writing). This warrants further analysis if similar rates of achievement are wanted across all three standards.
- Further clarity is required regarding expected achievement rates for each of the assessments in order to determine the appropriateness of the readiness indicators used in the pilot.
- The Numeracy CAA was criticised by many survey respondents for the amount of reading and writing involved. It will be important to ensure that any teaching and learning resources developed appropriately prepare students for these types of questions.

- The breakdown of achievement by ethnicity highlights again the wider inequity of the education system and the importance of maintaining, and potentially increasing, the wider capability building and support to remedy this. NZQA will continue to make further improvement to the assessments, however there is a limit to what the assessments can do without wider change occurring.
- It is important to understand more about the significant differences in achievement for low decile schools and tertiary/alternative education organisations. The second phase of this evaluation will seek the available e-asTTle data for these sub-groups, and further analysis will better determine whether the variation in results are reflective of the already known differences in achievement or if the CAA design is inequitable.

SECTION C: Findings for objective two

Objective two: understand how kura, schools and organisations are building their capability and what is needed going forward

The 2022 Pilot was scaled-up from the 2021 Pilot in terms of the numbers of schools, kura, and organisations taking part in the pilot. Capturing the range of experiences and the various perspectives about using the new standards and assessments is central to understanding what will best support the transition year in 2023 and full implementation in 2024.

Some guiding questions underpin objective two.

- How have schools understood student readiness and applied it?
- How have schools/organisations used the tools and resources to prepare students for the new Literacy and Numeracy standards and assessments?
- What has been the impact of the implementation of the new standards and assessments on teaching and learning (for schools/organisations and for kura)?
- How have schools/organisations and kura experienced the pilot, and what opportunities are there for strengthening implementation?

5 Understanding student readiness

Student readiness is a common theme in the survey findings. A recommendation from the 2021 pilot was that students should be entered for the assessment only when they are working at or above the upper level 4 of the curriculum. The analysis of available e-asTTle data for participating students indicate that a sizeable proportion of students who had e-asTTle scores below the minimum recommended curriculum level (NZC 4A) were put forward for the CAA. For Literacy (reading) the proportion was 16.6%, for writing 26.6%, for Numeracy 21.1%.

Many 2022 pilot participants made it clear they are still uncertain about the curriculum level boundaries and the standard required for students to pass or fail the CAA. Thus, these participants found it difficult to assess the readiness of students to sit the Common Assessment Activities (CAAs).

Other survey respondents linked understanding student readiness to student results and questioned how they could have a good sense of readiness without seeing and analysing student results in detail. Moreover, fifteen participants did not feel that diagnostic tools such as e-asTTle or the Progress and Consistency Tool (PaCT) were accurate forecasters of the curriculum readiness of their students. These reasons may explain why some pilot participant schools used the mid-year CAAs as a "pilot within a pilot" to test their own understandings of the new assessment and the curriculum knowledge and understanding of students.

A few pilot participants entered their extension class or classes as these students were deemed to be at the standard or "ready" through inclusion in this type of class, rather than teachers identifying readiness through a tool or professional judgement. In the case of an alternative education provider, students were allowed to choose whether they sat the

common assessment task or not. This provider noted that choice had a positive impact on student engagement.

The largest group of survey participants used the mid-year CAAs as an opportunity to test their own processes, understanding, and accuracy in identifying student readiness. This pilot participant explained, *"As this is a pilot, it was important for us to allow the whole cohort to try. This will give us an insight into which students to enter in the future"*, while another participant wanted to use the CAA to assess the readiness of their students and *"then use these results to triangulate our own data to see which is most useful."*

Within this largest group of survey participants were schools/organisations testing the curriculum progress tools. Seven participants commented on the difficulty of using PACT or e-asTTle to identify student readiness. A small number shared the view of one school who stated, *"we have not used PACT because it is too time consuming"*, while another participant commented they were mainly using teacher judgement as e-asTTle and PATs, *"which are not perfect indicators of readiness."*

The third group of schools/organisations admitted that assessing student readiness was difficult. One respondent explained that they had entered all students because "*we do not have a robust method for checking student readiness"* and noted that they had to *be "able to explain to students and parents why individual students were identified as ready and others were not."* Another respondent linked readiness to whether a student passed or failed the CAA and asked how they could "*determine readiness when there is no clear pass mark?"*

Two tertiary institutions noted that student readiness "*didn't matter"* in their context and one explained that:

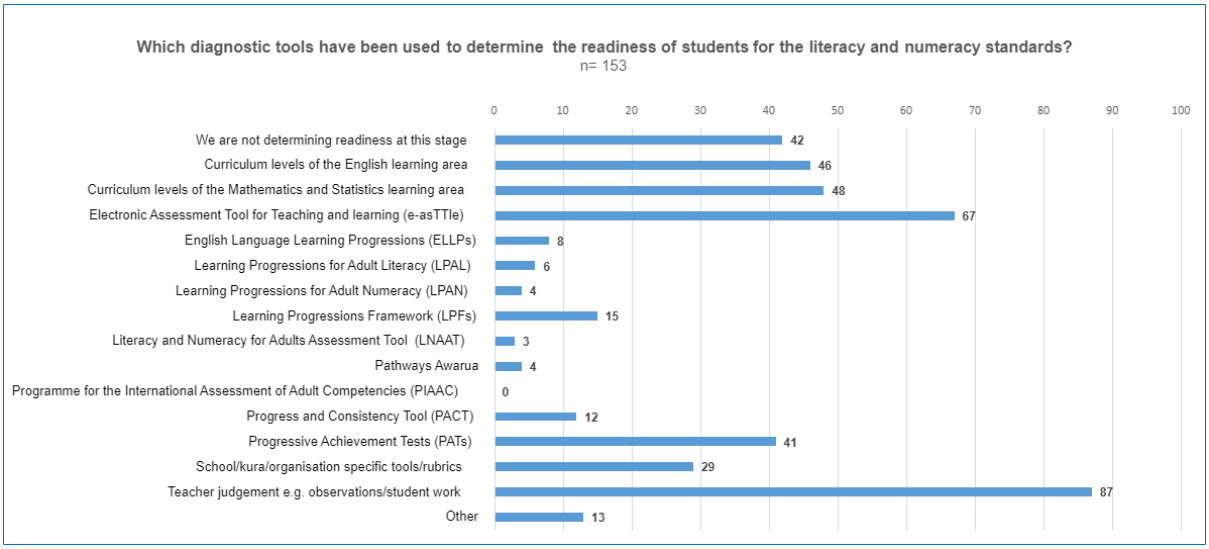
> *"Our course is 20 weeks, so if there are only two opportunities per year for the learners to do their assessment, then it doesn't matter if they are ready or not. They have to do it, otherwise they'll finish the course and leave without attempting their literacy and numeracy … The assessment needs to be more frequent than two times a year."*

Diagnostic tools for student readiness

Pilot participants were asked to consider the types of diagnostic tools they used to support their decisions about student readiness to sit the July Common Assessment Activities. They chose from a list of 14 different tools and benchmarks published in two official NCEA documents: *Unpacking Literacy 2022, Final version* and *Unpacking Numeracy 2022, Final version*. These documents are part of the NCEA Literacy and Numeracy Standards material available on the NCEA website. Participants could also disclose if they had not used any tools or could add other tools that were not listed in the questions.

All 153 pilot participants responded to the question and gave, in many cases, multiple responses with a total of 424 responses in all. These responses were categorised and are illustrated in Figure 23.

Figure 23: Diagnostic tools used to determine the readiness of students for the literacy and numeracy standards



Participants used a range of diagnostic tools. Figure 23 shows that teachers used fourteen out of the fifteen categories of diagnostic tools. One hundred and eleven (73%) of the participants who responded to this survey question were using a diagnostic tool or tools to assist them to identify student readiness. In judging student readiness to sit the CAAs, by far the greatest readiness indicator participants used was teacher judgement through observations and student work, which 87 (20%) respondents identified. This was followed by 67 (16%) who used e-asTTle, the second largest response. Almost equal numbers of pilot participants used the Curriculum levels Numeracy from the Mathematics and Statistics learning area (48) and Literacy reading and writing from the English learning area. (46).

Forty-two respondents stated that they did not use a readiness indicator for the mid-year CAAs. In explaining their reasons, many participants said that were conducting their own "pilot within a pilot" and wanted to gauge their students results before making any decisions about what tools might be useful or not. Several participants entered all their students whether they were ready or not. One participant explained the deliberate choices they made at their setting, for which student readiness was not the priority.

> *"We have knowingly entered all Year 10. We have used asTTle and teacher judgement extensively, but in some cases, we considered the experience of sitting in the gym in assessment conditions, logging in and coping was an important learning step that could be gained in a relatively low risk way."*

Recommendations for student readiness

- Further understanding of student readiness for the CAAs is required. Many teachers stated they did not understand the student level required, and some schools indicated they used the pilot to understand the level of readiness needed rather than use selection methods such as data from the recommended readiness tools.
- School, kura, and organisations could benefit from more focused PLD about the curriculum levels of the readiness indicators and the tools that can help them indicate student and ākonga readiness. Webinars and in-school PLD could support teachers' understanding and practice in this area.
- Some pilot teachers asked for specific support in unpacking the Big Ideas and Significant Learning areas that support the Literacy and Numeracy standards.
- More availability of resources and activities set at the CAAs levels of difficulty may also support teacher confidence and understanding of the levels required. To prevent teaching to the test, it would be necessary to identify activities and tasks that could broaden units of work and teaching and learning programmes (versus one-off activities aimed to "teach to the test").
- Many teachers indicated they would like the return of student papers as it would give them the opportunity to understand the marking of the papers and the standard that students need to reach to pass the exam.

6 Tools and resources

This section focuses on understanding how pilot participants used the literacy and numeracy tools and whether teachers have used the resources and tools to inform their teaching and learning. The previous section described the tools teachers were using to support decisions about student readiness to sit the assessments. This section looks at the resources available to support teacher and student learning and the feedback from participants about the usefulness of these resources. Additionally, pilot participants identified resources that would be useful to be developed in the future.

Key questions follow.

- What is the uptake of resources and tools by pilot participants?
- What resources are being currently used to support teachers and students and how useful have they been?
- What other resources have teachers identified that would be useful to support could support teachers and students to be ready?

6.1 Uptake of literacy and numeracy tools

Up-take of resources and their usefulness

Those survey participants who had used the resources from the NCEA Education website and the Te Kete Ipurangi (TKI) website were asked to rate how useful they had found them. The responses rating the usefulness of the NCEA (21 ratings) and TKI (34 ratings) website resources were overwhelmingly negative, with a rating of "not useful" being given by 18 (86%) respondents for the NCEA resource site and 33 (97%) respondents for the TKI resource site. One participant explained:

> *"We wanted specific resources for teaching and learning that the students could use. All the resources were more about the change and what it is, rather than teaching and learning actual resources."*

There were 106 responses from the 88 participants who identified the resources most useful to them: sample papers, exemplars, task/activities, and question banks based on the 2021 Literacy and Numeracy pilot. Table 11 illustrates the wide range of resources used with students as part of the Literacy and Numeracy pilot.

Table 15: Resources identified as useful by teachers for supporting student achievement

Types of Resource (collated from teacher responses)	Number of respondents who listed this type of resource	% of total responses received
Exemplars/Tasks/Pilot resources Resources such as questions, 2021 Pilot sample, Task activities	28	26%
Subject Associations All were Maths Association. 21 out of 22 named the NZAMT (New Zealand Association of Maths Teachers)	22	21%
Teaching and Learning programmes School programmes already made. Department run programmes. Resources for cross-curricula literacy and numeracy programmes	19	16%
Commercial resources 7 different resources. 12 named Education Perfect. Others were Mangahigh, Dr Frost Maths, Corbettmaths, Quindew, Secondary literacy resources	18	17%
Other School-based textbooks, syntax/grammar resources, ESOL resources, ARBs. Te Aka, Pathways, Google classrooms, speaking to other schools, Dyscalculia scanner	14	13%
Personnel PLD facilitators, 1 literacy and 1 Numeracy Co-ordinators, 1 Pilot Supervisor	5	5%

The resources survey participant (117 responses) most requested were CAA-related resources such as practical CAA tests, marking schedules, exemplars, and task activities that teachers could use directly in classrooms. Many participants (52) also requested marking schedules for these resources. This participant summed up the opinions of many explaining that "*the practice sample assessment was the best guideline for the style of questions and confirming the level.*"

Numeracy survey participants liked the NZAMT resources because the 45 task activities and numeracy questions closely mirrored the level of the Numeracy CAA and were useful for students and themselves. Some reading and writing participants identified Education Perfect as a useful source for students preparing for Literacy (reading) and Literacy (writing) CAAs.

One tertiary institution requested access to the same resources and support available to schools.

Useful resources participants want to use in the future

Participants were asked to consider what other resources they would like developed to support and improve student learning. One hundred and two participants responded, some with multiple responses.

Most participants (117) reiterated that they wanted more

> "*question banks, task activities, and previous CAA pilot papers at a comparable level of difficulty to the common assessment activities sat at mid-year, complete with marking schedules.*"

For some (27) participants there was an expectation that "ready-made" resources would be provided for them. While some participants asked for these to be developed as they "*were time-consuming, to do themselves*", it was unclear if some realised that at the beginning of the 2022 pilot, only one previous paper to use as a practice example was available. Nine responses wanted these resources to be digital, so students were able to practice and gain confidence in using and understanding the digital CAA style resources and using the digital assessment platform.

Some (17) survey participants made it clear that they are still uncertain about the curriculum level boundaries and the standard required for students to pass or fail the CAA. In addition, they found it difficult to assess the readiness of students to sit the CAAs. Many (28) wanted practical CAA material to improve their own understanding and practice and *"to see what assessment looks like and understand how they are graded."* Some teachers were anxious about the CAAs and said, "*it is impossible to improve our teaching practice when we have no idea as to what the assessments look like and how they are being marked."* Other respondents expressed concern about student understanding and engagement when they asked for "*parallel papers to prepare students in a realistic situation."*

Nineteen participants asked for the development of subject resources. Some of these suggestions were very specific and appeared to be similar to resources that should be available already in their own schools and organisations. These resources included syntax and grammar resources, resources about sentence structure, formal writing, writing conventions, and reading skills. Because of a lack of detail, it was sometimes difficult to tell from survey participants if these were cross-curricula skill resources or just resources

that would be useful and specific to the numeracy and literacy requirements of the standards.

A small (9) group of respondents expressed the need for professional development, and this was often linked to their belief that cross-curricula approaches, knowledge and skills, and resources needed more careful development with teachers. One numeracy school/organisation asked for "*examples of subject specific tasks and how they link to numeracy - especially in some subjects which tend to see this as being an inauthentic stretch – for example, Year 10 level languages*" and in one case the need for teachers *"to improve literacy pedagogical content knowledge*."

One school asked for

> "*a wide range of subject-related writing prompts, short video tutorials for teachers to use – in class or in professional learning – to increase knowledge and confidence with teaching writing, including the important mechanics/conventions of writing appropriately in that particularly subject area."*

Some (9) survey participants asked for more explicit training to understand the CAAs and how they are marked. There were requests from a few participants who asked for specific PLD on common pedagogical approaches to the types of literacy and numeracy used in the CAAs that will determine "best practice" in these areas. Another request came for workshops for principal nominees and support staff for PLD in digital assessment practice.

In the general question at the end of the survey, several participants reiterated their requests for more resources, particularly questions, activity tasks, and sample assessment papers with clear marking schedules. One participant succinctly summed up their purpose for this: "*the more we can prepare our students for the exam format, the less likely it is to pose a barrier to assessing their actual literacy abilities."*

Use of portfolios for Te Reo Matatini me te Pāngarau

There are two assessment options available for Te Reo Matatini me te Pāngarau, that is, ākonga can choose to demonstrate their achievement of the standards using a CAA or submitting a portfolio. In the first assessment event of 2022, all ākonga used the CAA. There has been a positive response from kura and secondary schools about the opportunity for ākonga to use portfolio, and while no portfolios were submitted in the first assessment event, the Ministry and NZQA remain committed to developing resources to support the use of portfolios.

Summary: tools and resources

This section provides the main observations about the perceptions of survey participants and the use of tools and resources that support teaching and learning.

There was an expectation that "ready-made" resources for the Common Assessment Activities would be provided for participants and the most-requested resources were CAA-related resources such as practical CAA tests, marking schedules, exemplars, and task activities that teachers could use directly in classrooms. The CAA-style resources were deemed valuable because these not only allowed students to focus on the style and level of the questions and tasks covered in the CAAs but also reduced student anxiety about a new assessment event. Teachers expressed uncertainty about the levels of difficulty, content, and style of questions asked in the CAA.

Many numeracy survey participants found the NZAMT the most useful source of resources for their schools/organisations. These resources have closely mirrored the CAA resources already provided.

Recommendations: tools and resources

- A range of case studies showcasing effective literacy and numeracy practice for students and tips for managing the CAA processes effectively would provide the opportunities for schools and other organisations to see and understand how the Literacy and Numeracy Standards and Common Assessment Activities are best applied in a variety of settings. These could be in the form of webinars, written examples, and materials.
- The development and sharing of different literacy and numeracy approaches should be encouraged. The setting up of spaces in which to collaborate and discuss resources and approaches, plus brainstorm solutions, may increase confidence levels and feelings of support. The use of online hubs and professional learning groups are good ways to encourage this and might improve students' levels of preparedness for the assessments. Likewise, the use of online hubs and professional learning groups are good ways to create effective professional practice.
- Illustrating different ways of incorporating activities into units of work would prevent "teaching to the test." A cross-curricula approach to incorporating literacy and numeracy skills will promote rich curriculum experiences while also possibly preventing the use of ability groupings and the narrowing of the curriculum.
- Further resources to support the preparation and understanding of the levels required would make teachers feel more supported. This could include marking schedules, exemplars and task activities, and question banks. This will require careful design and messaging so that these are used in ways that support the overall purpose of the standards.
- Regarding the option for portfolio assessment for ākonga in kura and secondary schools, NZQA has already begun developing more specific resources to assist kaiako and ākonga to understand how learning can best be captured and

presented using portfolios. Feedback from assessment experts notes the additional complexity to undertaking portfolios and the time this requires to grow and build capability, so this will be an area of ongoing development to monitor closely.

- For kura and organisations delivering Te Reo Matatini and/or Pāngarau programmes, further support, tools, and guidance are required to set these assessments within the context of their marau ā-kura, the redesign of Te Marautanga o Aotearoa (Te Tīrewa Marautanga), and the development and implementation of Hei Raukura Mō Te Mokopuna.

7 Reported impact of the implementation of the Te Reo Matatini me te Pāngarau | Literacy and Numeracy standards and assessments on teaching and learning

In the Literacy and Numeracy Survey, teachers were asked about the existing system and programme changes that they are making, or intending to make in the future, to support students to achieve the standards. Currently, there is a dichotomy between schools and organisations that have undergone or are undergoing a "systems shift" in literacy and numeracy in their schools and are actively preparing their students for Common Assessment Activities and those who appear to be waiting to see how changes or support work before implementing them.

It was unclear in many responses whether the changes were intended to be introduced, were being implemented, or had been implemented. It was evident from the responses that participants were at different points along the change continuum. This ranged from doing little or nothing different to prepare students for the CAA to introducing major changes that would affect their entire school/organisation.

The first part of this section describes the different levels of change to systems and programmes that participants reported. What follows this are comments about the impact on pedagogy and the impact of SAC (Special Assessment Conditions) changes.

7.1 Changes to systems and programmes

Participants were asked an open-ended question about changes being introduced in their schools/organisations; schools/organisations could give multiple answers. One hundred and twenty-three participants responded to this survey question, providing 162 changes that were either intended or had been made. These survey responses were coded in three main ways:

- in-depth system changes, which affect the whole school/organisation
- programme changes, which affect smaller units such as individual departments or faculties, often the result of individual teacher actions
- responses that indicate no or very little change at all.

Changes to systems

System change are those changes intended to affect the whole school/organisation to improve the reading, writing, and numeracy of their students. They often are changes not only to teaching and learning but also to processes and systems within the whole school/organisation.

Twenty-five (15%) of the changes reported were system-wide changes. These changes respond to the changes to literacy and numeracy and are multi-faceted. One school highlighted this when describing the school-wide literacy programme as a:

> *"School-wide audit at the beginning of 2022, school-wide PD – term 1 2022 – and plan to implement timed-writing tasks across the curriculum and a school-wide collection of writing data* [for moderation purposes]."

The following comment from another participant sums up a systemic approach to change in their school:

> *"We have begun a concentrated cross-disciplinary literacy programme, beginning with disciplinary specific vocabulary – front loading, low-stakes quizzing, embedded retrieval practice, sequenced literacy, and numeracy language/conventions from years 7-10 (a work in progress). This will have a big impact on increasing student competency in literacy and numeracy."*

One school described their "*move to a LPF progression approach across the Junior school*" but also identified the more substantial changes that will come with "*the new curriculum and the RAS as there are other things influencing our programme including the increasing lack of mathematics in the incoming Year 9s."*

Two schools talked about their school-wide literacy strategy, the embedding of literacy pedagogy into units of work combined with how they "*are working with our Kahui Ako to develop and reinforce literacy skills from year 0 upwards."*

One tertiary institution described the scale of change required to support students to prepare for the Numeracy common assessment task.

> *"Mapping of the Adult Learning Progressions against the NZ Maths resources and the NZC to ensure teaching is at the correct level. Development of differentiated activities - in progress. Increased use of Pathways Awarua at the learners' level and gap areas. Waiting on the results of the assessments to fully identify this. Planned PLD for a few areas that we know need attention."*

Changes to programmes

The largest group of intended or actual changes reported were those within individual year level programmes or individual departments, making up 66% (107) of the changes survey participants identified. Some were specific small-scale changes whilst others were in the process of being introduced and involved professional development over time.

One of the significant changes made to programmes has been a more deliberate focus on literacy and numeracy. This has included the appointment of literacy and numeracy co-

ordinators and, in some cases, committees. For some (17) schools/organisations, it has resulted in the introduction of diagnostic tools and more systematic collection of data across year levels. One school explained that after establishing a literacy co-ordinator role, they have "*implemented more systematic testing of student literacy levels and are developing and trialling interventions for students who are below expectation.*"

The largest group of programme changes reported was the deliberate introduction of specific literacy and numeracy skills. Thirty-one participants identified this deliberate introduction has targeting of skills at an individual student level, and across the whole school/organisation. Twelve participants used a combination of commercial resources such as Education Perfect and the Write that Essay programme, as well as using existing curriculum resources such as "*punctuation, sentence structures, fluency, sentence types, and writing for purpose and audience.*" Another common change has been the introduction of starter activities such as contextualised problems in Maths and practice tasks in literacy.

Changes yet to be made

Thirty (19%) of the respondents reported having not implemented any changes in their schools /organisations as they are waiting for the results of the CAAs. Their responses indicated that it was either too early to comment on the impact of the new Literacy and Numeracy standards on students, or there was none as they were yet to see results (or analyse student papers), or the questions were not relevant as they had not introduced any changes in their teaching and learning programmes yet. One participant explained that there were *"no specific actions. We do not teach to a test."*

Sixteen schools and organisations intended to analyse their results, identify gaps in student knowledge and skills, and then decide how they will improve their literacy and numeracy programmes. Waiting to see student results and student CAA papers when they are returned is a persistent theme in responses to other questions in the survey. One participant explained that they "*want explicit feedback on our school to identify trends in errors*", while another school requested individual feedback as *"it would be nice to know why our individuals are not achieving, as opposed to a country-wide pilot cohort."* One participant who was part of the 2021 literacy and numeracy pilot said that that they "*wanted our students' assessment scripts returned with justification for why they passed or failed."*

In the general question at the end of the survey, seven participants reiterated their comments about the importance of students getting their assessment papers back. One survey participant asked for a forum so they can provide feedback about the actual CAA questions. This knowledge of the CAA contrasts with others surveyed who had not seen the CAA but were requesting strongly that it is important; generic feedback is not enough.

7.2 Impact of changes on pedagogy and students

Comments from participants about the impact of the implementation on pedagogy were very limited. A tertiary institution commented that they had "*seen a noticeable improvement in the Adult Numeracy Assessment Tool results since introducing Prime Maths.*"

From one school, the introduction of cross-curricula literacy and numeracy programmes throughout the junior school – with significant subject and timetabling changes – have the potential to impact on teacher pedagogy and student literacy and learning, but they are not yet fully implemented. Another school commented that "*we can already see how much better our Year 10s are after two years of intentional and explicit language teaching: some are even better than our Year 13s.*"

As expected so early in a large change process such as the literacy and numeracy pilot, the impact of changes on students are tentative and are largely based on early observations the respondents made. There were 17 (out of 80) positive comments about the impact of the changes on their students. Five comments focused on the growing confidence and engagement of students because of the initiatives introduced in their schools/organisations. One pilot participant said, "*students have engaged with the technology and enjoy the specialised programs,*" while a tertiary institution reported that changes, they have made have led to a "*high percentage learner engagement with Pathway's Awarua and a strong movement to TEC threshold in the LNAAT Progress assessments.*"

Two participants commented tentatively on an improvement in student achievement, "*Despite not having official results yet, our internal data suggests improvements in reading comprehension.*"

Many responses indicated that it was either too early to comment on the impact on students or there was none as they were yet to see results, or analyse student papers, or the survey questions were not relevant as they had not introduced any changes in their teaching and learning programmes yet. Most of the seventeen participants who reported positive impacts on students had the proviso that it was too soon to see definite trends or patterns. However, almost all these responses described small pedagogical changes to students' learning.

There were two responses that provided a different perspective on the impacts of the new Literacy and Numeracy standards. Both expressed their fears of possible consequences of the CAAs. One said:

> *"We are very afraid that teaching and learning will be dropped in favour of desperate attempts to get the students through the test and the long-term outcomes will be even more detrimental than they are now."*

A school offering Numeracy was concerned about the possible impact on their mixed-ability classes. This participant said:

> *"There is some concern that once the examination has been completed mid-year, it will it divide the students into the groups of 'passed Numeracy' and 'not passed Numeracy'. Schools will need to find ways of supporting the 'not passed' by introducing deliberate actions to support their learning to enable them to pass Numeracy without making them feel singled out as a different (lower ability) group or affecting their mana and their self-confidence in Maths."*

7.3 Impact of changes on students who access special assessment conditions (SAC)

NZQA does not hold data on the number of students who were afforded SAC provisions (for example, readers and writers). Pilot schools were advised that a formal SAC application was not required for SAC support to be provided for the standards.

Seventeen participants described the implications of (SAC) students and on students and school processes. The changes teachers and teacher aides made in some schools appear to have better prepared their students to sit the CAAs. One participant described the "*growing confidence in our SAC students as they were allowed to sit the CAAs like everyone else in their class.*" However, the impact of preparing these students for the CAA has been considerable on school resources and time. One pilot participant school described the detailed changes it has introduced for supporting students who access special assessment conditions.

> *"We had to process a huge number of possible SAC applications (over 12 in Y10 and 12 in Y13). We made provision where we could (printed copies, special paper, special screens, separate accommodation, and so on) and applied for reader/writers where appropriate. Providing this many support staff was a huge challenge both to arrange and implement."*

Other pilot participants also gave more detailed comments about decisions to review their SAC programmes, use data to reconsider how they identify students who access special assessment conditions, and re-evaluate their school-wide data tracking system "*with respect to students at risk for [Literacy and Numeracy].*"

In the general question at the end of the survey, some comments were made about the time and difficulty of students who access special assessment conditions supporting students who access special assessment conditions with teacher aides and resources to sit the CAAs, but also there were a small (5) number of requests for more PLD in how to provide the right type of support for ELL students and those with learning difficulties.

Summary: changes and the impacts on teaching and learning

A small number of schools/organisations described more complex system changes they are introducing in their schools. Changes to teaching and learning programmes and school systems and processes characterise this change process. Examples of changes involved the introduction of cross-curricula literacy and numeracy programmes in junior programmes, the appointment of literacy and numeracy co-ordinators, dedicated literacy and numeracy lessons, and a rebalancing of junior schools' timetables to meet the demands of the Literacy and Numeracy programmes.

The most common changes introduced in schools/organisations in response to the pilot were small scale changes to teaching programmes. These were often dedicated teaching of specific numeracy and literacy skills, usually as “starters” to lessons and the use of CAA style practice questions and tasks where either they were available, or teachers had created them.

Some schools/organisations have begun to introduce systems and processes to support their SAC and learning support students sit the CAAs. Participants reported that this has been a time-consuming process leading up to the mid-year CAAs. As students had not received their assessment results, few responses have been received about the effects of the literacy and numeracy changes on students. Those schools/organisations waiting for the results of the CAAs intend to analyse their results, identify gaps in student knowledge and skills, and then decide how they will improve their literacy and numeracy programmes.

The impact of changes introduced within schools/organisations on students is small. Participants were tentative in their responses to describing these changes and their impact on students, which is to be expected so early in such a large change process as the literacy and numeracy pilot. However, the responses received from those who did comment are largely positive.

Recommendations: on schools/organisations' teaching and learning

- It is important to reinforce the vital role of teachers in supporting student success. This includes students being engaged in effective teaching and learning and being well prepared for the CAAs.
- Positive experiences should be highlighted within resources such as case studies. Despite this being early in the process for most, some teachers have commented on increased confidence and student engagement. Only engaging students when they have a good chance of achieving the standard should be a key message.
- Schools should be encouraged to consider what specific actions are needed to support student learning and prepare them in ways that gives them a good chance to achieve the standards.
- Webinars and case studies could provide a useful framework for schools/organisations implementing the standards in 2023 and 2024, and they could highlight the need for schools to prepare students for the assessment events.
- Deeper insight into the experiences and challenges of certain groups such as Realm country schools, tertiary and alternative education institutions, English language learners, and those requiring special assessment conditions is recommended. This will reveal any specific issues or supports that they may require.

8 How have kura, schools/organisations experienced the pilot and what opportunities are there for strengthening implementation?

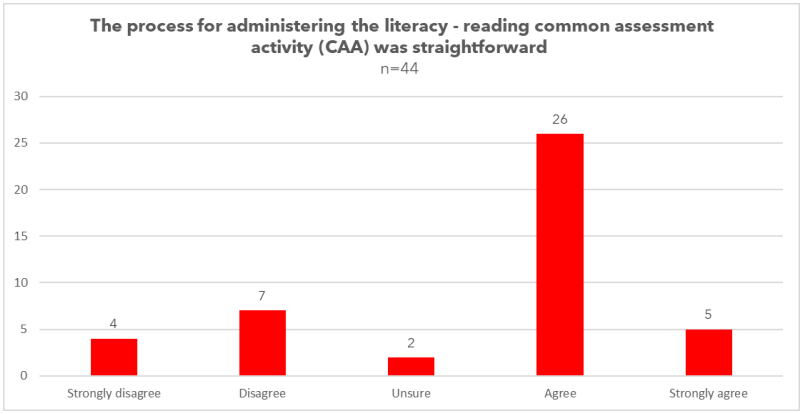
This final section captures participants experiences of piloting the new standards and administering the common assessment event in July. This helps understand the extent to which teachers and ākonga/students are ready for the new assessments and what might help them be successful in the future. The experiences of piloting the Literacy and Numeracy standards are discussed first, and this is followed by a discussion about the experiences of kura and schools who piloted the Te Reo Matatini me te Pāngarau standards.

8.1 Perspectives on piloting the literacy and numeracy standards

Participants identified a range of concerns and problems in the literacy and numeracy pilot, leading up to and following the assessment events. For many of these issues, they also offered solutions. These experiences should support the review and refinement of the assessments and guide the provision of the supports for teachers working with their students to achieve success.

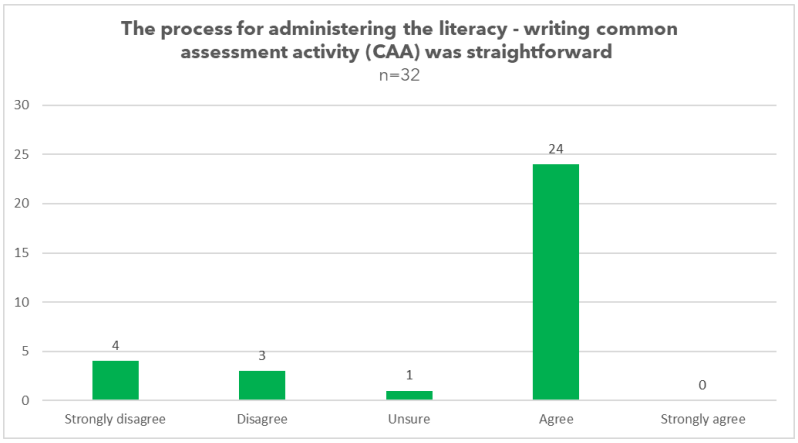
Reading, writing, and numeracy survey participants were asked to rate their level of agreement or disagreement with the statement that the process for administering the Common Assessment Activity was straightforward. The responses for each of the three assessment areas are detailed in the charts below.

Figure 24: Agreement ratings for "the process for administering the literacy (reading) CAA was straightforward"



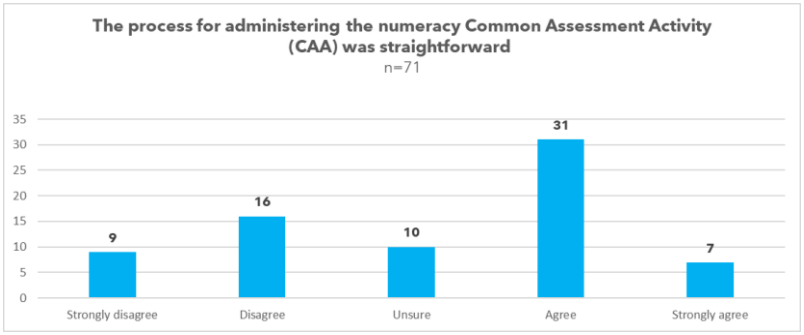
68% (31) of the 44 pilot participants interviewed in the survey agreed or strongly agreed that the process for administering the Literacy (reading) CAA was satisfactory. While two schools/organisations were unsure about the statement, 27% (12) disagreed or strongly disagreed with the statement.

Figure 25: Agreement ratings for "the process for administering the literacy (writing) CAA was straightforward"



75% (24) of the 32 pilot participants agreed that the process for administering the Literacy (writing) CAA was straightforward. While one school/organisation was unsure about the statement, 22% (7) disagreed or strongly disagreed.

Figure 26: Agreement ratings for "the process for administering the numeracy (CAA) was straightforward"



53 % (38) of the Numeracy participants agreed or strongly agreed that the process for administering the Numeracy CAA was straightforward compared to 34% (25) who disagreed or strongly disagreed with statement. 13% (10) of those surveyed were unsure about this statement, and four participants did not complete this question.

While there are some differences in the responses between the three assessment areas, there was overall agreement that the process for administering the Common Assessment Activities was straightforward. However, there were less favourable ratings from those who were piloting the Numeracy standard compared to Literacy (reading) and Literacy (writing). Nearly half of participants offering Numeracy either disagreed/strongly disagreed or were unsure about the statement.

8.2 Strengthening Implementation

Participants were asked for their perspectives on strengthening the implementation of the Literacy and Numeracy standards and assessments. Most responses focused on issues, concerns, and problems, which is understandable as participants grapple with the changes that are inherent in a literacy and numeracy pilot of this complexity. However, most respondents also suggested ways to solving these issues or problems.

Of the 90 participants who responded to this question, 30 identified multiple issues around the administration of the Common Assessment Activities. Their descriptions are useful in understanding why pilot participants may have not agreed that the administration of the CAAs was straightforward or were unsure about the statement. One participant said that schools were not ready for *"the enormity of the challenge"* in introducing the Literacy and Numeracy Standards and conducting the assessments. Others found implementing the assessments incredibly difficult and time-consuming. Some of these multiple administrative issues are described by one survey participant who said:

> *"The major issue is the administration of the test. Setting up rooms for over 150 students to take tests over three days is huge. Classes needed to be re-roomed, teachers were needed to supervise, then other teachers needed to cover their classes. Teacher aides were needed to be reader/writer (there were far too few available). Supervisors needed to be trained. SAC provision took weeks to set up, with batteries of testing completed on dozens of students and many teachers writing reports on students. A teacher aide timetable had to be arranged. Calculators and rulers needed to be handed out. Time needed to be taken in class to ensure every student had a login setup (this takes many lessons, as some students are absent on any given day). This all led to students missing hours of other classes, which is a problem when they have other upcoming assessments."*

Although many of the administrative concerns and problems are interconnected, they have been separated so technical issues, resourcing, workload, budget, and staffing issues can be clearly discussed.

Technical and Logistical issues

While most participants appeared to have few problems with students and ākonga accessing and using Assessment Master (AM), other schools, kura, and organisations, identified problems and suggested some changes to AM or the instructions being provided for AM. Issues identified follow.

- Logins not working or survey participants found them overly complicated.
- Accessing the NZQA emergency number was identified as difficult.
- In emergencies such as loss of power/no Wi-Fi, paper options were not allowed.
- Several supervisors locked out the day before with a long time spent trying to sort out the problems, which affected the set-up.
- Student instructions about the CAA were not received in time so the school had to improvise.

- A couple of schools experienced problems with the size of Chrome's small screen. Students had to scroll between the question and graph and retain information from the graph without having to see it.

Survey participants often offered solutions to specific problems they had experienced with using Assessment Master.

- Allow all assigned students in a particular class to a particular room at once rather than student by student. Having a filter field where form codes could be imported would make the allocation of exam rooms much less time consuming.
- Update filters in AM, for example, blank entries.
- Use the "backup" as the way for students to log in. Students could just be issued with their NSN number and a validation code once and only for the assessment time.
- Easier access to training on the use of AM and differentiation between the real platform and the training platform.
- Extra supervisors' logins took several days to get and could not be activated.
- Exam Supervisors logins meant they were "super users" so could inadvertently change the exam room of candidates. Give them a limited login.
- Cumbersome for exam supervisor resets to go through the principal nominee especially in large schools. Slow to come through if at all. Allow a technical lead position to be established.
- Need the ability to export data from AM (instead of just copying out). Good for seating plans.
- Notification data in AM. Need further identifying data such as a name instead of just the students' NSN number.
- Ability to download the students work as a PDF in bulk so allows students to receive individualised feedback on their work.
- Guide written for IT staff.

Resourcing

Comments were made about complicated and time-consuming administration leading up to the Common Assessment Activities. There were some (8) negative comments about the training videos for digital assessments and teacher supervision. One school found them inaccessible and felt they needed to be able to be "faster." Another school described the webinars as "*substandard"* with poor presentation that included far too much information. A suggestion from one school was that these videos need to have been made with "*non-computer confident staff in mind*." Another school requested "*more online opportunity to become familiar with the platform"*, which contrasted with another organisation that appreciated the opportunity to practice on the digital platform in preparation from the assessment.

For one school, the biggest challenge was students who did not have a device or a suitable device. There was an expectation for the school to provide this. It was suggested that funding is allocated to provide devices to students which would make implementation a lot smoother.

Workload, budget, and staffing

Some (8) schools and kura regarded the administration of the CAAs as challenging in staffing and time, and as one participant said, "*We want staffing for the actual assessments. This CAA was very disruptive for our staffing and timetable.*" Another respondent asked for resources so a small expert team outside of the school "*could run the whole show.*"

Some (9) participants commented on "massive" increases in workload and time that the administration of the common assessments entailed.

- Entering students for assessment.
- Training supervisors.
- Training students to use the digital platform.
- Preparing rooms.
- Changes in timetables.
- Increased staff to supervise (in some cases, two teachers per class).
- The difficulties in finding SAC supervisors and Learning Assistants.
- Systems for students who finish early.
- Systems for students who want extra time.
- Funding. Resourcing and going over the budget provided were identified as concerns for larger schools.
- IT staff available in case of technical issues.

One school reflected the comments of others (4) when they said that they "*may need to send home some groups just so we can have staff to be able to do this.*" Organising whole year-groups into assessments while other lessons continue was described as problematic, likewise for schools without large spaces like halls. A few (7) respondents suggested that limiting the CAAs to two-a-week or allowing more assessments throughout the year would contribute to alleviating or lessening some of these problems.

Participants discussed other issues in the general feedback question. These include suggestions for resources, issues regarding SACs, the return of student papers, and the need for professional learning and development. This evaluation report discusses these issues in other areas, taking into account the comments in this survey question.

8.3 Perspectives on the implementation of Te Reo Matatini and Pāngarau standards

It is critical that the development and ongoing refinement of assessment activities for the Te Reo Matatini and Pāngarau standards is more closely aligned to the strategic vision outlined in Hei Raukura Mō Te Mokopuna, the Ministry's strategy for Te Reo Matatini me te Pāngarau. Hei Raukura Mō Te Mokopuna and its accompanying draft action plan focus on promoting and enhancing "a mātauranga and kaupapa Māori understanding of te reo matatini and pāngarau across the curriculum" and ensuring teaching and learning is meaningful to ākonga, their whānau, hapū, iwi, and communities.

Feedback – that kaiako shared with external PLD providers and advisors from the Ministry and NZQA following the first assessment event for Te Reo Matatini me te Pāngarau standards – includes:

- ensuring parity and equity for mātauranga Māori
- reinforcing the value of te reo Māori and pāngarau within learning
- developing quality resources where and when needed to grow kaiako capability and ākonga learning success
- implementing robust and responsive PLD support
- affirming the place of whānau, hapū, and iwi as active contributors to and key influencers of ākonga learning, achievement, and progress.

Summary: experiences and opportunities for strengthening implementation

Respondents mostly agreed that the process for administering the Common Assessment Activities was straightforward. Numeracy was the assessment area that was less positively viewed, with nearly half of participants either disagreeing/ strongly disagreeing or were unsure about the statement.

Most participants feedback on how implementation of the standards could be strengthened, focused on issues, concerns, and problems, which is understandable as participants grapple with the changes that are inherent in a literacy and numeracy pilot of this complexity.

The problems most identified by participants were administrative concerns such as technical issues, resourcing, workload, budget, and staffing issues. While most participants appeared to have few problems with students accessing and using Assessment Master (AM), other schools/organisations identified problems and suggested some changes. These issues included: logins not working, accessing the NZQA emergency number, paper options not allowed. Some schools had to improvise because student instructions about the CAA were not received in time.

Some participants made negative comments about the training videos for digital assessments and teacher supervision. Other participants commented on significant increases in workload and time that the administration of the common assessments entailed. There were also suggestions that students may need to be sent home so that the CAAs can be adequately staffed, and enough classroom space can be provided.

Recommendations: opportunities for strengthening implementation

- Schools will need to consider logistics carefully for further assessment events. Aspects such as technical issues, resourcing, workload, budget, and staffing were identified as issues in their schools. Increased support for school administration and examples of effective administrative practice would be beneficial for all schools.
- It is important that leaders and teachers understand the purposes of the pilot and their role in preparing students for the Common Assessment Activities. Several participants did little to prepare their students or were using the assessments as a way of identifying students' strengths and weaknesses before changing teaching and learning programmes.
- Some participants indicated there was a requirement to administer the assessment at the same time for all students, and that it made logistics difficult. Some consideration could be given to allow the sitting of Common Assessment Activities in sessions, so adequate classrooms and staffing are available.
- Embedding the NCEA Te Reo Matatini and Pāngarau standards more firmly within this vision of Hei Raukura Mō Te Mokopuna will create greater cohesion with the overarching focus of Te Tamaiti Hei Raukura and Te Tīrewa Marautanga. This would support a necessary shift in thinking to consider more broadly the definition of what ākonga achievement and progress looks like, sounds like, and feels like. At present, the standards and assessments could be perceived more as an attainment measure and qualification step for ākonga.

Summary for Objective Two

The second objective for this evaluation aimed to understand how kura, schools, and organisations are building their capability and what is needed going forward. A range of experiences and perspectives on piloting the new standards and assessments were gathered through:

- the teacher survey
- direct discussions with kura pilot participants
- feedback gathered at a hui about Te Reo Matatini me te Pāngarau, involving multiple stakeholders.

The pilot schools and organisations reported different understandings of readiness and different approaches for determining student readiness for participating in the Common Assessment Activity. This is reflected in the numbers of students who participated in the assessments that had e-asTTle scores below the minimum recommended level. Many schools and organisations were using the mid-year assessments as an opportunity to test their own processes and accuracy in identifying student readiness. The diagnostic tools teachers most commonly used to assist with determining readiness were teacher judgement and e-asTTle.

Pilot participants were asked about the tools and resources available to prepare students for the new Literacy and Numeracy standards and assessments, and views about the resources on the NCEA and TKI website were largely negative due to lacking relevance to actual teaching and learning. The participants' most-requested resources were CAA-related resources such as practical CAA tests, marking schedules, exemplars, and task activities that teachers could use directly in classrooms. The responses provided some insight into the expectations that "ready-made" resources would be provided for them, not just to assist student learning but also to help teachers better understand the Common Assessment Activities.

The impacts on teaching and learning due to implementing the new standards and assessments varied. A small number of schools and organisations described relatively complex whole-school systems change focused on teaching and learning programmes and wider systems and processes. Smaller scale changes to teaching programmes were more commonly reported. Some schools and organisations were waiting for the results of the CAAs before starting on any internal changes. Impacts on students were minimally described, which is not unexpected given that schools and organisations are in the early stages of implementing this large change process.

The final part of this section reported how the pilot had been experienced and what the opportunities were for strengthening implementation. While there was relatively strong agreement that administering the CAAs was straightforward, a range of mainly logistical and resourcing issues were also raised.

Appendices

Appendix 1: List of Figures and Tables

List of Figures

Figure 1: Literacy and numeracy standards achievement rate compared to e-asTTle scores 25

Figure 2: Literacy and numeracy achievement rates for students below and above readiness indicators 25

Figure 3: Agreement ratings for "the literacy (reading) standard is easy to understand" 28

Figure 4: Agreement ratings for "the CAA for literacy (reading) allowed for valid and reliable assessment" 28

Figure 5: Agreement ratings for "the literacy (writing) standard is easy to understand" 28

Figure 6: Agreement ratings for "the CAA for literacy (writing) allowed for valid and reliable assessment" 28

Figure 7: Agreement ratings for "the numeracy standard is easy to understand" 29

Figure 8: Agreement ratings for "the CAA for numeracy allowed for valid and reliable assessment" 29

Figure 9: Achievement by ethnicity for literacy and numeracy assessments 32

Figure 10: Achievement by gender for literacy and numeracy standards 34

Figure 11: Achievement by gender for Pāngarau and Te Reo Matatini standards 35

Figure 12: Achievement by decile for literacy and numeracy standards 36

Figure 13: Achievement by decile for Pāngarau and Te Reo Matatini standards 38

Figure 14: Achievement for English language learners compared to overall achievement rates 39

Figure 15: Achievement for tertiary/alternative education compared to overall achievement rates 40

Figure 16: Achievement by Realm countries compared to overall achievement rates 41

Figure 17: Agreement ratings for "the CAA literacy (reading) is equitable for all learners" 49

Figure 18: Agreement ratings for "the CAA for literacy (writing) is equitable for all learners" 49

Figure 19: Agreement ratings for "the CAA for numeracy is equitable for all learners" 50

Figure 20: Agreement ratings for "the CAA for literacy (reading) was equitable for (ELL)" 53

Figure 21: Agreement ratings for "the CAA for literacy (writing) was equitable for English language learners (ELL)" 53

Figure 22: Agreement ratings for "the CAA for numeracy was equitable for English language learners (ELL)" 53

Figure 23: Diagnostic tools used to determine the readiness of students for the literacy and numeracy standards 62

Figure 24: Agreement ratings for "the process for administering the literacy (reading) CAA was straightforward" 75

Figure 25: Agreement ratings for "the process for administering the literacy (writing) CAA was straightforward" 75

Figure 26: Agreement ratings for "the process for administering the numeracy (CAA) was straightforward" 75

List of Tables

Table 1: Numbers of pilot schools, kura, and organisations who entered students in the June/July 2022 Common Assessment Activity 14

Table 2: Number of participating ākonga/students in CAA by assessment standard (June/July 2022).... 14

Table 3: Proportion of students participating in CAA by assessment standard combination type (July 2022) 15

Table 4: Type and numbers of respondents to the mid-year Literacy and Numeracy teacher survey 16

Table 5: Achievement by standard (July 2021 and July 2022 results) 21

Table 6: Achievement by standard and year level for secondary students 23

Table 7: Achievement by ethnicity for literacy and numeracy assessments 32

Table 8: Achievement by ethnicity for te reo matatini me te pāngarau assessments 33

Table 9: Achievement by ethnicity for Te Reo Matatini me te Pāngarau 34

Table 10: Student achievement of literacy and numeracy standards by decile (count and percentage) .. 36

Table 11: Relationship between e-asTTle scores and achievement of the numeracy CAA for ELL .. 40

Table 12: Achievement results for pilot sub-groups 42

Table 13: Response rate for the mid-year learner survey 44

Table 14: List and count of further student comments regarding the reading assessment 47

Table 15: Resources identified as useful by teachers for supporting student achievement 64

Appendix 2

Psychometric analyses Summary Report:

Background:

Statistical analyses have been completed on the July assessments for Numeracy, Literacy Reading, and Literacy Writing. These analyses include applying IRT approaches.

The purpose of the analyses is to investigate how the assessments performed from a statistical/psychometric viewpoint: Did each item contribute information to the measurement of student achievement?

The reports NZCER produced for NZQA summarise the statistical findings and make recommendations for consideration when developing future assessments. The reports are produced for the National Assessment Facilitators for Literacy and Numeracy and the contracted assessment developers.

The reports, together with the reports to be produced for the September assessments, will be used to inform the 2023 pilot assessments.

NZQA will provide a consolidated summary of the statistical findings from the July and September assessments by December, including commentary on any differences between the two assessments for each standard and strengths and weaknesses. NZQA will also provide information on any changes to the 2023 pilot assessments arising from the evaluation of the psychometric analyses at this time.

Psychometric analyses have not yet been completed for the Te Reo Matatini and Pāngarau assessments. The small numbers participating limit the usefulness of a statistical methodology; however, it is expected that these will be produced for the November assessments. Consideration is being given to the type of analysis that will be useful to inform the development of these assessments going forward.

Summary findings:

Overall, the analyses indicated the items in the Numeracy, Literacy Reading, and Literacy Writing assessments performed well.

Overarching recommendation:

Inherent in assessment is idea of measurement error and reliability. Reliability captures how likely a student would be to get the same score on consecutive assessments of the same standard. Measurement error captures how precise the assessment instrument is in locating achievement on a scale. The impact of measurement error should be considered when determining what score students are expected to score, overall, and on each separate outcome.

Below is a summary of the key points from the three Psychometric analyses:

Assessment	Key points	Recommendations for item developers for future assessments	Considerations / recommendations for future analysis
Numeracy	• The assessment (30 items) was divided into three Outcomes, each represented by 10 items. • The items in Outcome 1 were relatively easier (with one exception) than those in Outcomes 2 and 3. The items in Outcome 2 tended to be a little more difficult than those in Outcome 1, and the items in Outcome 3 tended to be relatively more difficult still (also with one exception). • To achieve the standard, students were required to correctly answer at least 17 items as well as correctly answer a minimum number of questions related to each Outcome. For those students who correctly answered at least 17 items but did not achieve the standard, the majority did not meet the required minimum level of achievement on Outcome 3. • The scores associated with each outcome are less reliable than the overall score. This means that care should be taken when using the outcome scores to make decisions about individual students. • Analysis of the data revealed that overall, the assessment performed well:	• When determining the required minimum scores for each Outcome and the minimum score required overall, ensure that the impact of measurement error is considered.	• Analyse the "distractor quality" for multiple choice items (Sometimes a response option can have characteristics – for example, wordiness, vocabulary – that might make it overly appealing to certain groups of students.) This could not be completed for the July assessments due to data retrieval issues. • Analyse time students spent on the assessment and scores. Early analysis of Assessment Master log-in data indicates that the median time students took to complete the assessment was 82 minutes (longer than the recommended 60 minutes); however, the data should be treated with caution given that students didn't consistently log off when they completed the assessments.

	○ The Cronbach's alpha (KR-20) was 0.90, indicating a high level of internal consistency among the items. ○ The difficulty measures of the items were spread across much of the distribution of student achievement. ○ There was evidence to suggest some lower-achieving students guessed their response to some of the multiple-choice items.		
Literacy Reading	• The assessment (35 multiple choice items) was divided into three Outcomes, with Outcome 1 represented by 15 items, Outcome 2 represented by 10 items, and Outcome 3 represented by 10 items. • The items in Outcome 2 tended to be more difficult than those in Outcome 3, which in turn tended to be more difficult than those in Outcome 1. • The items in Outcomes 1 and 3 covered a greater difficulty range than those in Outcome 2. • To achieve the standard, students were required to correctly answer at least 20 items as well as correctly answer a minimum number of items in each outcome. All students who correctly answered at least 20 items also met the required level of achievement on the outcomes.	• When determining the required minimum scores for each Outcome and the minimum score required overall, ensure that the impact of measurement error is considered.	• Analyse the "distractor quality" for multiple choice items. (Sometimes a response option can have characteristics – for example, wordiness, vocabulary – that might make it overly appealing to certain groups of students.) This analysis could not be completed for the July assessments due to data retrieval issues. • Analyse time students spent on the assessment and scores. Early analysis of the Assessment Master log-in data indicates that the median time students took to complete the assessment was 44 minutes.

	• The scores associated with each outcome were less reliable than the overall score. This means that care should be taken when using the outcome scores to make decisions about individual students.
	• The analysis showed that the assessment performed reasonably well.
	  o Cronbach's alpha (KR-20) was 0.81, indicating a reasonably high level of internal consistency among the items.
	  o Most items discriminated well.
	  o The difficulty measures of the items were spread across much of the distribution of student achievement.
	  o There was evidence to suggest some lower-achieving students guessed their response to some of the items.
	  o Question 5b showed negative discrimination.		

Literacy Writing	• The assessment consisted of two options with differing content but identical structure: two longer written compositions and six shorter written responses. The two longer composition tasks contributed to Outcome 1, which was made up of four aspects. One of the longer compositions was marked according to a rubric based on these aspects with each scored from 0 to 3. The other longer response was also marked according to a rubric based on these aspects but with each scored from 0 to 2. The six shorter written response questions were each scored 0 or 1, and all contributed to Outcome 2. • The data did not include the assessment option (A or B) the student had chosen; therefore, the analysis was not able to distinguish the psychometric properties of these options. • The scores associated with each outcome were less reliable than the overall score. This means that care should be taken when using the outcome scores to make decisions about individual students. • For both longer composition tasks, the first aspect on the respective rubric was more difficult than the remaining aspects. • Rasch analysis showed that the assessment performed reasonably well:	• Ensure that the impact of measurement error is considered when determining the required minimum scores for each Outcome and the minimum score required overall. • Consider how options are used in the assessment. If options are offered and are assumed to be very similar psychometrically, this assumption should be able to be tested.	• Perform the analysis based on option choice (A or B). Ensure this data is made available. • Analyse the "distractor quality" for multiple choice items. (Sometimes a response option can have characteristics – for example, wordiness, vocabulary – that might make it overly appealing to certain groups of students.) • Analyse time spent by students on the assessment and scores.

	○ Cronbach's alpha (KR-20) was 0.89, indicating a high level of internal consistency among the items. ○ Most items discriminated well; however, two items for Outcome 2 discriminated poorly relative to the other items. ○ The difficulty measures of the items were spread across much of the distribution of student achievement. ○ The relationship between the total score and achievement of the standard was not straightforward. There was a range of total scores where some students achieved the standard and others did not. This range was relatively large compared to the measurement error associated with the total score.		

Appendix 3

Tai Huki - Initial readiness report (first report of two)

Tai Huki Initial Readiness Report 202

References

Caygill, R. & Sok, S. (2008). PISA 2006: School context of science achievement: how ready are our 15-year-olds for tomorrow's world? Wellington. Ministry of Education. Retrieved from https://www.educationcounts.govt.nz/publications/schooling/31278/6

New Zealand Council for Educational Research. (2018). The NCEA review: findings from the public engagement on the future of NCEA. Wellington. Retrieved from www.nzcer.org.nz

Yuan, J., Turner, T.R., & Irving, E. (2010). Factors Influencing University Entrance Success Rate. Auckland: Starpath Project, The University of Auckland. Retrieved from https://cdn.auckland.ac.nz/assets/education/about/research/images/Factors%20influencing%20University%20Entrance%20success%20rate.pdf