

The new national literacy tests for post-primary students in Aotearoa New Zealand: How process and design issues undermine principles of a strong and fair qualification

Margaret Franken¹

Independent Researcher

From 2026 in Aotearoa New Zealand, post-primary school students will be required to achieve new literacy standards in reading and writing through national standardised tests called *Common Assessment Activities* or CAAs. The CAAs, in their pilot phase, are revealing dire results and fundamental flaws. This paper discusses the results of three pilot events in 2021 and 2022 trialling the CAAs and suggests some factors that are likely to have played a part in influencing the results – particularly for English Language Learners (ELLs), Realm nations' students, Pasifika, and Māori students. These factors include conceptual or construct issues, as well as test design, marking, and administration issues. This unfair assessment is also inequitable because it has been positioned as a co-requisite for New Zealand's senior years' qualification, the National Certificate of Educational Achievement (NCEA), thus effectively trumping other measures of school achievement. The paper suggests that if testing is to continue, the New Zealand Qualifications Authority (NZQA) and the Ministry of Education should rethink not only the design, marking and management of the tests to make them a fairer, more valid and reliable process, but also the tests' standing as a co-requisite for senior years' qualification.

Key words: Literacy testing, literacy standards, foundational literacy, literacy across the curriculum

¹ Email address for correspondence: margaret.franken@gmail.com

Introduction

In Aotearoa New Zealand, new national literacy and numeracy standards for post-primary² learners have been designed. The English medium standards include one for reading (Reading - US 32403 – Read written texts to understand ideas and information), one for writing (Writing – US 32405 - Write texts to communicate ideas and information), and one for numeracy. In 2024, the standards will be a compulsory part of the *National Certificate of Educational Achievement* (NCEA), New Zealand's main secondary school qualification that students most commonly work towards in Years 11, 12 and 13. The standards are assessed through national standardised tests called *Common Assessment Activities* or CAAs. While the CAAs are being formally piloted and trialled, alternative ways of assessing literacy and numeracy exist – until 2026 – when the CAAs will be the only way to achieve the standards (Ministry of Education, n.d.a).

The standards are deemed to be a *co-requisite* – meaning that achieving them is compulsory for the award of the NCEA. In essence, most students experience the standards like a prerequisite, given that the CAAs evaluating them are most likely to be sat in Year 10 – as evidenced by the fact that approximately 86% of the students sitting the tests in the 2022 pilots were Year 10 students (Evaluation Associates, 2022, p. 22; 2023, p. 45). This is because the standards are said (by the Ministry of Education) to be benchmarked against upper Level 4 and lower Level 5 of the New Zealand Curriculum (typically experienced by learners in Years 9 and 10). However, if students do not sit the tests in Year 10, they can sit them at any other time in the future, and they can also sit them as often as they wish without any financial cost.

These standards are for all post-primary students enrolled in NCEA in English medium settings in Aotearoa New Zealand – not all of whom are in schools. They are also for students in the bilingual settings of the Realm nations of the Cook Islands and

² The term “post-primary” has been used because, although the standards are primarily for secondary school students, students in alternative education settings (such as teen parent units or prisons), and at tertiary level can also sit them.

Niue. In the near future, Tokelau students will also participate in NCEA.³ Students from the Realm nations and Tokelau are, in essence, English language learners, though first language maintenance and bilingual provision in schools varies greatly across the nations.

Māori medium standards for literacy and numeracy (Te Reo Matatini me te Pāngarau) exist as an option for te reo Māori proficient students in Māori medium settings (which serve some 3% of the total school population) (Education Counts, n.d.). King and Cunningham's (2017) analysis of 2013 census data indicated that this amounted to 17,343 primary or secondary students (almost all of whom identified as Māori). The Māori medium standards are not within the scope of this paper, given that the method of assessment differs significantly, and that it has its own set of issues (see Evaluation Associates, 2023, pp. 30-41 for some discussion of these). There is no provision for other bilingual students within Aotearoa New Zealand to have their literacy and numeracy skills assessed in their first language. We might think here, for example, of the significant numbers of first language speakers of Gagana Samoa⁴.

While acknowledging that there are shared concerns across both the literacy and numeracy contexts, this paper focuses primarily on issues with the two literacy standards. These issues include conceptual or construct issues (such as lack of clarity around the construct of foundational literacy), as well as test design, marking, and administration issues (such as little or no control of text complexity and difficulty, lack of transparency around how marking rubrics are applied, and no feedback for students).

The primary purpose of this paper is to present these issues and to suggest a way forward. To this end, I briefly discuss the review of the NCEA and other drivers that led to the testing of literacy through two external and formal high-stakes tests. I present the troubling results from piloting the tests to date as these were the impetus for my interrogating the tests themselves and the processes associated with them -

³ The Cook Islands, Niue and Tokelau are all part of the New Zealand Realm, and their people are citizens of New Zealand Aotearoa. The Cook Islands and Niue are self-governing, but Tokelau is a dependent territory.

⁴ King and Cunningham's (2017) analysis of census data showed 15,897 school age speakers of Gagana Samoa in 2013.

which constitutes the “study” part of the paper. While I did this in my role as the Subject Matter Expert in Literacy in the Ministry of Education when the CAAs were in draft form, any teacher could have done this once the CAAs were launched during the assessment period.

The analysis and findings illustrate the issues which are likely to have a significant impact on learners – particularly, but not exclusively, ELLs. In order to evaluate how the process and design issues undermine the principles of a fair and “strong qualification”, I draw on a framework developed by a Ministerial Advisory Group to guide the changes to NCEA (described in more detail below, and in NZCER, 2018). Lastly I consider possible ways forward.

The review of NCEA

The new standards arose in the context of a full review of the National Certificate of Educational Achievement (NCEA) that began in 2018. A Ministerial Advisory Group was formed by the then Minister of Education, Chris Hipkins, at the beginning of 2018. The Ministerial Advisory Group came out with two major frameworks to guide the process of review and re-design, as set out in the review report published by the New Zealand Council for Educational Research (NZCER, 2018). The first of these was the five principles of a strong qualification:

- Wellbeing: NCEA should promote the wellbeing of learners and teachers through effective and fair teaching and assessment practice.
- Inclusion and equity: NCEA should facilitate high expectations for all learners and ensure that every learner has the opportunity to succeed.
- Coherence: NCEA should ensure learners access the powerful knowledge, skills, capabilities, and attitudes identified in the National Curriculum.
- Pathways: NCEA should make it easy for learners, their parents and whānau [family], and teachers to make informed choices to enable success in education and later life.
- Credibility: NCEA should be readily understood, widely supported, and validly measure achievement (NZCER, 2018).

These principles apply to the NCEA qualification as a whole and also to the subjects within it. The principles are strong, but also sound, and align with the National Educational Learning Priorities or NELP (The Statement of National Education and Learning Priorities (NELP) & Tertiary Education Strategy (TES), n.d).

The second framework was a set of high level “opportunities” for change, one of which was “to strengthen literacy and numeracy” (see NZCER, 2018, p. 43, for the list of all 6 “Big Opportunities” and an explanation of how they were used in the initial stages of the review).

The 2018 conceptualisation of a “big opportunity” to improve post-primary students’ literacy was transformed over a short period to be manifested as external tests that began to be piloted in 2021. In effect, big opportunities to focus on improving and supporting the teaching and learning of literacy – particularly from a more contextualised approach based on literacy practices and multiliteracies – became fewer and smaller as the change became focused on assessment and accountability. This can be seen in the shift in the examples of discourse in selected official and public texts over this time:

- “Strengthen literacy and numeracy” (NZCER Review, 2018).
- “Strengthen and clarify our expectation for literacy and numeracy attainment” (Hipkins, 2018).
- “Strengthen literacy and numeracy requirements” (Ministry of Education, 2019).
- “Strengthen literacy and numeracy requirements and assessments” (Ministry of Education, n.d.b).

A Technical Advisory Group (TAG) in 2019, looking at the development and implementation of a way of assessing literacy and numeracy, had recommended (amongst other things) that there be a digital adaptive tool that would provide evidence against the benchmark, and that such a tool be co-designed with the sector and thoroughly tested for any kind of bias (Technical Advisory Group, 2019, p. 38). Shortly after this, the Ministry of Education stated that the standards were to become a corequisite to the NCEA qualification: “Students will need to pass specific literacy

and numeracy external standards in order to be awarded an NCEA” (Ministry of Education, 2020, p. 1).

Throughout 2020, standards and the tests to assess them were developed in a process outlined by NZQA (n.d). The description of the process is broad and the nature and quantity of input provided by different sectors and stakeholders is unclear. In 2021, the New Zealand Ministry of Education began piloting the external and formal CAAs for the new standards. One pilot event occurred in 2021 and two in 2022. One of two planned events for 2023 took place in June, though the results from this have not yet been publicly released. The tests will become mandatory in 2026, and the method of assessment is likely to remain entrenched at that point, if not earlier.

Drivers of the change

There appear to have been two major drivers of the change to assessing literacy in a high-stakes, formal and decontextualised national test (Ministry of Education, 2020), rather than in a way that is more flexible and adaptive. These were: concerns about the previous contextualised manner of assessing literacy and numeracy in NCEA; and concerns about declining standards.

Literacy was previously assessed in the context of selected NCEA subject standards which were deemed to be “literacy rich” (Ministry of Education, n.d.a). In essence then, literacy was assessed indirectly. If students passed one of the selected literacy or “tagged” standards, they were credentialled with NCEA literacy. A study commissioned by the Tertiary Education Commission (Thomas et al., 2014) of students who had achieved NCEA Level 1 and NCEA Level 2 by means of this contextualised approach indicated that only 50% and 60% respectively of these students were deemed to have a literacy level sufficient for tertiary study (Thomas, et al., 2014). Thomas et al. used the Literacy and Numeracy for Adults Assessment Tool (LNAT) and the Adult Literacy and Life (ALL) survey in reading to assess alignment with NCEA achievement in literacy (see Thomas, et al., 2014, p.2, for a more detailed discussion of this process).

That there had been no “standard” way of measuring literacy, and that the identification and classification of such literacy rich standards appeared to be relatively subjective, added to concerns about declining standards (Hughson & Hood, 2022).

As in many other countries, over the last decade or so concerns have been loudly voiced about the decline in literacy standards in Aotearoa New Zealand by some individuals and organisations (e.g. the New Zealand Principals’ Federation; the New Zealand Initiative). Results from the Programme for International Student Assessment (PISA) between 2000 and 2018 showed a significant decline in literacy achievement: “The proportion of students with significant literacy issues in Aotearoa New Zealand has grown. . . [while] the proportion of advanced readers . . . has declined . . .” (Hughson & Hood, 2022, p. 23).

Not all agree that this is a situation requiring a crisis response. Brown (2021) for instance points to a number of factors that may negatively impact on literacy achievement, e.g. students’ relative inexperience of testing methods used in PISA, and lack of motivation to strive to succeed in such tests that are not of personal significance. However, the PISA results and other studies do highlight specific areas of significant concern. The average scores for ākonga Māori and Pasifika students (see below) were significantly lower than the Aotearoa New Zealand average, as were scores for boys. These scores have likewise exhibited a significant decline over the eighteen-year period. Brown (2021) comments that little has been done within the education system to close this gap and to address the more system-wide drivers of inequity that impact on literacy achievement.

The outcomes of the “opportunities” to date

There have to date been three iterations of the English medium literacy tests since piloting began in 2021. At the time of writing this paper, a fourth iteration has recently taken place, but the results and analysis (previously in reports produced by Evaluation Associates) have not been made public. Most of the discussion focuses on 2022 as the 2021 pilot was small (with only 2313 instances of completed assessments across

English medium and Māori medium contexts, and across literacy and numeracy). The results for 2022 have been analysed in terms of the following student subgroups:

- All students
- Ākonga Māori
- Pasifika students
- Students in Decile 1 and 2 schools
- English Language Learners (ELLs)
- Tertiary/Alternative Education students
- Realm nations students

Ākonga Māori is a te reo Māori term referring to learners of Māori ethnicity, most of whom are English dominant.

Pasifika students refers to students of Polynesian, Melanesian or Micronesian descent (including but not restricted to Samoan, Tongan, Cook Islands Māori, Fijian) living in Aotearoa New Zealand. (See King & Cunningham, 2017).

Decile 1 and 2 schools are those situated in low socio-economic communities. Schools are ranked on a number of variables including, but not restricted to, household income, occupation, educational qualifications, and then divided into 10 groups, called deciles. (Evaluation Associates, 2023, March, p. 34).

Tertiary/alternative education covers contexts outside of the regular school system. For example, it includes educational programmes for teen parents, for young people in youth justice/care and protection residential facilities and community campuses, and at-risk or disengaged students enrolled at Te Kura (the Correspondence School).

Realm nations students refers to learners who reside in the Cook Islands or Niue.

The numbers of students who participated in the literacy assessment events in 2021 and 2022 are given in Table 1. A breakdown of subgroups was not available for 2021.

Table 1. Number of literacy participants by group and by standard across two assessment events (compiled from Evaluation Associates, 2022, 2023)

	2021		2022 June		2022 September	
	Reading	Writing	Reading	Writing	Reading	Writing
All students	609	575	9,346	8,855	18,420	17,583
Ākonga Māori	NA	NA	1,371	946	2,463	2,597
Pasifika students	NA	NA	966	858	1,362	1,385
Students in Decile 1 and 2 schools	NA	NA	516	593	398	233
ELLs	NA	NA	84	84	27	30
Tertiary/Alternative Education students	NA	NA	40	33	17	13
Realm nations students	NA	NA	234	235	180	179

Note: NA means not available

The results from the three pilot assessment events for all students are shown in Table 2 (Evaluation Associates, 2022, 2023). The results across the board have been troubling. Over the course of three assessment events, the percentage of students achieving the Reading standards has steadily dropped, with the most recent event resulting in only 58% of students passing. For the Writing standard, no more than 35% passed in the first two events, but this saw an increase in the assessment event in September 2022. This may have been due to a decision to lower the *cut score* – the point that determines where the pass mark lies - in this case using the Angoff method (Angoff, 1971). However, whether a decision to do so was made has not been made public by NZQA.

Table 2. Literacy participants and rates across two assessment events by standard (compiled from Evaluation Associates, 2022; 2023)

	2021 June			2022 June			2022 September		
	All (n)	Achieved (n)	(%)	All (n)	Achieved (n)	(%)	All (n)	Achieved (n)	(%)
Reading	609	409	67%	9,386	6,016	64%	11,022	6,418	58%
Writing	575	205	35.5%	8,855	3,029	34%	12,299	5,688	46%

Selective results from the first assessment event of 2023 have been released in response to an Official Information Act request (NZQA, personal communication, September 29, 2023): 64% of students achieved the Reading standard and 56% achieved the Writing standard. While the results for the Reading standard were similar to the average of previous scores, the writing standard again shifted upwards – perhaps for the same reason as the previous rise.

Table 3 presents a breakdown of results for the subgroups in 2022. The data within the 2022 cohort shows a troubling pattern of disparities between certain groups of students. Realm nations students had the lowest percentage pass rates in Reading on both occasions (17% and 21.5%), and in Writing their results were low and relatively close (16% and 27.5%) to those of students in Decile 1 and 2 schools, who had the lowest percentage pass rates on both occasions (8.5% and 27%). For ākonga Māori, close to half of students passed Reading on both occasions, but no more than a third passed Writing. Pasifika students' rates for Reading were lower – close to a third of students passed Reading on both occasions, but for Writing, the results were the same as those of ākonga Māori. For ELLs and Tertiary/Alternative Education students, numbers of participating students were low and dropped further in the second assessment event, making it difficult to analyse with confidence – though their results are relatively similar to those of Pasifika students.

Table 3. Literacy CAA pass rates for subgroups across 2022 and 2023 assessment events (compiled from Evaluation Associates, 2022, 2023)

Sub-group	2022 June		2022 September	
	Reading	Writing	Reading	Writing
All students	64%	34%	58%	46%
Ākonga Māori	50%	24%	44.5%	34%
Pasifika students	34%	24%	34%	34%
Students in Decile 1 and 2 schools	27%	8.5%	28%	27%
ELLs	35%	21%	33.5%	40%
Tertiary/Alternative Ed students	47.5%	12%	35.5%	31%
Realm nations students	17%	16%	21.5%	27.5%

The recently acquired data from the Official Information Act request (NZQA, personal communication, September 29, 2023) included only the percentage achieved for

ākonga Māori, Pasifika students, and Realm nations students. The results for Reading were 49%, 39.5% and 18.5% respectively – not too dissimilar to previous results. For writing, they were 42%, 44% and 45%. While it is encouraging to see some improvement in Writing, for Realm nations students in particular, it is somewhat anomalous given the low rate of 18.5% in Reading. As discussed above, the results for Writing are able to be manipulated while those for Reading are not.

The reading results to some degree mirror those in PISA and PIRLS (Progress in International Reading Literacy Study). For PISA (2018), “the average score for ākonga Māori in reading was 463, . . . [and] the average score for Pacific students in reading was 442.” These scores are both significantly lower than the Aotearoa New Zealand average of 506 and the OECD average of 487. For PIRLS (2016), “there are large ethnic gaps, with Pākehā students gaining an average of 545 points, ākonga Māori an average of 479 points, and Pasifika students an average of 485 points” (Hughson & Hood, 2022, p. 19).

The present study: Interrogating the tests and the assessment processes

After seeing the 2021 and 2022 achievement results, I believed that an explanation of the dire pattern of achievement could not be solely laid at the feet of learners and their teachers. Firstly, there was the fact that the results were fairly consistently low across different teaching and learning contexts. Secondly, for the contexts in which there were ELLs, there are low results (not only for Literacy, but also Numeracy) suggesting that language issues might be at play.

In order to scope the issues that might lie in the CAAs themselves, I engaged in interrogating the tests – I surveyed documents, observed processes, and analysed the text, task and language demands of the CAAs over the period of some 18 months while I worked as the Subject Matter Expert in Literacy in the Ministry of Education. My expertise and experience as a language and literacy academic allowed me to interrogate the tests and assessment processes. It should be said that both the Ministry of Education and NZQA largely ignored my analysis, given that very few changes were made in the iterations of the CAAs over time.

Analysis and Findings

The analysis and findings firstly take a more general perspective, covering construct issues, issues to do with the validity of benchmarks, and process issues related to feedback. The section then focuses on analysis and findings related specifically to the Reading CAAs and then the Writing CAAs.

The constructs of literacy and foundational literacy

There are widely divergent views on what literacy is. This is true in the theoretical realm as well as in the practical – for example, for the purposes of designing an assessment that seeks to measure literacy. In the public consultation phase of the NCEA review, NZCER found that people’s views of literacy tended to be traditional and English-only, focusing on reading and writing - specifically on “spelling, grammar, sentence construction, reading text, and understanding ideas” (NZCER, 2018. p. 50), and it was this view that would come to be instantiated in the way the standards were worded by the Ministry and the way the CAAs were written by NZQA. The NZCER report commented that “very few talked about literacy in a wider context in relation to oral language, about literacy in te reo Māori, or about other modes of literacy” (NZCER, 2018. p. 50). It also appears that very few, if any participants, expressed a bilingual or multilingual view of literacy. There is no mention of conceptualising literacy in a way that recognises ELLs’ literacy proficiency in their first language or additional languages. This is surprising since we know that literacy in the first-language is strongly correlated to literacy development in English (Ríos & Castellón, 2018). A more contemporary and current perspective which goes beyond English, and reading and writing only can be seen in a multi-literacies perspective or a social practices view of literacy/literacies (Brandt & Clinton, 2002; Lotherington, 2007). Arguably such perspectives could possibly have positively influenced the way in which literacy has come to be assessed in NCEA so that certain learners were not so disadvantaged, and their skills and strengths in languages other than English were recognised. While acknowledging that this is a challenge to operationalise in a standardised assessment for even for some of the more commonly spoken languages other than English and te reo Māori (e.g. Samoan), the inclusion of such a perspective

might have supported teachers to exploit transfer of language and literacy skills into English for the benefit of their students. This is an opportunity lost.

A related issue lies in the fact that “foundational literacy” as used to describe the focus of the reading and writing standards is somewhat difficult to instantiate in the assessment. Foundational literacy is seen as “the base level required to support participation in one’s community, employment, and further learning” (Ministry of Education, 2022a, p. 9). In a sense, this definition is workable in teaching and learning as teachers are encouraged to seek a range of contexts that might be meaningful for their own students. However, it is not workable for nation-wide assessment; the contexts of community, employment and further learning set far too expansive a context brief for the assessment of the standards that are geared towards learners in Years 9 and 10. Based on this definition of foundational literacy, the CAAs attempt to incorporate community, employment and further learning contexts, primarily in the test writers’ selection of texts for reading and topics for writing (see <https://ncea.education.govt.nz/literacy-and-numeracy/literacy/reading/unit-standard> for samples of CAA texts and tasks, some of which have been partially redacted). This however means that problems can arise from writers’/markers’ bias and their world views – an issue discussed further in relation to text selection and difficulty below. One way to mitigate the effects of unfamiliar contexts might be to pre-determine a more restricted range of contexts in the test specifications, and/or to pre-announce topics within contexts which will feature in an upcoming assessment window.

Context is but one dimension of the definition of foundational literacy used by the Ministry of Education and NZQA. The other is “base level”. This is not clearly analysed in the specifications for the standards, and the Ministry of Education’s attempts to align the level with existing assessment tools have not provided clarity.

Another issue associated with the NZQA writers’ interpretation of foundational literacy is its alignment with subject English, and thus the markers’ appreciation of comprehension and production of stylistic or more literary aspects of text.

Alignment between the standards and assessment tools for benchmarking

The Ministry of Education published a number of “benchmarks” or tools which they deemed could operate as indicators of readiness or *readiness tools* (Ministry of Education, 28 January, 2020, p. 39). There are two major reasons why benchmarks for the literacy standards and, by proxy, the high-stakes external tests assessing them need to be identified and specified. The first of these, for designers of the CAAs in particular, is to check that the benchmarks are valid indicators of the construct of foundational literacy. The second, but related reason concerns the validity of the benchmarks as readiness indicators. In such assessments as these, which students can choose to sit as often as they wish, and at different times in the school year, teachers and students need to have confidence in their judgements about students’ readiness.

Despite the fact that the Ministry of Education commented that the benchmarks or tools would need to be “quality assured” (Ministry of Education, 2020, January 28), no systematic analysis had been carried out to verify the correlation between the benchmarks or tools and the results in the literacy and numeracy assessments before their promotion as readiness indicators in the pilot period. After the fact, that is, after the identification and promotion of a number of benchmarks and readiness tools by the Ministry of Education, one small set of data from the first of two 2022 pilot assessments seems to indicate a disappointing and worryingly low level of correlation between a particular widely-used tool, e-asTTle – previously standardised against a large secondary school population sample (see Ministry of Education, n.d.c) – and results in the literacy assessments (Evaluation Associates, 2023, pp. 25-27).

The literacy standards are deemed to be aligned with upper Level 4 and lower Level 5 of the New Zealand Curriculum. Curriculum levels have been mapped onto e-asTTle test results so that upper Level 4 is equivalent to 4A (Advanced) and lower Level 5 is equivalent to 5B (Basic) on e-asTTle⁵ (see <https://e-asttle.tki.org.nz/Help-with-e-asTTle/Reports/e-asTTle-norms-Reading-and-Maths>). Thus, we might expect that the majority of students who had scored these e-asTTle scores (4A and 5B) would pass the corequisites. However, from the data set mentioned above, Evaluation Associates

⁵ The order of progression of e-asTTle scores is: B (Basic), P (Proficient), then A (Advanced).

(2022) found that 72% of students who scored 4A in e-asTTle Reading achieved the literacy reading standard. The writing standard had a much weaker correlative relationship to e-asTTle as only 53% of learners who scored 4A achieved the writing standard (see Evaluation Associates, 2022, pp. 25-27). The percentages of successful students who scored 5B in e-asTTle Reading and Writing were respectively 88%, and 65%; and at 5P (Proficient) the percentages were 95% and 72%. We can possibly then say that 5B is a strong predictor of success in Reading, but in Writing this is not seen till 6P (where we see a 86% success rate). Evaluation Associates comment, “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” (Evaluation Associates, 2022, p. 26).

This analysis needs to be repeated with the results from subsequent assessments – the second assessment from 2022, and the two assessments that will be conducted in 2023. If the results are consistent, then it would appear that the assessments are not reflective of the curriculum levels that the Ministry of Education states they represent, despite the fact that Evaluation Associates claim, “the achievement results show that e-asTTle can be used as an appropriate indicator of readiness for sitting the Literacy and Numeracy standards” (Evaluation Associates, 2022, p. 26).

While the discussion above has focused on the validity of the identified benchmarks, of course we might just as well look at the strong possibility that the tests themselves are not set at the appropriate level. As the Ministry itself pointed out, “[T]he assessment can be validated by making sure its results correlate to the system-wide readiness results” (Ministry of Education, 2020, p. 39). Thus, we might question if the use of these tests has been validated.

Lack of feedback

Teachers and educators recognise that feedback is an integral part of the teaching and learning process and is the most important teacher practice in improving student learning (Hattie & Timperley, 2007; Wisniewski, Zierer, & Hattie, 2020). It is also consistent with the Ministry’s Effective Practices – one of the keystones of the teaching and learning support for the NCEA literacy standards (Ministry of Education, n.d.d).

A learner who has been judged by their teacher to exhibit readiness but still does not meet the standard needs feedback to understand how far they are from achieving. If such feedback is not given, the experience is demoralising for both learners and their teachers. Learners must be clear on what/where they need to improve.

However, to date NZQA has not returned the Reading or Writing scripts to students; they merely receive an A (achieved) or N (not achieved) judgement. Since the Reading CAAs are machine marked, they could be returned to students – as happens with some other NCEA scripts. For Writing, the feedback could perhaps be given by means of the rubrics used for marking the CAAs. There has been some discussion that feedback may be given to the students in 2023, but this has not to date occurred, and it is not clear how personalised that might be.

In the following section, I discuss some of the challenges that present: in the Reading CAAs, because of the conceptualisation of the construct of foundational reading, and the design of the test; and in the Writing CAAs, primarily because of the conceptualisation of the construct of foundational writing, the design of the marking rubrics, and the way in which the marking is managed. Essentially these are issues of both test and item bias. Bazemore-James et al. (2017) remind us of the importance of minimizing any bias and measurement error “because the predictions from these test scores have a massive impact on students’ lives” (p. 6).

Challenges in the Reading CAAs

Digital fluency and cognitive load

NZQA has a digital-first policy, which has meant that most learners sit the tests on NZQA’s assessment platform for NCEA, rather than using a paper-based version of the tests. Very few Year 9 and 10 students will have had any experience of using this platform before sitting the corequisites, given that this is likely to be their first experience of NCEA assessment.

It is also likely that some subgroups of students will have more limited access to digital tools in general and for learning (e.g. being exposed to and reading digital texts). Moore, Vitale, and Stawinoga (2018) found this when they analysed the type and

number of devices different students had and what they were used for, noting that those who only had access to a smartphone were significantly limited in accessing “school-related activities”. This reflects the well-documented phenomenon of the “digital divide”, whereby a digital-first approach excludes underprivileged learners from social, educational, and/or employment opportunities. While Moore et al.’s study was located in the US, similar findings have been seen in the New Zealand Digital Government’s (2022) study, for example. Using 2015 data from PISA, Digital Government examined internet access of 15-year-olds at school and at home. Pasifika students reported far lower internet access rates at home than did students of all other ethnicities; Pākehā students had the highest rate at home. Similarly, in terms of internet access at school, Pasifika students reported much lower access than did other students. Māori students’ rates were lower than those of Pākehā students (Digital Government, 2022).

Students from low-income households, in Realm nations and from migrant and refugee backgrounds are also likely to be affected by this digital divide as they too may have limited access to devices to practise on. In Aotearoa New Zealand, the Ministry provides devices for refugee families, but the current rules only allow for one device per family, irrespective of the total number of people in the family.

Even with experience and “digital fluency” (Wenmouth, as cited in Spencer, 2020), doing a reading test online poses a significant challenge to less able readers because of the extra cognitive load the medium places on them. This additional load is well documented (Clinton, 2019, Delgado et al., 2018; Kong et al., 2018), and arises from the need to manage the medium (such as scrolling) and attend to comprehension at the same time (Hahnel et al., 2017). Learners should be given the option of completing either a digital or a paper-based assessment.

The Reading CAA requires students to read up to eight texts, with some of these being made up of multiple small text excerpts. Switching between, concentrating on, and close reading of eight texts – some of which are composites of several mini texts (as in Assessment 1, 2022) – in an hour presents cognitive load challenges. The reading assessment should contain fewer texts which could be exploited more extensively,

rather than having many texts with few questions. This adds to the existing cognitive load resulting from reading texts in digital mode as opposed to on paper.

Text selection and difficulty

The texts in the Reading CAAs and the questions accompanying them have been designed by subject English teachers and an adult literacy educator.

Texts need to be chosen carefully from a narrow range of levels that as closely as possible represent the standard. Having texts that are either too hard or too easy compromises the ability to assess whether learners have met that standard. We might assume that most of the reading texts represent a reading age consistent with the level of the curriculum and the age of students at that level. Thus, we might expect most texts to require a reading age of approximately 13-14 years – consistent with Years 9 and 10.

The significant learning statement unpacking the Reading standard on the Ministry of Education website states: “Successful comprehension depends on understanding **most** [emphasis added] of the meanings of the words in the text.” This aligns with the fact that readers need to understand around 95% of running words in a text if comprehension is not to be impeded by a lack of fluency (Nation, 2001). Given that the standard is aimed at assessing foundational knowledge and use of vocabulary, we might assume that there are no more than about 5% of low frequency words in the texts.

I ran two types of analyses: Text Readability Consensus Calculator - a readability calculator that combines scores from seven well-known tools (see <https://readabilityformulas.com/free-readability-formula-tests.php>); and vocabulary levels using Lexical Tutor Vocabulary Profiler (<https://www.lextutor.ca/vp/>), to make an initial and broad assessment of the complexity and difficulty of the texts. Table 4 shows the results for the 2022 Reading CAAs.

Table 4. Results of readability and vocabulary level analysis for the 2022 Reading CAAs

Texts in CAA 1 2022			Texts in CAA 2 2022		
Readability: Reading level; age of reader	Vocabulary: Academic word list; Low Frequency		Readability: Reading level; age of reader	Vocabulary: Academic word list; Low Frequency	
fairly easy	AWL	3.39%	fairly easy	AWL	2.70%
10-11 yrs. olds	Low	3.39%	12-14 yrs. old	Low	4.28%
fairly easy	AWL	.76%	fairly easy	AWL	1.66%
10-11 yrs. olds	Low	19.47%	11-13 yrs. old	Low	16.90%
fairly easy	AWL	.66%	fairly difficult	AWL	4.53%
10-11 yrs. olds	Low	8.79%	14-15 yrs. old	Low	9.71%
standard/ average.	AWL	4.17%	standard/average	AWL	3.75%
14-15 yrs. old	Low	21.47%	12-14 yrs. old	Low	12.24%
fairly difficult	AWL	2.56%	fairly easy	AWL	4.85%
14-15 yrs. old	Low	17.18%	11-13 yrs. old	Low	19.09%
fairly difficult	AWL	5.14%	standard/average	AWL	2.85%
14-15 yrs. old	Low	8.38%	13-15 yrs. old	Low	12.20%
fairly difficult	AWL	1.25%	standard/average	AWL	2.82%
14-15 yrs. old	Low	21.95%	14-15 yrs. old	Low	10.34%
fairly easy	AWL	2.82%	fairly easy	AWL	1.09%
10-11 yrs. olds	Low	9.86%	11-13 yrs. old	Low	18.85%

Table 4 shows (in bold font) how many of the 16 texts arguably fall outside the standard in terms of reading age, and how many have a threshold of low frequency words that may impede comprehension. It also shows why it is important to analyse these texts in both ways, as reading age and vocabulary level are not mutually inclusive.

The words classified as low in frequency in Table 4 actually cover a broad range of frequency levels. In the absence of an operational definition of foundational literacy, there are no clear guidelines as to which words outside of the high-frequency category can reasonably be included in texts to be used for assessment purposes with Year 9 and 10 students in general. However, it is important to be alert to lower-frequency vocabulary that can create a significant barrier for English language learners, and those who have experienced less “reading mileage” in English. Such words should be minimised, especially if they are not essential for meaning and can be substituted with easier words. AWL words (Academic Word List words) are also important in

considering text difficulty; they include words necessary for successful tertiary learning, which many learners in Years 9 and 10 may not be familiar with.

Two texts in the first 2022 pilot assessment were on science topics. In one of the texts about insects, almost 22% of words were low frequency/topic-specific words that learners may well not be familiar with e.g., “ovipositor”, if they have not studied the topic. The low frequency words also include scientific names for insects. Arguably, this is the realm of NCEA Science – not NCEA Literacy.

Another issue associated with low frequency words is that a number of the texts include vocabulary such as colloquial New Zealand English (or “Kiwi”) expressions and Māori words (kupu Māori). The large extent of borrowing of Māori words in New Zealand English is well researched and documented (e.g. Calude, 2017; Macalister, 2006; Salzar, 2023), and can be seen in the *Dictionary of New Zealand English*, which now includes 746 words of Māori origin (with about 69 per cent being names of flora and fauna, 18 per cent being connected with social culture and 13 per cent with material culture) (“Māori One of Most Borrowed Languages”, 2014). While a number of the kupu Māori words in the literacy tests could be considered within the realm of foundational literacy, they should be controlled and possibly glossed if they are not so common.

Analysis of the draft CAAs for 2023 has shown that both assessments had texts which were too high in level relative to the standard, but that the texts to be used for the second assessment were easier than those for the first. To ensure a greater degree of reliability, the texts should have similar reading age and vocabulary level profiles over versions of the assessments. This issue reinforces the importance of controlling for linguistic complexity and difficulty.

Texts should also be checked for other aspects of complexity, which might not be accounted for in vocabulary level or readability measures. For example, one text in the first 2022 pilot assessment on “tramping boots” contained these complex collocations e.g., “shock absorbing upper layer, the breathable material, personalised cushioning, endurance athlete, a multi-day tramp, dual density rubber”.

In addition, when looking at the language of the questions, complex and difficult low frequency words, collocations, and idioms (e.g., “fuse”, “durable”, “aspirational”, “fly under the radar” – June 2022) were selected as items to be tested. There were also instances in the tests where the words in the list of multi-choice options were of lower frequency than the target word being tested. In this case, the item was in effect testing knowledge of words in addition to the target word – and ones that were more difficult than the target word.

In the section above, I have outlined issues to do with the difficulty of the reading texts used in the CAAs – both in terms of readability and in terms of vocabulary level. It is fundamental in language-related testing that confounding variables are addressed. A first step in designing a reading test is surely controlling for text difficulty.

Unfamiliar contexts, topics, and styles

In teaching and learning we strive to make texts and tasks interesting, motivating, and relatable. In assessment, however, the long-established research on “situational interest” and its effect on comprehension suggests that it needs to be managed. Situational interest (also called “seductive details”) can distract and unfairly disadvantage poorer readers who struggle to comprehend connected text (see for example, Ivanov, 2010; Schraw & Lehman, 2001; Wade, et al., 1993). Situational interest often strays into unfamiliar territory for readers – unfamiliar contexts, topics, and styles. Writers of the CAA Reading texts need be aware of and manage the demands of situational interest given this is an assessment, not a teaching and learning activity.

I contend that for an assessment to be as equitable as possible, the contexts and topics featured in the texts and test items should be those that are likely to have been widely experienced by as many students as possible. Several reading topics in the pilot assessments in 2022 are likely to have been outside of the range of experience of many learners e.g., reading a series of texts to choose expensive tramping boots. Kearns (2016, p. 137) explains this: “There are cultural, social, political and economic norms that exist within the logic of the test itself that exclude some youth from being successful...”.

It is imperative that NZQA writers selecting reading texts and setting writing topics conduct an “experience audit” when selecting contexts, topics, and text forms for all the standards. While experts agree that no text can express universally shared experiences and language, the writers do need to focus on selecting the least exclusive option, with a particular emphasis on equity for disadvantaged groups. The audit could encompass the following questions:

1. How many learners [*in a large south Auckland school] will have experience of this context/topic/text form?

*Substitute with different profiles of priority learner groups e.g., refugee learners, foundation tertiary learners, ELLs, Pacific learners, learners in small rural schools, in Realm nations.

2. Does the text represent experience from the learner’s point of view?

Challenges in the Writing CAAs

The Writing CAA requires students to write two “continuous” texts: one of up to 250 words and another of up to 350 words. These constitute the first two parts of the Writing CAA. The third part of the CAA largely involves “error correction” and “feature spotting” – item types critiqued by Cushing (2021). Unlike the Reading CAA, which is electronically marked (enabled by the fact that most of the responses are multiple-choice), the Writing CAA is marked by markers contracted by NZQA. There are two rubrics for the continuous texts and guidelines for the third question. The rubrics used for the two texts have four dimensions, which are elaborated in the following way:

1. Content: focus on what ideas/info etc., rather than how it is said. Quality of ideas.
2. Language Choices for purpose and audience e.g., sentence types and variety/word choice
3. Structure/organisation: overall flow of ideas across the text as a whole
4. Accuracy: e.g., sentence correctness, tense consistency, singular/plural, pronoun usage, verb forms

The writing tasks, prompts and marking rubrics are designed by subject English teachers and an adult literacy educator.

Marking the continuous texts

Extensive research has explored the biases that affect judgements of writing quality, and we know how important it is for good assessment practice to mitigate its effects. One highly problematic practice associated with marking the Writing CAAs is the allocation of scripts to markers in identifiable school batches. Expectations about students from particular schools could easily influence the way in which the rubrics are applied. Peterson et al.'s (2016) large scale study explored “teachers' implicit prejudiced attitudes for academic achievement” (p. 127) and pointed to possible negative effects of these for particular ethnic groups in New Zealand. To protect students from markers' expectations for academic achievement that might be prompted by association with particular schools, all scripts should be coded anonymously and have identifying information about writers and schools removed before marking.

What the rubrics reflect about what is valued

I have not participated in marking the scripts, and thus have not been in a position to analyse in depth the ways in which the rubrics have been applied, though I have observed some initial markers' meetings. Thus, the comments I make in this section are primarily related to the rubrics themselves. A version of these can be found on the Ministry of Education's website (Ministry of Education, 2022b). The rubrics have been designed by a subject English teacher, and the marking panel is made up of a large proportion of subject English teachers.

The rubrics are essentially an attempt to instantiate the construct of foundational literacy. The rubrics for the two questions asking learners to write connected texts cover fairly commonly used aspects or dimensions – content, language choices, structure/organisation, accuracy (but not complexity) – and each of these four aspects is assessed on the same scale. It is unclear whether this then means that each is accorded equal weight. Accuracy, which encompasses “sentence correctness, tense consistency, singular/plural, pronoun usage, verb forms,” should arguably be given

less weight than other aspects associated with communicating meaning because the definition is restricted and does not consider grammatical complexity.

The rubric for the longer text, on the accuracy dimension, shows little or no awareness of sociolinguistic variation demonstrated through grammatically non-standard texts, and what might be the standard of acceptability in structural terms. For example, what does “sufficient accuracy to communicate idea(s)/information clearly” look like?

0	1	2	2
Technical errors (punctuation, grammar, spelling) interfere with meaning, reader understanding/or require work from the reader.	Text conventions (grammar, punctuation, spelling) communicate idea(s)/information but require work from the reader.	Text conventions (grammar, punctuation, spelling) used with sufficient accuracy to communicate idea(s)/information clearly.	Text conventions (grammar punctuation spelling) used with accuracy and control to communicate idea(s)/information clearly, concisely, and effectively.

Figure 1. Section of marking rubric (Ministry of Education (n.d.d))

Foundational literacy needs only to be functional, not stylistically complex. However, the rubrics reflect an appreciation of stylistic effectiveness, for example: “Vocabulary, register, tone, sentence composition choices work together effectively for purpose and audience” – a descriptor more appropriate perhaps for a subject English assessment.

Another issue related to the use of a scaled rubric with a range of indicators (see Figure 1) is that this is suitable for NCEA *achievement standards*, where it is necessary to capture a range of performance at the four designated levels (not achieved, achieved, merit, excellence). However, the new literacy standards are deemed to be *unit standards*, which simply seek to determine whether a student’s performance is at a particular level or matches a pre-determined standard. It is important that the Ministry of Education and NZQA explicitly define what represents a pass.

Error correction and feature spotting as a way of assessing writing proficiency

Surface-level features in students’ writing are assessed in the rubric for the continuous texts, but also, in a decontextualized way, in the error correction and feature spotting item types. Examples of these from the first 2022 CAAs

(<https://ncea.education.govt.nz/literacy-and-numeracy/literacy/writing/unit-standard>) are as follows:

- Rewrite as one sentence without using: and, or, but, so. Correct punctuation, grammar and spelling.
i think sport are fun
you can be outside with your freinds
- Rewrite the sentence, making it clear that Daniel should play more sport.
Daniel mentioned to Andrew that he should play more sport.
- Rewrite the notice correcting punctuation, grammar and spelling.
there is a session tomorrow night learn how you can be more active bring long
your mates

These types of items lack ecological validity (Cushing, 2021); they are not authentic tasks or language samples, and are more difficult to complete successfully than if one is writing and checking one's own work. It would be more valid if students were evaluated on their ability to proofread the two texts they had written, but this would require a portfolio assessment.

Discussion: Evaluating the impact of the issues

A useful way to summarise the issues is to return to the principles of a strong qualification (wellbeing, inclusion and equity, coherence, pathways and credibility) and evaluate the standards and the CAAs against them given what we now know.

Wellbeing

“NCEA should promote the wellbeing of learners and teachers through effective and fair teaching and assessment practice.”

The assessments of literacy are neither effective nor fair. The analysis of problematic test design clearly shows how language complexity and difficulty is not controlled; and neither are the contexts and topics used in the CAAs.

In addition, students in the first three pilot events were not given feedback other than a “achieved” or “not achieved” judgement. This is not fair. While opportunities to re-sit are unlimited, a lack of personalised feedforward – to help students know how to improve – risks undermining their sense of self-efficacy and wellbeing. It also prevents teachers from supporting their learners in an effective way. In fact, a related issue is that there is very little transparency around how the rubrics are used to make judgements about achievement of the standards and how cut scores are set to determine the pass mark for Writing.

Inclusion and equity

“NCEA should facilitate high expectations for all learners and ensure that every learner has the opportunity to succeed.”

It is clear that not every learner has the opportunity to succeed in NCEA Literacy. Foundational literacy in reading and writing should not be so difficult for most students in Years 9 and 10 to achieve, as evidenced from lack of alignment with e-asTTle. It is not only a question of difficulty, but also of appropriateness and inclusiveness of student experience.

The test has reinforced the outcomes of international tests such as PISA. By replicating the long tail of non-achievement and doing so in a high-stakes and highly visible way, we run the risk of reinforcing low expectations of literacy achievement for particular subgroups of students. As Kearns (2016) comments, “Seemingly neutral and desirable educational goals, such as literacy, exist within discourses that operate by sorting, selecting and naming some students as deficient and others not” (p. 122).

Coherence

“NCEA should ensure learners access the powerful knowledge, skills, capabilities, and attitudes identified in the National Curriculum.”

Given the fact that there are issues with construct validity and the likelihood of teaching to the test, we cannot be assured that learners are accessing powerful knowledge, skills, capabilities, and attitudes related to literacy. It is likely that through the tests, students are experiencing a lack of meaningful interaction with texts and

tasks. Teachers have reported that schools are setting up special literacy catch-up classes for Years 9 and 10 students who have not achieved in the tests. Taking time from students in this way robs them of opportunities to engage in the meaningful and integrated learning experiences that are part of a junior secondary programme. Research is sorely needed at this time to expand our knowledge about the experiences and responses of teachers and learners as they trial the CAAs.

Pathways

“NCEA should make it easy for learners, their parents and whānau, and teachers to make informed choices to enable success in education and later life.”

The fact that the literacy standards are corequisites – compulsory for gaining the NCEA qualification -- means that those learners who do not achieve the standards will be excluded from gaining a secondary school qualification. What’s more, the assessments typically occur in Year 10 (although they can be sat at any year level), so they are perceived to some extent as prerequisites, to be sat before NCEA Level One begins. Given this, and the action of schools in setting up “special” literacy classes for non-achieving students, learners are likely to be disincentivised from continuing on to do NCEA if they have not achieved the literacy standards by the end of Year 10. The pathways have thus been limited and truncated. Given that a formal standardised test will be the only option available to diverse students, Kearns’s comment is apposite: “Arguably, the literacy testing mechanisms constrain marginalized youth’s possibilities, freedom, and diminish the value of literacies and knowledges they may possess” (2016, p. 122).

Credibility

“NCEA should be readily understood, widely supported, and validly measure achievement.”

The education sector’s response to the standards to date has been mixed, with successful schools (largely high decile) being the most supportive. A number of principals from low decile schools and/or schools with large ākongā Maori, Pasifika and ELLs have been vocal in sharing, with the Ministry of Education, NZQA and the

public through the media, the destructive effects that the tests are having and are likely to continue to have on their students. For them the tests are not seen to be valid or fair. One recent example has been reported in the media:

Poor results in trial NCEA tests have been heart-breaking for teachers and students in Pacific nations. . . . The report said people in the Realm countries had a positive attitude to the tests and put a lot of effort into preparing for them. "But the pilots have really taken their toll . . . on the well-being of staff and students. . . there have been aspects of the pilot that haven't been good enough," a teacher told evaluators. "It's battering. We had kids that were disillusioned. We've had examples of mental well-being really being impacted. . . . It's high stakes for them. . . . It's been massive and heart-breaking," said another.

(Gerritsen, 2023)

Thoughts on a way forward

Existing standards that have been successfully used by teachers for some time could be used as proxy measures for literacy reading and writing while the Ministry revisits the conceptual underpinning of the standards. This is fundamentally the responsibility of the Ministry of Education. They must:

- resolve where foundational literacy should be located,
- distinguish between subject English and Literacy; between teaching/learning and assessment,
- plan for and collect robust data to check the validity of the benchmarks,
- add details of the required level in language complexity and difficulty to the outcomes and performance criteria of the reading standard, and
- consider the negative consequences of making the standards a co-requisite, and thus reconsider this decision.

If tests continue to be the mode of assessing the standards, significant changes need to be made to ensure that the tests are fair, valid and reliable, and meet all the principles for a good qualification promoted by NZQA and the Ministry of Education. In terms of the design of the tests, this means NZQA must:

- control for readability and vocabulary levels of texts and questions in the Reading CAAs,
- reduce the number of texts and exploit this smaller number in a deeper way,
- address other issues that increase cognitive load,
- conduct an “experience audit” for topics, texts, and contexts,
- place less emphasis on accuracy, and remove the decontextualised questions for accuracy.

In terms of a valid and fair marking process, NZQA must:

- code all scripts anonymously and have identifying information about writers and schools removed before marking,
- be more transparent about how pass/fail decisions are determined, especially in writing, and
- include more teachers of and experts in English-language learning and/or testing on the writing and marking panels to balance the disciplinary perspectives of subject English teachers and experts.

In terms of the management and administration of the test in a fair way, NZQA must:

- allow for paper-based options,
- make significant improvements in digital format used in the tests, and
- return “not achieved” scripts to learners with personalised feedback.

Conclusion

The education sector and its teachers do not deny that improving the literacy skills of learners in Aotearoa New Zealand schools is an important priority, and many, if not most, are supportive of the opportunity to strengthen literacy (and numeracy). However, they do not accept that high-stakes external testing of reading and writing in its present design is fair or equitable. This claim is clearly supported from a detailed look at the tests, the results, and from using the principles of a strong qualification (NZCER, 2018, p. 2) as a touchstone. The tests violate many of the principles we have been working to embed in educational policy and best practice for the past several decades (see for instance, the National Educational Learning Priorities or NELP).

As Kearns comments, “Educators and policy makers ought to not work to reproduce and institutionalize inequities, through such practices as high-stakes standardized literacy testing” (2016, p. 138). The results from the pilots to date point to serious flaws in the way in which the tests are conceived, designed, marked and managed such that the results are damning – particularly for already disadvantaged students. The tests’ position as a co-requisite for the NCEA is particularly problematic, as this interferes with formally crediting students’ other strengths and learning, and may derail their future study and work opportunities, given the importance of NCEA as a credential (NZCER, n.d., pp. 23-24; NZQA, 2012).

References

- Angoff, W. H. (1971). Scales, norms, and equivalent scores. In R. L. Thorndike (Ed.), *Educational measurement* (2nd ed.) (pp. 508–600). American Council on Education.
- Bazemore-James, C. M., Shinaprayoon, T., & Martin, J. (2017). *Understanding and supporting students who experience cultural bias in standardized tests*. https://www.researchgate.net/publication/329427346_Supporting_Students_Who_Experience_Cultural_Bias_in_Standardized_Tests
- Brandt, D., & Clinton, K. (2002). Limits of the local: Expanding perspectives on literacy as a social practice. *Journal of Literacy Research*, 34(3), 337-356. https://doi.org/10.1207/s15548430jlr3403_4
- Brown, G. (2021, February 10). Test or invest? NZ’s sliding international student assessment rankings are all about choices. *The Conversation*. <https://theconversation.com/test-or-invest-nzs-sliding-international-student-assessment-rankings-are-all-about-choices-154729>
- Calude, A. (2017, October 4). Kia ora: How Māori borrowings shape New Zealand English. Royal Society Te Apārangi. <https://www.royalsociety.org.nz/research/kia-ora-how-maori-borrowings-shape-new-zealand-english/>
- Clinton, V. (2019). Reading from paper compared to screens: A systematic review and meta- analysis. *Journal of Research in Reading*, 42, 288–325. <https://doi.org/10.1111/1467-9817.12269>

- Cushing, I. (2001). Grammar tests, de facto policy and pedagogical coercion in England's primary schools. *Language Policy*, 20, 599–622.
<https://doi.org/10.1007/s10993-020-09571-z>
- Delgado, P., Vargas, C., Ackerman, R., & Salmerón, L. (2018). Don't throw away your printed books: A meta-analysis on the effects of reading media on reading comprehension. *Educational Research Review*, 25, 23–38.
<https://doi.org/10.1016/j.cedpsych.2020.101870>
- Digital Government (2022). *Digital inclusion and wellbeing in New Zealand*.
<https://www.digital.govt.nz/dmsdocument/161~digital-inclusion-and-wellbeing-in-new-zealand/html>
- Education Counts (n.d.). *Māori language in schooling*.
<https://www.educationcounts.govt.nz/statistics/maori-language-in-schooling>
- Evaluation Associates (2022, October). *2022 NCEA Te Reo Matatini me te Pāngarau/Literacy and Numeracy pilot evaluation: Evaluation report one*.
<https://ncea-live-3-storagestack-53q-assetstorages3bucket-2021xte0r81u.s3.amazonaws.com/>
- Evaluation Associates (2023, March). *2022 NCEA Te Reo Matatini me te Pāngarau/Literacy and Numeracy pilot evaluation: Evaluation report two*. Evaluation Associates.
- Gerritsen, J. (2023, April 2). 'Kids were disillusioned' - Pacific teachers worried about new NCEA test rates. Radio New Zealand.
<https://www.rnz.co.nz/international/pacific-news/487176/kids-were-disillusioned-pacific-teachers-worried-about-new-ncea-test-rates>
- Hahnel, C., Goldhammer, F., Kröhne, U., & Naumann, J. (2017). Reading digital text involves working memory updating based on task characteristics and reader behaviour. *Learning and Individual Differences*, 59, 149–157.
<https://doi.org/10.1016/j.lindif.2017.09.001>
- Hattie, J., & Timperley, H. (2007). The power of feedback. *Review of Educational Research*, 77(1), 81–112. doi: 10.3102/003465430298487

- Hipkins, C. (2018, May 27). *Big, bold ideas to change NCEA – do you agree?* [Ministerial Release] <https://www.beehive.govt.nz/release/big-bold-ideas-change-ncea-%E2%80%93-do-you-agree>
- Hughson, T., & Hood, N. (2022). *What’s happening with literacy in Aotearoa New Zealand? Building a comprehensive national picture*. The Education Hub. https://theeducationhub.org.nz/wp-content/uploads/2022/03/Ed-Hub_Long-literacy-report_v2.pdf
- Ivanov, I. V. (2010). *Predictors of recall and reading time for seductive and nonseductive text segments*. [Masters dissertation, University of Nevada, Las Vegas]. NLV Theses, Dissertations, Professional Papers, and Capstones. 768. <http://dx.doi.org/10.34917/2043727>
- Kearns, L. L. (2016). The construction of “illiterate” and “literate” youth: The effects of high-stakes standardized literacy testing. *Race, Ethnicity and Education*, 19(1), 121–140. <http://dx.doi.org/10.1080/13613324.2013.843520>
- King, J., & Cunningham, U. (2017). Tamariki and fanau: Child speakers of Māori and Samoan in Aotearoa/New Zealand. *Te Reo*, 60, 30-45. https://nzlingsoc.org/wp-content/uploads/2018/04/Te-Reo-60.pp29-46_KingCunningham.pdf
- Kong, Y., Seo, Y. S., & Zhai, L. (2018). Comparison of reading performance on screen and on paper: A meta-analysis. *Computers & Education*, 123, 138-149. <https://doi.org/10.1016/j.compedu.2018.05.005>
- Lotherington, H. (2007). From literacy to multiliteracies in ELT. In J. Cummins & C. Davison, (Eds.), *International Handbook of English Language Teaching* (pp. 891-905). Springer.
- Macalister, J. (2006). The Maori presence in the New Zealand English lexicon, 1850–2000: Evidence from a corpus-based study. *English World-Wide*, 27(1), 1-24. <https://doi.org/10.1075/eww.27.1.02mac>
- Māori one of most borrowed languages (2014, May 5). *New Zealand Herald*. <https://www.nzherald.co.nz/kahu/maori-one-of-most-borrowed-languages/5DLTIGEKH5L3HFWBNDSMFT32MA/>

- Ministry of Education (2019). *Changes to strengthen NCEA*.
<https://www.education.govt.nz/our-work/publications/budget-2019/changes-to-strengthen-ncea/>
- Ministry of Education (2020, January 28). *Education report: Update on changes to NCEA literacy and numeracy requirements*. [Ministerial report released under the Official Information Act].
- Ministry of Education (2022a). *External assessment brief: Literacy and Numeracy/Te Reo Matatini me te Pāngarau*.
<https://assets.education.govt.nz/public/Documents/our-work/information-releases/responses-to-Official-Information-Act-requests/1301655-Appendix-B-Documents-4-5.pdf>
- Ministry of Education (2022b). *32405 Literacy writing pilot marking: Term 2 CAA 1*. <https://ncea-live-3-storagestack-53q-assetstorages3bucket-2021xte0r81u.s3.amazonaws.com/s3fs-public/2022-10/Literacy%20Writing%2032405%20Marking%20Schedule.pdf?VersionId=8n3WUJWtdAt3u8q0HzSLWTmAz.HRRjAu>
- Ministry of Education (n.d.a). *Standards approved for NCEA co-requisite for 2024 and 2025*. <https://ncea.education.govt.nz/standards-approved-ncea-co-requisite-2024-and-2025>
- Ministry of Education (n.d.b). *What is the NCEA change programme?*
<https://ncea.education.govt.nz/what-ncea-change-programme>
- Ministry of Education (n.d.c). *E-asTTle*. <https://e-asttle.tki.org.nz/Help-with-e-asttle/Reports/e-asTTle-norms-Reading-and-Maths>
- Ministry of Education (n.d.d). *NCEA. Literacy/Writing*.
<https://ncea.education.govt.nz/literacy-and-numeracy/literacy/writing/unit-standard>
- Moore, R., Vitale, D., & Stawinoga, N. (2018). *The digital divide and educational equity: A look at students with very limited access to electronic devices at home*. ACT Research & Center for Equity in Learning.
<https://www.act.org/content/dam/act/unsecured/documents/R1698-digital-divide-2018-08.pdf>

- Nation, I. S. P. (2001). *How many high frequency words are there in English?*
<https://www.wgtn.ac.nz/lals/resources/paul-nations-resources/paul-nations-publications/publications/documents/2001-High-freq-Hakan.pdf>
- NZCER (New Zealand Council for Educational Research) (2018). *NZCER review: Findings from the public engagement on the future of NCEA.*
<https://conversation.education.govt.nz/assets/Uploads/NZCER-NCEA-Review-Report-FINAL4.pdf>
- NZCER (New Zealand Council for Educational Research) (n.d.) *Getting to grips with NCEA.*
https://www.nzcer.org.nz/system/files/press/abstracts/Understand%20NCEA_2nEd_ch1.pdf
- NZQA (New Zealand Qualifications Authority) (2012). *Understanding NCEA.*
https://www.schoolwebsites.school.nz/buller/wp-content/uploads/sites/104/2021/12/understanding_ncea_-_for_year_10.pdf
- NZQA (New Zealand Qualifications Authority) (n.d.). *Background to the development of the Literacy and Numeracy unit standards.*
<https://www.nzqa.govt.nz/ncea/subjects/literacy-and-numeracy/literacy-and-numeracy-unit-standards/background-to-literacy-and-numeracy-standards/>
- Peterson, E.R., Rubie-Davies, C., Osborne, D., & Sibley, C. (2016). Teachers' explicit expectations and implicit prejudiced attitudes to educational achievement: Relations with student achievement and the ethnic achievement gap. *Learning and Instruction, 42*, 123-140.
<https://doi.org/10.1016/j.learninstruc.2016.01.010>
- Ríos, C., & Castellón, C. (2018). Bilingual literacy development: Trends and critical issues. *International Research and Review: Journal of Phi Beta Delta, 7*(2), 85-96. <https://files.eric.ed.gov/fulltext/EJ1188732.pdf>
- Salzar, D. (2023). *Words from the land of the long white cloud: New Zealand English additions to the OED.* Oxford English Dictionary [Blog].
<https://public.oed.com/blog/words-from-the-land-of-the-long-white-cloud-new-zealand-english-additions-to-the-oed/>

- Schraw, G., & Lehman, S. (2001). Situational interest: A review of the literature and directions for future research. *Educational Psychology Review*, 13, 23–52.
<https://doi.org/10.1023/A:1009004801455>
- Spencer, K. (2020, March). *What is digital fluency?* Digital Learning Collaborative.
<https://www.digitalllearningcollab.com/blog/what-is-digital-fluency>
- Technical Advisory Group (2019, September). *Recommendations of the Technical Advisory Group on NCEA changes to literacy and numeracy (NZC). Prepared for the Ministry of Education.* [Document released under the Official Information Act].
- The Statement of National Education and Learning Priorities (NELP) & Tertiary Education Strategy (TES) (n.d).
<https://assets.education.govt.nz/public/Documents/NELP-TES-documents/FULL-NELP-2020.pdf>
- Thomas, G., Johnston, M., & Ward, J. (2014). *Aligning measures of adult literacy and numeracy.*
<https://www.tec.govt.nz/assets/Reports/3198200d93/Alignment-research-paper-Summary.pdf>
- Wade, S. E., Schraw, G., Buxton, W. M., & Hayes, M. T. (1993). Seduction of the strategic reader: Effects of interest on strategies and recall. *Reading Research Quarterly*, 28(2), 92-114. <https://doi.org/10.2307/747885>
- Wisniewski, B., Zierer, K., & Hattie, J. (2020). The power of feedback revisited: A meta-analysis of educational feedback research. *Frontiers in Psychology*, 10.
<https://doi.org/10.3389/fpsyg.2019.03087>