

Accommodating diversity in assessment

A snapshot of practice in 2022

A report produced for the Ministry of Education

Janet Lee and Rosemary Hipkins

Rangahau Mātauranga o Aotearoa | New Zealand Council for Educational Research
Te Pakokori, Level 4, 10 Brandon St
Wellington
New Zealand

www.nzcer.org.nz

© New Zealand Council for Educational Research, 2022

Accommodating diversity in assessment

A snapshot of practice in 2022

A report produced for the Ministry of Education

Janet Lee and Rosemary Hipkins

2022

Contents

1. Introduction	1
Special Assessment Conditions	1
Structure of the report	2
The research questions	2
The research methodology	3
2. A quantitative snapshot of the current situation regarding SAC	4
School resourcing challenges and patterns of access	4
For what conditions are SAC accessed?	6
What sorts of accommodations are requested?	7
Participation in digital assessments	8
3. A snapshot of SAC practices in schools	10
Is there an increasing need for SAC?	11
SAC practices vary considerably	11
Concerns about SAC provision for numeracy and literacy standards	13
Joined-up learning and assessment support	14
How RTLBs support inclusive assessment practice	15
Text-to-speech and other assistive technologies: Examples from elsewhere	16
Student control and agency	17
Is the term Special Assessment Conditions appropriate?	18
In summary	19
4. Why assessment design matters	20
A brief overview of Universal Design for Learning	20
The low-hanging fruit: Deploying UDL-informed accommodations to make quick gains	21
Eliminating “construct-irrelevant” barriers	23
Designing more flexible assessment tasks	26
A school-wide approach to enacting inclusive assessment	27
5. Are new NCEA materials more inclusive?	29
A nod to UDL	29
Taking task design a step further	31
Walking the talk of UDL?	33
Is “naturally occurring evidence” a feasible alternative?	33
Missing pieces	35
Concluding comment	36
6. The change potential in hybrid learning models	37
The scope of hybrid learning	37
Agreement from international sources	39
In summary	40

7. Short answers to the research questions	41
A snapshot of the current situation	41
What we can learn from good practice elsewhere	41
Establishing continuity between learning support and assessment support	42
The impact of the review of achievement standards	42
Potential for improving access through assessment design	43
Overall	43
8. Recommendations	44
Short-term opportunities for quick gains	44
Medium-term opportunities for more coherent system alignment	45
Long-term consolidation of sustainable practices	45
Thoughts on replacement of the acronym SAC	45
Concluding comment: Looking towards a system-wide view of change	46
References	47
Tables	
Table 1 Approved requests for SAC in 2020: Percentage of total cohort by school decile	5
Table 2 SAC by ethnicity as a percentage of overall assessment enrolments (2020)	6
Table 3 Summary of demographic data of schools interviewed about SAC	10
Table 4 How UDL can be used to modify traditional accommodation practices	22
Table 5 A comparison between NCEA processes and ECD, as outlined in the international literature	25
Table 6 Indicative alignment of naturally occurring evidence with specified understandings	35
Table 7 Hybrid learning features with the potential to support inclusive assessment practices	38
Figures	
Figure 1 SAC applications as a percentage of total assessment enrolments	5
Figure 2 Percentage of SAC applications in each of the four categories	7
Figure 3 SAC categories as a percentage of total applications	7
Figure 4 Number of digital assessment results recorded, by decile of school	9
Figure 5 One school's system for embedded inclusive practice	28
Figure 6 A hybrid learning model (Wenmouth, 2022, p. 7)	37

1. Introduction

Kei tēnā, kei tēnā, kei tēnā anō, atōnā ake āhua, tōnā ake mauri, tōnā ake mana.

Each and every one has their own uniqueness, life essence, and presence.

The New Zealand Council for Educational Research (NZCER) was asked to carry out research for the Ministry of Education¹ related to the future of Special Assessment Conditions (SAC) in the context of the National Certificates of Educational Achievement (NCEA). This report details our findings related to the current application of SAC entitlement within NCEA assessment and provides some short-term recommendations. It highlights areas of concern about the inclusiveness and sustainability of the current SAC model and explores options that could address these areas and provide a basis for new policy direction and solutions in the medium and long term.

Special Assessment Conditions

The NCEA system includes provision for what have become known as Special Assessment Conditions (SAC). This term (and its acronym) is widely used to describe accommodations and supports provided for learners in assessment contexts in an attempt to ensure that they have an unobstructed opportunity to show their true capability. SAC provision helps students fairly demonstrate their knowledge, skills, and understanding when being assessed, without providing an unfair advantage over other students.

SAC provision is administered through the New Zealand Qualifications Authority (NZQA) and is available for both internal and external assessments. Submissions for SAC are made by schools to NZQA on behalf of students. Following a successful application, SAC is provided as an entitlement to individual candidates.

In 2019, the Government announced an NCEA change package,² which included the following aims relevant to SAC:

- The design of assessments will be improved, so that SACs are required in fewer cases but are available as part of everyday teaching and learning.
- We will make some existing SACs more readily available for all where possible (e.g., larger-print papers).
- Where SACs cannot be extended to all conditions, we will work with users and experts to ensure the process is as simple and low-cost as possible.
- Through the review of achievement standards, we will modify assessment standards, teaching and learning guides, and exemplar programmes and assessment resources to ensure they exemplify inclusive practice (including for languages, cultures, identities, disabilities, genders, and sexualities). (p. 6)

¹ The Ministry of Education is sometimes shortened to “the Ministry” throughout this report.

² <https://conversation.education.govt.nz/assets/Uploads/NCEA-Change-Package-2019-Web.pdf>

SAC in their present form pose some issues. Applications for SAC entitlement are growing year-on-year. However, the system operates essentially as a “bolt on” to NZQA’s administration of NCEA external assessment processes. At face value, there appear to be inequities related to who accesses SAC, and in administration of the associated processes. Another issue is that the term “Special Assessment Conditions” is not popular with disability and learning support stakeholders, because it suggests that students who have an entitlement to SAC are either receiving special treatment, and therefore an assessment advantage, or that they require special conditions because they are otherwise not capable.

Structure of the report

This report is divided into eight sections:

- Section 1 outlines the research questions, agreed with the Ministry of Education, that have guided the research.
- Sections 2 and 3 provide a snapshot of the current situation informed by quantitative and qualitative analysis, respectively.
- Section 4 explores opportunities and challenges for developing more equitable opportunities to demonstrate learning, based on an assessment design perspective.
- In the light of the design considerations addressed in Section 4, Section 5 reports on our analysis of a small sample of the newly developed Level 1 pilot achievement standard assessment materials.
- Section 6 discusses the potential for inclusive assessment practices in hybrid learning models.
- Sections 7 and 8 conclude the report with our high-level summary and recommendations.

The research questions

In collaboration with the Ministry, the research team generated a series of questions to guide the research. These have been clustered to form five key themes as follows:

A snapshot of the current situation

- How are special assessment conditions currently being used in high-stakes assessment in Aotearoa New Zealand?
- What practices are perceived to be failing and why? Where there are examples of good practice, what does this look like?

What we can learn from good practice elsewhere

- How are Special Assessment Conditions addressed in other international education environments, including those serving indigenous students?
- Are there examples of inclusive practice that we could adopt in Aotearoa New Zealand?
- Are there examples of good practice that ensure equity for ākonga in hybrid learning contexts?

Establishing continuity between learning support and assessment support

- In the context of NCEA, how are RTLBs currently supporting inclusive practice in New Zealand secondary education?
- Is there a disconnect between learning support systems and assessment support in New Zealand secondary education? If there is, then how can the two be more closely aligned?

The impact of the review of achievement standards

- In what ways have the revised assessment models introduced at Level 1 for piloting in schools addressed inclusiveness and changed the likely requirements for SAC?

Potential for improving access through assessment design

- Can a standards-based assessment model such as NCEA be made more inclusive? If it can, what kinds of changes could enable this (e.g., changing ways that achievement can be evidenced)?
- What are perceived to be barriers to making these changes?

The research methodology

This investigation involved three parts:

- A quantitative study where we analysed statistical information from data about SAC applications in 2019 and 2020. Information for this analysis was supplied by NZQA and did not include 2021 data, which were not available at the time of analysis.
- A qualitative study where we conducted a series of semi-structured interviews and focus group interviews. These started with initial scoping interviews with NZQA representatives responsible for liaison with schools about SAC entitlements and MOE advisers in learning support advisory roles and roles involving pathways and transitions. The interactions with the Ministry and NZQA were followed by a series of interviews and focus group meetings involving staff across eight secondary schools, chosen to provide a broad spectrum of schools within the limited scope of this investigation. Finally, several Resource Teachers of Learning and Behaviour (RTLBs) were contacted from clusters associated with the school communities we visited, and a further three interviews and one focus group interview were held to understand the role of RTLB services in supporting schools with the SAC process.
- An investigative study into the principles and practice of Universal Design for Learning to explore the potential that the design of assessments could provide more equitable opportunities for all students and improve the current model for SAC. Questions concerning Universal Design for Learning were asked in each school interview and most RTLB interviews. As a result of this, one school was identified for a more in-depth case study.

The study also drew on input from the interviews with NZQA representatives and MOE advisers along with research of examples of good practice from Aotearoa New Zealand and overseas. It included a review of a selection of the Level 1 Review of Achievement Standards (RAS) pilot materials available on the NCEA website,³ using a Universal Design for Learning perspective.

The research methods used in this investigation were approved by and followed the guidance given by the NZCER Ethics Committee.

³ <https://ncea.education.govt.nz/>

2. A quantitative snapshot of the current situation regarding SAC

This section provides an outline of quantitative data related to SAC access and use. The data were provided by NZQA. A short final section introduces newly released data about participation in digital assessments for NCEA. This is an important part of the snapshot because digital accommodations have the potential to make some SAC universally available—without them needing to be seen as “special” at all. Research supporting this suggestion is introduced in Section 4.

School resourcing challenges and patterns of access

Currently, evidence to support a student’s application for SAC is provided by the school. It may be:

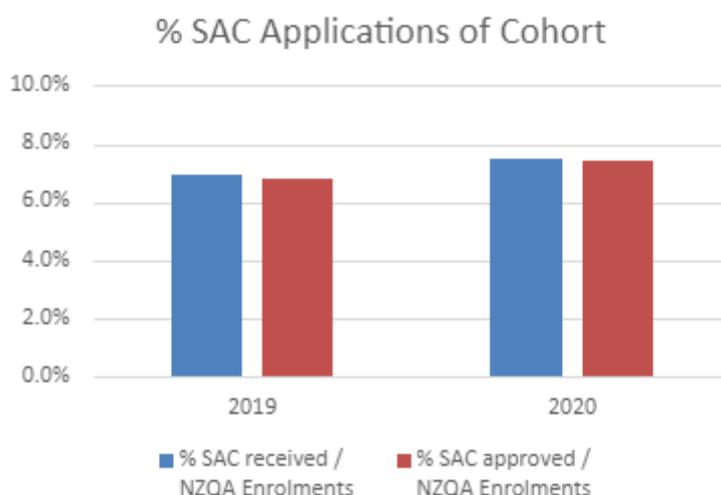
- appropriate documented evidence derived from school-based testing or observations undertaken during the student’s time at secondary school, or
- a report from an appropriately qualified independent registered professional who recommends SAC to address the student’s specified assessment needs.⁴

SAC entitlements need to be approved by NZQA only once, and schools are expected to review entitlements with their students annually through a needs analysis and rollover process. The SAC application process implies a considerable workload for schools.

Data suggest that this workload was somewhat higher in 2020 than in 2019. Figure 1 shows the total numbers of SAC applications in each of those years, as a percentage of the overall numbers of assessment enrolments. The actual numbers approved were 12,450 in 2019 and 14,417 in 2020. It remains to be seen if the data for 2021 show further growth or if 2020 was an anomaly.

⁴ <https://www.nzqa.govt.nz/providers-partners/assessment-and-moderation-of-standards/managing-national-assessment-in-schools/special-assessment-conditions/sac-information-for-schools/what-sac-is-available/>

FIGURE 1 SAC applications as a percentage of total assessment enrolments



Data source: NZQA

There are some indications that resourcing SAC is a challenge for lower decile schools. In 2019, 5% or fewer students who were enrolled for NCEA from schools with decile ratings 1 to 3 received SAC entitlements, whereas more than 11% of students enrolled in schools with decile 10 ratings received them. The next table shows the equivalent data for 2020.

TABLE 1 Approved requests for SAC in 2020: Percentage of total cohort by school decile

Decile	D1	D2	D3	D4	D5	D6	D7	D8	D9	D10	ND
Cohort %	4.8	5.0	4.8	6.5	7.4	7.2	6.5	9.1	8.7	11.2	4.9

The average across all deciles was 7.5%. Deciles 8, 9, and 10 schools had a higher proportion of successful applications, while for all other deciles the percentage was lower—and lowest overall for deciles 1, 2, and 3 schools.

A caveat to these data is that students from lower decile schools have lower involvement in external assessments, where the need for SAC could be seen as greater. These two things (entitlement and involvement in external assessments) could well be conflated. It is also likely that ethnicity data are conflated with decile data, given that proportionally greater numbers of Māori and Pacific students attend lower decile schools. Table 2 shows proportions of successful applications by ethnicity.

TABLE 2 SAC by ethnicity as a percentage of overall assessment enrolments (2020)

Ethnicity	Percentage of each group making SAC applications	Percentage of each group receiving SAC approvals
European	10.0	9.1
Māori	6.0	5.0
Pacific Peoples	4.0	3.3
Asian	2.3	2.0
All ethnicities	8.0	6.9

Source: NZQA NCEA Annual Report 2020

These data indicate that proportionally more applications were received on behalf of students of European ethnicity, compared to students of Māori and Pacific ethnicities. As already noted, this pattern could well be conflated with decile-related differences. However, the fewest applications were received on behalf of students from Asian backgrounds. This group is very diverse. Students from China, Japan, Korea, India, Pakistan, Sri Lanka, Indonesia, and so on are grouped together. It seems likely that other dynamics were also in play for this group.

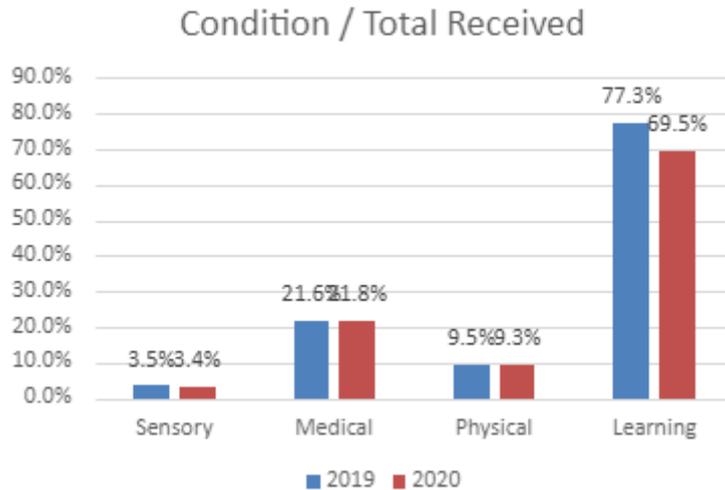
For what conditions are SAC accessed?

For the purposes of NCEA, SAC applications are categorised within four possible conditions: medical, physical, sensory impaired, and learning.

As Figure 2 shows, learning difficulties constitute by far the largest category—77% of all applications in 2019 and 70% in 2020. To get an idea of the size of the workload for schools, these percentages translate to 9,789 applications in 2019, and 10,082 applications in 2020, to support assessment when students have demonstrable learning difficulties.

Another indication of the workload for schools is provided by data that show the extent of their involvement in preparing specific applications for support for learning difficulties. The percentage of all these applications supported by school-generated evidence (as opposed to evidence-generated by relevant supporting professionals) was 64% in 2019, and 70.1% in 2020.

FIGURE 2 Percentage of SAC applications in each of the four categories

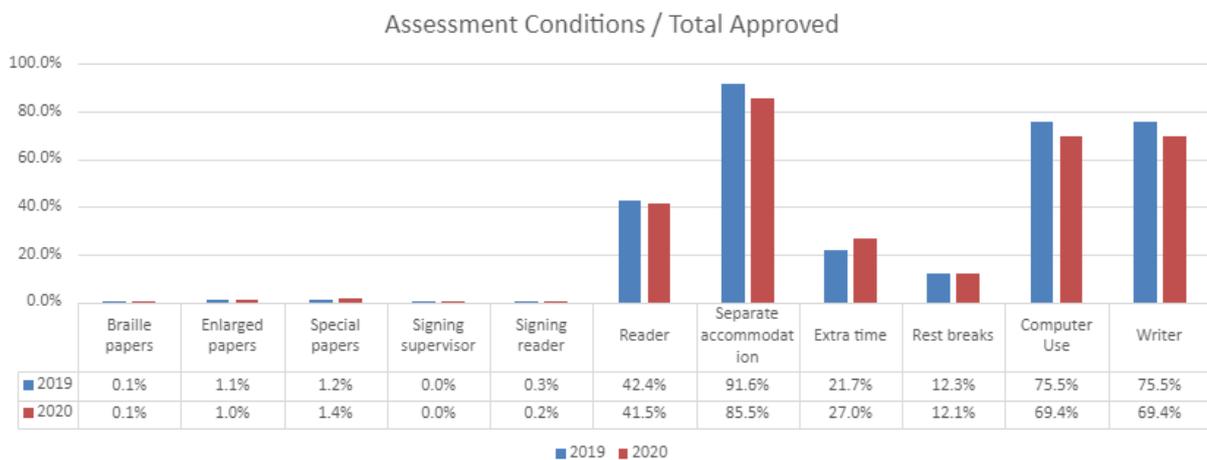


Data source: NZQA

What sorts of accommodations are requested?

Ten categories of SAC are listed on the NZQA website, with one additional category for using a computer in exams, which requires notification to NZQA only. Figure 3 shows these categories as a percentage of all applications.

FIGURE 3 SAC categories as a percentage of total applications



Data source: NZQA

The numbers in Figure 3 add to more than 100% because multiple accommodations are needed by many students. Separate accommodation (defined as a separate space to eliminate distractions) is the most frequently accessed condition. Around three-quarters of students who access SAC need a writer, while just under half need a reader. If this is a person reading aloud (as opposed to a text-to-speech digital accommodation) a separate space is very likely to be needed, so that other students are not disrupted. Digital text-to-speech can be delivered silently via headphones.

Earlier research by Burgon (2018) reported that the “playing field” is not level in terms of SAC provision. Although a student may receive SAC approval, whether or not they receive the actual support could be dependent on the resourcing their school is able to manage. This challenge is likely to be particularly acute for internal assessments and common assessment tasks, where both the assessment administration and the provision of SAC are the school’s responsibility. If the school encounters complex resourcing issues (see next section), students could miss out on receiving the support to which they are entitled. Guidance available through the NZQA website acknowledges this possibility in relation to digital exams and CATs.⁵

Participation in digital assessments

In June 2022, just as the preparation of this report was in its final stages, NZQA released data that describe aspects of students’ participation in digital assessments.⁶ Given the potential of digital delivery to support SAC, we provide some discussion of the data below.

On balance, the number of schools taking part in digitally delivered examinations appears to be increasing. In 2021, 323 secondary schools participated in NCEA Online Exams. Fifty-six of these were taking part for the first time but 17 schools that participated in digital exams in 2020 did not participate in 2021.

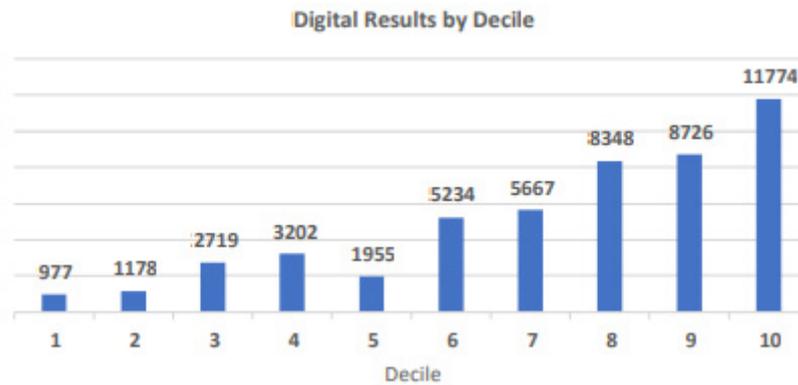
There were more opportunities to participate digitally in 2021. Compared to 2020, the number of digital exam sessions increased from 58 to 67. The number of subjects in which these assessments were offered increased from 21 to 24, over Levels 1–3, and the number of assessment standards (AS) assessed this way increased from 134 to 151. To give an idea of the scale of digital provision, these numbers translate to 6.5% of all external assessments in 2021. There is clearly a considerable way to go if the need for SAC is going to be scaled back by digitally delivered assessment accommodations. (But note that these data only cover externally assessed standards. In principle, any schools could design their own digitally delivered internal assessments if the design of the AS allows for this.)

As with access to SAC provision in general, there were decile-related differences in participation in digital exams. Figure 4 shows the much higher numbers of results recorded deciles 8–10 schools.

5 <https://www.nzqa.govt.nz/providers-partners/assessment-and-moderation-of-standards/managing-national-assessment-in-schools/special-assessment-conditions/sac-for-digital-external-assessments/>

6 <https://www.nzqa.govt.nz/assets/NCEA-Online-2021-Summary-Dashboard.pdf>

FIGURE 4 Number of digital assessment results recorded, by decile of school



As noted earlier, lower decile schools have tended to use more internally assessed standards and digital exams are, by definition, external assessments. Resourcing issues are also likely to have played a part (see Section 3). Interestingly, “void” results made up 8.1% of total digital entries. This means that either the students did not show up, or, if they did, they did not complete the assessment. NZQA noted that measures taken to contain the COVID-19 pandemic were likely to have impacted examination attendance. It also seems probable that the learning and attendance of students in lower decile schools was comparatively more severely impacted. Specific learning needs create equity challenges, but so do socioeconomic differences in opportunities to learn.

Familiarity with IT, and ability to carry out the tasks required, seems unlikely to be a barrier for most students. A clear majority of students use technology as part of their learning, at least quite often. In 2021, most students (85%) agreed or strongly agreed that sitting their examination digitally was a positive experience. Most (84%) did not encounter any problems with the technology during the examination. These responses suggest there is already a positive foundation on which to build increased access to digital accommodations.

3. A snapshot of SAC practices in schools

We now outline qualitative findings from the field work carried out in a sample of schools and RTLB clusters. The logistics involved in the provision of special assessment conditions for both internal and external assessments are discussed. This snapshot includes a reference to assistive technologies used by examination boards elsewhere, following feedback from schools about the potential for these to be used to support SAC here in New Zealand.

We canvassed schools in rural and urban areas, covering a range of deciles and with diverse student populations. Eight schools were included, four from rural/minor urban areas and four from major urban areas. The school demographics are summarised in Table 3.

TABLE 3 Summary of demographic data of schools interviewed about SAC⁷

Demographic		Range
Roll size (median = 829)		345–1,549
Isolation index*		0.04–2.7
Decile rating**		2–10
Ethnicity % in school	Pākehā	18–76
	Māori	12–87
	Pacific	1–16
	Asian	1–25
	Other	1–8

* Median isolation index 1.03 (based on a school's distance from the nearest population centres, where a higher value indicates greater isolation).

** Modal decile rating 2.

Interviews were carried out during May 2022. Initial contact with schools was made with the principal and the principal's nominee, if we could determine who held that role, as well as the SENCO and head of learning support. Of the eight schools where we carried out field work, four school interviews included either the principal or principal's nominee and the head of learning support or the SENCO, three school interviews involved only the head of learning support, and one school interview involved the head of learning support and RTLB attached to the school.

In at least two of the schools in our sample, NCEA assessment of courses set at Level 1 has been discontinued. High levels of assessment pressure, and increasing student anxiety, were given as contributing reasons.

7 Source for all demographic data was New Zealand Schools Directory, Education Counts, New Zealand Government.

Is there an increasing need for SAC?

The data presented in Section 2 showed a growth in applications for SAC when comparing 2019 data with 2020 data. We wondered if this was a trend or a change specific to 2020. Responses from teachers in the field provide possible explanations for the increase.

In the schools we visited with well-developed systems to transition students into the secondary system from their previous schools, learning support staff and SENCOs reported seeing increasing numbers of Year 9 students requiring learning support and potentially future SAC. They also reported higher numbers of students arriving at secondary school with anxiety issues. Some learning support staff saw these as actual increases, but others considered that the need had always existed, and they now had better processes for identifying specific learning needs of their students. They also acknowledged the work NZQA has done to make the application process for SAC more straightforward.

Some interviewees said that students are now less likely to feel stigmatised if they speak up about needing SAC and that this need is now more likely to be acknowledged. Perhaps acceptance of more diverse learning needs is increasing? This could be an interesting question to investigate more systematically.

Another possible reason for observed increases could be that Education Review Office (ERO) requirements of schools related to strengthening learning opportunities and analysing and reporting outcome data have contributed to increased identification of students who require SAC, when their visits result in strengthened teaching and learning systems.

SAC practices vary considerably

We found that the way in which schools organised SAC varied considerably and depended on the school's leadership and support, and on how well developed the systems used to manage SAC were. In all schools we visited, a major part of the responsibility for SAC lay with either the learning support department or the SENCO, with oversight from the principal's nominee. There were, however, clear differences in how SAC was organised and how much of the responsibility for it lay with learning support staff or was shared with other staff within the school. We found variation in four key areas:

- systems used for identification of eligible students
- how applications were managed
- how logistical issues are dealt with in providing SAC
- how SAC is resourced in the school.

Identifying students who require SAC

How schools identified students who may need SAC varied. Some schools routinely screened all new entrants for potential learning difficulties and the possible need for SAC support, following up as the student approached NCEA assessments. In other schools, the SENCO or head of learning support asked teachers to identify students in Years 10 and 11 who might require SAC.

Screening on entry appeared more certain to identify students and engage with them early. This ensured that support could be more tailored to student needs. However, some schools said they did not have sufficient resourcing for this approach, instead relying on teachers to identify and refer students with potential SAC needs. All the schools we approached agreed that identification was better as early as possible and at the latest by the time a student reached Year 10. They recognised that not having access to SAC could potentially be detrimental to student success

and that early identification allowed for practice and evaluation before high-stakes assessments. When identification happened early, students had the time to familiarise themselves with SAC and determine whether it would make a difference.

Collecting evidence to support a SAC application

Staff from the schools we approached described the application process as straightforward, and NZQA responsive and supportive. However, collecting evidence to support applications remained time consuming. Many schools relied on a suite of online tests known as Lucid Exact⁸ and LASS and completion of a PATOSS⁹ handwriting test. The tests provide evidence to support applications and help teachers to justify their own observations, particularly if they are unsure if students meet the SAC criteria. The alternative was to provide teacher observations and more time-consuming tests such as YARC,¹⁰ which few teachers were trained to use.

Several inequities became apparent during the evidence-gathering stage. The Lucid Exact test is now web-based, requiring the purchase of a licence, and payment for each test. A few schools could fund the subscription licence and per student fee (reported to be approximately \$35), but others relied on their RTLB service for the test. These schools needed to arrange appointments, check RTLB availability, and ensure students were present on the day of their test. In rural areas where RTLBs liaised with multiple schools at some distance, schools were being asked to ensure that students they referred were likely to be eligible and were working at Level 5 of the curriculum. These requirements can be problematic when schools count on the testing to confirm their teacher observations. Interviewees were variously concerned about: the loss of student learning time to testing; difficulties scheduling RTLB appointments, with sometimes reluctant student attenders; and the lack of available practical alternatives.

Some schools asked whether the costs for testing could be paid for by the MOE or NZQA. They would like to see some means of subsidising the test software or developing software that is designed specifically for use with New Zealand students, with NCEA accessibility in focus. Given that making NCEA more accessible was a key aim for the NCEA change package, these seem to be reasonable questions for schools to be asking. Funded support for testing would help remove the inequities we noticed between schools.

Logistical challenges

Equitable access to SAC can be impeded by logistical issues. Every school we contacted made provision for SAC support for internal assessments. Smaller schools relied on teachers and students to signal when SAC might be needed. Larger schools ran booking systems for SAC using the school administration system or a separate booking app.

All schools acknowledged difficulties in resourcing reader-writers for both the internal and external assessments. Given the varying need for reader-writers throughout term time, several schools relied on volunteers. This can mean that they need to manage last minute cancellations, or difficulties in providing a reader-writer who has already practised with a particular student. Some schools paid reader-writers because they were then less likely to cancel. These schools absorbed the costs in their budget. More commonly, schools used their existing teacher aides, taking them away from their roles

8 Lucid Exact and LASS diagnostic tests are produced by GL Assessment in the UK (<https://www.gl-assessment.co.uk/assessments/>) available through IT distributors in New Zealand.

9 PATOSS-Professional Association of Teachers of Students with Specific learning difficulties. <https://www.patoss-dyslexia.org>

10 YARC—York Assessment of Reading for Comprehension.

supporting students in the junior school. At least one school called upon senior students and past students and encouraged them to add the experience as community service into their CVs. All the schools recognised that matching suitable reader-writers to students was a priority: often students practise with two or three people to increase the chances of at least one being available. To ensure this availability, schools often assigned a teacher aide who was already working at the school.

Section 2 reported that separate accommodation was the largest special assessment condition approved by NZQA in 2019 and 2020. This is required when students suffer from anxiety or health issues or require a reader-writer or other assistive technology that cannot be used in a classroom. Many schools said they struggled to provide enough separate spaces when managing SAC. Schools with open plan learning spaces and schools experiencing rebuilds found this an acute problem. All schools we interviewed said it was a concern as the number of students requiring SAC increases year-on-year.

Variability in the resourcing of SAC

Equitable access to SAC can also be impeded by resourcing issues.

Time availability for SAC-related work was not consistent across schools. In our sample, the lower decile schools tended to have fewer available staff to co-ordinate SAC provision for students. Most staff with responsibility for SAC also managed teaching loads, some of them substantial. Only one of these schools employed a student services co-ordinator with flexibility to address SAC needs as needed. By contrast, in three schools we visited, additional funding of learning support and SENCO hours had been allocated to assist with the workload in making SAC applications and managing their resourcing.

Provision of SAC for students throughout the year is typically co-ordinated by the learning support department or the SENCO. In schools where there was awareness of the proposed changes to Level 1 NCEA achievement standards, staff expressed concern about how they would manage the increased demand for resources to support school-based assessments.

At least two schools expressed concern about access to appropriate information technology when resourcing SAC. They acknowledged that applying for computer use for their students was now much easier, but it was an additional resource to manage, and it could not be assumed that students had the computer skills or familiarity with the technology.

Several schools enquired about the availability of text-to-speech technology and asked why this was not available as a standard provision with digital assessments. Some schools that expressed concern about the potential increase in school-based assessment following on from the NCEA change programme suggested that text-to-speech technology would be a “game-changer” for many students requiring SAC. The technology was used by some schools for internal assessments where it was not judged to impact on the authenticity or validity of the assessment.

Concerns about SAC provision for numeracy and literacy standards

Several schools raised concerns about plans to introduce external numeracy and literacy assessments. These will soon become co-requisites for the achievement of an NCEA qualification. Schools that run support programmes for students with dyslexia are very concerned that access to reader-writer accommodation may be restricted, particularly where students need to demonstrate proficiency in

decoding written texts. These schools perceive that, if dyslexic students cannot meet the literacy co-requisite, they are effectively shut out of any realistic hope of gaining an NCEA qualification at all.

We heard that the co-requisite requirement has the potential to work against more equitable access to ongoing learning for students with specific learning needs. The concerns raised with us looked ahead to the consequences of not being able to gain an NCEA award, even when subject-based achievement standards and/or unit standards are successfully gained. Failure to meet the co-requisites, and hence not gain an NCEA award, will act as a barrier to students who wish to access vocational courses, such as those offered by trades academies. Schools that run special classes for dyslexic students anticipate that the successes they have been experiencing will be undone when students realise the futility of working towards a qualification that is effectively beyond their grasp.

Our enquiries of NZQA regarding these concerns indicated that the use of reader-writers has yet to be confirmed prior to the pilot assessments taking place. This is an area of high concern that we recommend is addressed urgently.

Joined-up learning and assessment support

We investigated whether there is a disconnect between learning support systems and assessment support in the schools we visited.

We found that the way in which SAC processes and logistics were managed and organised was different in every school we visited. As already noted, responsibility lay with either the learning support lead or the SENCO, either jointly with the principal's nominee or delegated fully to learning support. In some schools, support from senior leadership was evident, as were clear working relationships between the NCEA principal's nominee and the SENCO or head of learning support, with well-developed processes in place. However, this was not the case in all schools. Some SENCOs said they did not have enough allocated time and support for making applications, collecting supporting evidence, and dealing with issues when managing SAC resourcing.

It was reassuring to find that many schools had developed, or were in the process of developing, systems that made connections between support for students in their learning and support for NCEA assessments. This was evident in schools where processes for transitioning of students from primary and intermediate into secondary schools were established. Three schools had well-embedded systems where learning support departments routinely interviewed all new entrants and their whānau, identifying learning support needs and noting the potential requirement for SAC early. These schools had online systems to track students and manage both their learning and assessment needs. The learning support staff felt more confident that they would identify most students who might require SAC at least by Year 10, enabling a period of trialling and practice to ensure the entitlement was both necessary and likely to be successful. The online systems also provided a means of communication with subject teachers for signalling SAC requirements for internals and school-based assessments.

A smaller number of the schools we visited did not have established transitioning processes for new students joining from elsewhere. Transitioning processes allow existing learning support needs to be identified, and students monitored for future potential SAC entitlements. In these schools, it was harder for learning support staff and SENCOs to manage ongoing monitoring of students as they progressed towards NCEA assessment. In four schools, learning support leads or SENCOs requested referrals from teachers at different times of the year, starting in Year 10. Staff said they had identified

most students who needed SAC by their examination year. However, they were not fully confident that they had captured all student SAC entitlements. We heard about cases of late referrals, and student self-referrals occurring as late as Level 2 externals.

Some schools were able to rely on support from a learning support co-ordinator or RTLB, but this varied widely and depended largely on whether the RTLB or learning support co-ordinator was attached to the school or had relevant experience of NCEA and its systems. We turn to this issue next.

How RTLBs support inclusive assessment practice

We investigated how RTLBs support inclusive practice in New Zealand secondary education.

Our investigations began with a scoping meeting with representatives from the Ministry of Education. This was followed with a focus group meeting attended by members of an RTLB cluster supporting several rural schools and individual interviews with RTLBs supporting urban schools, kura, and an RTLB based in a secondary school.

We found that RTLB support at the senior secondary level in general tends to be minimal, and that support around SAC was varied.

- In some clusters, at least one RTLB took a more involved role. They carried out testing and they held the licences for the online tests that are commonly used to support a SAC application (e.g., Lucid Exact or LASS).
- In clusters where RTLBs were less directly involved, they worked with the SENCO or learning support staff, showing them how to complete the testing for SAC applications using online tests.
- Some RTLB clusters were not involved with SAC at all. They did not work with senior level students and had no capacity or funding to resource this area.

We heard that many RTLB clusters do not have an RTLB who is secondary trained. Hence no-one in the team has experience with NCEA. In one cluster, where schools in the area relied on the RTLB service to support them with SAC testing, a secondary-trained team member taught some of the primary-trained RTLBs to carry out the role. However, unless the RTLB was directly attached to the school, travel, appointment making, and navigating student attendance meant the RTLB role in SAC was less efficient than if the school managed the process themselves. As described earlier, some schools funded the costs of the testing themselves to remove these barriers, but reported this as a significant learning support cost, both monetary and in staff time.

Schools reported that they frequently contacted NZQA directly for advice around SAC matters and relied on the online and face-to-face training that NZQA provided.

It appeared that RTLBs do provide support for kura, but the extent to which this support addresses inclusive practice within the NCEA context was beyond the scope of our investigation. We were advised by an RTLB who did support kura that some kura are still not aware that there are testing instruments available in Te Reo Māori to support SAC. This issue needs further investigation.

Overall, the level of RTLB support for inclusive practice in the NCEA context appeared to be much reduced in this investigation compared to that reported in an earlier study (Burgon, 2018). Perhaps supporting the application for, and resourcing of, SAC is not the most practical use of an RTLB resource. A more holistic approach to inclusive practice may lie in strengthening the RTLB role in guidance and training to help schools build stronger connections between learning support and assessment support.

We heard from both schools and RTLBs we interviewed that the rapid turnover in SENCOs, and in some cases a lack of support for SENCOs new to their role, could be problematic. Those schools with well-developed systems in place to keep track of their students' learning needs from the start were much better placed to respond to student needs for assessments and manage SAC requirements effectively.

RTLBs could be well placed to work together to produce and deliver a professional development training and support package related to strengthening the connections between learning support and assessment support for the secondary schools within their clusters. This package would assist schools to set up robust systems to identify and track learning support needs, and potential SAC entitlements, of students entering secondary school at Year 9, or earlier for schools with Year 7 intakes. Once in place, these systems would support schools in requesting and effectively utilising transitioning information from primary and intermediate feeder schools. They would also connect learning support with potential future SAC needs at an early stage to help with planning and resource management.

The development of a systems support package that could be tailored to the needs of each school has the potential to deliver greater consistency, reduce the disconnect we observed between learning support and assessment support that schools are currently managing internally, and provide greater equity between schools.

RTLBs may already be aware of schools that have well developed systems for identifying and tracking learning support needs, including online systems, which could be used as examples of good practice for the benefit of all schools. Funding would be required to develop support packages and ensure online systems are accessible for all schools, as well as increasing RTLB training and capacity in the secondary schools' sector. This would help to make the current system less variable and more equitable between schools in the short and medium term.

Text-to-speech and other assistive technologies: Examples from elsewhere

Several schools we visited said that digital NCEA exams should include a text-to-speech function. This could reduce (but not eliminate) the need to resource reader-writers. We found examples of this technology being available in Scotland through the Scottish Qualifications Authority (SQA) and through the Education Quality and Accountability Office (EQAO) in Ontario, Canada.¹¹

The EQAO administers the Ontario Secondary School Literacy Test (OSSLT) which is a requirement to earn an Ontario Secondary School Diploma. The test is an e-assessment with a range of built-in tools to support success for students who require accommodations.

These include:

- a built-in text-to-speech function that reads the text on the screen out loud as well as supporting a further four text-to-speech function third party software types
- zoom in and zoom out capabilities
- a line reader that helps students focus on one line of text at a time
- a high-contrast view
- annotation tools (highlighter, line, eraser) and rough notes.

¹¹ Detail was sourced from <https://www.eqao.com/the-assessments/osslt/>

Additional accommodations are available, dependent on availability. These must be consistent with those in a student's Individual Education Plan and could be alternative formats of the test for students with special education needs. For example:

- a Unified English Braille version of the test (contracted or uncontracted)
- MP3s with audio descriptions
- an alternative version of the test in the e-assessment system with audio descriptions.

SQA operates a quality assurance system which audits the systems and procedures used by its examination centres to identify and verify candidates' need for assessment arrangements. It does not require centres to seek separate approval to produce adapted assessment material for internal assessments, provided the adaptations made do not change the AS or competence standards. The authority suggests that the:

inherent flexibility of some internal assessments will allow many candidates to achieve the standard for the qualification without the need for an assessment arrangement to be provided. For example, candidates can provide evidence in a variety of ways, using a range of assessment methods to demonstrate their knowledge and understanding.¹²

SQA allows centres to apply for the use of speech recognition as arrangements for external assessments. Learners with writing or spelling difficulties, visual impairment, or physical disability are eligible. SQA provides comprehensive guidelines to centres concerning the application and use of this technology. It notes that "candidates must be confident, competent users of the technology before it is used in an assessment or examination" (p. 1). It also advises centres about contingencies should there be issues with IT systems or hardware. The use of a physical reader is reserved for candidates who have "substantial difficulties with reading text and who cannot access material by any other appropriate means, for example, by using a text/screen reader"¹³

These two examples suggest that similar technology could apply to NCEA assessments and possibly it could be incorporated into a universal design approach to assessments.

Student control and agency

The sustainability of the type of model currently used for SAC in NCEA is summed up by Smith and Buchannan (2012), from Western Illinois University. Their research took place in a tertiary setting but is also relevant for the NCEA context. They say that the use of individualised accommodations as the sole method for resolving access barriers in the classroom is neither sustainable nor equitable and has implications for the students themselves:

This model requires constant administrative oversight on the part of the disability resource staff, places additional responsibilities on students with disabilities beyond what their non-disabled classmates experience, and often puts disability service providers at odds with faculty. In addition, the process itself creates a systemic barrier and serves to perpetuate the myth that persons with disabilities require assistance. The locus of control is placed within the disability resource department rather than with the primary constituents in a classroom setting, the students and faculty. (Smith & Buchannan, 2012, p. 259)

¹² https://www.sqa.org.uk/sqa/files_ccc/AssessmentArrangementsExplained.pdf, p. 6.

¹³ https://www.sqa.org.uk/sqa/files_ccc/AssessmentArrangementsExplained.pdf, p. 22.

They make two vital observations that resonate with the SAC “bolt-on” model that applies not only for NCEA but is also commonly used in other countries. First, application of a process such as SAC perpetuates the myth that students with disabilities require assistance so that they can achieve academically. Second, the locus of control does not sit with the student who requires SAC, as it does for other students who are taking the same assessment. SAC is intended to support students in their assessments, yet in many ways it may restrict them.

We asked teachers in schools we visited if they could give any examples where their students had taken a position of agency or greater control of their assessment needs. Whereas some schools felt their students would struggle to manage this, they frequently described the increased confidence their students experienced in knowing that SAC had been approved for them. Approval told them that their teachers and NZQA recognised they had the ability to achieve in NCEA.

We were given some specific examples where student agency played an important role in managing individual assessments. One example involved students in a composite Level 1 and 2 mathematics class who co-created a selection of their assessments with their teachers. Students, including those with additional learning needs, managed the assessment design relying on the advice and guidance of their teachers, using the standard, specifications, and the rubric. The teachers noticed that their students had a greater understanding of the assessment purpose and approached their tasks with greater confidence. This was standard practice for the class and provided us with an early window into a potentially new approach to assessment that embraces inclusiveness at the start of the assessment journey instead of accommodating it at the end.

Is the term **Special Assessment Conditions** appropriate?

Before our discussion moves to how the SAC “bolt-on” model could be changed, it is important to consider whether the process should be renamed so that the emphasis moves from adjustments and special conditions towards inclusiveness and barrier-free access that should come with entitlement. Our fieldwork did not provide us with a satisfactory solution for renaming SAC, but we return to this challenge in the final section.

In our milestone report submitted in April 2022, we recommended that an oversight panel with responsibility for reviewing and ensuring inclusiveness in NCEA assessment design should be established. This recommendation came from our research into how special assessment provisions are managed within the South Australian Certificate of Education (SACE).¹⁴

We reported that a Special Provisions Advisory Panel is appointed by the SACE board Chief Executive to monitor, evaluate, and conduct research into the appropriateness and effectiveness of the Special Provisions in Curriculum and Assessment Policy and its procedures. The panel is composed of a broad range of interested parties from schooling, the Equal Opportunities Commission, ministerial advisers for children and students with disabilities, and other disabilities advocates. This panel provides accountability for the special provisions policy and monitors how well it serves its purpose.

If the recommendation to establish an advisory panel is adopted, the renaming of SAC to reflect more inclusive terminology could be an important initial task for them. Meanwhile, we would suggest considering the use of terms that reflect the important concepts of assurance and inclusion, assuring greater equity within assessment. We have made tentative suggestions about this in our recommendations.

¹⁴ SACE Special provisions in curriculum and assessment: <https://www.sace.sa.edu.au/web/special-provisions/legislation>

In summary

Although interviewees in all the field work schools described SAC applications and provision as a priority for their senior school students, it was evident that the barriers faced by schools in resourcing SAC contributed to inequalities in: identification of student need; processes for making applications; and the actual provision of SAC.

SAC resourcing is time consuming and costly for schools. Those who have been able to develop effective management systems have become victims of their own success in that they have identified a greater level of need, and this has placed a greater strain on their resourcing. Some schools are very aware of the potential increase in SAC resourcing throughout the year should more flexibly school-based assessments be introduced from 2023.

Some suggestions for immediate improvement did emerge during our field work conversations. While SAC provision remains essentially a “bolt-on” to NCEA, possible changes identified by the schools include:

- Access to the testing software for all schools and kura to make consistent and equitable judgements about SAC requirements for their students. Ideally, this would be software designed for New Zealand students, including a version in te reo Māori, and would be funded for schools and kura.
- Inclusion of text to speech in all digital examinations so that it becomes available to all who may require it. This should be an urgent priority.
- Funding training and support packages for schools and kura designed to effectively identify, track, and manage their students’ learning needs and identify potential SAC support as early as possible. Ideally, schools and kura would receive access to an online system that could be tailored to their specific needs, leading to greater consistency and equity across schools.

In the medium term, however, it will be important to identify ways in which assessments can be designed and delivered to reduce the current level of need for assessment support. This would foster greater inclusivity for students with learning disabilities. We turn to this challenge in the next section.

4. Why assessment design matters

This section discusses the potential for the design of assessments to provide more equitable opportunities for all students to show what they know and can do. Early scoping conversations alerted us to the potential for Universal Design for Learning (UDL) principles and processes to make assessment more inclusive without necessarily needing to create special conditions for some students. Such a solution has obvious appeal in view of the growing demand for SAC, as set out in earlier sections of the report. There are, however, some substantive challenges to enacting UDL in a high-stakes assessment context. We discuss these challenges in this section.

Some challenges were raised in research papers we accessed during the literature sweep, and we heard about related concerns during the field work phase of the project. Relevant details from both sources are brought together as we describe the application of UDL principles to assessment, then outline our findings about the potential and challenges of this approach.

A brief overview of Universal Design for Learning

Universal Design for Learning opens the door to rethinking assessments—to ensure that the assessments themselves are not the barriers to improved learning. Applying UDL principles to assessment processes is a promising approach to providing appropriate assessment conditions for all students, giving each student a comparable opportunity to demonstrate achievement of the standards being tested (Thompson et al., 2004, p. 3).

The concept of UDL has its origins in neuroscience research of how individuals learn. It has parallels to, but is distinct from, the concept of Universal Design that emerged in the field of architecture. In architecture the principle is that buildings should be designed in ways that allow access for people with the widest possible range of mobility issues, without needing any special accommodations to be made. UDL similarly posits that learning (and assessment) experiences should be designed in ways that take account of neurodiversity in meaning-making, and hence in learning. Both UD and UDL draw on a socio-ecological theoretical framework. In this framing, diversity of abilities is seen as a part of human functioning and is not treated as atypical (Karvonen et al., 2020).

Universal Design for Learning is underpinned by three basic principles that take account of neurodiversity in three meaning-making neural networks (Glass et al., 2013). These principles aim to ensure that the design of learning should be such that students with specific learning needs have the same opportunities to participate as students who do not have such needs, without any sense of being singled out. The TKI guide to UDL¹⁵ summarises these three principles as follows:

- Students should be able to *engage* with the learning in multiple ways. Suggested strategies include focusing on affective aspects of engagement and being flexible in the provision of learning options; ways of sustaining effort; and opportunities for self-regulation.
- Students should be able to access multiple ways of representing learning. This principle supports *meaning-making* for students who perceive and understand information and ideas differently.

¹⁵ <https://inclusive.tki.org.nz/guides/universal-design-for-learning/>

- Students should be able to use a range of ways to *do things and to express their ideas*. This principle focuses on students' active role in learning.

Glass et al. (2013) categorise these three principles as addressing: the why of learning via the affective neural network; the what of learning via the representation neural network; and the how of learning via the strategic neural network. Neuroscience research has established diversity of meaning making in all three networks, making overall learning much more complex than was ever previously imagined.

While UDL focuses on the learner and the learning environment, UDL for assessment addresses the interaction of the student being assessed with the assessment task. From a UDL perspective, the “disability” lies in “the gap between the person’s abilities and capacities and the demands of the environment” (Karvonen et al., 2020, p. 3). The problem is *not* framed as belonging to the individual alone and the aim of UDL for assessment is to reduce the gap between the individual and the task (*ibid*).

This is a very brief overview of the theory. How it plays out in practice is another matter. Recently, one of the original champions of UDL has described “cracks in the foundations” of the overall initiative (Rose, 2021). With the benefit of hindsight that comes from a long career in education, two of the cracks he describes are as follows:

To design a curriculum in which all students will feel equally challenged, supported, and belonging will require more than making the existing curriculum more accessible, it will require universally re-designing the *mission* or goals of the curriculum. (p. 6)

A student in a wheelchair continually finds that many people, including teachers and other students, will generate implicit biases about their intelligence, ambition, education, future promise, etc., that are not based on their ability but on their identity as a disabled person. The same is true for learners with labels like autism, learning disability, ADHD, and so forth. They all face the barriers of identity and lowered expectations that are not addressed in the existing UDL Guidelines. (p. 12)

These cautions are important caveats to what UDL can be practically expected to achieve. While the issues raised relate to learning in the first instance, they have important implications for assessment. They also accord with concerns that we heard during the field work phase of the project, as outlined in what follows.

The low-hanging fruit: Deploying UDL-informed accommodations to make quick gains

In much of the literature we sourced, UD principles are applied to traditional formal assessments. In the NCEA context, similar opportunities would apply to external assessments completed under examination conditions and Common Achievement Tests (CATs). They would also include internal assessments completed under formal test conditions. The gap between the student and a traditional assessment task can be reduced by providing universally designed accommodations for completion of that task.

Using assistive technologies to accommodate reading challenges

Use of a reader–writer is one of the most common traditional SAC provisions. In Sections 2 and 3 we outlined *practical* challenges associated with the provision of a reader–writer. Further to these concerns, Dolan et al. (2005) summarised potential issues that could compromise the *validity* of assessments when human readers are used:

- Reading aloud can vary in quality. Some readers might mispronounce or misread words.
- Human readers might inadvertently influence students' attention or responses by their intonations or non-scripted responses.
- Students might be reluctant, or might not be allowed, to ask for portions to be reread if they didn't hear them the first time.
- The act of reading can impose a linear navigation path and a set pace through the work, both of which can negatively impact test performance.

In the field, we heard that those who support students with specific learning needs are keenly interested in the use of assistive technologies, and in particular text-to-speech functionality. As digital assessments become more prevalent, there is an expectation that this type of practical support will be designed into assessment tasks from the outset. Here, potentially, is a source of “low-hanging fruit” in UDL-informed approaches to SAC.

Other suggestions of making universal accommodations

Table 4 categorises one research team's suggestions for accommodations that might be comparatively easily made in more formal assessment contexts, where a written response is typically required. The left-hand column shows how this is done when SAC students are singled out for special treatment whereas the right-hand column suggests accommodations that could be made to the whole assessment, and hence be available for all students. They do not single out those with specific learning needs but instead design assessment conditions to reduce the gap between those needs and the assessment task.

TABLE 4 **How UDL can be used to modify traditional accommodation practices**

Category	Traditional accommodation	UDL changes to make the practice more inclusive
Presentation	Have a facilitator read test material to students with learning needs	Design testing materials at reading level accessible to all students
Setting	Private room for students who are easily distracted	Provide students with choices as to when and where tests are completed
Timing	Give students with learning needs extended time for tests	Design tests to be free of time pressure for all students
Response	Provide use of a scribe for students with learning needs	Provide students with assistive technology options for responses

Adapted from Goegan et al. (2018).

These suggestions are not without challenges. It is important that an accommodation does not make the assessment easier than intended (Cohen et al., 2019). However, as suggested by Rose (see second quote above) there are also likely to be “hearts and minds” issues to address. There is a risk of perceptions of unfair advantage when some students access the accommodation and others do not. A limited body of research on this point suggests that all students can potentially benefit from these types of accommodations, not just those with specific learning needs (e.g., Scalise et al., 2018, in the context of STEM subjects) but this challenge might need to be surfaced and addressed in professional learning and development (PLD) contexts. Ensuring authenticity is a different type of challenge again, as we next outline.

Addressing authenticity concerns

The “setting” example in the right-hand column of Table 4 above would almost certainly raise questions about authenticity: how do we know that the student actually did the assessment without additional assistance when they can choose where and when it is completed? This is not just a concern for more formal assessments. The question is also relevant to other types of internally assessed tasks. There are interesting parallels with authenticity concerns raised in the context of conducting NCEA assessments during the COVID-19 pandemic, when many students were working from home. NZQA has provided practical advice to principals’ nominees about how to ensure authenticity in such cases (NZQA, 2022). They suggest using a mix of authenticity processes, including:

- tracking student progress towards the completion of assessments using milestones or checkpoints rather than relying on a single point to measure authenticity
- using digital tools such as revision history to monitor work in progress
- talking with the student to check their understanding of what they have submitted before awarding a final grade. (p. 5)

All three practical suggestions could be applied to SAC accommodations, with the proviso that the judgement does not rest only on the result of a single formal assessment task. The relevance of the suggestions in more than one context (i.e., SAC challenges and hybrid learning challenges) points to their potential as another source of comparatively quick gains when the focus is on professional learning related to implementing UDL approaches. There could be “two bangs for the buck”. In Section 6 we elaborate on opportunities afforded by hybrid learning contexts.

Eliminating “construct-irrelevant” barriers

Specific learning needs such as reading and/or writing issues can be a barrier that prevents some students from successfully demonstrating their learning. If they cannot meaningfully read written questions, they can hardly be expected to answer. If they cannot write an adequate response in the given time, they again will not be able to show their learning. A question or task set in an unfamiliar context or culture might not cue the concept being assessed, even if the student knows this in principle (Perry et al., 2022; Rose, 2021). These are all examples of “construct-irrelevant” barriers—they are ostensibly not what is being assessed.¹⁶

As already outlined, reader–writers or the use of assistive technologies can help overcome some construct-irrelevant barriers to assessment. Rich computer resources can provide contextual cues that help students access prior knowledge, or simply give them a way in to tackling the question (Perry et al., 2022). However, a truly universal approach would suggest that it is better that construct-irrelevant barriers are not part of the assessment to begin with. Here the focus turns to the critical role played by assessment design.

In the literature we reviewed, we found some advocacy for a synthesis of UDL principles with a sharp curriculum focus on what is actually being assessed and why (Karvonen et al., 2020; Ketterlin-Geller, 2008). The process of Evidence-Centred Design (ECD) is seen as one way of creating a much sharper “assessment argument”. A paper that explains and justifies the ECD process¹⁷ defines an assessment argument like this:

¹⁶ Counter examples serve to strengthen the point. For instance, decoding would not be construct-irrelevant in the context of a literacy assessment—at least for that part of the assessment that sought evidence of this specific skill. If, however, lack of ability to decode leads an assessor to conclude that a student did not have comprehension skills, then the act of decoding would present a construct-irrelevant barrier.

¹⁷ It is of note that this paper was initially developed for the influential USA-based Gordon Commission on the Future of Assessment.

An ECD approach to educational assessment design considers which types of evidence would ideally be useful to reason about student learning and infer what students know and can do. Assessment is framed as the process of designing observational contexts (i.e., assessment tasks) that provide such evidence in service of some question, claim, or inference. (Gorin, 2014, p. 5)

Defined like this, an assessment argument clearly specifies the focal knowledge, skills, and abilities being assessed, and then ensures that the actual assessment, and its administration, match the purpose of the assessment as envisaged at the design stage.

An assessment argument, generated via a curriculum design process such as ECD, can be combined with UDL principles to make assessment much more equitable by eliminating construct-irrelevant variables from the initial design stage right through to making valid and reliable judgements about students' work. Table 5 on the next two pages briefly outlines the ECD process in the left-hand column. The middle column aligns the steps of ECD with relevant NCEA processes. The right-hand column suggests challenges that might be anticipated when the focus is on making the actual assessments more inclusive for all students.

Table 5 highlights specific challenges at every stage. For example, stages 1 and 2 of the ECD process develop a clear assessment argument. This is high-level curriculum design work. As Rose noted, the curriculum itself needs to be universally designed in order to build inclusive foundations for the work to follow (Rose, 2021). This has implications for the RAS timelines, which are currently running ahead of the Curriculum Refresh process.

Looking across the whole table, another challenge resides in the division of the overall NCEA assessment process between different groups at different stages. ECD was devised for national testing contexts where professional assessment design teams assure continuity of all stages. By contrast, different groups look after different equivalent stages of NCEA design and delivery. These split responsibilities increase the potential for miscommunication of design intentions, with an associated risk of construct-irrelevant choices creeping in at any stage. We will further discuss this design challenge in the next section.

TABLE 5 A comparison between NCEA processes and ECD, as outlined in the international literature

Stage of ECD	NCEA equivalent	Potential issues
Domain analysis: gathering information on important aspects of the domain, relationships between these aspects, and how they might be represented. Can include identifying standards, defining learning sequences, and grouping concepts that depend on each other.	Initial exploratory of work of each SEG during the RAS process—identifying the big ideas, etc. Oversight is provided by the Ministry of Education’s (MOE’s) internal RAS team.	Elsewhere, this analysis is typically conducted by curriculum experts. SEGS are mainly composed of teachers with expertise in NCEA, but not necessarily in high-level curriculum design.
Domain modelling: an assessment argument is developed. From the domain analysis, focal knowledge, skills, and abilities (FKSAs) are developed. These highlight the aspects of the domain that could be the focus of an assessment. The concepts are further defined, and potential evidence sources identified.	This process equates to the step when each SEG develops a matrix of achievement standards, ¹⁸ to represent the intended focus of assessment across each subject.	Do teachers on the SEGs have the deep curriculum knowledge needed to develop specific arguments about focal knowledge, skills, and abilities that could potentially be assessed? The emphasis on awareness of cognitive demand is a clear theme in the inclusion literature and this is where such demands should be made visible.
Conceptual assessment framework: the assessment argument is further refined and translated into task specifications, which take specific account of the FKSAs identified in the second stage. These are published as a “task template”.	This equates to each SEG developing the actual achievement standards (AS). Each AS specifies the type of assessment activity envisaged, with range notes that provide some guidance for making a judgement. Oversight is provided by a small team of MOE LALs (Learning Area Leads).	At the moment, potential criteria are implicit in the range notes and accompanying guidance, rather than explicit in the AS. There is no “assessment argument” as such because the AS do not include a purpose statement with an explicit curriculum link. ¹⁹
Assessment implementation: tasks and rubrics are developed matching the specifications developed in the task template.	Subject examiners develop assessments for externally assessed AS. QA processes are managed by NZQA’s NAFs (National Assessment Facilitators). Teachers in schools develop tasks for internally assessed AS. QA processes are managed by NZQA’s NAMs (National Assessment Moderators).	A layer of translation is inevitable because external assessments are developed by subject examiners, who may each only work with one AS. Similarly, teachers must interpret the intent of each internally assessed AS and develop their own task-specific criteria.
Assessment delivery: assessment tasks are administered and resulting data analysed to determine if revisions are required. The focus is on differential performance on FKSAs by groups with specific learning needs.	NZQA data analysts and subject moderators would be involved at this stage (we are unclear if a process for doing this already exists).	Is there a systematic process for identifying differential performance of student groups, so that inequalities of opportunity can be highlighted and the system can learn and improve on inclusive assessment design?

¹⁸ The review of NCEA specified that there should be four AS per subject at each curriculum level, but science has been allowed to develop more than this number at level 1, to accommodate the four discipline areas involved.

¹⁹ See Darr and Hipkins (2022).

Designing more flexible assessment tasks

Allowing students to show their learning in different formats, or via alternative assessment tasks, is another UDL strategy to accommodate diverse learning needs. We understand that Subject Expert Groups (SEGs) have been encouraged to model both strategies as part of the RAS process. Section 5 will provide examples from the Level 1 pilot materials created for the RAS. That section will illustrate practical challenges for making equitable judgements:

- How can different-looking forms of evidence be fairly compared? How are assessors cued into what to focus on and what to set aside as construct-irrelevant?
- Do alternative tasks present comparable levels of difficulty? What evidence would tell us that?²⁰

These are questions that have both validity and reliability implications. ECD advocates argue that it is easier to address these challenges if the design process outlined in Table 5 has been followed. A clear assessment argument makes it easier to strip away construct-irrelevant distractions. But again, this can never be an easily implemented cure-all as we now briefly outline.

Innovative assessment tasks

In principle, innovative assessment tasks provide opportunities for students to show their learning in different formats, reducing the need for SAC. The ultimate here is that the evidence needed for assessment is generated as a naturally occurring part of the learning process. It could be “any valid, authentic, and standard-specific evidence produced by a student” and might be generated via “written, digital, photographic, audio-visual and portfolio formats” (NZQA, 2022, p. 7).

NZQA’s advice implies that relevant types of evidence need to be *anticipated* in advance of the learning, and that this anticipation is informed by familiarity with the evidence requirements of the standard, supported by exemplars and benchmarks. We looked for this type of support in the pilot materials but, as the next section outlines, well-intentioned efforts to be inclusive could founder if signals about what evidence counts, and why, are not sharp and clear.

Alternative assessment tasks

Shaw et al. (2020) published a framework for analysing comparability between alternative forms of an assessment test. Their framework indicates 19 points for which specific comparisons need to be made. They divide these comparison points into four broad categories: content of the assessment; demands of the test; marking practices; and awarding (credentialling) practices. Many of these points of decision making seem obvious (e.g., the alternative tests should cover similar content). The decision-points that seem especially salient in the context of SAC accommodations include:

- The cognitive processes (as supported by tools) are the same across tests as far as we can tell [demand category].
- The possible effects of any response format are carefully considered (e.g., for on-screen tests, the effects of typing rather than writing on paper or of using a drop-down list rather than circling a response on paper) [demand category].
- The way student responses are presented to markers needs to give equal opportunity for accurate marking across tests [marking practices].
- Auto-marking (if used) and human marking are both sufficiently accurate and reward intended constructs (only relevant if comparing an on-screen test to a paper-based test) [marking practices]. (Shaw et al., 2020, p. 21)

²⁰ Designing examinations of comparable difficulty year after year is well known to be very difficult, which is why scaling practices are so often used.

This small sample from the 19 decision points in the paper gives the flavour of the complexity of the decision-making involved, even for tasks in essentially the same format, namely a test. It is, of course, possible to design a variety of tasks into a more formal test. In fact, Perry et al. (2022) report that computer adaptive tests, underpinned by a large bank of items, can accommodate task diversity without compromising reliability because psychometric checks allow task difficulty to be measured. But designing such an item bank is a task for assessment experts, and not something that teachers could be expected to do unilaterally, or in their own time. On the positive side, Perry et al. point out that, once the initial design investment has been made, tests can be administered at different times without compromising security. The student never knows what combination of items they are going to encounter. In principle, this sort of flexible timing would support a foundational intention of NCEA—that assessment should occur when the student is ready, not at a fixed time.

A school-wide approach to enacting inclusive assessment

The most common message we heard in the field was that UDL is a great idea, but until teachers see convincing examples in action, they don't know how to go about enacting it, or where to start. For most of our interviewees, it seemed to be in the “too hard” basket. This is hardly surprising given the complexities and challenges outlined above. There was one interesting exception. An interviewee named a specific school that appeared to be successfully enacting UDL, including in their assessment practices. We immediately contacted the senior leader with oversight of curriculum and assessment in this school, and subsequently interviewed them, alongside the school's leader of the special learning needs team.

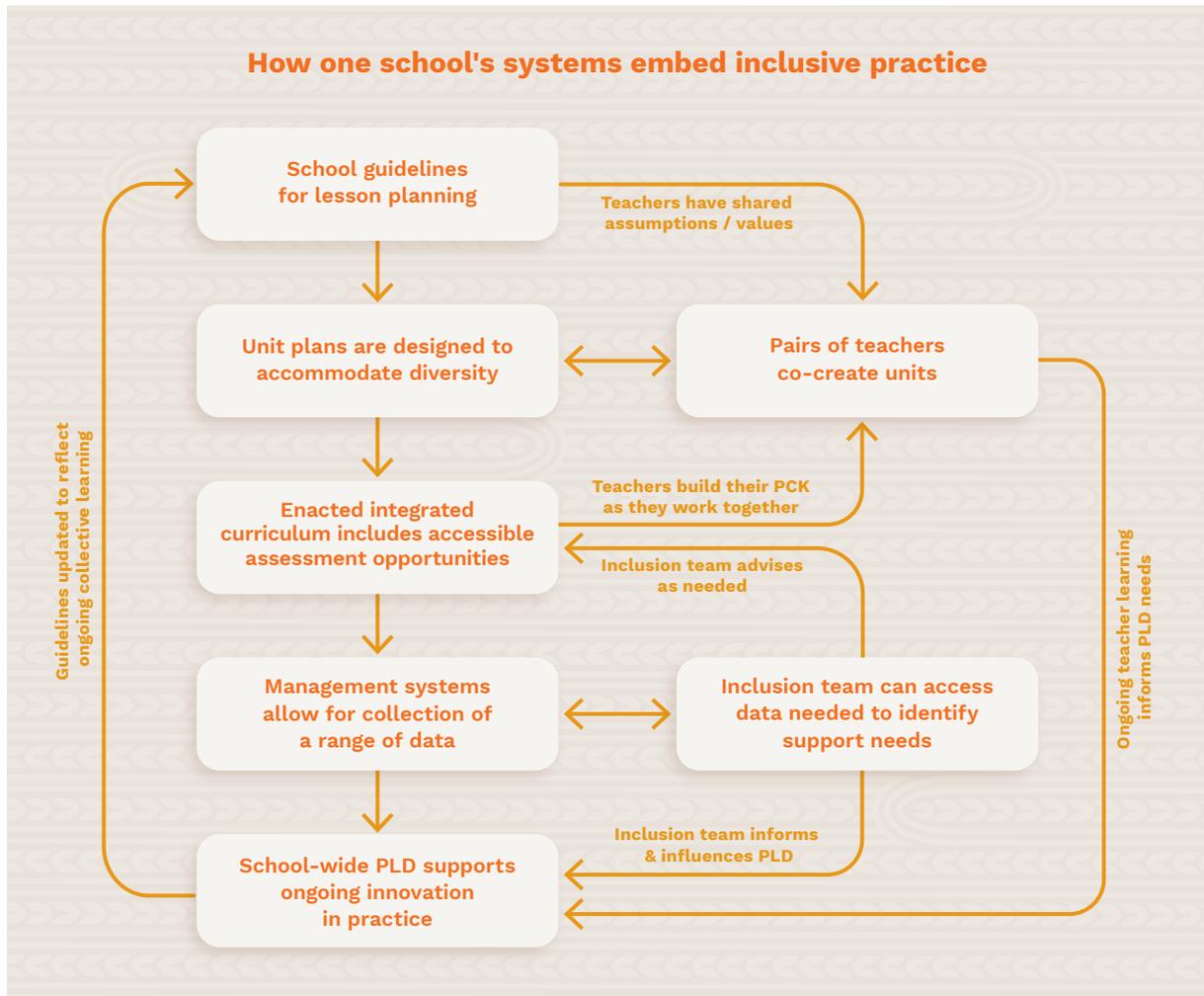
To be clear before we go any further, this school does not specifically enact a UDL model. Their practice is congruent with UDL principles, but this is a consequence of coherence that comes from systematic and aligned attention to every aspect of the learning programme, and to the sustained effort and attention given to meeting the needs of all their learners. Figure 5 on the next page is our attempt to catch the dynamics that were described to us during the interview.

The sketch we made clearly shows the multiple feedback loops and points at which teachers (and the overall system) learn together. The crack in UDL foundations, created by a lack of shared understandings and values, as described by Rose (2021), is minimised as far as humanly possible. Pairs of teachers work together in each unit, building their pedagogical content knowledge (PCK) as they go, including in the design and execution of innovative assessments. Apart from core literacy and numeracy requirements, the school does not assess for Level 1 of NCEA because they want to keep the focus on richly contextualised integrated learning. A strong inclusion team also has a highly visible presence, impacting school-wide practice and both student and teacher learning. Students are unlikely to fall through the cracks. One consequence is that, by the time students with specific learning needs reach the senior secondary school, the special needs systems will have gathered all the necessary evidence for making SAC applications.

The clear message that emerged from this interview was that a pedagogical solution such as UDL cannot be “bolted on” to existing practice. Any such innovation needs to become, in essence, part of the school's DNA—how we do things around here. One interviewee with expertise in UDL agreed strongly with this message. This interviewee was concerned about the potential for UDL to be “blamed when things go wrong” rather than looking at the overall system to try and work out where problems actually originated.

Any change to a complex system needs to start somewhere of course, and systems dynamics suggest that there are potentially many entry points for change. The caveat is that entry points are just that. They cannot single-handedly induce change across the whole system, unless perhaps the specific change is so compelling (and easy to make) that, in effect, it goes viral. Given the complexities of UDL outlined above, that hardly seems a likely scenario. Nevertheless, nothing will change if we don't try, and so, in the next section, we outline exploratory efforts to employ UDL principles when designing new achievement standards and supporting materials as part of the overall review of NCEA.

FIGURE 5 One school's system for embedded inclusive practice



5. Are new NCEA materials more inclusive?

Section 4 set out in-principle approaches for creating assessments that are more inclusive, and hence reduce demand for SAC when carrying out NCEA assessments. Now we turn our attention to a small selection of Level 1 materials generated for use in the pilot schools in the first instance, as part of the ongoing RAS.

We reviewed a range of internally assessed achievement standards, looking for evidence of the application of UDL strategies to the assessment process. We chose to focus on internal assessment because this is where professional learning about UDL might be most meaningfully applied by all teachers in the senior secondary school. This is also where we anticipated that it might be easier for SEGs to begin to apply UDL principles to the design of assessment materials.

The three achievement standards we focused on come from the science learning area of *The New Zealand Curriculum (NZC)*. This is where our own subject expertise lies. We felt we could carry out a better-informed critique in this learning area than would have been possible for us in other learning areas. We anticipate that those with the relevant expertise could readily search for similar opportunities and challenges in other learning areas. In any case, some issues are likely to be generic across the whole suite of draft AS at Level 1.

A nod to UDL

Our analysis begins with AS 92021: *Demonstrate understanding of a chemical reaction in the taiao*. This is one of four new AS created to assess an integrated course in biology and chemistry. For AS 92021, the focus is on a big idea drawn from chemistry: “Chemicals exist everywhere in the taiao and reactions rearrange these chemicals”.

Early indications are that UDL principles have at least been considered. Under the “Conditions” tab for this internally assessed AS, the website lists a range of possibilities for presenting evidence of learning. These are: oral presentation; written report; pamphlet, or article (digital or paper); poster; mural; infographic; cartoon or comic strip; a video or animation; a slideshow. Setting aside concerns about how to make consistent judgements across different task formats (see Section 4) this sounds promising. But how does it play out in practice?

Neutralisation is one of three chemical reactions to be studied. The suggested assessment activity sets neutralisation in one of two possible biological contexts: digestion processes in the stomach, or impacts of ocean acidification on marine life. So far, so promising. However, notwithstanding the advice given under the “Conditions” tab, the assessment task modelled is the completion of a worksheet. The teacher notes indicate that carrying out a simple “experiment” followed by completion of this task sheet could take “up to two hours of class time”. This advice cues a traditional formal test-type setting, which is a one-off event.

There is a nod to UDL in that the answers to the questions on the worksheet could be written or verbal (albeit formally recorded). But the design of the worksheet questions negates any possibility that a student with specific reading or writing needs could complete this task without SAC support.

Here is the introductory text from Worksheets A and B respectively:

Antacid tablets contain a base as their active ingredient, such as calcium carbonate, magnesium hydroxide, aluminium hydroxide, or sodium bicarbonate. They are used to treat heartburn, which occurs when hydrochloric acid rises into the oesophagus from the stomach and ‘burns’ the oesophagus.

Te Moana-nui-a-Kiwa holds a special place in the history, cultures, and identity of Māori and Pacific peoples, and continues to be a source of livelihood and food security for many in the Pacific Islands and Aotearoa New Zealand. Sadly, life in the ocean is being affected by ocean acidification. Oceans are becoming more acidic as CO₂ builds up in the atmosphere and dissolves in seawater to form carbonic acid. This reacts with the shells of plankton—tiny marine ‘drifters’ that are carried through the ocean by tides and currents.

Both opening statements are lexically dense. They contain long sentences with conditional and/or causal logic clauses. These are features of traditional science texts, but ability to read and unpack such texts is presumably not what is being assessed here: the reading demand constitutes a construct-irrelevant variable. We assume that students would have had opportunities to study the biological systems that provide the context for each version of the worksheet. If not, this contextual knowledge would also constitute a construct-irrelevant variable. This could potentially be ameliorated by adding a simple visual (e.g., a basic outline of the top of the digestive system for example A).

While it is beyond the scope of this report, we also have reservations about the substitution of complex concepts from mātauranga Māori for single-word concepts in English. For example, taiao is used to broadly mean environment. We acknowledge the positive intentions behind this idea, but if these ideas are unfamiliar, or only very superficially understood (or actively resisted in worse-case scenarios), their use could act as a construct-irrelevant barrier for some students. Paradoxically, there is at least a theoretical possibility that a Māori student with a deep-lived knowledge of their own culture might read too much into the intended meaning,²¹ and get lost in complexities. We suspect this particularly sensitive issue will apply to many other AS. It is also relevant to the next AS we consider, but we do not revisit the point.

Some of the worksheet questions appeared to invite students to guess what was in the mind of the writer(s). We found ourselves asking exactly what “assessment argument” underpinned the design of each set of questions. The closest we came to finding an answer was the final cell of the “Excellence” column of the marking guide provided for teachers. For the ocean acidification context, the content of this cell reads as follows:

Student’s report²² discusses the application of knowledge of neutralisation as it applies to acids in naturally occurring bodies of water such as mitigation of ocean acidification or pollution. By conveying knowledge of neutralisation products and the effect they have on the mauri of the taiao and the effect on living things, the discussion demonstrates understanding that chemicals do not simply ‘go away’ after a reaction. Rather, matter is conserved but in a different form and that concentration and/or location can have a significant bearing on the effect of the products when acids are present in the taiao.

21 The risk of recolonisation of complex and deep terms from te ao Māori is elaborated in a recent blog post titled “Taking care of our kupu”, on the E-Tangata website: <https://e-tangata.co.nz/reo/taking-care-of-our-kupu/>

22 Arguably, answering a series of short response questions does not constitute a “report”. This appears to be a design mismatch.

Here is a short statement of the “so what” thinking behind the design of the unit of work. Put this way, the somewhat obscure questions in the worksheet about the conservation of mass make more sense. A task designed with this argument right up front would surely look very different from the traditional format worksheets modelled here for teachers. We will revisit this challenge in our third example.

We should point out that the other two assessment activities suggested for this AS do not rely on completion of a worksheet. They do ask for some sort of concluding report, with similar issues to those discussed in the next example.

Taking task design a step further

The next standard we review is AS 92020: *Demonstrate understanding of a microorganism in the taiao*. Again, this is one of four new achievement standards created to assess an integrated course in biology and chemistry, but for this AS the focus is on a big idea drawn from biology: “All living things are interconnected within the taiao”. The first of two paragraphs that scope this big idea is as follows:

To understand the world, ākongā must understand the intricate interconnections within the taiao, including between themselves, other living things and the environment. Any change in a balance of these interconnections will resultantly impact the mauri of everything in the system, be it an ecosystem or a living organism.

This big idea is further elaborated as four learning objectives as follows:

- explore how mātauranga Māori and Pacific knowledges interact with microorganisms
- consider microorganisms and the taiao as an interconnected entity and explore the role of microorganisms in the taiao
- explore how microorganisms can be beneficial or harmful
- investigate how changes in conditions in the taiao affect microorganism growth and the mauri of the taiao, including the role of defences in protecting living organisms against harmful microorganisms.

Presumably, there are reasons for the emphasis on microorganisms, but these reasons are not mentioned in the notes that elaborate the big idea in the context of this standard. The notes section of the AS does outline a range of understandings that must be brought together to achieve the standard. These include: the role of microorganisms; knowledge of their life processes; their population dynamics; interactions between microorganisms and the environment; and two complex concepts from mātauranga Māori: taiao and mauri. Why students should understand this particular combination of concepts is implicit in the notes and accompanying tasks but, again, is not clearly focused into an assessment argument.

Three suggested assessment activities are offered. One is set in the context of fermentation of food. Students produce “a resource that shows the method for making the food”. A list of things that should be covered in the resource mirrors the range of understandings outlined above. Another activity is set in the context of kauri die-back. The assessment task specifies knowledge of various aspects of kauri die-back. It appears to be a research exercise. (The formats suggested in the student materials refer to presenting “your findings”.) The third activity is set in the context of human health and disease. This is also a research exercise, carried out for the semi-specific purpose of providing “information to your community” [about the chosen topic, presumably to aid awareness and/or prevention].

All three examples variously suggest a range of potential formats for presentation of evidence of learning. These include: an oral presentation; a poster; via video or animation; a written story or

report; an infographic; a pamphlet; or a slideshow. The purpose of informing the public about the dangers of kauri die-back is implied in the use of the term “public service pamphlet” but is otherwise missing from the suggestions for this exemplar. The fermentation exemplar indicates that the resource could be created by a group, but each student must identify their contribution.

All three tasks have been briefly outlined in this case to illustrate how the team developing the resources for this AS has attempted to enact two of the three UDL principles (see Section 4):

- Students should be able to access multiple ways of representing learning. This principle supports meaning-making for students who perceive and understand information and ideas differently.
- Students should be able to use a range of ways and to express their ideas. This principle focuses on students’ active role in learning to do things.

Arguably, what is not clearly focused is the high-level “why”—in ECD terms, the “assessment argument”. Again, the final cell of the “Excellence” column of the marking guide came closest to clearly setting this argument out for each individual resource:

Student’s report is a concise discussion that recognises that mauri of unsterilised food will be diminished and that conditions must be in a careful balance to produce the product we seek to eat. Links between environmental conditions such as temperature and nutrition of the yeast relate to the quality of the products that result from the fermentation process.

Student’s report demonstrates understanding of the mauri of the ngāherehere and the effect this has on kauri dieback. Knowledge of the interconnected nature of micro-organisms and their kauri host is used to link the presence or absence of *Phytophthora* to environmental conditions, recognising that *Phytophthora* life processes are linked to presence of nutrition, competition with other species of microbe, and soil composition, for example.

Student’s report is a concise account of the human response to a pathogen and links response to conditions in the body and to the life processes of the pathogen. The interconnected nature of the microorganism, the overall health of the body and the defences, internal and/or external that are employed to control pathogen infection relate to the mauri of the body and overall health.

All three statements have been provided here to demonstrate inevitable differences in emphasis and reasoning when the suggested contexts are different. The explanation of the big idea, provided in the introductory materials, is closer to being an overarching assessment argument, against which these variations might be more reliably interpreted in practice, and in very different presentation formats. However, as already noted, this overarching statement does not say why the focus is specifically on microorganisms and this seems to be a lost opportunity to “join the dots” related to an overarching purpose for studying this topic.

It took a lot of clicking between different tabs to bring these pieces together when constructing our critique. How many busy teachers would have the time to make a similar careful comparison, to get to the “essence” of the AS, as might be conveyed by a carefully constructed assessment argument. There is a clear intention to make the assessment tasks more inclusive, but we wonder how readily teachers will be able to make reliable judgements when so many different combinations of variables are in play.

Walking the talk of UDL?

Promisingly, the introductory content for the two internally assessed AS for agricultural and horticultural science includes more specific advice about inclusive assessment practice. This advice goes further than suggesting diverse methods for demonstrating learning; that specific advice is included and is similar to that outlined for the two AS above. Additional advice covers a range of topics that were discussed in Section 4. These topics include:

- designing assessments that are congruent with the ways students have been learning, and that provide valid and fair opportunities for a diverse range of students
- using different methods to collect evidence from those used in external assessments, particularly if the latter are time-bound examinations
- collecting naturally occurring evidence during learning rather than relying on a one-off event or task, along with practical suggestions for ensuring authenticity. Potential evidence sources are suggested: notes from teaching and learning; observations from field trips; observations from experiments; research; interviews; discussions with whānau or kaumatua; and tuakana-teina discussions.

Sadly, this in-principle advice is not readily apparent in the supporting materials. Our example here is drawn for AS 91928: *Demonstrate understanding of life processes and how they are managed in a primary production system*. Three assessment activities are provided. They are set in different primary production contexts: venison farming; bee-keeping; and growing apples. In each case, students show their learning by creating a presentation on the production of the chosen food, with a focus on life processes and management practices.

A presentation is, by definition, a one-off event. It is not readily apparent how the different evidence sources suggested could be factored into the assessment judgement, as the schedules currently stand.

Is “naturally occurring evidence” a feasible alternative?

The failure to translate good intentions into worked examples left us asking whether the current NCEA model, and practices that have evolved over time to support AS as they are currently structured, can actually accommodate the valid and reliable use of naturally occurring evidence.

Authenticity concerns were outlined in Section 4 and will not be revisited here but they could be one “mindset” barrier. We also wonder to what extent the mindset that assessment comes at the *culmination* of learning is getting in the way. We did not talk to any of the designers of these resources so we cannot answer this question, but it has obvious implications for PLD if it is indeed correct.

As we saw in Section 4, recent critical commentary about UDL emphasises the importance of being very clear about the focus of an assessment—in effect, using an evidence-centred assessment design process. As it stands, the design teams seem to be trying to signal all the subject “content” that should be learnt (and assessed) for each new, bigger, AS. In the absence of an overarching assessment argument, grounded in clear curriculum design, this seems to leave teams feeling obliged to try to design a one-off task that covers as much of this content as they can squeeze in. The challenge might be addressed via the Curriculum Refresh, but this currently seems unlikely given that the timing of the two initiatives is not aligned.

We now draw on another of the three suggested assessment activities for biology/chemistry AS 92021 to explore a potential way out of this conundrum. The left-hand column of Table 6 on the following page lists all the specific concepts students must show they understand as they complete the suggested activity. The right-hand column draws from the list provided for agriculture and horticulture to make indicative suggestions of naturally occurring evidence sources for each.

The table includes a range of suggestions for showing learning in ways that do not rely on a formally produced piece of writing. In principle, students with diverse learning needs could show their learning without accessing any help in addition to that already provided for them day to day. Anxiety issues for some students (see Section 3) would also be lessened if assessment was integrated into the flow of learning.

This approach illustrates one way of meeting the advice provided under the “Conditions” tab for the agriculture and horticulture suite of achievement standards: namely that assessment should be undertaken in conditions similar to those already experienced in learning. Critics could say that a collection of parts does not demonstrate a coherent understanding of the whole. This is why we have suggested some sort of concluding activity that allows students to demonstrate a dynamic understanding of why it matters that we all understand the significance of conservation of mass. We note that such a demonstration would constitute an “Excellence” award of the AS when using the marking schedule provided. This schedule already awards an “Achieve” pass for demonstrating some, but not all, of the concepts/skills listed in the left-hand column of Table 6 (i.e., it is arguably subject to the same critique as a collection of evidence from different sources).

One of the ECD papers already cited suggests that collection of evidence from a range of sources will allow assessors to make more robust and defensible judgements about learning that really matters. There is one important proviso—these decisions must be guided by a clearly developed assessment argument (Gorin, 2014).

TABLE 6 Indicative alignment of naturally occurring evidence with specified understandings

Conceptual understandings and skills to be demonstrated for AS 92021 (paraphrased from student information)	Potential evidence sources (naturally occurring during learning) drawn from suggestions for AS 91928
Make observations of complete and incomplete combustion (describe ways they look different).	Laboratory notes (verbal or written).
Identify examples of complete and incomplete combustion on the basis of these observed differences.	Additions to lab notes. Annotated examples (perhaps with photos?).
Show how mass is conserved in complete and incomplete combustion, including writing or at least interpreting balanced chemical equations.	Writing and/or explaining equations for complete and incomplete combustion of the same material. Annotations or a verbal explanation convey understanding of how balancing reflects conservation of mass. Chemistry modelling sets could also be used to show this understanding in a “hands on” matter.
Outline how complete and incomplete combustion contribute to climate change.	Research and/or interview notes. Graphic depictions of impacts (cycle diagrams, flow charts, concept maps are all possibilities). Interaction with computer animations or games is also a possibility.
Discuss how impacts are affected by relative concentrations of combustion products or by locations where they are released.	Research and/or interview notes. A short written, verbal, or visual explanation.
Show how carbon is conserved through all the dynamics discussed during the unit.*	This could take the form of an innovative concluding activity.

* This is rather cryptically described as demonstrating understanding of the “significance of conservation of mass”—we think this is what the notes are hinting at.

Missing pieces

It is important to note that the analysis in this chapter draws on published materials that are being tried out in the pilot stages of the RAS. They are drafts that will no doubt evolve with use and feedback.

There is also an important omission. Given the early stage of the pilot, we did not have access to any commentary about how the materials have been interpreted in practice by teachers, or by subject moderators who have a strong influence on how AS are applied in practice. Such records will hopefully come with time. Continuing PLD support for UDL will be critical to how professional understandings unfold.

Concluding comment

Throughout the field work for this project, we often heard in-principle support for the use of UDL to reduce demand for SAC. But we also heard that this was not really happening and/or seemed to be in the “too hard” basket. In Section 4, we looked at potential difficulties through the lens of published research. In this chapter, we have reframed the analysis to look at challenges through the lens of well-meaning early attempts by SEG teams to bring UDL into their work. Applying UDL is clearly not going to be a quick fix. In fact, one expert in UDL indicated that the approach would need to permeate every aspect of the education system before it could be successfully embedded in NCEA. The leaders of the school that came closest to modelling a through-going approach to inclusive assessment would no doubt agree.

Looking across Sections 4 and 5 as two sides of a whole, we recommend that the Ministry of Education/ NZQA:

- Take steps to better align the Curriculum Refresh and the RAS, so that clear, defensible assessment arguments can be built for each AS. Valid and reliable assessment decisions rest on this condition, which assumes additional importance when UDL considerations demand innovative assessment practices.
- More clearly describe multiple evidence sources that could contribute to assessment decisions, and link these to the assessment arguments that have been developed. This sequenced process would have the dual benefit of helping to identify construct-irrelevant variables while providing clear signals to assessors about what to look for in different assessment tasks and/or ways of demonstrating learning are deployed to better accommodate student diversity.
- Provide ongoing well-supported professional learning related to UDL for everyone involved in the NCEA system. Reliance on examples developed as part of the pilot will almost certainly fall short of stimulating a system-wide shift in practice that will be needed if the demand for SAC is to be meaningfully and equitably reduced by means of innovative assessment design and associated more inclusive practices.

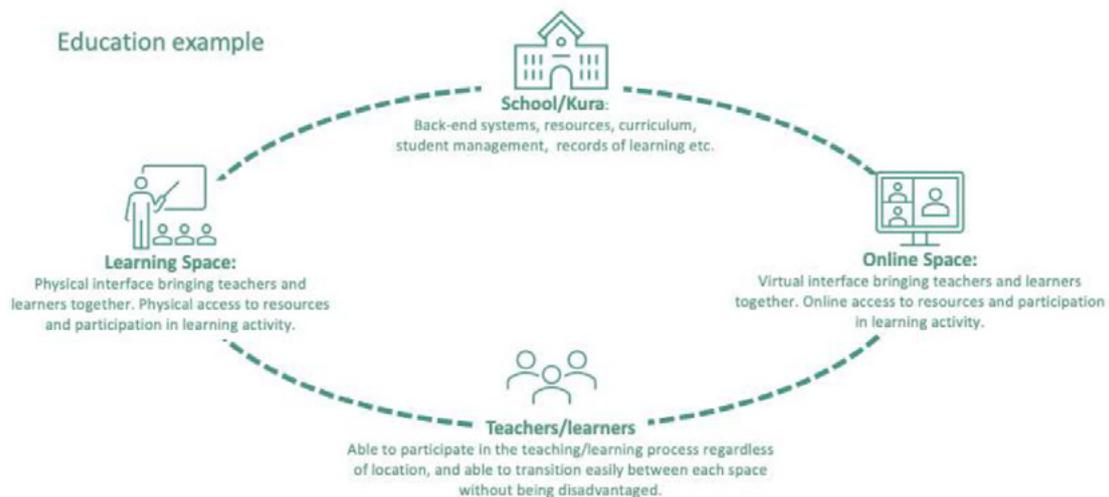
6. The change potential in hybrid learning models

This report was written during a stage of the COVID-19 pandemic when schools were struggling to maintain learning programmes with an ever-changing mix of students and teachers at home, either sick with the virus, or isolating because others in their home were infected. We were asked to search for any reports of successful practice in these circumstances. The most practical advice we found was sourced locally, in the form of a conceptual guide for establishing and embedding a hybrid learning model in New Zealand schools (Wenmouth, 2022).

The scope of hybrid learning

In what follows, we endeavour to highlight the potential for developing a robust hybrid learning model to support inclusive assessment practices for diverse learners. Simply put, a hybrid learning model allows both teachers and students to “continue their roles regardless of where they are physically located” (Wenmouth, 2022, p. 5). Figure 6 below is Wenmouth’s model.

FIGURE 6 A hybrid learning model (Wenmouth, 2022, p. 7)



The critical conceptual move here is the separation of the functions of the school as an organisation from the provision of physical spaces for learning. Teachers and students can access the curriculum and its resources from *either* the physical space or from the bespoke digital learning space designed and maintained by the school. Wenmouth stresses that this does not mean students at home simply follow along with what is happening in the classroom. Rather, a comprehensive audit and redesign of the school’s systems and processes supports learning experiences that can be taken up in either the physical or the digital space. This is transformational work, not a quick fix.

Wenmouth’s resource makes many detailed and practical recommendations for undertaking the transformational change needed to establish a robust hybrid learning process. Many of these suggestions have direct relevance to the challenge of reducing the need for SAC. Table 7 summarises key points from the report in the left-hand column. The right-hand column adds our thoughts about the implications for SAC and for NCEA in general, informed by the previous sections of our report. Notice the multiple points of resonance between this advice and the model of one school’s system depicted in Figure 5.

TABLE 7 **Hybrid learning features with the potential to support inclusive assessment practices**

Features of hybrid learning (paraphrased from Wenmouth, 2022)	Alignments and considerations in NCEA/SAC context (as outlined in Sections 4 and 5)
Student-centred learning design should make intentions as transparent as possible, both for students and whānau who support them.	<p>Transparency of intentions also supports the development of a sharply focused assessment argument to underpin the design of matrices and individual AS.</p> <p>These intentions need to be understood by students and whānau as well as by teachers.</p>
Continuity between learning in physical and digital spaces can be ensured by use of curriculum design templates shared by the whole teaching team.	Curriculum design should precede assessment (AS) design and should be evident within the content of the AS.
Application of UDL principles will help meet the learning needs of diverse students.	Professional learning to support UDL should be available and ongoing for all teachers.
<p>The intended learning should be active rather than passive, with an emphasis on sense-making and reflection rather than remembering and recall.</p> <p>Among other benefits, evidence generated during active learning is less likely to be directly copied from other sources when students are learning in a digital space.</p>	<p>Assessment criteria for each AS should focus on evidence that is actively generated during learning.</p> <p>Collecting evidence from active learning sources should be modelled in activities developed to support each AS.</p>
<p>Assessment should be ongoing during learning and made as transparent as possible.</p> <p>Robust support for self-assessment (e.g., by using carefully designed rubrics) reduces the time burden for teachers and provides a record of learning progress.</p>	<p>This feature implies the collection of naturally occurring evidence.</p> <p>It also has the potential to contribute to attestations of authenticity.</p>
<p>Use of e-portfolios allows for systematic collection of evidence from diverse sources.</p> <p>Shared practice uses the same tools and processes in both hybrid and classroom contexts, and at all levels of the school. This build familiarity, aids transparency, and reduces workloads over time.</p> <p>This evidence can be accessed by students, whānau, and teachers.</p>	<p>Marking schedules should convey specific information about the types of evidence required.</p> <p>Reference to a clear assessment argument allows pieces of evidence to be presented as a coherent whole.</p>
Systems used for data storage etc. should support a consistent approach across the school, as should shared professional learning.	

Agreement from international sources

We did not find peer-reviewed research that evaluates assessment accommodation practices in hybrid learning contexts. It is probably too soon for that. Wenmouth himself stresses that his advice is work in progress. The introduction to the resource includes this caveat:

This guide has been produced with haste given the current circumstances schools are facing. It is based on the personal views and experiences of the author and does not attempt to represent the complexity of our wider education system and the specific actions that may be required in every circumstance. It is offered here as a conceptual guide, to be engaged with, argued about, pulled apart and re-constructed to suit the needs of local contexts. (2022, p. 5)

We did find informal commentary in international contexts that provides assurance that others are thinking along the same lines. For example, a UK university has recently made the following observations about tertiary assessment and course design:

Developing and deploying flexible and innovative assessment supports the need to be responsive to the requirements of a hybrid model of learning and teaching and the possibility of a combination of on-campus and online delivery.

At the level of the curriculum, the most effective strategies for achieving such flexible assessment arrangements will utilise a variety of accessible and inclusive approaches and tools, employing a carefully designed and balanced range of authentic assessment tasks that enable all students to demonstrate what they know, understand and can do.

For such strategies to be successful there is need to create environments wherein assessment, learning and teaching are not artificially separated, but where assessment and feedback are fully and seamlessly integrated into an holistic, course-focused, view of the learning process. Crucially, such a view helps to frame curriculum and assessment design to fully consider the learning journey and experience of each student and to critically evaluate what needs to be assessed and how within a hybrid model.²³

The first quote argues that flexible assessment supports hybrid learning. This is the same argument as saying that hybrid learning requires more flexible assessment practice. The second quote underscores the importance of curriculum-informed evidence-centred approaches when designing learning and assessment in hybrid contexts. The third reinforces key messages about gathering evidence from naturally occurring learning. Drop-down additions provide practical advice and support, much along the lines proposed by Wenmouth.

Is hybrid learning/assessment fair?

Section 4 noted that perceptions of unfairness could constitute a “hearts and minds” barrier to adoption of more inclusive NCEA assessment practices. It seems that this same issue arises when considering hybrid learning and assessment in other high-stakes contexts. Professor Velda McCune from The University of Edinburgh tackles concerns about unfairness when some students learn face-to-face and others learn remotely, ostensibly having a different experience of the intended learning:

I think the first thing I want to say about the concern over students having the ‘same’ experience is that this has never been the case. A gay white man who grew up in poverty near Glasgow, a black male international student, and a middle class white woman student with dyslexia might sit in the same classroom but they will not be having the same experience. Depending on the teaching and the course

²³ The post is called “Flexible assessment for a hybrid model” and is linked to Teeside University: <https://blogs.tees.ac.uk/teonline/learning-and-teaching/flexible-hybrid-assessment/>

context, each of these students will feel more or less valued and included and will have more or fewer opportunities to thrive educationally ...²⁴

This comment has interesting resonances with Rose’s reflections about what he sees as cracks in the foundations of UDL (see Section 4). Teachers’ perceptions about students’ abilities and potential can get in the way of enactment of inclusive intentions. Similarly, their perceptions about what is fair will impact their willingness to enact new practices, such as hybrid learning. Again, Wenmouth addresses this challenge. His recommendations for effective professional learning include adopting a concerns-based approach that begins by surfacing attitudes and beliefs that will support or impede the intended shifts in practice (Wenmouth, 2022, p. 31).

In summary

The necessity to create more robust and sustainable models of hybrid learning for uncertain times offers opportunities to enact more inclusive assessment practices. Although only a limited number of sources are cited, none of them creates dissonance with the findings set out in previous chapters, or indeed with NZQA’s own advice (NZQA, 2022). Quite the opposite, in fact. This seems to be another context where there is a chance to get “two bangs for the buck”, provided that the challenges for enacting more inclusive assessment are systematically identified and addressed at every level of the system.

²⁴ The post is hosted on The University of Edinburgh’s *Teachers Matter* blog. It is called Equality, equity and social justice in hybrid education. <https://www.teaching-matters-blog.ed.ac.uk/equality-equity-and-social-justice-in-hybrid-education/>

7. Short answers to the research questions

This section returns to the research questions that were shaped at the very beginning of the project. We now present a summary overview of the findings, along with indications about where each question has been addressed in more depth within the report.

A snapshot of the current situation

- How are SAC currently being used in high-stakes assessment in Aotearoa New Zealand?
- What practices are perceived to be failing and why? Where there are examples of good practice, what does this look like?

Special Assessment Conditions (SAC) are used to support students who have been identified as having difficulties likely to impact their achievement opportunities. These difficulties fall into four categories: sensory; medical; physical; and learning difficulties, with learning being significantly the largest.

Provision of SAC operates in a variety of ways and depends very much on how well an individual school can resource it. This situation creates inequities for students who rely on SAC to manage their assessments. The number of students needing SAC is increasing and the model is becoming unsustainable. A change in approach is needed to create a longer-term solution.

Meantime, some shorter-term solutions could help make SAC more equitable and sustainable. These include more equitable access to consistent testing resources to determine entitlement, the use of text-to-speech in digital exams as well as other assistive technologies, and the funding of a training and support package for schools to aid with the management and closer connection between students' learning needs and potential SAC support.

What we can learn from good practice elsewhere

- How are special assessment conditions addressed in other international education environments, including those serving indigenous students?
- Are there examples of inclusive practice that we could adopt in Aotearoa New Zealand?
- Are there examples of good practice that ensure equity for ākonga in hybrid learning contexts?

Examples of practice elsewhere that could enhance the current model used in Aotearoa New Zealand were difficult to find. The report discusses examples of the use of assistive technology, in particular the use of text-to-speech in digital examinations, and the management of support processes.

Examples of inclusive practice we could adopt to ensure greater equity in hybrid learning contexts have been addressed as relevant throughout the report. Section 2 reports recently released data about students' participation in digital assessments. Section 4 addresses the challenge of attesting authenticity when assessments are completed in contexts other than the classroom. Section 5 discusses the feasibility of gathering naturally occurring evidence of learning, whether in the

classroom or elsewhere. Section 6 draws on a future-focused model of hybrid learning to outline the considerable change potential in a hybrid-learning approach, with the proviso that the relevant changes are co-ordinated across all aspects of the assessment system.

With this system-wide challenge in mind, we recommend the establishment of an independent oversight panel for assuring inclusivity for ākonga in NCEA assessment.

Establishing continuity between learning support and assessment support

- In the context of NCEA, how are RTLBs currently supporting inclusive practice in New Zealand secondary education?
- Is there a disconnect between learning support systems and assessment support in New Zealand secondary education? If there is, then how can the two be more closely aligned?

We found that a disconnect exists between learning support systems and assessment support. This is apparent externally, where learning support and RTLB services tend to be focused towards primary and early secondary school learning needs and are less involved in senior assessment unless there is a high-level medical or physical need. Internally, schools in several cases are working to ensure this is addressed for their ākonga, with strong transitioning systems that help identify both learning support needs and potential SAC entitlement at an early stage in their secondary schooling. SAC resourcing appears to be managed through learning support budgets unless schools have additional means.

In the context of NCEA, RTLB involvement in supporting inclusive practice in secondary education is variable. Such support was minimal or non-existent in most of the schools we visited. Where it does exist, RTLB support can include testing support for SAC applications. Some RTLBs provide training, advice, and guidance for SENCOs who may be new in their roles.

The impact of the review of achievement standards

- In what ways have the revised assessment models introduced at Level 1 for piloting in schools addressed inclusiveness and changed the likely requirements for SAC?

We reviewed the new internally assessed Level 1 achievement standards and assessment resources from the perspective of inclusion. We looked across the range of subjects but ultimately selected three standards from the science learning area to review in-depth. Our analysis revealed some evidence that the design teams had been introduced to the idea of Universal Design and had made some attempt to apply its principles. Despite their good intentions, the design teams did not make changes that would actually make a real difference to inclusiveness. Most of their efforts seemed token, and problematic to apply fairly (there were validity and reliability challenges). Drawing on our own science education backgrounds, we concluded that the AS themselves needed to convey a much clearer assessment argument, grounded in the curriculum (see design section below). More thoroughgoing change seems unlikely until this type of clarity is established and embedded.

Potential for improving access through assessment design

- Can a standards-based assessment model, such as NCEA, be made more inclusive? If it can, what kinds of changes could enable this? (For example, changing ways that achievement can be evidenced.)
- What are perceived to be barriers to making these changes?

Section 4 outlines the principles of Universal Design for Learning (UDL) as applied to assessment, and draws on published research to demonstrate that this approach has considerable potential, but also some substantive challenges when applied to assessment. As noted above, valid and reliable application of UDL principles to assessment processes requires that a clear curriculum-based assessment argument is in place before actual assessments are designed. A number of researchers advocate for the use of Evidence-Centred Design models (ECD), in combination with UDL principles, to achieve the necessary clarity, and to identify and remove construct-irrelevant variables from the assessment itself, and from the marking schedules.

Our analysis in Section 5 suggests that deep and sustained PLD will be important if UDL is to be used to create a more inclusive assessment system that reduces the need for SAC. How teachers and others involved in the overall assessment process understand the purpose of learning, and the types of evidence they see as valid and robust indicators of learning, will influence both design and application of new AS. The standards themselves will be insufficient to prompt system-wide change (and are not yet fit for this purpose). As well as provision of PLD, the overall system will need to work towards more evident alignment of the national curriculum for Levels 6–8 and the suites of AS. This is problematic because the Curriculum Refresh is running behind the RAS process.

Overall

Our investigations indicated that the current “bolt-on” model used to address SAC has drawbacks that even moderate tweaking could not fix. These have been documented in our findings in Sections 2 and 3. From this position, we considered it vitally important to explore thoroughly how our NCEA assessment model could be made more inclusive and what would be needed to enable this to happen. We have explored the perceived barriers to achieving this from a UDL perspective and discussed what we understand would be needed to overcome these barriers and enact meaningful change.

8. Recommendations

We undertook this research in the context of concern about unsustainable growth in the SAC process for NCEA. A complex mix of factors has potentially contributed to that growth. These factors could include:

- increasing numbers of students presenting at secondary school with anxiety issues
- increased awareness of the impacts of brain injuries (especially concussion)
- improved systems in schools for identification of students who require SAC support
- provision of classes for students with dyslexia in some schools (and increased awareness and diagnosis of this condition)
- NZQA provision of enhanced support for making SAC applications
- attention paid by ERO reviewers to inclusion.

Looking at this list, it seems that SAC has become the victim of its own success. Applications for support have increased to the point where they are not sustainable. Nor are the current processes necessarily the most desirable, given that they single out students for special treatment rather than being more universally inclusive.

With this high-level challenge in mind, we drew on our findings from the previous sections to shape three sets of recommendations: short-term; medium-term; and longer-term. This concluding section also briefly discusses the vexed question of terminology that could be used to replace the SAC acronym with something less deficit-sounding.

Short-term opportunities for quick gains

Elsewhere in the report we have used the term “low-hanging fruit” to describe changes that could be made quickly, to make existing SAC processes more manageable and equitable. We have identified opportunities to:

- facilitate the addition of text-to-speech assistive technology in all digital assessments
- develop PLD support packages that would allow RTLBs to assist schools to effectively manage their students’ learning and assessment needs in a manner that is consistent and equitable across schools
- provide access and funding that allows all schools to carry out diagnostic testing to support SAC applications on behalf of their students
- review materials produced to support the Level 1 standards in the RAS pilots and either remove or update those with counter-inclusive signals
- consider ways to exemplify the potential for students themselves to take greater agency over their specific learning needs.

Medium-term opportunities for more coherent system alignment

As we worked to see SAC and associated practices from the perspectives of those we talked with, we were cognisant of the “hearts and minds” challenges that arise when considering system-wide change in uncertain times. The short-term opportunities suggested above have a focus on modest resourcing changes and the provision of convincing exemplars as part of the RAS. Such changes are necessary but not sufficient. Ultimately, teachers must be convinced that more inclusive assessment practices will work and will not disadvantage those who do not need this type of support. They will resist anything they see as not fair, not robust, or that sends mixed messages about what matters. We saw hints of this resistance in some of the RAS pilot materials. This cluster of recommendations is made with the challenge of shifting practice in mind:

- Invest in a programme of ongoing professional learning about UDL, and its application to inclusive assessment design, for all secondary teachers.
- Foster a more visible alignment between curriculum and assessment design, to ensure inclusive practices are connected and embedded into assessment at the earliest opportunity.
- Continue to exemplify UDL principles in materials that support NCEA assessments and establish a quality assurance process to ensure these materials are fit for purpose and signal consistency of approach across subjects.
- Establish an oversight panel with responsibility for reviewing and ensuring inclusiveness in NCEA assessment design. The brief could include both quality assurance and strategic oversight of assessment design from an inclusivity perspective.

Long-term consolidation of sustainable practices

The suggested medium-term changes would take some time to work through. The ultimate goal is to make inclusive practice so thoroughly embedded in the overall education system that the number of students needing specific SAC support is reduced to a minimum, compared to the current situation. Once all the systems involved are fully aligned, it should be possible to develop an NCEA assessment model that is equitable for ākonga, manageable for schools, and largely barrier-free.

Thoughts on replacement of the acronym SAC

Although it was not one of the research questions, we were asked to give some thought to a term to replace SAC that is less discriminatory and does not send a signal that SAC is, inevitably, a bolt-on to “normal” assessment practice. The term “accommodation” still signals that SAC is different and special.

To really walk the talk of UDL principles, the terminology used needs to be more sensitive to diversity and more inclusive in the signals it sends. We gave consideration to the term “managing diversity”, as per the title of this report. However, this could still suggest something that is “done to” students—and that they are somehow different in ways that need managing. There is no signal that greater student agency in the process could be fostered.

At this point, our tentative suggestion is “assuring inclusive assessment” (AIA). Assurance suggests a more conscious and visible effort to be inclusive, and the term “inclusive” signals that students with specific needs are part of the rich spectrum of human diversity, rather than being deficient in some way.

Determining the suitability of any replacement terminology could be a task for an oversight group, should the medium-term recommendation to create such a group be adopted.

Concluding comment: Looking towards a system-wide view of change

Our investigation of the current enactment and projected future demand for SAC has documented a system that is clearly becoming increasingly unsustainable. We heard that NZQA has been working hard to make the system more streamlined and practical to apply. But it also became very evident that while SAC remains a bolt-on to traditional assessment practices, it is unlikely that more transformative change can be anticipated.

There are important “hearts and minds” issues to address, for parents and others in the wider community, as well as for teachers. It is not so many years ago that high-stakes senior secondary school assessments were expected to be a gatekeeper, used to select and ration access to higher education. The very possibility of more inclusive assessment, supporting continuing learning for everyone who is able, is at odds with long-established sorting practices. It is important to note that we did not address specific questions in this area. However, our analysis of teachers’ beginning efforts to take account of Universal Design for Learning principles was telling. They did try to incorporate UDL in the design of assessment activities to support the new level 1 achievement standards, but these efforts were neither consistent nor effective.

There is undoubtedly rich potential in the more systematic application of a UDL approach to NCEA assessments. However, it is equally clear that this type of change cannot be transformative if it is applied to the current system “as is”. What has emerged from our study is the clear need to take the whole assessment/curriculum nexus back to first principles. More inclusive assessment cannot be achieved by tinkering. It must begin with inclusive curriculum design, which then feeds into more inclusive assessment practices—both of which are supported by an ongoing, comprehensive programme of professional learning support for teachers.

The recent imperative to introduce hybrid learning pedagogies is another area of rich potential for making NCEA assessment more inclusive. This change has been imposed on the system by the impacts of the ongoing COVID-19 pandemic. NZQA has provided timely leadership, with advice about how traditional assessment practices might be adapted to cope (NZQA, 2022). But again, the more transformative potential lies in going back to first principles, to redesign a hybrid system that takes account of the need for change in all aspects of the work of secondary schools (Wenmouth, 2022).

References

- Burgon, J. (2018). *Special Assessment Conditions—Barriers to use*. New Zealand Council for Educational Research.
- Cohen, D. J., Zhang, J., & Wothke, W. (2019). Effects of item modifications on test accessibility for persistently low-performing students with disabilities. *Applied Measurement in Education*, 32(4), 269–280. <https://doi.org/10.1080/08957347.2019.1660343>
- Darr, C., & Hipkins, R. (2022). *Improving the design of achievement standards: A discussion paper*. New Zealand Council for Educational Research.
- Dolan, R. P., Hall, T. E., Banerjee, M., Chun, E., & Strangman, N. (2005). Applying principles of universal design to test delivery: The effect of computer-based read-aloud on test performance of high school students with learning disabilities. *Journal of Technology, Learning, and Assessment*, 3(7). <https://ejournals.bc.edu/index.php/jtla/article/view/1660>
- Glass, D., Meyer, A., & Rose, D. (2013). Universal design for learning and the arts. *Harvard Educational Review*, 83(1), 98–273
- Goegan, L. D., Radil, A. I., & Daniels, L. M. (2018). Accessibility in questionnaire research: Integrating universal design to increase the participation of individuals with learning disabilities. *Learning Disabilities: A Contemporary Journal*, 16(2), 177–190.
- Gorin, J. (2014). Assessment as evidential reasoning. *Teachers College Record: The Voice of Scholarship in Education*, 116(11), 1–26. <https://doi.org/10.1177/016146811411601101>
- Karvonen, M., Kingston, N. M., Wehmeyer, M. L., & Thompson, W. J. (2020). New approaches to designing and administering inclusive assessments. In M. Karvonen, N. M. Kingston, M. L. Wehmeyer, & W. J. Thompson (Eds.), *Oxford research encyclopedia of education*. Oxford University Press. <https://doi.org/10.1093/acrefore/9780190264093.013.1182>
- Ketterlin-Geller, L. R. (2008). Testing students with special needs: A model for understanding the interaction between assessment and student characteristics in a universally designed environment. *Educational Measurement: Issues & Practice*, 27(3), 3–16. Education Research Complete. <https://search.ebscohost.com/login.aspx?direct=true&db=ehh&AN=34081212&site=ehost-live>
- New Zealand Qualifications Authority. (2022). *Assessment in uncertain times*. Author.
- Perry, K., Meissel, K., & Hill, M. F. (2022). Rebooting assessment. Exploring the challenges and benefits of shifting from pen-and-paper to computer in summative assessment. *Educational Research Review*, 36, 100451. <https://doi.org/10.1016/j.edurev.2022.100451>
- Rose, D. (2021). *Cracks in the foundation: Personal reflections on the past and future of the UDL guidelines* [A collation of six blog posts]. CAST. <https://www.cast.org/binaries/content/assets/common/news/cracks-foundation-whitepaper-20211029-a11y.pdf>
- Scalise, K., Irvin, P. S., Alresheed, F., Zvoch, K., Yim-Dockery, H., Park, S., Landis, B., Meng, P., Kleinfelder, B., Halladay, L., & Partsafas, A. (2018). Accommodations in digital interactive STEM assessment tasks. *Journal of Special Education Technology*, 33(4), 219–236. <https://search.ebscohost.com/login.aspx?direct=true&db=ehh&AN=132790276&site=ehost-live>

- Shaw, S., Crisp, V., & Hughes, S. (2020). A framework for describing comparability between alternative assessments. *Research Matters*, 29, 17–22.
- Smith, R. E., & Buchanan, T. (2012). Community collaboration, use of universal design in the classroom. *Journal of Postsecondary Education and Disability*, 25(3), 259–265.
- Thompson, S., Thurlow, M., & Malouf, D. B. (2004). Creating better tests for everyone through universally designed assessments. *Journal of Applied Testing Technology*, 6(1), 1–15.
- Wenmouth, D. (2022). *Getting started with hybrid learning: A teacher guide*. FutureMakers Ltd. <https://futuremakers.nz/getting-started-with-hybrid/>

