Making NCEA assessment more equitable

Can UDL principles help?

ROSEMARY HIPKINS AND JANET LEE

Assessment News outlines findings from an investigation of inequalities in the current use of special assessment conditions (SAC) in NCEA assessments, along with rising demand for their use. The research also explored the potential of Universal Design for Learning (UDL) to help improve assessment accessibility, and to reduce the overall demand for SAC.

Embedding UDL principles in assessment design has the potential to make assessments more accessible, and, in doing so, reduce the demand for SAC and improve students' agency over their assessments. But UDL principles should be applied from the start of the assessment design process, and they challenge traditional thinking about curriculum and assessment.

Introduction and context

When the Ministry of Education released the National Certificates of Educational Achievement (NCEA) Change Programme in 2020, making NCEA more accessible was listed as the first of seven changes aimed at strengthening NCEA:

> 1. Make NCEA more accessible—zero fees, fewer barriers for learners with disabilities and learning support needs.[1]

One aspect of the change programme involved a review of the provision of special assessment conditions (SAC) for students who need additional support to show what they know and can do. SAC support has traditionally entailed: the use of a reader–writer; being given modified exam papers; provision of separate accommodation during an assessment; use of assistive technologies for computer-based assessments; or some combination of these.

One known problem was that SAC applications were being made more often by schools in higher socioeconomic areas. For example, in 2019, 5% or fewer students who were enrolled for NCEA from schools with decile ratings 1–3 received SAC entitlements, whereas more than 11% of students enrolled in schools with decile 10 ratings received them. There was a similar pattern in 2020, as the following table shows.

Another problem was that demand for SAC was growing year on year, to the point where it was becoming unsustainable, both for schools to administer, and for the New Zealand Qualifications Authority's

https://doi.org/10.18296/set.1534

TABLE 1. APPROVED REQUESTS FOR SAC IN 2020: PERCENTAGE OF TOTAL COHORT BY SCHOOL DECILE (HIPKINS & LEE, 2022)

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

(NZQA) approval systems to manage. Work to streamline the application process might make access more equitable but would also be likely to increase demand to even more unsustainable levels. Some different thinking was needed. The Ministry of Education commissioned the New Zealand Council for Educational Research (NZCER) to investigate the current application of SAC in schools and how it could be improved. Alongside that, the Ministry asked us to investigate possibilities for making greater use of Universal Design for Learning (UDL) principles in the design of assessments, with the aim of reducing demand for SAC. Our full report can be found on NZCER's website.[2] In this column, we focus on what we found out and suggested regarding the UDL challenge for assessment design.

The potential seen in UDL

A further goal of the Ministry of Education's NCEA Change Programme entailed a review and substantive reworking of all the exiting suites of achievement standards used for NCEA assessments:

> 4. Fewer, larger standards—new achievement standards and resources will be developed to replace existing standards and ensure the qualification achieved credentials the most significant learning in a learning area or subject. (See endnote 2.)

Some members of the Ministry's team who are managing the NCEA changes saw the potential to apply principles from UDL to the new achievement standards specified in the review. These "fewer, larger" achievement standards are being redeveloped by subject expert groups (SEGs) made up of teachers and other subject education experts in each curriculum learning area. As part of our overall investigation of SAC we investigated both the opportunities and challenges of making this sort of change to the achievement standards. We began our investigation with an initial scan of UDL research literature and followed this up with a brief analysis of the extent to which we could see the influence of UDL principles in the design of the first Level 1 NCEA internally assessed pilot standards and their accompanying materials.

What has UDL got to do with assessment?

UDL arose from neuroscience investigations into the brain activity of students with specific types of learning disabilities. These investigations resulted in the development of three principles for designing learning activities to make them more accessible to all students. This potentially involves three types of modifications to planned learning:

- making sure there is a range of ways to engage with the learning in the first place
- using a range of meaning-making strategies to enable as many students as possible to take an active role in their own learning
- giving students options for demonstrating what they are learning or have learnt.[3]

While the third of these is most obviously related to assessment, we found that all three ways of adapting learning can potentially be applied to the design and conduct of assessments for NCEA. Self-evidently, the internally assessed achievement standards are a promising starting place for individual teachers to apply UDL principles, but the design of the standards and their accompanying materials can smooth the path—or make it much harder to do this.

How design of NCEA structures and processes can support equity in assessment

When UDL principles are applied to the design of assessment standards, one key aim is to remove potential barriers to participation. Every student should have a fair chance to show what they know and can do. This might seem obvious but "construct irrelevant variables" can get in the way. For example, if a student has difficulties in decoding formal written text, they might not be able to answer written questions, even though they could give an answer if they were invited to do so in a different way, or with appropriate supporting technologies. Ability to read fluently is a construct irrelevant variable if what is being assessed is a practical capability.

To be effective, the task design and specifications need to be underpinned by a design process that clarifies what is being assessed and matches the demands of the assessment to that focus. Construct irrelevant variables should not be allowed to get in the way of demonstrations of learning. In the literature, we read about evidence centred design (ECD). This is an assessment design approach used by large test design agencies in countries that have national testing programmes (e.g., the USA). It begins with a curriculum analysis that clearly specifies the focal knowledge, skills, and abilities to be assessed. Once these have been identified, an "assessment argument" is shaped. This assessment argument can be used to eliminate construct irrelevant variables at the initial design stage,

and then continues to be applied right through to making valid and reliable judgements about student work.

On a challenging note, we read that the curriculum itself needs to be universally designed to build inclusive foundations for any assessment design work to follow. The literature is clear that UDL is most effective in eliminating inequalities when it is applied across the whole learning system. At the time we did this SAC investigation, SEGs were developing new achievement standards before the relevant learning areas of the national curriculum had been refreshed. This meant that the SEGs did not necessarily have access to new or clarified curriculum thinking.

One of the people we interviewed as part of the overall SAC investigation named a specific school that took a school-wide approach to enacting inclusive assessment. We looked into this and included a brief case study of this school's practices in our final report. Though not specifically a UDL model, what they did was congruent with UDL principles and embodied the coherence that comes from systematic and aligned attention to every aspect of their learning programme. In this school, teachers learn together as they design richly contextualised units of learning that integrate at least two curriculum learning areas. As they work in different collaborative pairs for the diverse units they design, teachers share and build rich pedagogical content knowledge. One focus of this growing body of shared professional knowledge is the design and execution of innovative assessments that are inclusive for a wide range of students. The clear message emerging from this case study is that UDL and ECD cannot be "bolted on" to existing practice. Instead, they need to become an integral part of how each school's learning environment functions. For UDL to be successful and beneficial, each school needs support and guidance through professional development and learning to enable the model to function uniquely for the school and its community.

How NCEA processes could support greater study agency

In the preliminary part of our research, we found great variability between schools in the systems and processes they used to administer SAC. One difference related to the extent to which students were involved in making key decisions about their need for SAC, and how and when the support might best be provided. We continued to think about opportunities to help students build greater agency in their learning and assessment as we explored the application of UDL to the new assessments.

We have already noted that building a clear and focused assessment argument can enable the elimination of construct irrelevant variables from the assessment process. A clear assessment argument also has the advantage of being readily communicated to students. When they are clear about why the learning matters, and what assessors will be looking for to judge the standard they have reached, students are empowered to take greater agency for their own learning and achievements. We acknowledge that some teachers could worry that provision of this level of clarity (maybe in the form of well-designed rubrics) constitutes a type of cheating and enables too many students to do well. We return to this "hearts and minds" challenge at the end of this column.

How are SEGs going so far?

We reviewed a range of internally assessed achievement standards, prepared as part of the Level 1 Review of Achievement Standards (RAS). We were informed that the SEGs had undertaken some professional learning about the use of UDL principles and so we looked for evidence of the application of UDL strategies to the assessment process. We could see that SEG teams had tried to apply some UDL principles, but also that this did not appear to be an easy or coherent change for them to make to current practice. None of the three science standards that we chose to review in depth successfully modelled an inclusive but also valid and reliable assessment design.

Both during this work, and subsequently, we have noticed a strong trend to design achievement standards where some form of summative report is judged as evidence of learning. It might take a variety of forms (a poster, a presentation, a pamphlet, etc.) but it is basically a summing up statement about what has been learnt and is typically produced by an individual student, even when the learning has taken place more collaboratively. This poses several challenges. The likelihood of construct irrelevant variables creeping back in, is one. The "authenticity" (for each individual) of work generated interactively is another. We also noted the lack of take-up of opportunities to document "naturally occurring evidence" captured over time as the learning sequence unfolded. Use of such evidence was recommended when NCEA was first designed so it would be interesting to investigate what is getting in the way of its use, especially now with the emphasis on UDL.

One challenge that could potentially loom very large in the near future wasn't even on our radar when we did this work. New natural language artificial intelligence (AI) tools such as Chat GPT can potentially generate well-written reports in seconds. We include an example at the end of this column. There are many ways in which these reports can be problematic, not the least being their accuracy or veracity. But the temptation is clear. Some students will use this shortcut unless the finished product is personalised in some way that precludes use of AI at the reporting stage.

Addressing "hearts and minds" challenges

Before more sustained and widespread use can be made of UDL, 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 UDL 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 overall.

There is undoubtedly rich potential for 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.

We conclude this column with a small demonstration of Chat GPT in action. The box below shows an AI-generated report on barriers to implementation of UDL. On the internet there is a lot of open-access commentary on UDL to draw on. We think this is a pretty good summary of the issues we also encountered.

Notes

1. https://ncea.education.govt.nz/what-ncea-change-programme#:~:text=The%20NCEA%20Change%20Programme%20is%20a%20work%20programme,or%20work%207%20Keep%20NCEA%20Level%201%20optional
2. https://ncea.education.govt.nz/change-4-fewer-larger-standards
3. Helpful materials that explain these principles can be found on the Inclusive Education section of Te Kete Ipurangi: https://inclusive.tki.org.nz/guides/universal-design-for-learning/

Reference

Hipkins, R., & Lee, J. (2022). *Accommodating diversity in assessment: A snapshot of practice in 2022*. New Zealand Council for Educational Research. https://www.nzcer.org.nz/research/publications/accommodating-diversity-assessment-snapshot-practice-2022

Chat GPT question: What are the barriers to the application of UDL principles in assessment contexts?

While Universal Design for Learning (UDL) principles promote inclusive and equitable education, there are several barriers to their application in assessment contexts. Some of these barriers include:

1. Traditional assessment practices: many assessment methods are deeply rooted in traditional models that rely heavily on standardised testing and one-size-fits-all approaches. These practices often prioritise memorisation and regurgitation of information, which can be challenging for students with diverse learning needs and preferences.
2. Lack of awareness and training: educators may not be familiar with UDL principles or may not have received sufficient training on how to apply them in assessment design and implementation. Without this awareness and knowledge, it can be challenging to incorporate UDL effectively into assessment practices.
3. Time and resource constraints: designing assessments that align with UDL principles may require additional time, effort, and resources. Some educators may face constraints in terms of limited time for assessment design and administration or lack of access to appropriate tools and technologies.
4. Resistance to change: resistance to change within educational systems and institutions can hinder the application of UDL principles in assessments. Some stakeholders may be hesitant to adopt new approaches or may be skeptical about the effectiveness and validity of UDL-based assessments.
5. Assessment validity and reliability concerns: there may be concerns about the validity and reliability of assessments that incorporate UDL principles. Ensuring that UDL-based assessments accurately measure student learning outcomes while maintaining fairness and rigor can be a challenge.
6. High-stakes assessment pressures: high-stakes assessments, such as standardised tests with significant consequences for students, schools, or districts, can create additional barriers to the application of UDL principles. The focus on standardised outcomes and limited flexibility in assessment formats can limit the adoption of UDL approaches.

Addressing these barriers requires a combination of professional development opportunities, policy changes, resource allocation, and a shift in assessment paradigms to create more inclusive and equitable assessment practices that align with UDL principles.

—This "report" was generated in approximately 10 seconds.

Rosemary Hipkins is a kaihautū rangahau/chief researcher for Rangahau Mātauranga o Aotearoa NZCER.

Email: rose.hipkins@nzcer.org.nz

Janet Lee is a kairangahau matua/senior researcher for Rangahau Mātauranga o Aotearoa NZCER.

Email: janet.lee@nzcer.org.nz