# Exploring tangled interrelationships between assessment and curriculum

### **Bronwen Cowie and Rosemary Hipkins**

https://doi.org/10.18296/am.0054

### A brief overview of the collection

The articles in this edition of *Assessment Matters* draw on both international and national research to explore the dynamics of the interplay between curriculum and assessment. The collection we present here begins with an assessment focus, drawing on science education as the context that introduces a curriculum element. The dilemmas explored in the various articles are not limited to science education of course. We hope the collection invites readers from a range of curriculum subjects to ponder implications for their own disciplines.

The articles represent the collective effort of the Science Education Special Interest Group (SIG) of the NZARE. In all, eight members of the SIG (along with some associates) explore aspects of the complex, dynamic space between assessment and the science curriculum. These articles come from varying perspectives and are set in different educational contexts.

## Why choose a subject-specific focus?

The collection had its genesis in a number of interesting developments and dilemmas that arise in the highly liminal space between curriculum and assessment. We briefly expand on some of these, as a prelude to the articles that follow.

On one level, it is obvious that curriculum and assessment must be closely interrelated in practice. In the context of high-stakes assessments, fairness demands that the focus of an assessment should align with student learning experiences; the focus of assessment should be communicated in advance. There is considerable research-informed evidence that high-stakes assessments can unduly influence, and indeed hamper, intended curriculum or pedagogical reforms (e.g., Black, 2013; Black & Wiliam, 1998; Hayward,

2015). The article by Thomas Everth in this collection explores aspects of this dilemma as it relates to changing curriculum design and enactment in response to the urgent challenge of climate change.

Arguably, however, there is an even closer relationship between assessment and curriculum when Assessment for Learning (AfL) is being enacted, because it "encompasses those everyday classroom practices through which teachers, peers and learners seek/notice, recognise and respond to student learning, throughout the learning, in ways that aim to enhance student learning and student learning capacity and autonomy" (Cowie et al., 2013, p. 10). The article by Taylor and Whyte in this collection provides a snapshot of the AfL dynamics just outlined.

Notwithstanding these close connections between curriculum and assessment, many important developments in assessment thinking have proceeded in a largely generic way. In much of the relevant research, very little attention has been given to the affordances of different curriculum contexts (Coffey et al., 2011). This dilemma impacts both AfL and high-stakes summative assessment, albeit with somewhat different dynamics. Given the increasing curriculum interest in how knowledge is generated, legitimated, and communicated in different subject areas, we see the lack of assessment attention given to the epistemological differences between disciplines as a problem.

Cowie and colleagues noted this problem almost a decade ago, saying that "assessment for learning also needs to reflect, be responsive to, and build on from how particular disciplines generate and legitimize meaning" (Cowie et al., 2013, p. 10; see also Heritage & Wylie, 2020). More recently, Quinlan and Pitt (2021) have argued that "fulfilling the educational potential of AfL suggests that at least some assessment tasks and processes should reflect the deep and implicit structures of a discipline and its knowledge generation practices" (p. 192). For example, they note that AfL aims to build students' capabilities in self-assessment by building their understanding of what might count as quality in whatever aspect of the discipline is being assessed. In a similar vein, peer feedback (another common AfL practice) implies the need to build a community of practice that models making judgements that are underpinned by discipline-specific norms and practices. All these observations imply a need for teachers to have sophisticated and nuanced

pedagogical content knowledge (PCK), including epistemic aspects of the relevant subject. These aspects are strongly signalled in the Nature of Science strand of the science learning area of *The New Zealand Curriculum* (*NZC*) (Ministry of Education, 2007) that was current at the time of writing this article. The article by Jared Carpendale and Mairi Borthwick in this collection picks up on aspects of this challenge.

In the context of high-stakes summative assessments, sciences and history stand out as two curriculum areas that have included some discipline-specific assessment discussions, with their respective efforts to assess students' epistemic understandings via the inquiry practices of science and historical thinking. This trend has not been without challenges. For example, Yates et al. (2017) point out that the "impact of testing and neoliberal thinking" (p. 138) can hamper teachers' best efforts to interest students in demanding epistemic learning if the students' primary interest lies in gaining the best test scores they can muster. The context for this comment is a deep exploration of similarities and differences in the epistemic treatment of physics and history as both school and university subjects. Johnson et al. (2018) reported that the epistemic practices of history and biology were not well represented in externally assessed achievement standards for the National Certificates of Educational Achievement (NCEA) for these subjects at the upper secondary level. Internally assessed NCEA standards for history better reflected the epistemic practices for the discipline, while those for biology did not. A statistical analysis of learning success in subsequent years suggested that this emphasis had conferred an ongoing learning advantage for history students, but no such advantage was detected for biology students. In this current collection, articles by Rosemary Hipkins and Charles Darr, and by Thomas Everth, explore ideas for revising/modifying NCEA structures and processes to better reflect the epistemic practices of the various science disciplines.

### Standing on the shoulders of giants

In our assessment thinking, we owe a legacy to Terry Crooks (1988) in terms of thinking about the impact of classroom assessment on students' learning—their motivation, what they come to view as important, and how they see themselves as learners. Crooks' thinking influenced the 1993 *New Zealand Curriculum Framework* and informed the original shaping of New

#### Cowie & Hipkins

Zealand's national monitoring programme (formerly known as NEMP, then NMSSA).<sup>1</sup> At the time of writing, it is morphing again into a reshaped curriculum, progress, and insights study.

NEMP assessment design exemplified the value of providing students with multiple different opportunities to express what they know and can do. Because teachers were (and still are) involved in administering the assessment tasks and making judgements about student responses, the programme also did, and continues to, affirm teachers as competent partners in assessment (Gilmore, 2002).

Most recently, NMSSA science assessments have focused on the idea of "science capabilities", giving teachers experience of using a range of assessment tasks that integrate epistemic (Nature of Science) and conceptual elements of the curriculum (NMSSA, 2017). In this collection, the article by Carpendale and Borthwick traces this influence beyond NMSSA itself (which has a focus on Years 4 and 8) exploring how the concept of science capabilities translates into assessment in secondary classrooms.

We also owe a debt of gratitude to Roger Osborne and Peter Freyberg for their Learning in Science Programme (LISP) of research into student explanations of natural phenomena. The LISP team chose the name "alternative conceptions" (rather than misconceptions) to identify student ideas that contrasted with those of scientists. They positioned students as active meaning makers, setting the stage for teachers' assessment attention in the classroom. The ideas about AfL cited above (Cowie et al., 2013) were expansions of the LISP (Assessment) project (Cowie & Bell, 1999). These projects arguably would not have happened without this credit view of student thinking. It is interesting to note that the LISP focus on encouraging students to actively explore their own ideas seeded the substantial international effort that went into encouraging teachers to use AfL practices routinely.

The DANZ report (*Directions for Assessment in New Zealand*) is a more recent beneficiary of the legacy thinking handed down from the LISP team (Absolum et al., 2009) and from international arguments for active involvement of students in making informed judgements about their learning as it unfolds in the classroom (e.g., Black & Wiliam, 1998).

DANZ argues for the development of "assessment capability" for everyone involved in the education system, including students. This advocacy again explicitly positions learners as capable participants in both learning and assessment. A more recent report updates this advocacy, again noting that fostering student agency via self and peer assessment is a key trend in the assessment research literature (Hipkins & Cameron, 2018). In the senior secondary context, recent classroom-based research has illustrated how one teacher was able to use the flexibility within *NZC* and NCEA to offer Year 12 secondary school students guided and structured choices of meaningful learning and assessment within an "alternative" science programme for non-science majors (Trask & Cowie, 2022a). Hipkins et al. (2016) outline earlier research that illustrates how innovation can be achieved through strategic and creative use of the NCEA achievement standards to aspects of the curriculum that are worthy of assessment attention.

Like Everth, the article by Carrie Vander Zwaag points out that the issues which face students in the Anthropocene cross the traditional learning areas of the curriculum, as well as being emotionally engaging for them. Her article introduces aspects of students' emotional involvement in their learning into the discussion. We have known for some years that assessment is fundamentally an "emotional practice" (Steinberg, 2008). Typically, the negative role of emotions in assessment is seen as problematic for both students and teachers. More recently, Rowe (2017) has argued that a better understanding of the role of emotions in enabling or hampering students from acting on feedback is key to scaling up AfL. How we might reframe both curriculum and assessment processes to capture positive emotional engagement would be an interesting challenge to explore further.

All these pieces of research identify increased motivation and interest in science as a benefit of greater student agency over learning and assessment. They also imply a need to carefully weigh the research attention given to formative and summative purposes of assessment. Both have an essential role to play in learning, in classroom-based curriculum development and implementation, and in system-level monitoring.

# Aligning purposes for learning with assessment processes and tasks

Clarity about both immediate and longer-term purposes for learning draws attention to students' knowing, doing, and being—how they take their learning into life contexts beyond school. This implies an expanded focus in learning intentions that should be reflected in how the learning unfolds, and in the nature of the evidence generated by both informal and formal assessments (Boud, 2000). In this collection, the article by Suzanne Trask and Bronwen Cowie expands on this point with a focus on the potential for assessment to complement a curriculum focus on science education for a social justice and social good agenda. The article by Carrie Swanson and two colleagues tackles the same challenge in a very different context. They explore ways that assessment experiences during initial teacher education might be designed to reflect and enhance early career teachers' sense of what it means to be a teacher and a colleague, at the same time as assessing actual curricular learning and associated lesson planning capabilities.

An interesting "chicken and egg" dilemma arises when curriculum innovations are intended to change the focus of what learning is for, and hence, by implication, the focus of assessments. As one example, when key competencies were added to NZC (Ministry of Education, 2007) there was a reasonably common belief that they would not be taken seriously if they were not assessed (see Hipkins, 2007). With hindsight, it would have been preferable to insist on setting the assessment challenge aside until there was more clarity about the curriculum work the key competencies were expected to do (Hipkins, 2009). Understanding their potential as curriculum change agents turned out to be an extended learning journey for teachers and researchers alike (McDowall & Hipkins, 2018). The question arises: Would teachers have more readily understood the key competencies as curriculum change agents if they had been able to access assessment exemplars? But then how could such exemplars be developed until those potential changes were clearly understood? This is what one research team in Australia referred to as a chicken and egg dilemma (Scoular & Heard, 2018).

Recent science curriculum development work in New Zealand has tackled this chicken and egg dilemma by elaborating on the idea of "enduring competencies" (Hipkins et al., 2022). In the context of the science curriculum, these are competencies that all students could be expected to develop if we take seriously the injunction that their science learning should support them to be and become "critical, informed and responsible citizens in a society in which science plays a significant role" (Ministry of Education, 2007, p. 17). The article by Hipkins and Darr speculates on the impact the idea of enduring competencies, as elaborated in this recent curriculum work, might have on potential assessment targets in senior secondary (NCEA) assessments. This discussion highlights the potential of standards-based assessment to meet the challenges of seeking new types of evidence of learning. However, it also emphasises the critical importance of clearly defined and elaborated assessment criteria, with a clear line of sight that helps teachers to navigate high-level curriculum and assessment constraints, while leaving them room to build a rich local learning and assessment programme that is meaningful and engaging for their students.

The notion of tight-loose framing highlights the complexity of the dynamics within and between high-level assessment structures and teachers' local curriculum and assessment decision making (Wiliam & Thompson, 2008). Trask and Cowie (2022b) explore the variability that the tight-loose framing of the NZC curriculum and NCEA assessment allows in secondary science learning. They outline the trade-offs teachers and students in their study made according to their felt accountabilities and priorities within the tightness and looseness of the curriculum-assessment duo. Zohar and Hipkins (2018) compare several tight/loose dilemmas in two quite different national contexts (New Zealand and Israel). Their analysis concludes that looseness is associated with a lack of clear epistemic criteria for designing appropriate learning experiences and assessments when the intention is to foster complex outcomes such as higher-order thinking. These articles highlight the need for teachers to develop strong epistemic understandings of the pertinent subject for them to be able to take advantage of the flexibility in both NZC and NCEA. It is noteworthy that the importance of developing teachers' epistemic PCK is a theme that also threads through most of the articles in this collection. Internationally, supporting and enhancing teachers' professional knowledge growth has been identified as the critical component when scaling up new curriculum and/or assessment initiatives, if change is to be sustained beyond pockets of early innovation (McNaughton, 2021; Zohar, 2023).

# **Concluding thoughts**

For this collection we invited our contributors to draw on both international and national research to explore the dynamics of the interplay between curriculum and assessment. We encouraged our contributors to consider how science is distinguished from other disciplines in ways that could or should have an influence on both assessment and curriculum thinking.

Looking back over what we have collectively achieved, we can see that a focus on inclusion is missing, along with any discussion of the dilemmas that arise when there is a serious intention to accommodate student diversity within assessment design. The process known as evidence-centred curriculum design (ECD) (Mislevy et al., 1999) was developed in part to meet this challenge. The initial stages of ECD focus on careful curriculum analysis to develop a clear "assessment argument". Advocates for the use of ECD point out that it is well suited to the creation of assessments that intend to collect evidence of complex and multifaceted curriculum outcomes (e.g., Newton et al., 2021). Its use is also advocated in high-stakes contexts where students with some types of specific learning needs are likely to be unfairly disadvantaged by construct-irrelevant variables inherent in traditional assessment designs (e.g., Gorin, 2014). It is interesting that the ECD process, and specifically the idea of developing a clear curriculumbased assessment argument, has not been adopted in New Zealand's national assessment processes, or accompanying research (Lee & Hipkins, 2022, explore one small exception). Might doing so help ameliorate the tight/loose dilemma outlined above? This is a question for policy makers to consider in the first instance. The universal entitlement to access relevant and engaging learning, and to be provided with accessible opportunities to share and assess that learning (e.g., Vander Zwaag, this issue) is a dilemma that will only grow in urgency as our student population becomes ever more diverse and the imperative to pay more than lip service to inclusion builds momentum (e.g., Rose, 2021). We note that there seems to have been greater recent research attention to this challenge in tertiary education than in the school sector (e.g., Ajjawi et al., 2022).

For us as co-editors, the most significant gap in the collection involves consideration of the imperative to decolonise the curriculum and its associated assessment practices (Kiddle et al., 2020; Nayeri & Rushton, 2022; Te Maro & Averill, 2023). This is emerging as an international concern with the

double worry that assessment thinking is lagging behind curriculum (see, for example, Crossouard & Oprandi, 2022; Mueller, 2021; Parker, 2023). This is an especially pertinent issue in New Zealand, where the current curriculum refresh has an intention to include consideration of different worldviews and knowledge systems, specifically mātauranga Māori (see Stewart, 2022). This intention poses especially acute challenges for science and science education, where the perhaps confronting focus is layered onto the already existing dilemma of lifting teachers' understanding of the nature of "science" (Parker, 2023). Again, we note that the tertiary sector appears to be ahead of the school sector in exploring and discussing this challenge.

One theme that does come through very clearly is the key role of teacher knowledge and capabilities in relation to desirable classroom-based assessment interactions, and in relation to teachers' choices and responses to assessments that take place in contexts beyond the classroom. Creating a national curriculum that meets espoused needs to foster a range of sophisticated and complex student capabilities is not enough. Nor would it be enough to align such a curriculum with appropriate high-stakes assessment design and practices at either the primary or the secondary level. The need for teachers to have a robust understanding of the epistemic practices of the relevant discipline is a strong theme of this collection. Such understanding could arguably serve as a bridge between curriculum and assessment intentions and the curriculum and assessment practices that are enacted in classrooms. However, in the context of science education, decades-long research efforts suggest that building teachers' epistemic knowledge is a goal that will not be easily achieved (e.g., Lederman et al., 2023). We can see glimpses of what could be effective for this purpose in the article by Carpendale and Borthwick in this collection. The need for informed, sustained, and well-supported professional learning is clear.

### Note

1. NEMP: National Education Monitoring Programme; NMSSA: National Monitoring Studies of Student Achievement

### References

Absolum, M., Flockton, L., Hattie, J., Hipkins, R., & Reid, I. (2009). Directions for assessment in New Zealand. Learning Media.

- Ajjawi, R., Tai, J., Boud, D., & Jorre de St Jorre, T. (2022). Assessment for inclusion in higher education: Promoting equity and social justice in assessment. Routledge. https:// doi.org/10.4324/9781003293101
- Black, P. (2013). Formative and summative aspects of assessment: Theoretical and research foundations in the context of pedagogy. SAGE Publications. https://doi. org/10.4135/9781452218649
- Black, P., & Wiliam, D. (1998). Assessment and classroom learning. Assessment in Education: Principles, Policy & Practice, 5(1), 7–74. https://doi. org/10.1080/0969595980050102
- Boud, D. (2000). Sustainable assessment: Rethinking assessment for the learning society. *Studies in Continuing Education*, 22(2), 151–167. https://doi.org/10.1080/713695728
- Coffey, J., Hammer, D., Levin, D., & Grant, T. (2011). The missing disciplinary substance of formative assessment. *Journal of Research in Science Teaching*, 48(10), 1109–1136.
- Cowie, B., & Bell, B. (1999). A model of formative assessment in science education. *Assessment in Education*, 6(1), 101–116.
- Cowie, B., Moreland, J., & Otrel-Cass, K. (2013). Expanding notions of assessment for learning: Inside science and technology primary classrooms. Springer Science & Business Media.
- Crooks, T. J. (1988). The impact of classroom evaluation practices on students. *Review of Educational Research*, *58*(4), 438–481.
- Crossouard, B., & Oprandi, P. (2022). Decolonising formative assessment. In J. Huisman & M. Tight (Eds.), *Theory and method in higher education research* (Vol. 8, pp. 181–196). Emerald Publishing. https://doi.org/10.1108/S2056-37522022000008010
- Gilmore, A. (2002). Large-scale assessment and teachers' assessment capacity: Learning opportunities for teachers in the National Education Monitoring Project in New Zealand. Assessment in Education: Principles, Policy & Practice, 9(3), 343–361.
- 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
- Hayward, L. (2015). Assessment is learning: The preposition vanishes. Assessment in Education: Principles, Policy & Practice, 22(1), 27–43. https://doi.org/10.1080/096 9594X.2014.984656
- Heritage, M., & Wylie, E. C. (2020). Formative assessment in the disciplines: Framing a continuum of professional learning. Harvard University Press.
- Hipkins, R. (2007). Assessing key competencies: Why would we? How could we? Ministry of Education.
- Hipkins, R. (2009). Determining meaning for key competencies via assessment practices. Assessment Matters, 1, 4–19. https://doi.org/10.18296/am.0070
- Hipkins, R., & Cameron, M. (2018). *Trends in assessment: An overview of themes in the literature*. New Zealand Council for Educational Research.
- Hipkins, R., Johnson, M., & Sheehan, M. (2016). NCEA in context. NZCER Press.

- Hipkins, R., Tolbert, S., Cowie, B., & Waiti, P. (2022). Enduring competencies for designing science learning pathways. Rangahau Mātauranga o Aotearoa / New Zealand Council for Educational Research. https://doi.org/10.18296/rep.0025
- Johnson, M., Hipkins, R., & Sheehan, M. (2018). Building epistemic thinking through disciplinary inquiry: Contrasting lessons from history and biology. *Curriculum Matters*, 13, 80–102. https://doi.org/10.18296/cm.0020
- Kiddle, R., Jackson, M., Elkington, B., Mercier, O. R., Ross, M., Smeaton, J., & Thomas, A. (2020). *Imagining decolonisation*. Bridget Williams Books.
- Lederman, N. G., Zeidler, D. L., & Lederman, J. S. (Eds.). (2023). Handbook of research on science education (Vol. 3). Taylor & Francis.
- Lee, J., & Hipkins, R. (2022). Accommodating diversity in assessment: A snapshot of practice in 2022. Rangahau Mātauranga o Aotearoa / New Zealand Council for Educational Research. https://www.nzcer.org.nz/research/publications/ accommodating-diversity-assessment-snapshot-practice-2022
- McDowall, S., & Hipkins, R. (2018). *How the key competencies evolved over time: Insights from the research*. Rangahau Mātauranga o Aotearoa / New Zealand Council for Educational Research. https://www.nzcer.org.nz/system/files/Paper%202%20 KCs%20research%20 final.pdf
- McNaughton, S. (2021). The conundrum research–practice partnerships face with system variability. *Studies in Educational Evaluation*, 70, 101048. https://doi.org/10.1016/j. stueduc.2021.101048
- Ministry of Education. (2007). The New Zealand curriculum. Learning Media.
- Mislevy, R., Steinberg L., & Almond, R. (1999). *Evidence-centered assessment design*. ETS.
- Mueller, B. (2021) Decolonising assessment within higher arts education. *JUICE: Journal of Useful Investigations in Creative Education*, (4). https://juice-journal. com/2021/11/23/decolonising-assessment-within-higher-arts-education/
- Nayeri, C., & Rushton, E. (2022). Methodologies for decolonising geography curricula in the secondary school and in initial teacher education. *London Review of Education*, 20(1), 4. https://doi.org/10.14324/LRE.20.1.04
- Newton, S., Alemdar, M., Rutstein, D., Edwards, D., Helms, M., Hernandez, D., & Usselman, M. (2021). Utilizing evidence-centered design to develop assessments: A high school introductory computer science course. *Frontiers, Secondary Assessment, Testing and Applied Measurement, 6.* https://doi.org/10.3389/feduc.2021.695376
- NMSSA. (2017). Science 2017: Key findings. Educational Assessment Research Unit, University of Otago, and New Zealand Council for Educational Research. https:// www.educationcounts.govt.nz/publications/series/nmssa/science/nmssa-2017-science
- Parker, K. (2023, 16 February). Do we need to decolonise the science curriculum? TES Magazine.
- Quinlan, K., & Pitt, E. (2021). Towards signature assessment and feedback practices: A taxonomy of discipline-specific elements of assessment for learning. Assessment in

*Education: Principles, Policy & Practice, 28*(2), 191–207. https://doi.org/10.1080/0 969594x.2021.1930447

- 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
- Rowe, A. (2017). Feelings about feedback: The role of emotions in assessment for learning. In D. Carless, S. M. Bridges, C. K. Y. Chan, & R. Glofcheski (Eds.), *Scaling up assessment for learning in higher education* (pp. 159–172). Springer.
- Scoular, C., & Heard, J. (2018, 5 June). Teaching and assessing general capabilities. *Teacher Magazine*. https://www.teachermagazine.com.au/articles/teaching-and-assessing-general-capabilities?utm source=CM&utm medium=bulletin&utm content=June5
- Steinberg, C. (2008). Assessment as an "emotional practice". *Waikato Journal of Education*, 7(3), 42–64.
- Stewart, G. (2022). Māori science curriculum. In M. Atwater (Ed.), *International handbook of research on multicultural science education* (pp. 871–893). Springer. https://doi.org/10.1007/978-3-030-83122-6
- Te Maro, P., & Averill, R. (2023). Ki te hoe! Education for Aotearoa. NZCER Press.
- Trask, S., & Cowie, B. (2022a). On their own terms? Opening up senior science learning for non-specialist science students. *International Journal of Science Education*, 44(4), 674–693. https://doi.org/10.1080/09500693.2022.2050489
- Trask, S., & Cowie, B. (2022b). Tight–loose: Understanding variability, trade-offs and felt accountability across the curriculum-pedagogy-assessment dynamic. *The Curriculum Journal*, 33(4), 587–601.
- Wiliam, D., & Thompson, M. (2008). Tight but loose: A conceptual framework for scaling up school reforms. *Tight but loose: Scaling up teacher professional development in diverse contexts*. Educational Testing Service.
- Yates, L., Woelert, P., Millar, V., & O'Connor, K. (2017). *Knowledge at the crossroads? Physics and history in the changing worlds of schools and universities*. Springer.
- Zohar, A. (2023). Scaling up higher order thinking: Demonstrating a paradigm for deep educational change. Springer.
- Zohar, A., & Hipkins, R. (2018). How "tight/loose" curriculum dynamics impact the treatment of knowledge in two national contexts. *Curriculum Matters*, *14*, 31–47. https://doi.org/10.18296/cm.0028

### The authors

**Bronwen Cowie** is a professor of education at the University of Waikato. ORCID ID 0000-0003-3578-0791

Email bronwen.cowie@waikato.ac.nz

**Rosemary Hipkins** is a kaihautū rangahau / chief researcher at Rangahau Mātauranga o Aotearoa New Zealand Council for Educational Research.

Email rose.hipkins@nzcer.org.nz