

Obstacles and opportunities in the New Zealand curriculum refresh: An illustrative case study of the enactment of formative assessment in a New Zealand primary school

Thomas C. Pearce

<https://doi.org/10.18296/cm.0212>

Abstract

In the context of significant reforms to the curriculum for New Zealand schools, this article presents an early line of findings from an ongoing ethnographic multi-case study investigating New Zealand primary school teachers' enactment of informal formative assessment. Data were gathered through observations, interviews, and document collection in one Auckland primary school classroom. The findings illustrate the value of informal formative assessment and provide a perspective on the use of formal assessment tools grounded in classroom practice. They suggest some barriers to greater student involvement in assessment, and identify trust, perceptions of validity, control, and definitions of quality as key considerations in how teachers enact formal assessments.

Introduction

New Zealand is in the process of updating the national curriculum for English-medium primary and secondary schools. The update includes a review of assessment guidelines and practices, including the introduction of a national testing regime for primary schools. This article contributes to the discussion by presenting findings from a case study of the enactment of formative assessment at a New Zealand primary school. The case study illustrates ongoing tensions between curriculum and assessment, and highlights opportunities and obstacles in the implementation of the updated curriculum from a perspective grounded in everyday classroom practice.

The article begins by considering the interrelationship of curriculum and assessment. It then locates assessment in research, curriculum, and practice in New Zealand primary schools, before considering specific issues relating to the enactment of assessment. Next, data from an ethnographic case study of the enactment of formative assessment in a Year 4 and Year 5 classroom at an Auckland primary school are presented. Finally, implications are discussed for the curriculum update, including considerations for policy makers in the development of the curriculum and considerations for schools developing their local curricula and assessment practices. The findings suggest that enactment—how policy and curriculum is translated, interpreted, and mediated into classroom practice—must be a key consideration in education reform.

Curriculum and assessment in New Zealand

The primary curriculum document in New Zealand English-medium primary schools is *The New Zealand Curriculum* (NZC) (Ministry of Education, 2007). The findings and discussion here relate to the enactment of formative assessment as guided by this curriculum. A replacement called *Te Mātaiaho* was set to be implemented in all schools by 2025, but was recently withdrawn for redevelopment and implementation is now planned in stages through to 2027. The most recent draft of *Te Mātaiaho* was released in March 2023 (Ministry of Education, 2023). That draft is the document used as a basis for discussion, and to consider the implications of these findings for the ongoing curriculum update in New Zealand. Elsewhere in this article the term *curriculum* refers more broadly to the constellation of documents, policies, plans, and lessons that make up the experiences of students within the education provided to them (Kelly, 2009).

The stance taken by the New Zealand Ministry of Education on the relationship between curriculum and assessment is that curriculum should underpin assessment (Ministry of Education, 2011). *NZC* is intended to provide the framework from which schools develop their local curricula, which then determines the assessment needs and approaches.

The reality is often contrary, and many researchers have investigated the phenomenon of curriculum “backwash”, in which assessments influence or determine what is taught (Hipkins & Cameron, 2018). Much of the literature on backwash considers it to be detrimental, causing teachers to focus on superficial and disconnected skills and knowledge (Daly et al., 2012). This

is a longstanding critique of high-stakes assessments, but recent research and commentary suggest that low-stakes assessments can also have a backwash effect, particularly when there is an emphasis on testing (Baird et al., 2017; Carter, 2020; Darr & Cosslett, 2024). Other research argues that backwash should be a neutral term, because good assessment practices can have positive effects on curriculum and pedagogy (de la Fuente Fernández & Calvo Pascual, 2022).

Regardless of the stance, curriculum and assessment have a more mutually constitutive relationship than the Ministry's stance suggests. What teachers assess and how they assess both impact significantly on the taught curriculum. When curriculum and assessment are out of alignment there are conflicting demands placed on teachers, as has been the case in New Zealand in the past (Poskitt, 2018; Thrupp, 2017).

Assessment in New Zealand primary schools

The New Zealand Government's recent announcement of twice-yearly standardised testing of all primary school students was a significant departure from the teacher-based formative assessment that has been the focus of curriculum, professional development, and government policy for more than 40 years (Mutch, 2012). Formative assessment, in its broadest sense, is the use of assessment to improve teaching and learning. The new national testing regime is intended for formative purposes, as well as for summative purposes such as reporting to parents, school boards, and the Ministry of Education (Stanford, 2024).

Formative assessment can range from formal to informal practices. More formal approaches can involve things like the standardised assessment tools soon to be compulsory in New Zealand, less formal approaches include things like teacher-made quizzes or written feedback on student work, while informal formative assessments involve unplanned on-the-fly judgements made and acted on in the moment (Shavelson et al., 2008). In New Zealand, an approach to formative assessment known as assessment for learning (AfL) has been the focus of policy and research. An orthodox view of AfL is as an interdependent set of practices that promote active student engagement in assessment and ownership of learning (Dixon et al., 2020). Dixon et al. include processes of collaboration, goal setting, and criteria discussions, eliciting and sharing of evidence from both formal and informal assessments, provision of feedback, and student peer- and self-assessment.

AfL is included in *NZC* and *Te Mātaiaho*, both of which discuss assessment in terms of formal and informal “gathering, analysis, interpretation, and use of information” to plan next steps for students and to inform adjustments to teaching (Ministry of Education, 2007, pp. 39–40). *Te Mātaiaho* integrates AfL into its “essential pedagogies” and highlights the central role of students in effective assessment practice.

While most New Zealand teachers affirm the use of assessment to improve teaching and learning (Harris & Brown, 2009), and 94% of teachers in a 2019 survey ($n = 620$) reported daily use of formative assessment (Wylie & MacDonald, 2020), there is little known about how closely teachers’ assessment practices match what is described in the curriculum documents. The Education Review Office (ERO), that evaluates and reports on schools, indicates “considerable improvements” in the collection and use of assessment information to improve learning, but notes that this remains inconsistent between and within schools (Education Review Office, 2018, p. 49). Older surveys show that conferences, observations, and oral and interactive modes of assessment dominate teachers’ self-reported assessment practices (Brown & Harris, 2009; Dunn et al., 2003). Other researchers similarly suggest that most formative assessment in New Zealand classrooms is informal and interactive, with information remembered rather than recorded. (Cowie et al., 2018; Crooks, 2011; Hill, 2002). Some small-scale studies highlight examples of highly effective AfL practice (Hipkins & Cameron, 2018), while others suggest that the use of AfL as a coherent set of assessment practices remains highly inconsistent in New Zealand classrooms (Dixon et al., 2020; Poskitt, 2018; Sinnema et al., 2018).

While limited by a paucity of recent studies, the available literature suggests that New Zealand is similar to other countries in that teachers’ formative assessment practices are inconsistent and hindered by practical considerations involved in a typical classroom (Schellekens et al., 2021).

Issues in the enactment of formative assessment

Despite the close connection between curriculum and assessment, the evidence above indicates that high-level curriculum policy is not reflected directly in classroom practice. Instead, assessments are enacted through the interaction between teacher agency and curriculum. Enactment is the “diverse and complex ways in which sets of education policies are ‘made sense of’, mediated

and struggled over, and sometimes ignored” (Ball et al., 2012, p. 3). Teachers translate, interpret, and operationalise the abstract ideas of curriculum and assessment policy into concrete and contextualised practices. This is influenced by a range of psychological, personal, social, cultural, discursive, material, and pragmatic factors (Willis & Klenowski, 2018).

For example, many New Zealand teachers associate informal assessment with improvement and learning, while formal assessment tools are associated with summative assessment purposes, negative perceptions of accountability, and teachers’ own negative experiences of being assessed (Brown, 2009; Brown & Harris, 2009; Smith, 2014). Informal formative assessment is also central to the responsive teaching that is valued highly in New Zealand pedagogic discourses, because it informs productive responses to student thinking and leaves space for divergent and more genuinely student-directed outcomes (Cowie et al., 2018; Kerr & Averill, 2021).

Teachers need data that are recognisable, timely, finely grained, and interpretable in a way that informs action (Crooks, 2011; Wiliam, 2018b). However, data must also be accurate, and factors like anxiety, self-efficacy, competition, and motivation can distort assessments (Kluger & DeNisi, 1996; Sadler, 2010; Zhao et al., 2022). Informal assessments are fraught with misunderstandings and biases that arise from human factors like gender, culture, language ability, or other characteristics of teachers and students (Baird et al., 2017; Torrance & Pryor, 1998). New Zealand teacher judgements and ability group placements of marginalised students are often lower than their peers with similar results on standardised assessments (Meissel et al., 2017).

Teachers and students need high assessment capability or assessment literacy to improve the accuracy and validity of assessments (Tai et al., 2018; Xu & Brown, 2016). This involves dialogic assessment practices that negotiate concepts of quality and allow for students to engage with the “tacit”, or socially and culturally embedded, aspects of learning and assessments (Beck & Jones, 2023; Pryor & Crossouard, 2008; Sadler, 2010). However, accountability pressures and personal or cultural expectations around teacher authority pose barriers to dialogic practice (Heng et al., 2021).

It is left to teachers to balance these and more considerations in their enactment of formative assessment and curriculum, reconciling “what are often competing

policies and practices in ways that are both pragmatic and idealistic.” (Willis & Klenowski, 2018, p. 32). This article presents some of these issues through a case study of one teacher’s enactment of formative assessment.

Research methodology and theoretical orientation

The findings reported here arose from one case in an ongoing qualitative ethnographic multi-case study research project. Case-study research is typified by the unit of analysis, emphasising the delimitation and bounding of a research subject (Merriam & Tisdell, 2015); in this case, at the level of the school. Merriam (1998) distinguishes ethnography from other qualitative methods through reference to ethnography’s sociocultural interpretive frame. For this research, the sociocultural frame includes the culture of the school and its community, as well as the attitudes, beliefs, and values of the participating teacher.

The theoretical orientation of this research is social constructivist. Social constructivism holds that our realities are contingent and created or shared by individuals through interaction within their social and physical worlds (Creswell, 2013). An epistemological axiom of social constructivism is that individual knowledge and experience are a valid basis for inquiry (Spencer et al., 2014).

Methods

The research was designed to explore teachers’ enactment of informal formative assessments in their day-to-day practice. The research questions that guided this project were:

- What are the participating teachers’ conceptions of informal formative assessment, and why?
- How are these teachers enacting informal formative assessment?
- What factors affect the ways in which these teachers enact informal formative assessment?

Participation criteria are summarised in Table 1. Criteria sought to include information-rich, “typical” cases in the sense of “not in any major way atypical, extreme, deviant or intensely unusual” (Patton, 2015, p. 284). A socioeconomic criterion was not used because socioeconomic information was not readily available for schools at the time of recruitment. Using these criteria, one teacher at each of three schools was recruited for close collaboration on

the project, along with one member of the senior leadership team to provide information about the broader context of each teachers’ practice.

Table 1. Inclusion criteria for the research project

School	• Serving an urban community within the Auckland area • Either contributing (Years 0–6) or full (Years 0–8) English-medium state primary school • More than 100 students on the roll • Fewer than 80% of students identifying as Pākehā as identified by the Ministry of Education • Most recent ERO review indicating a follow-up period of 3 or more years
Teacher	Five years of teaching experience in New Zealand state primary schools

Following approval for the research project from The University of Auckland Human Ethics Committee, consent was obtained from school principals, the participating teacher and member of the senior leadership team, and parents or guardians of students. Students provided assent to participate. Any students not participating were selectively excluded from data-collection practices and were unaffected by the research project. Pseudonyms were assigned to all participants, as well as the school.

Table 2. Data-collection sources and data collected at each research site

Data source	Data collected
Classroom teacher	• Artefacts and memos from one pre-observation meeting to discuss the teacher’s initial conceptions of formative assessment • Descriptive field notes from three observations lasting roughly 1 hour of classroom practice in reading, writing, and/or mathematics • Audio recording of three semistructured interviews following each observation, to discuss the observation as well as the teacher’s beliefs, values, understandings, and experiences • Audio recording of one small-group lesson • Audio recording of one semistructured exit interview for member checks and to discuss initial findings

Table 2. Data-collection sources and data collected at each research site

Senior leadership team member	• Audio recording of one semistructured interview discussing the school’s assessment policy, as well as the history, culture, and community of the school
School documents	• Teacher planning • Assessment data • School policies • Student work

Audio data were transcribed using the Microsoft Word dictation feature, then corrected and anonymised manually by the researcher. Transcripts were verified by participants and an early summary of findings was presented to the participating teacher prior to the final interview to allow for member checks. Transcripts, field notes, memos, artefacts, and documents were added to the NVivo 14 software for coding. An eclectic approach was taken to coding, primarily using process/action codes, in vivo codes and values codes to try to faithfully represent each teacher’s perspective and practice (Saldaña, 2016; Stake, 1995). The NVivo software allowed for a constant comparison and refining as new codes arose, checking if they supported, contradicted, or expanded on existing codes (Corbin & Strauss, 2008). A second cycle of coding involved cleaning, further refining, and (re)categorising codes to reveal themes within the data.

This data collection and analysis produced a number of findings relating to both formal and informal formative assessment. A group of related themes arose from one case study that were of relevance to the ongoing curriculum refresh in New Zealand. These findings are presented here.

The case study: Formative assessment in a Year 4 and Year 5 class at Hill Street School

Hill Street School is a diverse Auckland school serving roughly 400 students from Years 0–8. The school community is primarily working class, and most students are Māori and Pasifika, with smaller but significant numbers of New Zealand European/Pākehā, Indian, and Filipino students. Most families were reported to be very supportive of the school, but not closely engaged.

School assessment policy closely reflected *NZC*. The school had a top-down, data-driven approach to assessment via school “trackers”, as well as various formal assessment tools (“anchor tools”). These are summarised in Table 3. The trackers were descriptors of skills in reading, writing, and maths, broken down by curriculum level. They were developed by a school curriculum team through a synthesis of curriculum documents and the skills required by the anchor tools. The trackers were the framework of the school curriculum in reading, writing, and mathematics, and teachers’ planning was required to have a link to a specific tracker for each lesson.

Table 3. Prescribed assessment practices at Hill Street School

Anchor tools	• Progressive Achievement Tests of mathematics, reading comprehension, and vocabulary, early in Term 1 and in Term 4 • e-asTTle assessments of reading, writing, and mathematics at the end of Terms 1 and 3 • PROBE test of reading comprehension at the end of Terms 2 and 4, and ongoing as needed • Mathematics basic-facts tests twice per term • BSLA phonics assessments for students in Years 0–2
School trackers	• Teachers gather formal and informal evidence of students using skills independently • Teachers check a box in a spreadsheet for each “sighting” of evidence, for every student • After three “sightings”, a date is added to the particular tracker spreadsheet and a student is deemed to have “achieved” the skill

Bey taught a Year 4 and Year 5 class at Hill Street School and was also the team leader. She had 5 years of teaching experience, all of which had been at Hill Street School. She put relationships at the centre of her practice, making regular reference to students’ backgrounds, interests, culture, engagement, agency, social and emotional wellbeing. The centrality of relationships flowed on to her pedagogy, where she believed in individualisation of learning and valuing student input. She cited the work of Bishop (2019) and Rubie-Davies (2014) as particular influences, informing decisions like flexible grouping, reciprocal learning, conscious rejection of deficit thinking, and attention to students’ social and emotional wellbeing alongside their academic needs.

She aligned herself closely with the approach taken by the school, in terms of policy, pedagogy, and values.

At the start of the research project, Bey was knowledgeable about assessment and reflective on her assessment practice. In line with the conceptions of many New Zealand teachers, she associated formative assessment with informal and interactive practices, while identifying the anchor tools as primarily summative (Brown, 2009). She also understood assessments in terms of generating evidence to inform teaching, and for ensuring accountability of both teachers and students.

During observations, Bey used a range of informal formative assessment practices that aligned well with the notice, recognise, and respond framework promoted by the Ministry of Education (Te Poutāhū Curriculum Centre, 2023). Interviews and document collection gave insights into her more formal practices.

Findings

The value of informal assessments

In interviews, Bey discussed informal assessment practices as more individualised, more supportive of the students' social and emotional needs, and more useful in developing deeper and more interconnected learning. They aligned well with the central position of relationships in her pedagogy. Bey perceived strong relationships as adding validity to her informal and teacher-based assessment practices, because they allowed her to account for the personal, emotional, social, and contextual factors that affect students' performances. Informal assessments allowed Bey to supplement her judgements of students' learning with consideration of these contextual factors. If an assessment did not align with her expectations, she could "see it and question it and confirm it".

Bey also identified relationships and student behaviour or personality as potential points of bias in informal assessments. She said that teachers, herself included, can end up "making excuses for a kid" because they do not want to see them fail. She suggested this could lead to inflated judgements of student ability, which contradicts other research that suggests New Zealand teachers underestimate most students (Meissel et al., 2017). She also acknowledged that time spent with each student was a limiting factor, impacted by student personality and behaviour.

Validity, trust, and definitions of quality in assessments

Throughout the research project, Bey regularly acknowledged the importance of the anchor tools, particularly for accountability: “[y]ou can’t just go with teacher judgements ... assessment is important, so that we can, yeah, have evidence, as well”. Despite this, Bey primarily discussed anchor tools in negative terms; as time consuming, challenging to administer, and less useful for her teaching than other assessments.

A key issue for her was a lack of trust in the data from some anchor tools, particularly Progressive Achievement Tests (PATs) and e-asTTle. This was a point of conflict, because she also acknowledged the subjectivity and unreliability of teacher judgements. However she explained that, with the anchor tools, “there’s no relationship with the kids there”, meaning they were critically decontextualised. Construct validity issues further eroded her trust in their accuracy, including student rapid-guessing, problems for students with specific learning needs or English language learners, the decontextualised use of skills in a test environment, and situational factors like a student not having had breakfast the morning of a test. This did not lead her to disregard the data, but the mistrust made it difficult for her to feel ownership of it.

For leadership data perspective, they would see it as the black and white results. Whereas we see it with our kids ... And we obviously know that they can do these things, but then it’s not seen on their results. So then leadership will say often, ‘Well, you need to cover this’, but you would be like ‘I’ve covered this!’ ... They haven’t been in the classroom to see the whole teaching, so I feel like they’re just trusting the data, not trusting me.

She did not view this as leadership being unsupportive, stating that it is their job to play “devil’s advocate” with the data. However, it reflected a perception that these assessments externalise definitions of quality teaching by placing all responsibility on the teacher for student progress on these metrics. Comparisons and judgements made on this basis were seen as unfair and “disheartening”.

Formative assessment or AfL?

Bey’s formative assessment practice drew on many of the features of AfL, including collaboration, criteria discussions, eliciting of evidence from both formal and informal assessments, provision of feedback, and student peer- and

self-assessment. However, similar issues to those discussed in the previous section extended to students, and Bey highlighted these as obstacles to the authentic student ownership of learning promoted by AfL research, NZC, and included in *Te Mātaiaho*.

Only student writing goals and data from mathematics basic-facts tests were shared with students, though Bey recognised that more student ownership of these data was desirable, saying that she would like to “do better”. However, she found that test data created narrow and shallow definitions of success for students. She gave the example of the basic-facts tests establishing among students a hierarchy of who was “good” at mathematics that did not align with broader student ability or progress in other areas of mathematics. Sharing the data also invited comparisons that harmed students’ self-efficacy, and Bey observed students disengaging as a result. This conflicted with her regular use of informal assessments to allocate students to flexible and mixed-ability groups, an approach that she said had been key to engaging some students.

[O]ne of his comments was like, he loves reading now because he’s not in a ‘bottom group’ anymore. And like, that was like a huge thing. He’s one of the boys who always wants to read now and never used to. He used to hide his book instead of taking it home.

School trackers and explicit learning intentions

The school curriculum drove Bey’s informal formative assessments and her planning for assessment via the trackers. The school required each lesson to have a link to the trackers, which formed “the bones” of her planning. Bey reworded the trackers into “kid speak” and used them to develop broader, more open-ended learning intentions that allowed for mixed-ability grouping with outcomes across multiple curriculum levels.

Bey valued student agency very highly. She discussed this in terms of independence and choice, and a desire to enhance student agency influenced her enactment of formative assessment. Explicit criteria were a key to this, because without them students could struggle to understand the purpose of their work, limiting their independence. Explicit criteria were also central to Bey’s feedback, scaffolding of self-regulation skills, and broader use of AfL processes. In observations, the learning conversations and feedback between Bey and students were tightly focused on the learning intentions and success

criteria. Students were observed discussing and reflecting on learning intentions and success criteria independently, particularly in writing. To support this, Bey emphasised the need for criteria that were learning or strategy focused, as opposed to task or outcome focused.

Discussion

The case study presented here shows that assessment is not carried out prescriptively. There was a close alignment between the teacher and the school, and the school and policy. However, the teacher's enactment of assessment was often quite different from what was set out in high-level curriculum documents and the integrated AfL promoted in New Zealand. This has a number of implications for the curriculum refresh in New Zealand.

One implication relates to the formative use of PAT and e-asTTle in the curriculum refresh. The Government plans to mandate these tools, but this case study echoes earlier findings that there are significant barriers to teachers using the data from these tests formatively (Brown & Harris, 2009). At present, a lack of trust in the data may relegate these tools to summative purposes or use only by management in many schools. To some extent, this is inescapable, because formative assessments have to be useful to inform decision making (William, 2018a). This case study illustrates how, at the level of the classroom, decisions must be made on a timescale of minutes or days around things like content, grouping, pacing, and feedback. The time taken to administer and analyse data from e-asTTle or PATs is unfit for many classroom-level decisions. To ensure these tests serve genuine formative purposes in schools, the refreshed curriculum could indicate how they can support high-level or longer-term planning.

This case study also illustrates a view, common among New Zealand teachers, that strong holistic relationships are needed to support learning (Mutch & Peung, 2021). Testing risks displacing these relationships in favour of “external contingencies” (Ball, 2003, p. 222). A lack of trust in the data from tests coupled with a view of tests as management tools can therefore intensify feelings of guilt, anxiety, or resentment when data are presented to a teacher by school leadership who have little knowledge of what and how content has been covered in class, more distant relationships with the students, and little knowledge of students' personal circumstances. To enable effective formative use of formal assessment tools, the refreshed curriculum could allow flexibility in approaches

to assessment that ensure teachers have ownership over assessments and the data they produce. Interactive teacher-based assessment tools, such as the Neale Analysis of Reading Ability (Neale, 1999), may be better suited to this than standardised tests. The curriculum could also include guidelines for teachers and schools to help students understand and use assessment data in constructive ways. This might avoid harm to student self-efficacy, as well as avoid assessment data narrowing what students value in the taught curriculum.

The findings of this study also illustrate the value of informal formative assessments in supporting broad curriculum outcomes. Clear criteria were key to the informal assessments observed in this study, which is well supported by the literature (Ruiz-Primo & Brookhart, 2018). By ensuring clarity in learning outcomes, the curriculum refresh could support more effective informal assessment practice. The study also shows how this informal assessment practice can create space for negotiation and discussion of these criteria, which is key for developing students' self-regulation skills and giving students the ownership over their learning that is central to AfL practice (Dixon et al., 2020; Sadler, 2010). This all requires significant curriculum expertise (Cowie et al., 2018), so time and professional development will be crucial to the curriculum implementation.

A final obstacle may be the curriculum progression model recommended by the Ministerial Advisory Group recently appointed to inform the curriculum refresh (Ministerial Advisory Group, 2024). This case study illustrates a commitment to deep transformational learning that is common among New Zealand teachers (Brown, 2009). Progress is conceptualised as deepening or increasing the sophistication of a student's understanding. This is at odds with the Advisory Group's recommendation that the curriculum adopt linear teaching sequences with associated teaching methods.

It remains to be seen how the Ministry of Education will respond to the Advisory Group's recommendations, but this research suggests that there is a balance to be struck in the curriculum refresh that achieves greater clarity but avoids a level of prescriptiveness that may add significant workload for teachers in adapting teaching sequences to suit their teaching context and students' needs. Consultation with the education sector will be key in achieving this balance, because even when there is close alignment between the teacher, school, and government policy, teaching is not a purely technical practice conducted with the "fidelity" called for by the Advisory Group's report.

Conclusion

The present case study illustrates some of the broader opportunities within the curriculum refresh, like greater clarity of curriculum learning outcomes. It demonstrates how the informal modes of assessment common in New Zealand classrooms are well-aligned with the pedagogy described in the latest draft of *Te Mātaiaho*. However, it also points to likely tensions around the use of standardised tests and sharing data with students, and some paths to overcoming these. Ultimately, teachers are the end-users of curriculum and assessment policy, so the success or failure of policy initiatives hangs on how teachers enact that policy. The experience of teachers is a valuable resource for curriculum reform.

Conflict of interest statement

The author acknowledges a conflict of interest because he was co-editor of the journal volume to which this manuscript was submitted. He excused himself from all editorial decisions around the manuscript.

References

- Baird, J.-A., Andrich, D., Hopfenbeck, T. N., & Stobart, G. (2017). Assessment and learning: Fields apart? *Assessment in Education: Principles, Policy & Practice*, 24(3), 317–350. <https://doi.org/10.1080/0969594X.2017.1319337>
- Ball, S. (2003). The teacher's soul and the terrors of performativity. *Journal of Education Policy*, 18(2), 215–228. <https://doi.org/10.1080/0268093022000043065>
- Ball, S., Maguire, M., & Braun, A. (2012). *How schools do policy: Policy enactments in secondary schools*. Routledge.
- Beck, S. W., & Jones, K. (2023). Fostering agency through dialogue in classroom writing assessment. *Teaching and Teacher Education*, 124, 104012. <https://doi.org/10.1016/j.tate.2022.104012>
- Bishop, R. (2019). *Teaching to the north-east: Relationship-based learning in practice*. NZCER Press.
- Brown, G. T. L. (2009). Teachers' self-reported assessment practices and conceptions: Using structural equation modeling to examine measurement and structural models. In T. Teo & M. S. Khine (Eds.), *Structural equation modelling in educational research: Concepts and applications* (pp. 243–266). Sense Publishers.
- Brown, G. T. L., & Harris, L. R. (2009). Unintended consequences of using tests to improve learning: How improvement-oriented resources heighten conceptions of assessment as school accountability. *Journal of MultiDisciplinary Evaluation*, 6(12), 68–91.

Obstacles and opportunities in the New Zealand curriculum refresh

- Carter, J. (2020). The assessment has become the curriculum: Teachers' views on the phonics screening check in England. *British Educational Research Journal*, 46(3), 593–609. <https://doi.org/10.1002/berj.3598>
- Corbin, J. M., & Strauss, A. (2008). *Basics of qualitative research: Techniques and procedures for developing grounded theory* (3rd ed.). SAGE Publications. <https://doi.org/10.4135/9781452230153>
- Cowie, B., Harrison, C., & Willis, J. (2018). Supporting teacher responsiveness in assessment for learning through disciplined noticing. *The Curriculum Journal*, 29(4), 464–478. <https://doi.org/10.1080/09585176.2018.1481442>
- Creswell, J. W. (2013). *Qualitative inquiry and research design: Choosing among five approaches* (3rd ed.). SAGE Publications.
- Crooks, T. J. (2011). Assessment for learning in the accountability era: New Zealand. *Studies in Educational Evaluation*, 37(1), 71–77. <https://doi.org/10.1016/j.stueduc.2011.03.002>
- Daly, A. L., Baird, J.-A., Chamberlain, S., & Meadows, M. (2012). Assessment reform: Students' and teachers' responses to the introduction of stretch and challenge at A-level. *Curriculum Journal*, 23(2), 139–155.
- Darr, C., & Cosslett, G. (2024). Considering the assessment landscape in 2024. *Set: Research Information For Teachers*, (1), 47–50. <https://doi.org/10.18296/set.1551>
- de la Fuente Fernández, A., & Calvo Pascual, M. A. (2022). Washback of Spanish university entrance examination on chemistry teaching in upper secondary education. *Assessment in Education: Principles, Policy & Practice*, 29(4), 422–440. <https://doi.org/10.1080/0969594X.2022.2118664>
- Dixon, H., Hill, M., & Hawe, E. (2020). Noticing and recognising AfL practice: Challenges and their resolution when using an observation schedule. *Assessment Matters*, 14, 42–62.
- Dunn, K., Strafford, E., & Marston, C. (2003). *Classroom assessment practices in English and mathematics at years 5, 7 and 9*. New Zealand Council for Educational Research.
- Education Review Office. (2018). *Evaluation at a glance: A decade of assessment in New Zealand primary schools—Practice and trends*. Author. <https://ero.govt.nz/our-research/evaluation-at-a-glance-a-decade-of-assessment-in-new-zealand-primary-schools-practice-and-trends>
- Harris, L. R., & Brown, G. T. L. (2009). The complexity of teachers' conceptions of assessment: Tensions between the needs of schools and students. *Assessment in Education: Principles, Policy & Practice*, 16(3), 365–381. <https://doi.org/10.1080/09695940903319745>
- Heng, T. T., Song, L., & Tan, K. (2021). Understanding the interaction of assessment, learning and context: Insights from Singapore. *Educational Research*, 63(1), 65–79. <https://doi.org/10.1080/00131881.2021.1874248>
- Hill, M. F. (2002). Focussing the teacher's gaze: Primary teachers reconstructing assessment in self managing schools. *Educational Research for Policy and Practice*, 1(1–2), 113–125. <https://doi.org/10.1023/A:1021172303647>
- Hipkins, R., & Cameron, M. (2018). *Trends in assessment: An overview of themes in the literature*. New Zealand Council for Educational Research.

- Kelly, A. V. (2009). *The curriculum: Theory and practice* (6th ed.). SAGE Publications.
- Kerr, B. G., & Averill, R. M. (2021). Contextualising assessment within Aotearoa New Zealand: Drawing from mātauranga Māori. *AlterNative: An International Journal of Indigenous Peoples*, 17(2), 236–245. <https://doi.org/10.1177/11771801211016450>
- Kluger, A. N., & DeNisi, A. (1996). The effects of feedback interventions on performance: A historical review, a meta-analysis, and a preliminary feedback intervention theory. *Psychological Bulletin*, 119(2), 254–284. <https://doi.org/10.1037/0033-2909.119.2.254>
- Meissel, K., Meyer, F., Yao, E. S., & Rubie-Davies, C. M. (2017). Subjectivity of teacher judgments: Exploring student characteristics that influence teacher judgments of student ability. *Teaching and Teacher Education*, 65, 48–60. <https://doi.org/10.1016/j.tate.2017.02.021>
- Merriam, S. B. (1998). *Qualitative research and case study applications in education* (2nd ed.). Jossey-Bass Publishers.
- Merriam, S. B., & Tisdell, E. J. (2015). *Qualitative research: A guide to design and implementation* (4th ed.). Wiley.
- Ministerial Advisory Group. (2024). *Initial report: Redesigning the English and mathematics & statistics learning areas in the refreshed New Zealand curriculum for primary, intermediate and early secondary school students: Years 0 to 10*. Ministry of Education.
- Ministry of Education. (2007). *New Zealand curriculum for English-medium teaching and learning in years 1–13*. Learning Media.
- Ministry of Education. (2011). *Ministry of Education position paper: Assessment (schooling Sector)*. Author.
- Ministry of Education. (2023). *Te mātaiaho / The refreshed New Zealand curriculum—draft for testing / March 2023*. Author.
- Mutch, C. (2012). Assessment for, of and as learning: Developing a sustainable assessment culture in New Zealand Schools. *Policy Futures in Education*, 10(4), 374–385. <https://doi.org/10.2304/pfie.2012.10.4.374>
- Mutch, C., & Peung, S. (2021). ‘Maslow before Bloom’: Implementing a caring pedagogy during Covid-19. *New Zealand Journal of Teachers’ Work*, 18(2), 69–90. <https://doi.org/10.24135/teacherswork.v18i2.334>
- Neale, M. D. (1999). *The Neale analysis of reading ability* (3rd ed.). ACER.
- Patton, M. Quinn. (2015). *Qualitative research and evaluation methods: Integrating theory and practice* (4th ed.). SAGE Publications.
- Poskitt, J. M. (2018). Sustaining assessment for learning by valuing partnerships and networks. *Assessment Matters*, 12, 80–104. <https://doi.org/10.18296/am.0030>
- Pryor, J., & Crossouard, B. (2008). A socio-cultural theorisation of formative assessment. *Oxford Review of Education*, 34(1), 1–20. <https://doi.org/10.1080/03054980701476386>
- Rubie-Davies, C. (2014). *Becoming a high expectation teacher: Raising the bar*. Routledge.
- Ruiz-Primo, M. A., & Brookhart, S. M. (2018). *Using feedback to improve learning*. Routledge.

- Sadler, D. R. (2010). Beyond feedback: Developing student capability in complex appraisal. *Assessment & Evaluation in Higher Education*, 35(5), 535–550. <https://doi.org/10.1080/02602930903541015>
- Saldaña, J. (2016). *The coding manual for qualitative researchers* (3rd ed.). SAGE Publications.
- Schellekens, L. H., Bok, H. G. J., de Jong, L. H., van der Schaaf, M. F., Kremer, W. D. J., & van der Vleuten, C. P. M. (2021). A scoping review on the notions of Assessment as Learning (AaL), Assessment for Learning (AfL), and Assessment of Learning (AoL). *Studies in Educational Evaluation*, 71(Article 101094). <https://doi.org/10.1016/j.stueduc.2021.101094>
- Shavelson, R. J., Young, D. B., Ayala, C. C., Brandon, P. R., Furtak, E. M., Ruiz-Primo, M. A., Tomita, M. K., & Yin, Y. (2008). On the impact of curriculum-embedded formative assessment on learning: A collaboration between curriculum and assessment developers. *Applied Measurement in Education*, 21(4), 295–314. <https://doi.org/10.1080/08957340802347647>
- Sinnema, C., Alansari, M., & Turner, H. (2018). *The promise of improvement through and of the Teacher Led Innovation Fund*. Auckland UniServices Limited. <https://www.educationcounts.govt.nz/publications/schooling/evaluation-of-the-teacher-led-innovation-fund-final-report>
- Smith, L. F. (2014). Preparing teachers to use the enabling power of assessment. In C. Wyatt-Smith, V. Klenowski, & P. Colbert (Eds.), *Designing assessment for quality learning: Volume 1* (pp. 303–323). Springer.
- Spencer, R., Pryce, J. M., & Walsh, J. (2014). Philosophical approaches to qualitative research. In P. Leavy, *The Oxford handbook of qualitative research* (pp. 81–98). Oxford University Press.
- Stake, R. E. (1995). *The art of case study research*. SAGE Publications.
- Stanford, E. (2024). *Govt delivers consistency for assessing Kiwi kids*. Beehive.govt.nz. <https://www.beehive.govt.nz/release/govt-delivers-consistency-assessing-kiwi-kids>
- Tai, J., Ajjawi, R., Boud, D., Dawson, P., & Panadero, E. (2018). Developing evaluative judgement: Enabling students to make decisions about the quality of work. *Higher Education*, 76(3), 467–481. <https://doi.org/10.1007/s10734-017-0220-3>
- Te Poutāhū Curriculum Centre. (2023). *Notice, recognise, and respond*. Tāhūrangi—New Zealand Curriculum. <https://newzealandcurriculum.tahurangi.education.govt.nz/notice-recognise-and-respond/5637155863.p>
- Thrupp, M. (2017). New Zealand’s national standards policy: How should we view it a decade on? *Teachers and Curriculum*, 17(1), 11–15. <https://doi.org/10.15663/tandc.v17i1.171>
- Torrance, H., & Pryor, J. (1998). *Investigating formative assessment: Teaching, learning and assessment in the classroom*. Open University Press.
- Wiliam, D. (2018a). *Embedded formative assessment* (2nd ed.). Solution Tree.

- Wiliam, D. (2018b). Feedback: At the heart of—but definitely not all of—formative assessment. In A. A. Lipnevich & J. K. Smith, *The Cambridge handbook of instructional feedback* (pp. 3–28). Cambridge University Press.
- Willis, J., & Klenowski, V. (2018). Classroom assessment practices and teacher learning: An Australian perspective. In H. Jiang & M. F. Hill (Eds.), *Teacher learning with classroom assessment: Perspectives from Asia Pacific* (pp. 19–37). Springer. <https://doi.org/10.1007/978-981-10-9053-0>
- Wylie, C., & MacDonald, J. (2020). *What's happening in our English-medium primary schools: Findings from the NZCER national survey 2019*. New Zealand Council for Educational Research.
- Xu, Y., & Brown, G. T. L. (2016). Teacher assessment literacy in practice: A reconceptualization. *Elsevier Teaching and Teacher Education*, 58(August), 149–162. <https://doi.org/10.1016/j.tate.2016.05.010>
- Zhao, A., Brown, G. T. L., & Meissel, K. (2022). New Zealand students' test-taking motivation: An experimental study examining the effects of stakes. *Assessment in Education: Principles, Policy & Practice*, 29(4), 397–421. <https://doi.org/10.1080/0969594X.2022.2101043>

The author

Thomas C. Pearce (Tom) is a doctoral student at The University of Auckland's Faculty of Education and Social Work. He has worked for many years as a primary school teacher, both in Aotearoa New Zealand and abroad, and at the time of writing is working part-time at a school in Tāmaki Makaurau Auckland. He holds a MEd from Victoria University of Wellington, where he investigated the transformative power of student voice, as well as language and literacy acquisition and issues in sociolinguistics. Tom's current research interests are formative assessment and the enactment of education policy, from critical educational and social constructivist lenses.

Email: tom.pearce@auckland.ac.nz