Moving past the dichotomy: Toward transformative evaluation in early childhood education

Lisa Winiata

If you're an early childhood education professional in Aotearoa, you've likely felt the pressure of evaluations that don't quite fit the reality of your day-to-day work. This commentary cuts through the outdated internal versus external evaluation debate and shows why the current system isn't keeping up with the complexity of today's early childhood education environment. The article highlights why current approaches can fall short and argues for evaluation approaches that are useful, empowering, and grounded in practice and culture.

Introduction

It's Monday morning at a bustling early childhood centre in Aotearoa New Zealand when the message comes in: ERO is coming next month. A quiet panic ripples through the team. Within hours, the centre shifts into high alert. From the back corners of forgotten shelves, long lost frameworks and policies are unearthed like buried treasure. Planning documentation suddenly appears laminated, colour coded, and pinned proudly to the walls. New resources, once considered too expensive or impractical, are hastily ordered and set out with care. The curriculum is dusted off and internal evaluations are freshly written, hastily edited and polished ready for judgement. Behind the scenes, centre leaders scramble. Stress simmers just beneath the surface as kaiako are asked to juggle business as usual with the unspoken expectation to put on a show.

It's a familiar performance and one that begs a critical question: Are our evaluation practices truly helping early learning services grow and improve, or have they become just another high-stakes compliance routine?

This question is particularly pressing in a bicultural context where evaluation must not only drive improvement but also honour Te Tiriti o Waitangi and reflect the diverse realities of tamariki, whānau, and communities.

In Aotearoa early childhood education (ECE), evaluation is intended to serve a dual purpose: To ensure accountability and to foster continuous improvement. At the heart of this system are two key approaches: internal and external evaluation. Internal evaluation, as defined by the Education Review Office (ERO, 2016) is a formative, service-led process used to assess and improve curriculum implementation and teaching practice. While primarily formative, internal evaluation can also include summative elements as services assess their own progress and outcomes. It is intended to help services identify strengths, set goals, and drive context-specific improvement.

External evaluation, by contrast, is a structured, summative process led independently by ERO, typically within a three-year review cycle. Guided by frameworks such as *Ngā Ara Whai Hua Quality Framework for Evaluation and Improvement* (ERO, 2021a) and *Te Ara Poutama Indicators of Quality* (ERO, 2020a), external evaluations assess the value and effectiveness of learning and organisational conditions within ECE services. In 2025, the Government announced a proposed shift of regulatory functions, including licensing and certification of ECE services, from the Ministry of Education to ERO. This change, prompted by the Ministry for Regulation's ECE Regulatory Sector Review (Ministry for Regulation, 2024), aims to streamline oversight, reduce duplication, and modernise the regulatory framework. ERO will take on core regulatory responsibilities such as monitoring and enforcement, while the Ministry retains responsibility for curriculum development. The transition is intended to reduce compliance burdens for providers and create a more coherent system of accountability and support. While some sector leaders welcome the move for its potential to simplify processes and strengthen accountability, others have raised concerns about whether efficiency gains might come at the expense of quality, equity, and adequate resourcing for ERO's expanded role.

Ideally, internal and external evaluation work in tandem: Internal evaluation fosters a culture of ongoing improvement, while external evaluation provides independent oversight and alignment with national priorities. But in practice, this balance is not always achieved. As Moretti (2021) has noted, many ECE services face significant barriers to effective internal evaluation, such as limited time, expertise, and resourcing. These barriers often lead to internal evaluations that are surface level or driven more by compliance than by authentic inquiry.

Despite clear legislative frameworks and decades of ERO reporting, concerns about limited internal evaluation capacity remain widespread. Persistent variability in how services respond to tamariki and support learning raises questions about whether current evaluation practices are truly fit for purpose. This points to a deeper tension at the heart of the system: Evaluation is expected to serve both accountability and improvement, yet these purposes can pull in different directions. When compliance pressures dominate, internal evaluation risks becoming a procedural exercise rather than a tool for genuine learning. External evaluation, while providing independent oversight, cannot fully compensate for these gaps if internal evaluation processes lack depth and authenticity. This interdependence highlights a systemic challenge: Without robust internal evaluation, the potential of external review to drive sustained improvement is constrained.

This commentary examines how internal and external evaluation are currently understood and enacted in ECE. It explores the tension between accountability and improvement, identifies what supports meaningful evaluation, and proposes principles to reimagine evaluation as a transformative tool. To build an argument for change, the commentary is presented through three interrelated themes: Current practice, purpose versus reality, and future possibilities. Central to this argument is the view that compliance-focused models of evaluation are increasingly misaligned with the complexity and diversity of the sector; there is a need for evaluation approaches that are culturally grounded, equity-focused, and contextually relevant.

Current practice

While internal and external evaluation are intended to work together to support quality and accountability in early childhood education, the reality tells a different story. Instead of working as interconnected mechanisms for improvement, internal and external evaluation are treated as separate processes.

ERO (2016) defines internal evaluation as a service-led inquiry process aimed at continuous improvement, yet Moretti's (2021) research reveals that in many services, this process remains largely compliance driven. Drawing on interviews with kaiako across Aotearoa, Moretti found that internal evaluation is often procedural rather than reflective, disconnected from children's learning, and shaped by external accountability pressures. Practical constraints such as limited time, staffing pressures, and variable evaluative capability frequently lead to fragmented or performative evaluation practices. Moretti emphasises that authentic evaluation must be context responsive,

improvement focused, and linked to each ECE service's strategic priorities. Leadership, shared values, and ongoing professional learning and development (PLD) are central to sustaining this work, supported by flexible, strengths-based external guidance.

External evaluation, meanwhile, is often experienced as a short-term, high-stakes event (New Zealand Annual Review of Education, 2022). Although guided by frameworks such as *Ngā Ara Whai Hua* and *Te Ara Poutama*, the format of an external evaluation can reduce it to a snapshot in time. This snapshot may not accurately reflect the day-to-day realities of practice and may quickly become outdated due to staffing or ownership changes. This disconnect can reinforce a compliance mindset, where services focus on preparing for ERO, rather than engaging in sustained, self-driven evaluation for improvement.

ERO itself has acknowledged the perception that external evaluation can feel compliance focused rather than improvement oriented, noting the need for models that strengthen evaluative capability and trust. While this concern is most clearly documented in the school sector through the *Principal Perspectives—Schools: Evaluation for Improvement Approach* report, which found that school principals valued improvement but still felt accountability pressures (ERO, 2021b), similar themes are evident in the ECE sector. ERO (2016) has highlighted that when internal and external evaluation processes are misaligned or insufficiently supported, their potential to drive meaningful improvement is limited. Strengthening alignment and support for evaluation practices is therefore critical to unlocking capacity to foster authentic change.

While ERO (2021a) has signaled an intention to build data literacy and strengthen capability across the ECE sector, this vision has not yet been fully realised. The absence of clear exemplars continues to create uncertainty about the purpose and practice of evaluation. ERO's own reports note challenges in evaluative capability, highlighting the need for ongoing support and clarity. Some of this uncertainty reflects limited visibility of resources and frameworks, and initiatives such as *Akamatua: Evaluating Excellence and Innovation* (ERO, 2020b), designed to showcase excellence through case studies, have yet to be implemented. As a result, the sector lacks concrete examples to guide practice. This is compounded by inconsistent terminology, including language that blurs distinctions between "review" and "evaluation" (Wansbrough, 2004). Without strengthened capability and clearer guidance, sector confidence may be affected. Current practices reveal systemic misalignments, and the gap between evaluation's intended purpose and its lived reality continues to pose challenges, as explored next.

Purpose versus reality

Despite widespread agreement on the importance of evaluation, the way it is experienced in ECE settings often undermines its intended purpose. Instead of driving reflective inquiry or professional growth, evaluation is frequently shaped by external pressure, unclear expectations, and a lack of meaningful engagement. Moretti's (2021) research highlights the tension many kaiako and leaders experience when engaging with evaluation. While most recognise its importance, internal evaluation is often approached as a compliance task completed in preparation for ERO rather than as a tool for growth and improvement. With a compliance mindset, opportunities for genuine reflection and improvement may be lost, and evaluative energy redirected toward performative rather than transformative and innovative functions. The result is an evaluation culture that feels imposed, not owned. More concerning, internal evaluation has the potential to unintentionally deepen inequities by favouring services with the time, stability, and evaluative skills to navigate the system effectively, while those working in more complex communities are left at a disadvantage. Ultimately, the gap between evaluation's purpose as a tool for growth and its reality as an imposed compliance task not only stifles innovation but also perpetuates inequities.

To move forward, the sector must treat evaluation as more than a mechanism for accountability. It must be actively repositioned as a shared, values-based process; one that builds capability, reflects diverse realities, and strengthens the conditions for equitable learning. This means moving beyond the internal–external binary, and investing in systems that are not only rigorous, but also relational, adaptive, and grounded in the lived experiences of communities.

Future direction

To move beyond compliance, evaluation could be reframed as a practical, embedded tool for continuous improvement. To implement this reframing, capacity-building must be tailored to Aotearoa's diverse ECE sector, ensuring all kaiako, leaders, and whānau have equitable access to evaluative tools. Embedding tailored approaches, as outlined in the *Early Learning Action Plan* (ELAP, Ministry of Education, 2019) or ERO guidelines can reduce inequities intensified by marketisation (Mitchell et al., 2025), making evaluative capability a universal enabler and empowering ECE kaiako and leadership to move from compliance to curiosity, enhancing evaluation's transformative potential nationwide. Such reframing requires a shift in mindset from treating evaluation as a

reporting requirement to embracing it as a core part of daily professional inquiry. The ELAP supports this direction by promoting reflective practice, professional capability, and system-wide accountability as foundations for quality improvement. However, despite its promise, the ELAP's implementation appears to have stalled, limiting its influence on sector-wide progress.

In ECE settings, reframing starts with a shared understanding of why evaluation matters, not only for accountability, but for the wellbeing and learning of tamariki. When teams are aligned in purpose, they are more likely to engage deeply with evidence in ways that inform and enhance their practice. This also aligns with the ELAP, which emphasises the importance of sustained and planned professional learning and development, signalling that one-off training is insufficient to embed leadership and practice changes. Ongoing structures for learning and support are required to truly embed evaluative thinking in daily work.

Consistent, high-quality leadership development is a powerful lever for building sector-wide capacity. Sustained professional learning, rather than one-off initiatives, can foster reflective practice and drive meaningful, evidence-informed change across services. Leadership development that prioritises evaluative capability enables leaders to model inquiry, encourage reflection, and integrate evaluation into everyday routines without adding unnecessary workload. It should build on existing practices such as curriculum planning, observations, and whānau engagement to remain relevant and manageable. Research indicates that intensive, ongoing PLD for ECE leaders strengthens capability and confidence, ultimately improving teaching practice and curriculum implementation (McLaughlin et al., 2024). Approaches such as leadership coaching have shown promise, but the broader message is clear: investing in sustained, high-quality leadership development creates the conditions for stronger evaluation and continuous improvement. Future initiatives could also focus on cultural competence in evaluation, ensuring leaders embed Te Whāriki principles and Te Ara Poutama indicators in ways that reflect community aspirations.

Internal transformation cannot occur in isolation from the broader policy environment. Recent reforms risk reframing early childhood education as a market commodity rather than a public good. As Mitchell et al. (2025) caution, the marketisation of ECE can erode public accountability and professional trust, shifting focus from children's rights, relationships, and wellbeing toward efficiency and competition. Weakening licensing and regulatory requirements may further entrench this trend, reducing quality and increasing risks for tamariki. If education and care are treated as products rather than public responsibilities, the very foundations of equity and quality are placed in jeopardy.

Systemic change is therefore essential. ERO's model of external evaluation no longer reflects the complexity, diversity, and relational dynamics of contemporary ECE environments. When external review remains centred on periodic events and compliance (whether intentionally or not), it can undermine the development of evaluative capability it is meant to foster. To be truly effective, external evaluation must also adapt, becoming more collaborative, responsive, and closely aligned with the inquiry processes led by services themselves. As Cousins et al. (2014) argue, building evaluation capacity involves more than technical competence; it requires organisational structures that support the meaningful use of findings.

A key challenge is to support services to do more "meaning-making" and to ensure the system around ECE services enables, rather than hinders, a shift to meaningful evaluation, both internal and external. Meaning-making includes recognising that evaluation does not operate in isolation. Its effectiveness depends on alignment across organisational layers, including governance, leadership, and teaching, where the distribution of evaluative authority and capability significantly influences outcomes. Without structures that support shared decision making and sustained inquiry, efforts to build capacity may remain fragmented.

Embedding culturally grounded evaluation is not an optional extra: It is central to achieving equity and meaningful improvement. Frameworks like Te Whāriki and Te Ara Poutama, alongside ERO's recent reports, provide some guidance on culturally responsive evaluation in practice. Yet, the challenge lies in ensuring these approaches are consistently embedded across all services and supported by system-level structures that honour Te Tiriti and reflect the diversity of our communities. Strengthening culturally grounded evaluation requires more than policy statements; it calls for research and practice that examine how evaluation is interpreted across roles and contexts, and how global methodologies can be adapted, not simply imported, to fit Aotearoa's unique contexts. Evaluation that is relational, place-based, and grounded in local values is essential if it is to drive equity, innovation, and sustained change.

Conclusion

To achieve genuine equity and excellence in early childhood education, there is value in rethinking how the sector approaches evaluation. Evaluation's power lies not in compliance or outcomes alone, but in its capacity to generate positive, sustained improvement. Moving beyond the internal–external dichotomy, evaluation could be

reimagined as deeply embedded across all levels of organisations, fostering continuous dialogue, reflection, and responsive action.

For ERO and the broader sector, this means adopting contemporary evaluation theories that recognise complexity and context, while prioritising capacity building that equips services with practical skills, knowledge, and confidence to evaluate their own practice. Without this shift, external evaluations risk remaining isolated events that fail to catalyse sustained improvement. An evaluation culture that is clear, relevant, meaningful, and tied to shared values and strategic goals could transform how the sector views and uses evaluation. Achieving this is likely to require strong leadership, trust and sustained PLD, supported by system-level structures that enable collaboration and innovation. When evaluation becomes a tool for learning rather than a mechanism for compliance, it can truly enhance outcomes for tamariki and whānau.

References

Cousins, B., Swee, C., Catherine, J., & Elliot, I. (2014). Framing the capacity to do and use evaluation. In D. L. Stufflebeam & C. L. S. Coryn (Eds.), *Evaluation theory, models, and applications* (pp. 45–60). SAGE Publications.

Dickinson, P., & Adams, J. (2012). Building evaluation capability in the public health workforce: Are evaluation training workshops effective and what else is needed? *Evaluation Journal of Australasia, 12*(1), 28–39.

ERO. (2020a). *Te Ara Poutama: Indicators of quality for early childhood education*. https://www.ero.govt.nz/how-ero-reviews/early-childhood-services/akarangi-quality-evaluation/te-ara-poutama-indicators-of-quality-for-early-childhood-education-what-matters

ERO. (2020b). *Akamatua: Evaluating excellence and innovation*. In *Ngā Ara Whai Hua: Quality framework for early childhood services*. Education Gazette. Retrieved from https://gazette.education.govt.nz/articles/nga-ara-whai-hua-quality-framework-for-early-childhood-services/

ERO. (2016). *Effective internal evaluation for improvement: A guide for early childhood services*. https://evidence.ero.govt.nz/documents/effective-internal-evaluation-for-improvement

ERO. (2021a). *Ngā Ara Whai Hua: Quality framework for evaluation and improvement.* https://www.ero.govt.nz/how-ero-reviews/early-childhood-services/nga-ara-whai-hua-quality-framework-for-evaluation-and-improvement

ERO. (2021b). *Principal perspectives—Schools: Evaluation for improvement approach.* https://ero.govt.nz/sites/default/files/2021-09/External%20Evaluation%20of%20the%20School%20Improvement%20Approach%20Phase%203%20Principal%20Perspectives.pdf

McLaughlin, T., Clarke, L. R., Gifkins, V., & Carvalho, L. (2024). *ECE leadership coaching: A valuable investment* [Final report of the Leadership Coaching Programme]. Massey University. https://tewhariki.tahurangi.education.govt.nz/leadership-coaching-case-stories-and-final-evaluation-report/5637307326.p

Ministry for Regulation. (2024). *Regulatory review of early childhood education*. [Final report]. https://www.regulation.govt.nz/regulatory-reviews/early-childhood-education-ece-regulatory-sector-review/

Ministry of Education. (2019). *He taonga te tamaiti—Every child a taonga: Early learning action plan 2019–2029*. https://web-assets.education.govt.nz/s3fs-public/2024-07/Early%20Learning%20Action%20Plan%202019-2029.pdf?VersionId=h4KPtFh7nlSCvJlIeo9IbUx5fzan6YXz

Mitchell, L., Botes, V., & Kamenarac, O. (2025). Early childhood education as a public good: Challenges and possibilities. *Early Childhood Folio, 29*(2), 8–14. https://doi.org/10.18296/ecf.1158

Moretti, E. F. (2021). Exploring what motivates evaluation capacity building in early learning services: "What are the children getting out of this?" *Evaluation Matters: He Take Tū Te Aromatawai, 7*, 71–95.

Wansbrough, D. (2004). Self review: The hot topic. *Early Education INNZ, 35*, 23–30.

AUTHOR PROFILE

Lisa Lorraine Winiata, Kaitiaki Kindergartens

Lisa Winiata holds a Master of Specialist Teaching, endorsed in Early Intervention, and a Master's degree in Evaluation. She works as *Kaiarotake Matua – Evaluative and Specialist Support Lead* for Kaitiaki Kindergartens, where she leads the strengthening of evaluative capability across all levels of the association. Her work focuses on advancing improvement-focused evaluation, equity-driven practice, and culturally responsive systems that align with both national and international priorities in education and evaluation.

Email: Lisawiniata@outlook.com; Lisa@kaitiakikindergartens.org.nz