

To cite this article: Edwards, F. (2022). Novice teachers developing capabilities in the constructive use of summative assessment. *La Revue LEeE*, 6. <https://doi.org/10.48325/rleee.006.03>

NOVICE TEACHERS DEVELOPING CAPABILITIES IN THE CONSTRUCTIVE USE OF SUMMATIVE ASSESSMENT

Frances EDWARDS

Publication version: June 2022

Open and collaborative peer review

Reviewers: Gavin BROWN, Andrew COOMBS

Abstract

Preservice and beginner teachers need to build and develop summative assessment knowledge and skills in order to best gauge and measure student learning, as required by the education system. It is also beneficial if they can use the information they collect through summative assessment and grading to inform pedagogical decision making. The focus of this paper is the nature and development of novice teachers' capabilities to use summative assessment constructively. This paper reports on a study that tracked eight teachers over an 18-month period as they completed an initial teacher education program and then started working as qualified teachers. Findings describe the features of teacher beliefs and capabilities as they moved to being able to use summative assessment and grades to inform adjustments to their teaching programs in constructive ways for the benefit of their students. Fundamental beliefs about assessment, as well as careful reflection, particularly during their early teaching experiences notably influenced their approach to using summative assessment constructively.

Keywords: constructive assessment, summative assessment, preservice teacher, novice teacher

Résumé

Les enseignant·es en formation initiale et débutant·es doivent acquérir et développer des connaissances et des compétences en matière d'évaluation sommative afin d'évaluer au mieux l'apprentissage des élèves. Ils doivent également être en mesure d'utiliser les informations qu'ils recueillent par le biais de l'évaluation sommative et de la notation pour éclairer la prise de décisions pédagogiques. Le présent article porte sur la nature et le développement des capacités des enseignant·es novices à utiliser l'évaluation sommative de manière constructive. Cet article rend compte d'une étude qui a suivi huit enseignant·es sur une période de 18 mois alors qu'ils terminaient un programme de formation initiale et commençaient à travailler comme enseignant·es diplômé·es. Les résultats décrivent les caractéristiques de leurs croyances et capacités alors qu'ils passaient d'une compréhension élémentaire et naïve de l'évaluation

sommative à la capacité d'utiliser l'évaluation et les notes pour informer les ajustements de leurs programmes d'enseignement de manière constructive, au service des apprentissages des élèves. Les croyances fondamentales sur l'évaluation, ainsi qu'une réflexion approfondie, en particulier au cours de leurs premières expériences d'enseignement, ont notamment influencé leur approche de l'utilisation constructive de l'évaluation sommative.

Mots-clés : évaluation constructive, évaluation sommative, enseignant·e en formation, enseignant·e novice

1. Introduction

In this paper I focus on summative assessment, conceptualised by Scriven (1967, 2015) as a role played by evaluation in order to achieve evaluative conclusions. In this sense summative assessment has been likened by Stake (2003) to a guest tasting their soup when it reaches the table, whereas formative assessment is the process of the cook tasting the soup in the kitchen. This means that the difference in summative and formative assessment type is a matter of timing not of format, in that summative assessment occurs at the end of a learning process. For example, such assessment occurs at the end of a teaching unit in which students were learning about the carbon cycle, in order for student knowledge/skill and competence to be judged against success criteria designed to match the learning outcomes of that block of teaching.

Research and reviews of evaluation and assessment policy and practice across a number of countries reveal common patterns and trends in the use of assessment. The Organisation for Economic Co-operation and Development (OECD) reports highlight the increasing focus placed by governments and education policymakers on formal assessment for a range of feedback and accountability purposes (OECD, 2013). This aligns with moves motivated by a “top-down, test-based educational accountability” (Lingard & Lewis, 2016, p. 387) found within education systems around the world. Alongside this focus on externally mandated and managed summative assessment there is an increasing focus on involving classroom teachers in summative assessment of learning including in school leaver qualifications, given that teachers are better placed to design and use a wide range of assessment tasks appropriate for the diversity of learners they teach. Juxtaposed to this is a growing awareness of the benefits of formative assessment, with accompanying promotion of its use.

Recently, an international trend has emerged to involve teachers in national assessment that contributes to qualifications, giving them some control over and more responsibility in the assessment and grading process. In many countries where such assessment was historically managed by external assessment boards (e.g., Scotland, Wales, Northern Ireland, England, New Zealand, Australia) there has been a shift to more assessment and grading being completed by classroom teachers (Crooks, 2011; Gardner et al., 2010; Panizzon & Pegg, 2008). This does not mean that standardized or external testing has disappeared entirely from any of these countries' assessment landscapes, particularly for exit qualifications, but it does mean that a wider range of assessment tasks and practices are now considered as valid and reliable options. Such assessment practice enables more agency on the part of the teacher and a more robust assessment of students' learning (Assessment Reform Group, 2006; Harlen, 2005; Mansell et al., 2009), at least in terms of validity. However it also allows for an increasing range of innovative summative assessment strategies to be used; for example, the incorporation of digital technologies into assessment (Darling-Hammond et al., 2013; Katz & Gorin, 2016; Otrell-Cass et al., 2009;



Venville et al., 2008), the use of novel assessment such as vignettes with assessment, problem based assessment, viva voce, e-portfolios, collaborative assessment (Struyven & Devesa, 2016), computer based tests, simulation-based assessment and games based assessment (Dragow, 2016). The professional judgment of teachers is relied upon in these contexts, usually with systems set in place for quality assurance (Flockton, 2012; Wyatt-Smith et al., 2010), although issues exist around reliable consistent marking and scoring (Locke, 2007; Ogino, 2011; Tewkesbury, 2017). It is argued that an advantage of internal assessment is that it allows demonstration or performance of learning which is not possible in written examination format (e.g., a dance performance). Learning outcomes that relate to such learning can be observed directly, increasing validity of the assessment. Additionally, teachers are potentially better placed to make informed judgements because of their greater knowledge of students, and so the validity and reliability of results is enhanced (Black, 2013).

Another important trend internationally is the increasing prioritization in policy for the use of formative assessment, otherwise known as assessment for learning, by classroom teachers because of its capacity to support learners and raise student achievement (Black & Wiliam, 1998a; Cowie, 2013; National Research Council, 2010; OECD, 2013). A number of jurisdictions around the world now demand that classroom teachers assess students formatively and use the information they gather to better inform their ongoing planning and teaching. A recent review of international standards and measures for teacher assessment literacy standards (DeLuca et al., 2016) found that in early documents (1990–1999) summative assessment was the primary focus, whereas central foci on assessment purposes, assessment processes, communication of results and fairness were evident in standards for assessment literacy from the 1990s until the present time. From 2000 onwards explicit emphasis on formative assessment became evident in a number of countries' assessment standards for teachers.

For teachers, a professional knowledge of assessment and an understanding of its various purposes and practices is essential in this complex context. There are a number of agendas at work in the educational assessment policy arena in the current international environment that contribute to teacher practice, and additional to this there are the effects of teachers' own perceptions about assessment, which, as Bonner (2016) outlines in her review, vary between countries. On one hand, political pressures place a heavy accountability burden on schools and teachers to demonstrate impact, and lead to calls for standardization of assessment tools to allow for easy comparison between students, teachers and schools, for example in the US, the No Child Left Behind (NCLB) and Rise to the top (RttT) policies (Nichols & Harris, 2016; OECD, 2013). Countering this pressing force is the growing body of research evidence pointing to the benefits of teachers being more centrally involved in assessment, and the pivotal role of formative assessment (Black et al, 2011; Dolin et al., 2018; Moss, 2013). Internationally a number of countries are responding by developing standards for teachers that promote support and to account for teacher learning, and that involve the use of both formative and summative assessment by teachers; Ministry of Education policy in Turkey (Büyükkarçı, 2014) and New Zealand (Absolum et al, 2009; Hill & Eyres, 2016) are just two examples.

There have been calls to make better use of the information gained through summative assessment and grading practices (Allal, 2011; Carless, 2011; Laveault & Allal, 2016; Pasquini, 2021). A number of these authors argue that much formal and summative assessment information can be used to provide opportunities for student feedback, with the caveat that this depends on the educational environment and its constraints. The constructive use of summative assessment enables better informed support of further



learning for students as teachers make decisions with a greater range of information (Pasquini, 2021). Their reflections, judgements, decisions and actions are related to the nature of the tasks and learning cycles (Laveault, 2013) and the assessment culture of the school in which they are working (Birenbaum, 2016). Investigation into how teachers can use summative assessment constructively in a range of jurisdictions is a potentially rich area for research, as it has the potential to lead to better understandings of what is possible.

2. Developing teacher assessment practice

Teacher assessment practice and the development of teacher knowledge and skills have been the focus of considerable research over the past 20 years. Studies have documented teachers' development in assessment literacy through initial teacher education (ITE) and subsequent teaching experience (DeLuca & Klinger, 2010; Edwards, 2017; Edwards & Cooper, 2012; Evers, 2014; Graham, 2005; Lyon, 2013; Smith & Lev-Ari, 2005; Smith et al., 2014; Volante & Fazio, 2007). The development of assessment capability is affected by more than the amalgam of knowledge bases they garner throughout ITE. It has been found that teachers hold beliefs and perceptions of assessment and its uses that affect their decision making and practice (Brown, 2008, 2011; Edwards, 2020). The models of learning that teachers hold have been found to influence their assessment task design, practice and use in other contexts (Abell & Siegel 2011; Pelligrino et al., 2001; Penuel & Shepard, 2017; Shepard, 2000). Personal factors such as emotion have also been shown to influence teachers' development of assessment literacy over time (Edwards, 2020).

Tsui (2009) considered the reasons why some teachers become experts and others do not, and found that engagement and experimentation in teaching and learning as well as disposition to engage with tasks that stretch their skill set allowed teachers to extend their competence. The development of expert knowledge through conscious deliberation and reflection has been described as "theorizing practical knowledge and practicalizing theoretical knowledge" (Tsui, 2009, p. 429). This allows for the development of a more flexible and responsive application of professional knowledge, in this case the use of summative assessment for constructive purposes.

Over time preservice teachers have been found to develop increased awareness and understanding of their assessment capability, as well as of the requisite skills of designing and utilizing assessment (Edwards, 2017; Evers, 2014). In Edwards' (2017) study, development of summative assessment literacy in New Zealand secondary teachers was evident and measurable across time in ten dimensions, summarized within the following three categories: knowledge of assessment, understanding the context of assessment, and recognizing the impact of assessment. In work which focused on assessment literacy for New Zealand primary preservice teachers, teachers were found to demonstrate a shift in beliefs and understandings as well as an increase in their confidence from entry to exit from ITE programmes (Hill et al., 2013), although their understandings about giving children agency in this process was still emergent.

Research that focuses on how and why preservice teachers learn about assessment over time identifies key factors that influence a shift towards teachers becoming more assessment capable (Edwards, 2020; Hill & Evers, 2016; Hill et al., 2017). Key influences in helping preservice teachers develop assessment capabilities were firstly, university courses and course assignments (DeLuca & Klinger, 2010; Edwards, 2017; Evers, 2014; Hill et al., 2017). Secondly, practicum was found to be a critical element within ITE that can lead to a



better developed confident and knowledgeable teacher or a less confident teacher, depending on the experience (Edwards, 2017; Smith & Jang, 2011; Smith & Lev-Ari, 2005; Taber et al., 2011; Winterbottom et al., 2008). Through reflecting on their experiences to see links to theory, in much the same way as described by Tsui (2009), preservice teachers are able to develop their assessment capabilities. Edwards (2017) demonstrated, through the use of the SALRubric, that growth in teachers' summative assessment literacy was seen to be particularly evident over practicum placements. Preservice teachers observe their mentors (whose influence is particularly powerful), gain assessment-related feedback, and reflect on their own practice (Edwards, 2020; Edwards & Edwards, 2016; Graham, 2005; Hill & Evers, 2016). A third influence is the preservice teachers' experiences of being assessed themselves, as this was found to cause reflection and development of ideas (Edwards, 2017; Hill et al., 2017).

Other personal and contextual influences found to influence development include beliefs and conceptions, including cultural beliefs (Edwards & Edwards, 2016), affective factors such as emotion (Edwards, 2021; Shoffner, 2009), the sociocultural context reflected in school priorities, embedded practices and so on (Edwards, 2020) and specific relationships with colleagues and lecturers. The importance of the teacher-learner relationship has been acknowledged in all sectors of education including ITE (Giles, 2010; Korthagen, 2010; Loughran, 2013; Shoffner, 2009). Giles (2010) explains the need for teacher educators to be attuned to the "play" of the interactions and relationships they have with student teachers showing phronesis in the way they relate. Overall, these influences can speed up, slow down or direct development in particular ways, causing unpredictable and idiosyncratic implications (Edwards, 2020).

Teachers are often unskilled in the use and analysis of aggregated data (Datnow et al., 2007; Earl & Katz, 2006; Mandinach & Gummer, 2016; Pierce & Chick, 2010; Schildkamp & Kuiper, 2010). Given that expectations are increasing that teachers are data literate, this aspect of their learning and development is also argued to be important (Pierce & Chick, 2011; Schildkamp & Poortman, 2015). Studies have shown that this is an area which requires explicit and ongoing professional development as teachers face challenges to use data to inform their instructional decision making (Brown & Harris, 2009; Edwards & Ogle, 2021; Mandinach & Gummer, 2013; Schildkamp et al., 2013; Unger, 2013).

3. New Zealand context

This study was based in New Zealand, a country in which the education policy framework is distinctive in that the majority of decision making is devolved to self-managing schools. Because of the nature of New Zealand's self-managed schools, there are a number of standardized assessment tools available for teachers to use but no nationally mandated specific assessments, tests or examinations in primary or secondary education. This allows schools to tailor assessment regimes and design specific assessment tasks that suit their students (Brown et al., 2014). Primary and middle schools in New Zealand range from schools with very integrated formative assessment policies to schools that run formal summative assessments at the end of learning units and end of year. However no whole cohorts of students in New Zealand are subject to achievement tests nationally (Crooks, 2010).

New Zealand's assessment approach has been characterized by a high level of trust in schools and teachers' professionalism (Nusche et al., 2012). Teachers are entrusted to teach



and assess and grade in ways that best meet the needs of their students. New Zealand national assessment policy emphasises teachers' use of standards based summative assessment and of formative assessment practices (Crooks, 2002, 2011; Nusche et al., 2012), and considerable investment has been made into the professional development of teachers in both regards. Additionally initial teacher education programs provide support for teachers developing assessment capabilities. The National Certificates of Educational Achievement (NCEA) are national qualifications for senior secondary students. NCEA utilizes an extremely flexible system in which students can be assessed through any combination of internally and externally assessed standards, meaning that for some students most of the assessment that contributes to NCEA is teacher-designed and marked and for others more credit is achieved through external examinations. One further distinctive feature of the NCEA qualification system is its flexibility within internally assessed standards, to ensure fair assessment (NZQA, 2001). Teachers are charged with giving all students the opportunity to provide evidence of the learning, and this may at times involve designing unique assessment tasks that give students the best opportunities to provide evidence of what they know and can do. NZQA provides approval for special assessment conditions so that "entitled candidates can demonstrate their knowledge, skills and understanding, without providing unfair advantage over other candidates" (NZQA, 2016b).

At the senior secondary school level students are able to complete NCEA (levels 1, 2 and 3) qualifications that are offered usually in the final three years of students' secondary schooling, and used as a school leaving credential (e.g., university entry is granted to all students who have achieved NCEA Level 3, plus literacy and numeracy requirements). NCEA is of a formal nature as it is a qualifications-oriented summative assessment. As a qualification NCEA utilizes standards-based summative assessment, in which subjects offer a number of credits bearing standards and teachers are free to choose which standards to assess for any given class or student. Grading for internally assessed students is completed by classroom teachers, and for external assessment teachers are contracted to mark examinations or assess portfolios. Grades are given at four levels: Achieved (A) for a satisfactory performance; Merit (M) for very good performance; Excellence (E) for outstanding performance; and Not achieved (N) if students do not meet the criteria of the standard. At the time of this study slightly more than half of the standards that contribute to NCEA were internally assessed, i.e., classroom teachers assessed and graded student work and a moderation system was used to assure quality and consistency in the assessment.

There are formative characteristics built into NCEA, for example criteria and exemplar rubrics are provided, and for internally assessed standards there is some opportunity for limited feedback to be provided by teachers and peers. Additionally, students who fail individual internally assessed standards are sometimes provided the opportunity for a re-sit. However, the purpose of the assessment tasks is summative in that they are used at the end of a learning sequence to judge whether evidence of learning meets the standard. Externally assessed standards are mostly examined formally once at the end of the year. To assure quality, schools each must implement an internal moderation system. Additional external moderation is under the control of NZQA, although it is rare that data about the moderation quality is published.

For students in their final year of school (Year 13) NCEA is seen as high stakes as success opens doors to further education and career prospects. Year 11 and 12 students may or may not see NCEA as a high stakes assessment overall, but for many students their progress into higher level classes is dependent on a school-set minimum performance in Level 2 NCEA standards; for example, most schools have minimum requirements in Level 2 mathematics for those who want to complete Level 3 calculus. For senior secondary



students, non-NCEA summative assessment that occurs throughout the year includes end of unit tests, essays, portfolios, reports, seminars for stakeholders etc. In the junior levels of secondary schooling (Years 9 and 10), classroom teachers are responsible for all assessment and grading used in their classes. Although there are some standardized tests available, their use is not mandated and the results of these standardized tests are not necessarily used by classroom teachers to inform their teaching or for reporting (Heron et al., 2001). In their study of New Zealand secondary teachers, Irving, Harris, and Peterson (2011) found that the NCEA program dominated teacher thinking about assessment and that it had obvious washback effects into Years 9–10. As mentioned earlier, New Zealand assessment policy also encourages the ongoing use of formative assessment through all levels of education. All assessment is under classroom teachers' control, therefore, in ways not commonly seen elsewhere in the world, New Zealand teachers play a central role in all assessment processes.

New Zealand ITE programs are varied, given that tertiary providers develop their programs individually and these are then approved by the Teaching Council of Aotearoa New Zealand (Teaching Council of Aotearoa New Zealand, 2019). In order to gain approval there needs to be evidence that student teachers meet the *Standards for the Teaching Profession: Nga Paerewa mo te umanga whakaakoranga* (Teaching Council of Aotearoa New Zealand, 2017) in a supported environment. Capabilities and understanding of assessment are an important feature in three of the six Standards. Research on ITE within New Zealand shows that beginner teachers do leave ITE programs understanding a range of assessment principles and processes, ready to assess students (Edwards, 2017; Evers, 2014; Hill et al. 2013). However, most research in New Zealand to date has focused on formative assessment with primary and middle school teachers (e.g., Dixon & Hawe, 2018; Evers, 2014; Hill et al, 2020; Hill et al. 2013; Smith et. al. 2014), with far less attention on senior secondary school, where summative assessment features more strongly.

New Zealand's assessment policy focus on formative assessment has been found to affect teacher conceptions of assessment. When comparing his studies with those from other scholars, Brown (2011) commented that "New Zealand teacher conceptions of assessment reflect the much greater emphasis in New Zealand's educational culture for assessing students to improve learning, even in the secondary school qualifications system" (p. 63). Brown (2011) theorized that "teachers' conceptions of assessment are ecologically rational in that they reflect the legal, cultural or social priorities placed on assessment for their work environment" (p. 65).

The particular ITE program that participants in this study experienced did not have a large focus on assessment within its teaching content. The program was a graduate one-year one. Participants had a one-week focus on general assessment as part of a professional practice paper, and all other assessment content was provided within curriculum papers, and learnt while on practicum. Participants had a common understanding of summative assessment as that used to evaluate cumulative knowledge, completed at the end of a topic/unit or time point (e.g., end of year).

4. Study description

The evidence presented in this paper comes from previous research (Edwards, 2017) which investigated the development of teacher summative assessment literacy through an exploration of the experiences of beginner secondary teachers in New Zealand. A



qualitative approach underpinned by an interpretivist research paradigm was used to investigate summative assessment literacy development in eight beginner secondary teachers. The study tracked the teachers from their entrance to a graduate one-year initial teacher education program until they had completed six months teaching in their first schools. Each teacher was interviewed at five points through this approximately 18-month period. At each interview they were asked to present artifacts that they felt provided evidence of their understanding and use of summative assessment. These interviews helped establish the beginner teachers' understandings of summative assessment and allowed them to reflect on what and how they were learning.

In this paper I am examining evidence from that study as illustrations of novice teachers' developing capabilities in using summative assessment for useful or beneficial purposes i.e., the constructive use of summative assessment. Interview transcripts and documentation that illustrated this teacher practice were utilized. The data and findings from the larger study have been reviewed and reanalyzed for this paper. In particular, I analyzed the data from each teacher to elucidate the nature of teachers' constructive use of summative assessment. I was interested in how teachers made decisions and took action based on the information they gained from assessing their students. I also asked:

- What are the features of the constructive use of assessment evident in novice teachers?
- What elements contribute to novice teachers' constructive use of assessment?

Quotes which illustrate the findings are coded using a pseudonym for each teacher and a code I1 for interview one, I2 for interview two and so on. All arrangements for the conduct of the research was consistent with the University of Waikato's *Ethical Conduct in Human Research and Related Activities Regulations* (2008) including participants' rights to anonymity and their rights to withdraw from the research project.

5. Findings

The findings are presented in three sections. Firstly, there is a description of the participants' views about the primary use of summative assessment. The teacher action that emerged because of this constructive view of summative assessment is described. Secondly, findings describe the use and importance of reflection in teachers' constructive use of summative assessment. Thirdly, contributors to and development of the participants' use of summative assessment over time are identified.

5.1 An improvement-focused view of summative assessment

All novice teachers in this study expressed a similar central conception regarding the use of formal summative assessment: that it should somehow contribute to an improvement in teachers' teaching and students' learning. They emphasized that summative assessment and grading were used as a final measurement for accountability purposes, but that it could fulfill other purposes as well. Their improvement-focused view and responsive stance towards summative assessment meant the teachers felt it was very important to spend time considering the evidence of learning and grades of their students, thinking about how they had taught lessons and how this could be improved, and then taking actions based on their reflections. The effect of this improvement-focus influenced a number of teacher



assessment decisions and actions. Improvement focus as it relates to teaching and learning, and to student readiness for assessment is further explained below.

5.1.1 Improvement-focused on teaching and learning

All teachers in this study believed that they should take actions to improve their teaching, based on their interpretations of summative assessment. Kate, for example, identified this as one of her underpinning beliefs early in the study, and then referred to it again at the end of her ITE program, highlighting the unchanging nature of this belief over the period of the study:

It's a way to gauge students on what they have learnt, what they know at the start, what they have learnt at the end, and then it can relate back to your teaching. Was my teaching, was I good enough? How could I do it differently? Looking at teaching as inquiry, was my learning ... Was I teaching to the wrong learners? The right learners? Which ones did, and which ones didn't? Was my classroom management correct? Et cetera, et cetera ... formulating assessment is to know what the students know or have learnt, and also as a way to direct back at yourself as to what you can change and what you can make sure. [Kate, I1]

It's a way to gauge a student's whereabouts. So it would be gained throughout the course, are they getting what you are saying? Is it common in the classroom...? Are the students all gauging at the same level? Are there are a few students not getting it, a few students getting it? Some getting it better than others? So it is a gauge throughout the teaching and learning, and then it is the final assessment ... It is also a way of reflecting back as teachers ... If the students are all getting Question 1B wrong ... Did I teach 1B or did I forget to? [Kate, I4]

Similarly, Benjamin focused on endeavoring to gain a deep understanding of the students' learning, in order to be able to make changes to teaching plans.

.... to gain an understanding of where they're at and how well the students are doing and the teachers are doing, and I guess from there you can make improvements. I guess that would be the overarching thing, so where to make changes to improve for the teacher, and how to make things improve for the students. [Benjamin, I1]

But I would still go over it on the board with them all afterwards. But actually talking to them helped me know if they understood. [Benjamin, I2]

All novice teachers related stories which demonstrated how these beliefs amplified their actions. They were motivated to adjust or change their program to either re-teach or re-emphasize important content or skills where they identified areas of student misunderstanding, regardless of whether their schools promoted this approach. They also discussed their decisions to speed up or move on to next topics when they were convinced that their students had a good grasp of what they had taught. In these ways, they focused on reflecting on summative assessment to improve learning outcomes for their students, as illustrated in these quotes:

I was actually surprised by how much you can actually learn from the test. You see it and you see what is consistently being got wrong by students. You see where students are using incorrect terms to answer things, and so in that sense, just giving an assessment is not the end of the story. You've got a bit to reflect on, and a bit to think 'where to from here'? Essentially, now that we've got this sort of data, how do I then proceed from that? Rather than just 'they've done that, cool, move on'... That is a good amount of data to



think about, to process, and try to incorporate it into the 'why?' for doing things. [Wiremu, I2]

You have to have assessment to see how well the students are doing. ... for qualifications, for next year and all the rest of it. But I think the main purpose is assessments for teachers to see where the main gaps are, where our teaching could improve, or where the students just need more learning. [Benjamin, I4]

Both of these quotes were from interviews that occurred after Wiremu and Benjamin had been working in examination-oriented schools that encouraged very little reflection on summative assessment results from the final assessment at the end of a series of teaching a topic, when teachers were about to move on. As Wiremu alludes to with "Rather than just 'they've done that, cool, move on'..." However, their beliefs about the importance of an improvement focus amplified their motivation to operate in this way. Five of the participant teachers discussed their need to examine their teaching approaches rather than reteach the topic in the same way. They felt the students' results showed them that the way they had taught the topic originally was not altogether effective:

So it was good to be able to say, well this is a question the students didn't understand so next lesson I will be able to ... Some didn't really get it so we'll go over it a little bit more ... We might change their thought direction and try a different approach. [Kate, I2]

So just being able to alter your teaching to direct what the content is to what their needs are, basically. Just finding ways to approach it from different angles. [Elisa, I4]

The provision of feedback was discussed by most of the participant teachers. They valued the dialogic nature of discussion they could have with students based on the summative assessment evidence. Such feedback and discussion implied an active involvement from both teacher and student. They explained some of the strategies they used to provide this. This included specifically going over test questions, giving praise, showing them exemplars, and clarifying ideas through longer discussions.

We sort of broke it down bit by bit because they were a smaller class and I actually gave them a lot more feedback about their results... Lots of emphasis on the positive rather than 'tick this, tick this, tick this'[Benjamin, I4]

Emma felt that providing feedback straight away was important as it reduced the number of misconceptions students might be holding.

If it's wrong then obviously you're going to have to address that. And so it kind of makes you clarify things as you go rather than figuring out it's wrong, way down the track. [Emma, I1]

The provision of feedback to students after mock-exams was seen by many teachers as a very important element for student learning.

I took the year 12 foundation maths class, and again they only had one practice exam so I spent plenty of time going over that and plenty of build-up for that, and then went over the grades at the end. We sort of broke it down bit by bit because they were a smaller class and I actually gave them a lot more feedback about their results. [Benjamin, I4]



Participant teachers also talked about using aggregated summative assessment outcomes (class-wide NCEA grades across different standards) to inform their actions for the following year, rather than for immediate actions. This was especially the case for end-of-year external NCEA examinations for senior students (Years 11–13), as illustrated in these quotes:

I think it is something that if I was teaching I would look over the mock exams and the end of year exams, if you get to see them, and see where the struggles were. And then make that a priority next year to teach it better. Organize more personal reflection like, obviously that did not get taught as well as I could have, what can I change? What did work? What can I bring and to make it more interesting? Or more relevant for the students to learn? [Mary, I4]

However, early as career teachers they had little experience of using aggregated data, and therefore they did appear less confident and knowledgeable about its use.

The effect of teachers' improvement-focused beliefs was evident in their approach to NCEA. They understood the importance of providing quality feedback for school-designed NCEA summative assessment tasks, as they deemed this would contribute to further learning for NCEA re-sits (for internally assessed standards) and for final examinations (for externally assessed standards). NCEA summative assessment is by nature high stakes, which added its own pressure. In the following quote, Elisa articulates the internal pressure she felt to make changes to her teaching in response to students' NCEA assessments:

So if they haven't met the specific criteria then you have to... obviously, since you've assessed them... you have to figure out how to fix it. So what they haven't figured out. So if you've assessed yourself and what you've taught you have to be able to identify things to change as well. [Elisa, I1]

Another action that followed reflection on summative assessment in some cases was in regard to the design of the assessment tasks that were used, especially when teachers thought that the tasks needed improvement. Over time the teacher participants realized that the tasks used for assessment could help or hinder the students in providing evidence of their knowledge.

I think for more ways to assess, mathematics has to be taught in a different way. So if I teach in a different way I will assess in a different way. [T2, I4]

The freedom teachers had to implement assessment task design that made things more accessible for their students was seen by them all as a strength in the New Zealand education system, however this is potentially challenging for beginning teachers.

5.1.2 Improvement-focus on student readiness for assessment

Being improvement-focused meant participant teachers in this study thought carefully about students' readiness for final assessments. This included their consideration of students' literacy and numeracy levels, students' understanding of the requirements of standards, as well as student motivation, as the teachers saw these all being factors that influenced final summative grades and marks. Elisa used elements of backward design when planning her teaching because she saw the links between gaps in students' science knowledge and skills and their more general literacy and numeracy skills:



And so then we worked backwards from there. To work back to the very basics of stuff that wasn't even in the question but they have to be able to know in order to be able to answer the question. [Elisa, I2]

A number of participant teachers also talked about adding the teaching of assessment and examination techniques to their teaching program. Some took action and taught generic assessment skills in order to help students improve their future summative assessment. The participant teachers recounted a number of strategies they used. In particular, they mentioned the need to help students understand the requirements of specific achievement standards. To this end a number of teachers spent time showing students exemplars of graded work, providing revision lessons tailored to the standard including writing practice, explaining the literacy techniques best employed in various situations, and providing formative tasks that mimicked the requirements of the standards. Practice activities were marked and feedback was provided. One teacher talked about the need to remind students to memorize specific scientific facts. When talking about examination preparation the participant teachers explained the need to cover both content and exam-taking skills. In particular, some of them talked about the use of feedback to alert students to assessment-taking skills that are useful but that were not in evidence in their work. The literacy skills of students were mentioned frequently. For example, vocabulary and writing was seen as a weakness in many students, as illustrated in the following quotes:

I got them to highlight the main points of the questions, the main words. And then for the words they highlight, they have to write any ideas that come to mind, just write the word next to it just all the different words that come to mind. Then they can incorporate that into their answer. That was good. [Bella, I4]

I learnt a lot more about how to prepare the students for it, especially for them achieving well and summative assessments is not about them just knowing the content well, but how to write it. So I tried to focus on how to construct a good answer, and how to plan your answers. [Elisa, I4]

This sort of feedback was seen to help students with the next assessment opportunity they would be given.

One major issue raised by all teachers was the issue of 'teaching to the test' or giving the students too much guidance. Participants identified that such behavior would mean that teachers could not guarantee that students were able to provide sufficient evidence of their learning independently. The internal assessment system used with NCEA does mean that the teachers are also the assessment designers, and markers, and this high trust model does raise issues of credibility in some circles. The teachers in this study were strongly opposed to teaching to the test, although a number of them did witness this practice when they were on practicum. They identified the negative effects of this practice: for example the focus moves away from learning, and students can become demotivated. Some practicum schools seemed assessment-driven and some participant teachers in these schools were scared they would "fall into this trap". All of the teachers spoke against 'teaching to the test' as they thought it was unethical and unfair.

I don't want to teach to an assessment, but I want them to know, the holistic approach I guess, but the stuff I cover has to relate to the assessment in a way. But I don't want to be teaching to it. [Bella, I1]

I just think you have to prepare. Not prepare them for the test but... prepare them for the test but not to the test.. [talking about an osmosis experiment where the classroom



activities were very similar to the actual summative assessment task].... So you kind of do teach to it, but you don't write the discussion for them or anything like that. You just give them pointers, but you can't give too much away. [Bella, I2]

They are given a test and it might be handed in to be marked, and then they are almost like fed the answers. So they might have the working there, but I don't think they really understand they're doing at or the reasoning behind it. And that's not really being assessed, because it is just on a piece of paper. [Benjamin, I2]

Some participant teachers in this study seemed to still be struggling with knowing what was 'over the line' with respect to preparing students for assessment tasks. Most were determined not to fall into the trap of "teaching to the test", but were worried that they might do this subconsciously, given their involvement throughout the entire assessment cycle.

It was interesting to note that all of the specific examples used when teachers talked about preparing students for summative assessments related to senior classes i.e., classes being assessed for the formal qualification NCEA. No mention was made of preparing younger students for assessment tasks.

5.2 The importance of reflection

All participant teachers in this study were highly reflective and, when applied to their assessment experience, this enabled them to integrate their assessment thinking with their pedagogical practice. They did reflect on their assessment practice, and specifically on the answers/evidence of learning produced by students, as well as the grades their students achieved. They saw this act of reflection as an important element in their growth as teachers. They used the assessment results of their students as a gauge to check for their effectiveness when teaching specific content, and to inform adjustments they needed to make to their teaching methods. Here we see teachers talking about what they can gain from this reflection. The following are quotes that illustrate the use of reflection to help participant teachers use summative assessment constructively:

...What areas they struggled in, because that is a good reflection for my own learning as a teacher. [Kate, I2]

It is also a way of reflecting back as teachers... If the students are all getting question 1B wrong... Did I teach 1B or did I forget to?... [Kate, I4]

Teachers spoke of reflecting on students' grades as a type of feedback they were able to use to ascertain their effectiveness as teachers and to help them become better teachers.

...And assessing how they are learning but how you are teaching yourself.... It's just finding out the knowledge and how, and assessing yourself in how the students are actually learning, and what they're learning. So you're kind of assessing the students' ability, but your own ability to teach that certain aspect of the topic. Cause if you're not explaining it right the students just aren't going to get it. So it's a good way of understanding and learning what you are doing, and what the students are actually taking from your class. [Susie, I2]



...So I was just doing internal assessments so it is just getting grades, but it is also how you get feedback. It is kind of your way of getting feedback on your teaching as well. If multiple students aren't answering the questions it's probably not their fault... [Susie, I4]

And I think it's a way of formulating assessment is to know what the students know or have learnt, and also as a way to direct back at yourself as to what you can change and what you can make sure... [Kate, I1]

Like okay you're testing the kids but you're kind of testing yourself at the same time. It's quite amazing... [Susie, I2]

Benjamin explained that he reflected on the responses his students made to assessment tasks, and used this to directly alter his classroom practice.

Like use it to assess where students are struggling, and look at my own teaching. Was I teaching it right? Should I explain that again? That's what I'd use it for... [Benjamin, I2]

He went on to explain his strong beliefs regarding the need for teachers to also be assessed by students, so they can get the feedback they need to improve their practice.

I think it's all good, the teachers assessing the students, but I think that student should assess the teachers just as much almost.... we need feedback. [Benjamin, I2]

Interestingly, in this study the process of reflecting on students' summative assessment products and grades was done almost entirely at an individual level by participant teachers. Only two teachers mentioned that they reflected on the summative assessment results with their mentors and none with their peers. In these two cases, mentors were instrumental in helping the teachers have a more 'balanced' view as to the responsibility that teachers bear with respect to students' results. Overall, the eight participant teachers were found to be more likely to link students' results to their own practice after the experiences of their school-based practicum.

5.3 Influences on the development of the participant teachers' constructive use of summative assessment

Participant teachers' approaches to summative assessment showed sophistication they were able to explicitly discuss the source of these. A number of influences on the development of participant teachers' constructive use of summative assessment were evident.

5.3.1 Personal beliefs and experiences of being assessed

Participant teachers' beliefs influenced the development of their assessment and grading practices. Additionally, all participants related stories of themselves being assessed as students, and they talked about how their decisions were influenced by the emotion they remembered when being assessed. For example, Steve described his performance in formal examinations resulting from rote learning, and how he felt this did not lead to deep learning.

Well, from my own personal experience, mainly exams, and as I say I hated them. I got stressed out, wouldn't sleep, all that stuff. [Steve, 1]



As a result, Steve declared he would not use formal examinations unless it was absolutely necessary. He did not think students or teachers gained anything meaningful with respect to learning from this assessment format. The central belief of summative assessment being improvement-focused was obvious in Steve's and all other teachers' descriptions of their practice. When practices were mentioned that did not align to an improvement focus, participant teachers tended to dismiss them as not being worthwhile.

5.3.2 Practicum and classroom experiences

Over time in this study, and particularly after practicum experiences, teachers talked about their increasing confidence in constructively using summative assessment information. In particular, the experience of teaching blocks of work or a unit that incorporated summative assessment helped them see the links between curriculum, teaching and assessment. Time spent on practicum was identified as a key reason for their increased confidence and understanding the "big picture" of assessment. They started seeing assessment as a more natural part of teaching sequences, and the usefulness of assessment became more obvious to them. When talking about their developing capability the teachers tended to have a focus on informal assessment, as this was obviously at the forefront of their mind. However, their realization that diagnostic, formative and summative all linked together became obvious in the later interviews.

If you just looked at the final assessment as a stand-alone, not thinking about what went on before, it could be quite daunting. It's like if the kids fail this it will be my fault. But by going through the topic and throughout the lessons, by a reflecting process of what have I done and what to do differently to implement that in my next lesson, it's not as daunting because you kind of gauge that they will be fine for it [the assessment], or you will delay the assessment another week because you know they need to do it [learning] a bit more. So it's not really set in stone. You can adjust it. [T4, 12]

The effect of their first practicum experience was obvious, as between the three and six month interviews (i.e., before and after their first practicum experience) there appeared to be a clear shift for almost all participants. They moved from being able to describe the constructive use summative assessment information in general terms to being able to describe much more precisely how they considered summative assessment evidence as an important source of information for teachers, for feedback on their teaching, for feed forward for further teaching and planning. At the end of their initial teacher education program the teachers were able to talk about their summative assessment capabilities and indicate the areas they have identified for improvement more specifically and accurately. Improvement in areas such as marking, strategies useful for assessing deeper learning, and interpreting data, were identified by participants.

5.3.3 Teacher agency and autonomy as learning levers

As part of their classroom experience, participant teachers designed and used assessment tasks, tests and examinations, and then reflected on the effectiveness of these tools for allowing students to best show what they knew and could do. A number of the teachers appreciated the freedom they had to experiment with a range of assessment strategies during the early stages of their teacher education.

I'm glad... that you've got a bit of leeway as to how you want to design it. For example in the senior Biotech one, basically there are these two options, which one do you want to do? And we came up with a report supported with a poster, well it ended up being



just supported with diagrams because it ended up being easier that way. But actually having a bit more freedom [in design]... So it's not just this.... You don't just have to do this... There were more options than I was aware of.

The agency the participating teachers experienced allowed them to try this out and learn from their experiences. Another strategy tried by a number of participants was the use of project-based tasks which allowed a range of knowledge and skills to be assessed through one project. Designing such tasks was motivated by the teachers' desire that students also learn through completing assessment tasks.

Teachers attributed their development to what they learnt from specific lecturers and workshops within their initial teacher education program, as well as to their careful reflection on their observations of students' learning and assessment.

Firstly, university lecturers from the ITE program were identified as contributors to beginner teachers' summative assessment literacy as well as to their general confidence and willingness to take risks when teaching. However, from a large group of possible lecturers it was clear that three lecturers were particularly influential, as these same three lecturers were identified by every beginner teacher in the study. They valued the relationships that these three university lecturers built with them and they talked of the positive ways the lecturers related to them showing a genuine interest in them. Based on this relationship, the content that these lecturers taught and the way they modeled good pedagogy was well received and appeared to create real and memorable impact. These lecturers challenged the teachers' beliefs and practices, and left space for students to puzzle about things, as illustrated here:

...yes but in [lecturer's] class he always told us about how you need to assess like you teach. And I don't think you see that very often in schools. To me it almost feels like cheating, but I don't know whether that's the way it should be, or whether it's just a way to get easy grades. [Mary, I2]

Secondly, mentor teachers were influential. On practicum participant teachers highlighted the advice from mentors which they held in higher regard than advice and feedback from the visiting university lecturers. Mentor teachers were mostly found to be helpful and supportive. However, it was clear from the participant teachers' comments that the levels of trust involved in their relationships with mentors dictated their openness and the sense of support they received. Some mentors closely supervised beginner teachers and gave them little freedom with respect to summative assessment decisions. This type of mentoring relationship was found to be less helpful as the teachers preferred to have the freedom to manage summative assessment for their own classes.

Thirdly, participant teachers identified their classmates as contributors to their summative assessment development. During university class activities, participant teachers found it very fruitful to work collaboratively with their classmates, discussing their knowledge and understanding of summative assessment. They also appreciated learning about each other's practicum experiences and sharing concerns and worries in a non-judgmental space. This sort of discussion and collaborative learning was an important element in their development.

Therefore, the combination of their own beliefs about improvement, freedom to experiment while on practicum, combined with non-judgmental support by lecturers, other teachers and peers appeared to support participant teachers' development over time.



Because they were driven by an improvement-focused view, they tended to look for ways to improve their teaching, rather than accepting that summative assessment and grading were primarily for accountability and judgements. These ideas are discussed in the following section.

6. Discussion

Participant teachers in this study showed an improvement focused approach to summative assessment from an early stage in their development as teachers. Understanding this may inform the support provided for teachers' professional learning in assessment in both ITE and in-service professional development.

Participant teachers were able to identify and talk in detail about a range of purposes for summative assessment with some nuance. The prominent finding from this study was that rather than viewing summative assessment as a final process at the end of a teaching unit, all participant teachers considered that summative assessment enabled teachers to gauge students' learning so that they could use the information to inform their future pedagogical decisions. They reported that summative assessment could inform teachers about how they might best teach by adjusting their teaching, re-teaching topics and/or planning for future changes. This is consistent with other studies that have found that teachers do respond to student grades by considering their instruction (McMillan, 2005; Siegel & Wissehr, 2011; Timperley & Parr, 2005). Given the nature of education policy and practice in New Zealand, which has promoted formative assessment for many years, this finding is not entirely surprising (Brown, 2011), although it has not been widely documented.

Findings show that teachers were committed to using summative assessment in constructive ways very early in their teacher education program, and over time they were able to more articulately and specifically describe what actions could be taken to help student learning occur in response to summative assessment and grading. Their practicum experiences appeared to help them develop a deeper awareness of the complexities of assessment, as described by Edwards (2017). Other New Zealand studies (e.g., Edwards & Cooper, 2012; Evers, 2014) have also described the growing awareness of these complexities, particularly for formative assessment. Teachers' understanding of the nuances of assessment has been found to develop over time (Lyon, 2013) and practicum has been identified as an important factor in teacher growth (Smith & Lev-Ari, 2005; Smith et al., 2014). As has been found previously (Winterbottom et al., 2008; Taber et al., 2011), teachers in this study felt more able to participate in the full role of teacher, including as an assessor, through the knowledge and confidence they gained from teaching experiences as they linked theory to practice (Tsui, 2009).

In their ITE program the participant teachers were exposed to a strongly constructivist tradition whose principles are reflected in formative assessment (Black & Wiliam, 2009) and to which all the beginner teachers in this study subscribed. Research does show that models of learning can influence teachers' assessment practice (Abell & Siegel 2011; Pellegrino et al., 2001; Penuel & Shepard, 2017; Shepard, 2000). This being the case, it would be reasonable to expect that this could contribute to teachers being more likely to view summative assessment as having formative purposes as well as summative purposes. One aspect of this is feedback, which all preservice teachers discussed. They appreciated the feedback they received from lecturers, mentor teachers and colleagues. They also highlighted the need to provide feedback to students within a dialogic discussion, rather



than as a one-way process from teacher to student. This aligns with results from a national survey which found that New Zealand teachers focused strongly on improving learning in their feedback (Brown et al., 2012).

Teachers' concerns about teaching to the test did seem to play a role in their summative assessment practice. The teachers' own observations as well as research evidence (Hume & Coll, 2010; Moeed & Hall, 2011) indicates that teaching to the test can be an issue within the NCEA system in New Zealand, particularly for internally assessed standards given the teacher control of the whole process. This was something that the teachers worked actively to avoid. They were committed to responding to the information gained through summative assessment, but were very aware of the risks that the system involved.

Beginner teachers spoke of the potential of summative assessment having a learning purpose and observed that project-based summative assessment activities caused students to draw together knowledge from a number of domains and synthesize what they had learnt. Using summative assessment to provide students with opportunities to learn provides a counterpoint to the finding that it can be used to enhance teaching. Assessment tools used as a means of supporting learning has been a topic of interest for a considerable time and research shows that assessment can be used successfully in this way (Carless, 2007; Earl, 2013; Torrance, 2007). The use of summative assessment as learning is a useful strategy for the New Zealand context given the extreme time pressure under which some teachers of NCEA classes find themselves, being beneficial for students' learning and taking the focus of the assessment away from it being purely for accountability and reporting.

Interestingly, the beginner teachers in this study also spoke of the formative use of aggregated summative assessment information, although they appeared less knowledgeable about this use. Their reported lack of opportunity to learn about analysis and the use of aggregated summative assessment data echoes findings from research elsewhere in which teachers struggle to make good use of aggregated summative assessment data (Datnow et al., 2007; Mandinach & Gummer, 2016; Pierce & Chick, 2010; Schildkamp & Kuiper, 2010). This need for teachers to become more able to collect and analyze summative assessment data could inform ITE program development, moving forward.

Teachers in this study explained the importance and value of reflection for their own practice. As described by Tsui (2009) the act of linking practical and theoretical knowledge can lead to deeper understanding of the focus of the reflection, in this case the possible actions these teachers may take as a result. Their improvement focus view meant the preservice teachers spent time reflecting on the work and grades of their students as well as on their own teaching, and then taking actions to support student learning based on what they learnt from their reflections. Their focus on examining their own practice rather than reteaching the topic in the same way is in contrast to a much earlier study where New Zealand teachers did not take this view (Brown, 2002). If this is indicative of a change in teacher practice in New Zealand this would indicate a positive shift.

In this study genuine personal encounters between teacher educators, mentor teachers, peers and their students were seen to open up learning spaces and encourage engagement with what was being taught. From the descriptions offered by the participant teachers in this study, it was evident that some lecturers and teachers were particularly skilled in doing this, exemplifying the importance of genuine relationships within the ITE space (Giles, 2010; Korthagen, 2010; Loughran, 2013). These relationships can impact the



affective domain for preservice teachers, which is another important aspect contributing to teacher development (Shoffner, 2009).

In summary the teachers in this study were aware of the constructive uses that could be made of summative assessment and grades, firstly in a more general sense and then, as they gained teaching experience, in a more explicit sense. Through input from others, practice and reflection, they developed a growing awareness of the coherence between learning, curriculum and assessment, and strategies to support learners in more nuanced ways.

7. Implications

Based on the findings of this study, there are a number of implications worth considering, especially for those involved in the development of teachers.

1. Teacher beliefs and their own early assessment experiences as learners can have a profound effect on their assessment practices and decision making. This being the case it is important that these are made explicit early in ITE programs, so that novice teachers can reflect and adjust their thinking.
2. Teachers need to be able to reflect deeply, given the right conditions, and to respond in their planning and teaching because of reflection. For beginner teachers, the provision of schools for practicum that allow teachers to have agency to reflect, articulate and describe actions, provide effective feedback and justify decisions is important.
3. Relationships with open, encouraging lecturers and mentor teachers help beginner teachers establish themselves authentically, building on their beliefs and knowledge to better support their students. Time spent building these professional relationships needs to be valued as it leads to deep dialogue.
4. Teachers are able to develop more sophisticated views and skills for using summative assessment and grading constructively over time. Those responsible for the development of early career teachers need to acknowledge this and encourage growing reflective practice leading to action.

8. Limitations

It is important to consider the limitations of this study alongside the conclusions and implications drawn from it. This was a small qualitative study involving eight beginner secondary teachers who completed a one year initial teacher education program at one university in New Zealand. Its aim was to give rich, in-depth accounts of the constructive use of summative assessment for these teachers. A major issue for this sort of study is generalizability, as findings cannot necessarily be generalized from a qualitative study to apply to a whole population. Nonetheless, readers of the findings of this study who are interested in the development of the constructive use of summative assessment for beginner teachers are able to consider the findings and whether they may or may not be transferred to their own contexts.



References

- Assessment Reform Group. (2006). *The role of teachers in assessment of learning*. Institute of Education, University of London.
- Birenbaum, M. (2016). Assessment culture versus testing culture: The impact on assessment for learning. In D. Lavault & L. Allal (Eds.), *Assessment for learning: Meeting the challenge of implementation* (pp. 275-292). Springer.
- Black, P. (2013). Formative and summative aspects of assessment: Theoretical and research foundations in the context of pedagogy. In J. H. McMillan (Ed.), *Sage handbook of research on classroom assessment* (pp. 167-178). Sage Publications.
- Black, P., Harrison, C., Hodgen, J., Marshal, B., & Serret, N. (2011) Can teachers' summative assessments produce dependable results and also enhance classroom learning? *Assessment in Education: Principles, Policy & Practice*, 18(4), 451-469. <https://doi.org/10.1080/0969594X.2011.557020>
- 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>
- Black, P., & Wiliam, D. (2009). Developing the theory of formative assessment. *Educational Assessment, Evaluation and Accountability*, 21(1), 5-31. <https://doi.org/10.1007/s11092-008-9068-5>
- Brown, G. T. L. (2002). Student beliefs about learning: New Zealand students in year 11. *Academic Exchange Quarterly*, 6(1), 110-114.
- Brown, G. T. L. (2008). *Conceptions of assessment: Understanding what assessment means to teachers and students*. Nova Science Publishers.
- Brown, G. T. L. (2011). Teachers' conceptions of assessment: Comparing primary and secondary teachers in New Zealand. *Assessment Matters*, 3, 45-70. <https://doi.org/10.18296/am.0097>
- 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.
- Brown, G. T. L., Harris, L. R., & Harnett, J. (2012). Teacher beliefs about feedback within an assessment for learning environment: Endorsement of improved learning over student well-being. *Teaching and Teacher Education*, 28(7), 968-978. <https://doi.org/10.1016/j.tate.2012.05.003>
- Brown, G. T. L., Irving, S. E., & Keegan, P. (2014). *An introduction to educational assessment: Measurement & evaluation* (3rd ed.). Dunmore Press.
- Büyükkarcı, K. (2014). Assessment beliefs and practices of language teachers in primary education. *International Journal of Instruction*, 7(1), 107-120.
- Cowie, B. (2013). Assessment in the science classroom: Priorities, practices, and prospects. In J. H. McMillan (Ed.), *Sage handbook on research on classroom assessment* (pp. 473-488). Sage Publications.
- Crooks, T. (2002). Educational assessment in New Zealand schools. *Assessment in Education: Principles, Policy & Practice*, 9(2), 237-253. <https://doi.org/10.1080/0969594022000001959>
- Crooks, T. J. (2010). Classroom assessment in policy context (New Zealand). In B. McGraw, P. Peterson, & E. L. Baker (Eds.), *The international encyclopedia of education* (3rd ed., pp. 443-448). Elsevier.
- Crooks, T. (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>
- Darling-Hammond, L., Herman, J., Pellegrino, J., Abedi, J., Aber, J. L., Baker, E., ... Steele, C. M. (2013). *Criteria for high-quality assessment*. Stanford Center for Opportunity Policy in Education.



- Deneen, C. C. & Brown, G. T. L. (2016). The impact of conceptions of assessment on assessment literacy in a teacher education program, *Cogent Education*, 3(1), 1225380. <https://doi.org/10.1080/2331186X.2016.1225380>
- DeLuca, C., & Klinger, D. A. (2010). Assessment literacy development: Identifying gaps in teacher candidates' learning. *Assessment in Education: Principles, Policy & Practice*, 17(4), 419-438. <https://doi.org/10.1080/0969594X.2010.516643>
- DeLuca, C., LaPointe-McEwan, D., & Luhanga, U. (2016). Teacher assessment literacy: A review of international standards and measures. *Educational Assessment, Evaluation and Accountability*, 28(3), 251-272. <https://doi.org/10.1007/s11092-015-9233-6>
- Carless, D. (2007). Learning-oriented assessment: Conceptual bases and practical implications. *Innovations in Education and Teaching International*, 44(1), 57-66. <https://doi.org/10.1080/14703290601081332>
- Datnow, A., Park, V., & Wohlstetter, P. (2007). *Achieving with data. How-high performing school systems use data to improve instruction for elementary students*. Center on Educational Governance University of California.
- Dixon, H., & Hawe, E. (2018). Developing assessment-capable teachers through engagement in assessment for learning (AfL): a New Zealand study. In H. Jiang & M. F. Hill (Eds.), *Teacher Learning with Classroom Assessment* (pp. 59-76). Springer.
- Dolin, J., Black, P., Harlen, W., & Tiberghien, A. (2018). Exploring Relations Between Formative and Summative Assessment. In J. Dolin & R. Evans (Eds.), *Transforming Assessment* (pp. 53-80). Springer International Publishing.
- Earl, L. M. (2013). *Assessment as learning: Using classroom assessment to maximize student learning* (2nd ed.). Corwin Press.
- Earl, L. M., & Katz, S. (2006). *Leading schools in a data-rich world. Harnessing data for school improvement*. Corwin Press.
- Edwards, F. (2017). *The development of summative assessment literacy: An exploration of the experiences of beginner secondary science teachers in New Zealand* [Doctoral thesis, University of Waikato, New Zealand]. The University of Waikato Research Commons. <https://hdl.handle.net/10289/11567>
- Edwards, F. (2020). The effect of the lens of the teacher on summative assessment decision making: the role of amplifiers and filters. *The Curriculum Journal*, 31(3), 379-397. <https://doi.org/10.1002/curj.4>
- Edwards, F. (2021). The influence of emotion on preservice teachers as they learn to assess student learning. *Australian Journal of Teacher Education (Online)*, 46(4), 16-34. <http://dx.doi.org/10.14221/ajte.2021v46n4.2>
- Edwards, F., & Cooper, B. (2012, August). *Developing assessment capable secondary teachers* [oral presentation]. European Association for Research on Learning and Instruction: Sig1, Brussels, Belgium.
- Edwards, F. C. E., & Edwards, R. J. (2017). A story of culture and teaching: the complexity of teacher identity formation. *The Curriculum Journal*, 28(2), 190-211. <https://doi.org/10.1080/09585176.2016.1232200>
- Edwards, F., & Ogle, D. (2021). Supporting teachers' data informed decision-making: Data informed leadership by mathematics lead teachers in New Zealand. *Teacher Development*, 25(1), 18-36. <https://doi.org/10.1080/13664530.2020.1837217>
- Eyers, G. (2014). *Preservice teachers' assessment learning: Change, development and growth* (Doctoral thesis, University of Auckland, New Zealand). <http://hdl.handle.net/2292/23757>
- Flockton, L. (2012). *The development of the student assessment system in New Zealand*. (SABER - Systems approach for better education results No. 71517). World Bank.



- Gardner, J., Harlen, W., Hayward, L., Stobart, G., & Montgomery, M. (2010). *Developing teacher assessment*. Open University Press.
- Giles, D. (2010). Developing pathic sensibilities: A critical priority for teacher education programmes. *Teaching and Teacher Education*, 26(8), 1151-1519. <https://doi.org/10.1016/j.tate.2010.05.007>
- Graham, P. (2005). Classroom-based assessment: Changing knowledge and practice through preservice teacher education. *Teaching and Teacher Education*, 21(6), 607-621. <https://doi.org/10.1016/j.tate.2005.05.001>
- Harlen, W. (2005). Trusting teachers' judgement: research evidence of the reliability and validity of teachers' assessment used for summative purposes. *Research Papers in Education*, 20(3), 245-270. <https://doi.org/10.1080/02671520500193744>
- Heron, M., Hucker, J., Rooney, C., Robinson, V., Lai, M. K., & Mose, K. (2001). Progressive achievement testing at Southern Cross Middle School. *Set: Research Information for Teachers*, 2, 34-38. <https://doi.org/10.18296/set.0768>
- Hill, M. F., Dixon, H. R., & Hawe, E. (2020). *Observing assessment for learning (AfL) in action: Piloting an observation schedule to inform teacher assessment learning and research*. Teaching & Learning Research Initiative Nāu i Whatu te Kākahu, he Tāniko Taku.
- Hill, M. F., Ell, F. R., & Evers, G. (2017). Assessment Capability and Student Self-regulation: The Challenge of Preparing Teachers. *Frontiers in Education*, 2(21). <https://doi.org/10.3389/feduc.2017.00021>
- Hill, M. F., & Evers, G. (2016). Moving from student to teacher: Changing perspectives about assessment through teacher education. In G. T. L. Brown & L. R. Harris (Eds.), *The Handbook of Human and Social Conditions in Assessment* (pp. 57-76). Routledge.
- Hill, M. F., Smith, L. F., Cowie, B., Gilmore, A., & Gunn, A. (2013). *Preparing initial primary and early childhood teacher education students to use assessment Final Summary Report Key findings Major implications*. Teaching & Learning Research Initiative Nāu i Whatu te Kākahu, he Tāniko Taku.
- Hume, A., & Coll, R. (2010). Authentic student inquiry: The mismatch between the intended curriculum and the student-experienced curriculum. *Research in Science & Technological Education*, 28(1), 43-62. <https://doi.org/10.1080/02635140903513565>
- Irving, S. E., Harris, L. R., & Peterson, E. R. (2011). 'One assessment doesn't serve all the purposes' or does it? New Zealand teachers describe assessment and feedback. *Asia Pacific Education Review*, 12(3), 413-426. <https://doi.org/http://dx.doi.org.ezproxy.waikato.ac.nz/10.1007/s12564-011-9145-1>
- Korthagen, F. A. J. (2010). Situated learning theory and the pedagogy of teacher education: Towards an integrative view of teacher behavior and teacher learning. *Teaching and Teacher Education*, 26(1), 98-106. <https://doi.org/10.1016/j.tate.2009.05.001>
- Laveault, D. (2013). Summative assessment for learning: How it may impact task design [oral presentation]. *5th Conference of the European Association for Research on Learning and Instruction*, Munich, Germany.
- Laveault, D. & Allal, L. (Eds.). (2016). *Assessment for learning: Meeting the challenge of implementation*. Springer International.
- Lingard, B., & Lewis, S. (2016). Globalisation of the Anglo-American approach to top-down, test-based educational accountability. In G. T. L. Brown & L. R. Harris (Eds.), *The Handbook of Human and Social Conditions in Assessment* (pp. 387-403). Routledge.
- Locke, T. (2007). The grail of a reliable, standards-based qualifications system. In T. Locke (Ed.), *Resisting Qualifications Reforms in New Zealand* (pp. 105-123). Brill Sense.
- Loughran, J. (2013). *Developing a pedagogy of teacher education: Understanding teaching and learning about teaching*. Routledge.



- Lyon, E. G. (2013). Learning to assess science in linguistically diverse classrooms: Tracking growth in secondary science preservice teachers' assessment expertise. *Science Education*, 97(3), 442-467. <https://doi.org/10.1002/sce.21059>
- Mandinach, E. B., & Gummer, E. S. (2013). A systemic view of implementing data literacy in educator preparation. *Educational Researcher*, 42(1), 30-37. <https://doi.org/10.3102/0013189X12459803>
- Mandinach, E. B., & Gummer, E. S. (2016). What does it mean for teachers to be data literate: Laying out the skills, knowledge, and dispositions. *Teaching and Teacher Education*, 60, 366-376. <https://doi.org/10.1016/j.tate.2016.07.011>
- Mansell, W., James, M., & Assessment Reform Group. (2009). *Assessment in schools: Fit for purpose? A commentary by the Teaching and Learning Research Programme*. Economic and Social Research Council.
- McMillan, J. H. (2005). *The impact of high-stakes test results on teachers' instructional and classroom practices*. Metropolitan Educational Research Consortium, Virginia Commonwealth University, Richmond, VA. ERIC Document Reproduction Service No. ED490648.
- McMillan, J. H. (2013). Why we need research on classroom assessment. In J. H. McMillan (Ed.), *Sage handbook of research on classroom assessment* (pp. 3-16). Sage Publications.
- Moeed, A., & Hall, C. (2011). Teaching, learning and assessment of science investigation in Year 11: Teachers' response to NCEA. *New Zealand Science Review*, 68(3), 95-102.
- Moss, C. (2013). Research on classroom summative assessment. In J. McMillan (Ed.), *Sage handbook of research on classroom assessment* (pp. 235-256). Sage Publications.
- National Research Council. (2010). *Preparing teachers: Building evidence for sound policy*. National Academies Press.
- New Zealand Qualifications Authority. (2001). *Learning and assessment: A guide to assessment for the national qualifications framework*. New Zealand Qualifications Authority. <http://www.nzqa.govt.nz/assets/Providers-and-partners/Assessment-and-moderation/Assessment-of-standards/learningassessment.pdf>
- New Zealand Qualifications Authority. (2016). *Statement of Intent 2016/17-2019/20*. Wellington, New Zealand: New Zealand Qualifications Authority.
- Nusche, D., Laveault, D., MacBeath, J., & Santiago, P. (2012). *OECD reviews of evaluation and assessment in education: New Zealand, main conclusions*. <http://www.oecd.org/education/school/49681563.pdf>
- Ogino, M. (2011). The challenge of consistency in NCEA examination papers: The case of level 3 Japanese reading over a five year period. *The New Zealand Language Teacher*, 37, 28-38.
- Organisation for Economic Co-operation and Development. (2013). *Synergies for better learning: An international perspective on evaluation and assessment* (OECD reviews of evaluation and assessment). OECD Publishing.
- Otrell-Cass, K., Cowie, B., & Glynn, T. (2009). Connecting science teachers with their Māori students. *Set: Research Information for Teachers*, 2, 34-41. <https://doi.org/10.18296/set.0475>
- Panizzon, D., & Pegg, J. (2008). Assessment practices: Empowering mathematics and science teachers in rural secondary schools to enhance student learning. *International Journal of Science & Mathematics Education*, 6(2), 417-436. <https://doi.org/10.1007/s10763-007-9084-9>
- Pasquini, R. (2021). *Quand la note devient constructive. Évaluer pour certifier et soutenir les apprentissages*. Presses de l'Université Laval.
- Pierce, R., & Chick, H. (2010). Interpreting literacy and numeracy testing reports: What do teachers need to know? In C. Reading (Ed.), *Data and context in statistics education: Towards an evidence-based society*. Proceedings of the Eighth International Conference on Teaching



- Statistics (ICOTS8, July, 2010), Ljubljana, Slovenia. Voorburg, The Netherlands: International Statistical Institute.
- Pierce, R., & Chick, H. (2011). Teachers' intentions to use national literacy and numeracy assessment data: A pilot study. *Australian Educational Researcher*, 38(4), 433-447. <https://doi.org/10.1007/s13384-011-0040-x>
- Schildkamp, K., & Kuiper, W. (2010). Data-informed curriculum reform: Which data, what purposes, and promoting and hindering factors. *Teaching and Teacher Education*, 26, 482-496. <https://doi.org/10.1016/j.tate.2009.06.007>
- Schildkamp, K., & Poortman, C. L. (2015). Factors influencing the functioning of data teams. *Teachers College Record*, 117(4), 1-42. <https://doi.org/10.1177/016146811511700403>
- Shoffner, M. (2009). The place of the personal: Exploring the affective domain through reflection in teacher preparation. *Teaching and Teacher Education*, 25(6), 783-789. <https://doi.org/10.1016/j.tate.2008.11.012>
- Siegel, M. A., & Wissehr, C. (2011). Preparing for the plunge: Preservice teachers' assessment literacy. *Journal of Science Teacher Education*, 22(4), 371-391. <https://doi.org/10.1007/s10972-011-9231-6>
- Smith, L. F., Hill, M. F., Cowie, B., & Gilmore, A. (2014). Preparing teachers to use the enabling power of assessment. In Claire Wyatt-Smith, V. Klenowski, & P. Colbert (Eds.), *Designing assessment for quality learning* (pp. 303-323). Springer.
- Smith, D. C., & Jang, S. (2011). Pathways in learning to teach elementary science: Navigating contexts, roles, affordances and constraints. *Journal of Science Teacher Education*, 22(8), 745-768. <https://doi.org/10.1007/s10972-011-9257-9>
- Smith, K., & Lev-Ari, L. (2005). The place of the practicum in pre-service teacher education: The voice of the students. *Asia-Pacific Journal of Teacher Education*, 33(3), 289-302. <https://doi.org/10.1080/13598660500286333>
- Stake, R. E. (Ed.). (2003). *Standards-based and responsive evaluation*. Sage publications.
- Taber, K. S., Riga, F., Brindley, S., Winterbottom, M., Finney, J., & Fisher, L. G. (2011). Formative conceptions of assessment: Trainee teachers' thinking about assessment issues in English secondary schools. *Teacher Development*, 15(2), 171-186. <https://doi.org/10.1080/13664530.2011.571500>
- Teaching Council of Aotearoa New Zealand. (2017). *Standards for the Teaching Profession: Nga Paerewa mo te umanga whakaakoranga*. Teaching Council of Aotearoa New Zealand.
- Teaching Council of Aotearoa New Zealand. (2019). ITE Programme approval, monitoring and review requirements. Teaching Council of Aotearoa New Zealand.
- Tewkesbury, M. (2017). *An analysis of National Certificate of Educational Achievement chemistry courses in terms of curriculum, pedagogy and assessment: With comparison to the International Baccalaureate Diploma* [Doctoral thesis, Victoria University of Wellington]. <http://hdl.handle.net/10063/6610>
- Timperley, H., & Parr, J. (2004). *Using evidence in teaching practice: Implications for professional learning*. Hodder Moa Beckett.
- Torrance, H. (2007). Assessment as learning? How the use of explicit learning objectives, assessment criteria and feedback in post-secondary education and training can come to dominate learning. *Assessment in Education: Principles, Policy & Practice*, 14(3), 281-294. <https://doi.org/10.1080/09695940701591867>
- Tsui, A. B. M. (2009). Distinctive qualities of expert teachers. *Teachers and Teaching*, 15(4), 421-439. <https://doi.org/10.1080/13540600903057179>
- Unger, J. (2013). Flex your school's data muscles: Leadership strategies strengthen data's impact. *Learning Forward*, 34(4), 50-54.



- Venville, G., Sheffield, R., Rennie, L. J., & Wallace, J. (2008). The writing on the wall: Classroom context, curriculum implementation, and student learning in integrated, community-based science projects. *Journal of Research in Science Teaching*, 45(8), 857-880. <https://doi.org/10.1002/tea.20245>
- Volante, L., & Fazio, X. (2007). Exploring teacher candidates' assessment literacy: Implications for teacher education reform and professional development. *Canadian Journal of Education*, 30(3), 749-770. <https://doi.org/10.2307/20466661>
- Winterbottom, M., Taber, K. S., Brindley, S., Fisher, L. G., Finney, J., & Riga, F. (2008). Understanding differences in trainee teachers' values and practice in relation to assessment. *Teacher Development*, 12(1), 15-35. <https://doi.org/10.1080/13664530701827723>
- Wyatt-Smith, C., Klenowski, V., & Gunn, S. (2010). The centrality of teachers' judgements practice in assessment: A study of standards in moderation. *Assessment in Education: Principles, Policy & Practice*, 17(1), 59-75. <https://doi.org/10.1080/09695940903565610>

Frances Edwards

Frances Edwards (PhD) is Senior Lecturer at the University of Waikato, in New Zealand. Her research interests focus on Assessment, Teacher education, Secondary education, Assessment literacy, Pacific education, Curriculum and Science education.

