

Assessing ELLs in New Zealand primary schools: Gaps between the literature, policy, and practice

Sue Edwards

Waikato Institute of Technology

Abstract

In February 2015, all New Zealand schools moved to assessing English Language Learners (ELLs) using the English Language Learning Progressions (ELLP) to determine eligibility for additional funding to support these learners. This paper firstly provides the background to the current assessment situation, and summarises the literature regarding key principles of assessment. It then describes key guidelines made available to schools by the Ministry of Education for using the new assessment system, particularly the use of Overall Teacher Judgements (OTJs). The paper then presents findings from interviews with three primary school English language specialist teachers regarding their experiences with using the new system, known as 'ELLP assessment'. The gaps that exist between the literature, Ministry guidelines, and ESOL teacher practice are described, and recommendations are made for bridging these gaps. Currently little is known regarding teacher practice in regard to ELLP assessment, so this study fills a gap in the literature relating to the assessment of young ELLs in the New Zealand context.

Key Words: English language learners, Primary, ELLP assessment, Overall Teacher Judgements (OTJs)

Introduction

The steady increase in the number of English language learners (ELLs) in the New Zealand primary school population over the last two decades has also resulted in greater awareness of the need to adequately cater for their learning needs. As a result, the Ministry provides additional funding for schools to support the teaching and learning of ELLs who fall below given benchmarks. In order for schools to receive the funding, teachers are now required to make Overall Teacher Judgements (OTJs) about learners'

language skills in relation to the English Language Learning Progressions (ELLP) document (Ministry of Education, 2008), a process which is known as ‘ELLP assessment.’

This paper begins by providing a summary of ELLP assessment, and also the wider assessment context, including the use of OTJs. It then summarises key concepts from the assessment literature, followed by a description of Ministry of Education ‘policy’, or guidelines regarding assessment in general, OTJs and ELLP assessment. The paper then reports on findings from a recent study of assessment practices of English language specialist teachers, known as ESOL (English for Speakers of Other Languages) teachers, and compares these practices with statements in the literature and with Ministry expectations in regard to ELLP assessment. This is followed by a summary of the apparent gaps between the literature, Ministry policy, and teacher practice. The paper concludes with recommendations for ways in which these gaps might be closed.

What is ELLP assessment?

English language learners in New Zealand primary schools are assessed for a number of reasons. They are usually assessed on entry to a school, to gather information about their general English proficiency. As part of the New Zealand school system, ELLs must be assessed against National Standards in Literacy and Numeracy, and in curriculum areas. ELLs are also assessed twice yearly against the English Language Learning Progressions (ELLP). This document was originally written to “help teachers to choose content, vocabulary, and tasks that are appropriate to each learner's age, stage, and language-learning needs” (Ministry of Education, 2008, p. 2). However, schools were required from the beginning of 2015 to also use the ELLP to rate ELLs’ English language skills, in order to apply for ‘ESOL funding’. This funding is “available to schools for the provision of English language support for migrant and refugee background students with the highest English language learning needs” (Ministry of Education, 2014). Support usually takes the form of assistance from teacher aides in mainstream classrooms, or withdrawal classes with ESOL teachers.

The assessment of ELLs to determine funding eligibility can therefore be regarded as high-stakes assessment, as learners may or may not receive additional support

depending on their assessed level, and schools may or may not be able to provide support, depending on the funding received. McKay (2006) states that ‘high-stakes’ decisions are those that “are likely to affect students’ lives and decisions which are difficult to correct.” (p. 20). She also notes: “Many assessment procedures are more high-stakes for students than we think, since many decisions that teachers and schools make have a cumulative effect on students’ futures” (p. 20).

The wider assessment context: Standards and OTJs

Beginning with trials in 2013 and become mandatory in 2015, the move from an earlier ‘comparison with cohort’ assessment process to the current ELLP assessment has come about in the wider context of the introduction of literacy and numeracy standards in New Zealand schools (Ministry of Education, 2010). As Poskitt & Mitchell (2012) state, “Critical to the implementation of National Standards in New Zealand is the notion of standards and the centrality of the OTJ” (p. 54). Klenowski and Wyatt-Smith (2010) note that “the word *standard* is ubiquitous yet difficult to define”, and that a distinction needs to be made between content standards, referring to “the knowledge and/or processes that are taught”, and achievement standards, which is seen to “apply to students, and refer to what they have learnt” (p. 109-110).

The Ministry of Education (n.d.a) state that an OTJ “involves drawing on and applying the evidence gathered up to a particular point in time in order to make an overall judgment about a student’s progress and achievement.” Klenowski & Wyatt-Smith (2010) point out that teacher judgements “can be made dependable if standards are promulgated in appropriate forms and teachers have the requisite conceptual tools and professional training” (p. 113). Concerns about OTJs have also been summarised by Poskitt & Mitchell (2012), who note that OTJs can be problematic unless teachers are clear about what constitutes an OTJ, teachers have common understandings of standards, such understandings are supported by clear criteria and exemplars of student work, and teachers engage in moderation processes (p. 61).

Key assessment concepts

A central concept found in the literature is that the ultimate purpose of assessment is to improve learning (e.g. Fairtest, 2009), and this is often described as ‘Assessment for Learning’. The literature also distinguishes between assessment for formative and for summative purposes. The former refers to occasions when information about student learning “is collected during teaching”, that is, “while the student’s language skills are being formed” (Richards, 2015, p. 676). Formative assessment also “leads to feedback that is used by students to improve their learning” (Richards, p. 677). The latter refers to assessment which “measures the product of a student’s learning” (Harmer, 2015, p. 408), and is “given at the end of a learning period” (Murray & Christison, 2011, p. 181). Assessment for formative purposes is often closely associated with Assessment for Learning, and the terms are often used interchangeably (e.g. Education Services Australia, n.d.). However, it has been argued that assessment designated as ‘summative’ can also be used formatively (e.g. Darr, 2011). As Poskitt and Mitchell (2012) note, “New Zealand values the central role of formative assessment in improving learning and teaching, and the professionalism of its teachers” (p. 55).

Another distinction is made between formal assessment, which involves systematic, planned sampling techniques, and informal assessment, which is conducted as part of classroom activities (Brown & Abeywickrama, 2010, p. 6). There has been increasing recognition of the important role that informal assessment plays in teaching and learning. McKay (2006) points out: “Many assessment procedures for younger learners are embedded in classroom teaching” (p. 145), and these often include informal strategies such as ‘incidental observation’ and ‘on-the-run assessment’, or informal, instruction-embedded assessment. However, McKay notes that formal or planned assessment activities can and should also be used in the classroom, including strategies such as planned observation, conferences, self-assessment, and classroom tests, and also that “Keeping records is an integral part of classroom assessment” (p. 169).

Underlying all assessment decisions, a number of principles are referred to in the literature, including Validity, Reliability, Practicality, Authenticity, Fairness, Washback, Interactiveness, and Impact, and others. (e.g. Bachman & Palmer, 1996; Brown & Abeywickrama, 2010; Farhady, 2012). Although these principles may be prioritised differently, three are most commonly identified as being the most important. For example, Brown (2001) states:

“If in your language teaching you can attend to the practicality, reliability and validity of tests of language, whether those tests are classroom tests...or final exams, or proficiency tests, then you are well on your way to making accurate judgements about the competence of the learners with whom you are working” (p. 389).

Ministry of Education assessment guidelines

In line with the wider assessment literature, the Ministry of Education (n.d.b) state on their ‘Assessment online’ website that “the primary purpose of assessment is to improve students’ learning and teachers’ teaching as both respond to the information it provides.” Similarly, there is ample information for teachers on the web site regarding formative and summative assessment, with a clear focus on formative assessment. For example, the Ministry states: “A good teacher practises formative assessment constantly on an informal basis through classroom observation and interaction.” Further advice is that “both the teacher and the student will gain information from the assessment and use it collaboratively to plan future learning activities.” Regarding assessment principles, the Ministry of Education (2005) states that “All assessment tools and processes ... should be reviewed against three criteria – validity, reliability and usability” where usability is defined as “the extent to which an assessment tool is practical and yields results that users can easily understand, interpret, and make generalisations from” (p.10).

Clear statements and guidelines have also been provided by the Ministry of Education (2015) relating to ELLP assessment. Three key guidelines are as follows: Teachers will “use a wide range of assessment tasks, activities and observations to make an OTJ (overall teacher judgment) with reference to the various descriptors on the ELLP matrices”; these tasks will... “include formative and summative assessments, standardised tests and both formal and informal observations”; and the process of formulating an OTJ is... “based on your school’s usual age-appropriate assessment tools, activities, and observations”, and “should not be seen as additional to the school’s normal assessment schedule but as an integral part of it.”

The Ministry has provided a number of resources to assist teachers to complete the ELLP assessment. Workshops for ESOL teachers were held in 2013 and 2014 to enable trialling of the new system, and it was expected that ESOL teachers would conduct professional development for mainstream teachers. An online professional learning module has been provided which gives instructions for completing the ELLP assessment. There is also ongoing support for teachers through an email discussion forum. However, these resources are located on the ESOL Online website (Ministry of Education, n.d.c), which may not be known to mainstream teachers.

The current study

Participants

In the second half of 2015, the author carried out a small qualitative study, using semi-structured interviews, investigating three ESOL teachers' practices and thinking regarding ELLP assessment. As seen in Table 1, below, the participants were all experienced ESOL teachers, and they were responsible for reasonable numbers of ELLs. Two of the teachers reported that they continue to provide ESOL support to ELLs whose funding allocation (either three or five years depending on learner backgrounds) has ended. The teachers' role in ELLP assessment varied slightly, with Teacher A carrying this out together with mainstream teachers, Teacher B proactively assisting them (e.g. suggesting suitable assessment activities, offering assistance) and Teacher C, who was happy to collaborate with mainstream teachers if required.

Table 1: ESOL teacher participants

	Teacher A	Teacher B	Teacher C
Years of ESOL experience	16	20	29
No. of funded ELLs 2015	30	50	67
Role in the assessment of ELLs for ESOL funding	Assesses ELLs together with	Proactively assists	Collaborates with mainstream

	mainstream teachers.	mainstream teachers.	teachers as required.
--	----------------------	----------------------	-----------------------

The teachers reported that they preferred ELLP assessment to the previous cohort-based assessment, for two main reasons: firstly, because mainstream teachers learn more about their learners than they did in the past, as they now have the primary responsibility for completing the assessment (in the past the ESOL teacher completed the funding assessment); secondly, completing the assessment helps teachers focus on the next teaching and learning steps for their ELLs. The following comments were made:

- “The new system gives us a model for pulling the teachers in. The huge advantage is we’re sharing information.”
- “It’s not so much the document, it’s more involving mainstream teachers more with the assessment. The impact for these students on teaching is surely going to be far more positive in terms of knowing those learners and how we go from there with that knowledge.”
- “I think this system helps you plan more for ‘where to next’”

When asked whether they had completed any professional development regarding ELLP assessment, all three teachers reported that they had attended a workshop in 2013 about using the ELLP for funding eligibility, but that this workshop had not provided guidance about the number and type of assessments to use to arrive at an OTJ, or how to form an OTJ; it had focused only on the scoring system for ELLP assessment. When asked if they were familiar with the online professional learning module for completing ELLP assessment, one of the teachers reported that she had used it when leading professional development for mainstream staff in her school. However, none was using the ESOL Online website regularly; one teacher stated that it was “difficult to navigate through,” and another reports: “Whenever I go there, it’s not very satisfying”.

Research questions

The key questions below, accompanied by appropriate follow-up or probing questions, were asked in the interviews, which took just over an hour. The questions reflect the efforts of the researcher to uncover teachers’ practices in relation to key assessment

concepts found in the literature and in Ministry policy summarised above i.e. teachers' knowledge or awareness of the purposes of assessment and assessment principles, as well as the range and types of assessment measures used for ELLP assessment.

- 1) How do you see the purposes of ELLP assessment?
- 2) Which assessment principles do you take into account or recommend when planning and delivering ELLP assessment?
- 3) Which assessment measures do you use or recommend for ELLP assessment?
- 4) What are the positive aspects of using the ELLP for assessment?
- 5) What are the challenges with using the ELLP for assessment?

Teachers' responses to the questions above are grouped in the Findings below into three areas, corresponding to the first three questions. Teachers' responses to the last two questions are also included under these headings, as appropriate.

Findings

1) Teachers' perceptions of the purposes of ELLP assessment

Teachers were presented with a list of possible purposes for assessment, and were asked to say which were relevant to them when completing ELLP assessment. As can be seen from table 2, below, they agreed that they are rating overall language proficiency as well as the specific language skills. However, they saw the other possible assessment purposes slightly differently.

As seen in Table 2, none of the teachers were of the opinion that ELLP assessment has a formative purpose. This may be because they see the ELLP assessment as informing their own teaching (e.g. for placement, monitoring or diagnosis), but not something that would enable them to give feedback to a learner about their next learning steps. Teacher C commented: "I don't think ELLP [assessment] is used to give student feedback." Teacher B's comment reveals another reason for this view: "I see more formative as what I'm doing during the lesson...when I'm observing how they're managing."

Table 2: Teachers' perceptions of the purposes of ELLP assessment

	Teacher A	Teacher B	Teacher C
Rating of overall language proficiency	✓	✓	✓
Rating of specific language skills	✓	✓	✓
Monitoring/progress	✓	–	✓
Placement	✓	–	✓
Diagnostic	–	✓	✓
Summative/ Achievement		✓	✓
Formative	–	–	–

2) Teachers' awareness of assessment principles

When interviewing the teachers, the phrase 'assessment principles' was not used, to avoid the implication that teachers were expected to know a set of principles. Instead, teachers were asked what 'important considerations' were for them as they planned and delivered ELLP assessment. Their comments indicate that they are aware of the core assessment principles of validity, reliability and practicality, although only one of the teachers used these words. Table 3, below, shows teacher comments relating to these principles, based on features of assessment principles described in Brown and Abeywickrama (2010).

Table 3: Teachers' awareness of assessment principles when planning and delivering ELLP assessment.

	Comments relating to validity	Comments relating to Reliability	Comments relating to Practicality
Teacher A	<p>"Must be needs-based."</p> <p>"Must use effective tools e.g. videos"</p> <p>"Must tap into what</p>		<p>"Should be able to be done quickly."</p> <p>"Should be done after other assessments."</p>

	<p>teachers are already doing.”</p> <p>“Focus on assessment of academic, not social, language.”</p>		
Teacher B	<p>“Does it tell us what we’re wanting to know?”</p> <p>“Ask the right questions.”</p>	<p>“Is it reliable – does it give the same results as other classroom-based assessment, relative to cohort?”</p> <p>“Strive for consistency by the assessor.”</p> <p>“Assess ELLs away from the mainstream class.”</p>	<p>“Must be manageable for the classroom teacher, in terms of time.”</p> <p>“Is it part of what teachers already do?”</p>
Teacher C	<p>“Start with prior knowledge about learners.”</p> <p>“Select assessments known to be useful.”</p> <p>“Eliminate assessments that are too difficult.”</p> <p>“Make the assessment within the reach of the learner.”</p>	<p>“Ensure students won’t be stressed by assessment.”</p> <p>“Give encouragement and positive feedback for all attempts.”</p> <p>“Provide a quiet, private environment.”</p> <p>“Turn assessment into a game.”</p> <p>“Not in front of their peers.”</p>	<p>“Don’t plan to assess too much in one session.”</p>

When teachers were asked about challenges encountered with ELLP assessment, all three teachers reported a concern about the consistency of mainstream teacher judgements i.e. rater reliability (Brown and Abeywickrama, 2010), as reflected in the following comments:

- “There might be different interpretations – it’s not standard across the school.”
- “I don’t trust the teachers... they’d have everyone on Stage 2.”
- “Mainstream teachers rate too highly, in general.”

Another concern for the teachers, also relating to rater reliability, was the challenges that mainstream teachers may encounter with the descriptors in the ELLP document, which serve as the ‘benchmarks’ against which ELLs are rated. One reported that even as an ESOL teacher she had “had to learn how to interpret the descriptors.” Other comments indicating potential sources of rater unreliability connected to the descriptors included:

- “Some of the descriptors are confusing for mainstream teachers.”
- “Some things seem to be positive and some things seem to be negative – it doesn’t make sense to me.”
- “The descriptors on the matrices I don’t think are thorough - I guess they wanted to make it user friendly and not too onerous, but in a way maybe that’s made it hard to define between one stage and another.”

Also related to reliability is the process of moderation, which the literature indicates is essential for consistency of teacher judgements. All three teachers reported that moderation was being conducted only very informally in their schools, for example if a mainstream teacher was unsure about a judgement and approached the ESOL teacher for guidance.

3) Assessment activities used or recommended for ELLP assessment

a) Types of assessment activities

Asked whether they use or recommend formal or informal assessment activities for ELLP assessment, teachers indicated that a combination of these are used. They reported that the more formal mainstream assessment activities of ‘running records’ and ‘writing exemplars’ produced for National Standards are being used for OTJs about

ELLs' reading and writing skills, whereas informal observations and conversations are the basis for teachers' judgements about listening and speaking skills. Their comments included:

- “A lot of schools... are relying on the running record, and for the writing, we've got the writing exemplars”
- “Where it might be open would be assessing Listening and Speaking...I think teachers tend to rely on observation in the classroom.”
- “The Listening is the trickiest one, and it's all guesswork really.”

b) Range of assessment activities

In order to ascertain whether ‘a wide range’ of assessment activities was being used for ELLP assessment, teachers were asked to list the assessment activities that they use, or recommend that mainstream teachers use. They were also asked whether they preferred sourcing their own assessment tasks (there are no prescribed assessment activities for ELLP assessment), or being provided with a list of suggested assessment activities (as was the case with the previous system). One teacher commented: “I remember this feeling of reinventing the wheel... why didn't they (MOE) trial tools and recommend tools?” Another offered the opinion that it was “a cop-out” on the part of the Ministry of Education.

As seen in Table 4, below, there was a relatively short list of assessment activities provided by the teachers. They reported that mainstream teachers also rely on informal classroom observations of ELLs to arrive at their OTJs, particularly for Listening and Speaking. Although two teachers suggested that teachers conduct an oral interview as evidence of an ELL's listening and speaking skills, the third reported that in her school these OTJs were based solely on classroom observations. All three were concerned about the assessment of listening and speaking skills because they are not assessed in mainstream classes. One teacher commented: “It's the Oral [language assessment] - that's the tricky one”, another asking “What do we (schools) do for listening and speaking?” (the implied answer being “very little or nothing”).

Teachers also reported that one writing sample would likely be the main evidence of ELLs' writing skills, and this would usually be a sample completed for National

Standards assessment. When asked if one sample was enough, one of the teachers responded: “Well it has to be. If I was to do an OTJ – what would I be doing – three samples for (x) kids, no, no...” This comment perhaps indicates that although she is aware that several samples of language are preferable for an OTJ, neither she nor the mainstream teachers have time to do a ‘proper’ OTJ.

Table 4: Assessment activities used or recommended for ELLP assessment

	Teacher A	Teacher B	Teacher C
Listening	Oral Interview on curriculum topic, video-recorded	Oral Interview on curriculum topic, video-recorded	Teacher observations
Speaking			
Reading	Running Records; match with ‘PM Reader’ levels	Running records; match with ‘Ready to Read’ levels	Running records; other assessments done for National Standards
Writing	Independent Writing sample, mainstream class	Unassisted writing done for English curriculum	Unassisted writing sample/s

Conclusions: Gaps between the literature, policy, and practice

The findings indicate that there may be a number of gaps between the literature and teacher practice in regard to ELLP assessment, as well as between Ministry policy and teacher practice. However, as the data obtained from the study is limited, and cannot be generalized, the gaps described below are best thought of as tentative conclusions, with further research needed.

Firstly, while the literature indicates that the main purpose of assessment is to improve learning, it appears that the teachers do not see ELLP assessment as having a formative purpose. While they acknowledge that mainstream teachers will find out more about their ELLs from completing the assessment, which may indirectly lead to improved teaching and learning, there seems to be no direct use of the information obtained from

the assessment to plan further learning for ELLs. Secondly, there appears to be heavy reliance on evidence obtained from informal classroom observations of ELLs, particularly for OTJs about oral language skills. Although the literature acknowledges that informal assessment is appropriate for younger learners, the evidence obtained from informal assessment may not be as reliable as that obtained from formal assessment, making a balance of the two preferable. Thirdly, teachers' comments indicate that time is a key consideration in deciding on assessment activities, with a preference for those that are quick and easy for teachers to use. This may indicate that practicality is being prioritised at the expense of validity and reliability.

The literature also indicates that for teachers to arrive at sound OTJs, they need to have a clear understanding of the 'standards' that they are judging learners' performance against. However, it seems that this may not be the case, as teachers reported that mainstream teachers have difficulties with understanding the ELLP descriptors, some of which contain linguistic terminology likely to be unfamiliar to them. Finally, while the literature indicates that moderation of OTJs is essential for ensuring consistency of judgements between teachers, moderation was not being carried out by the teachers interviewed.

There also seem to be gaps between the Ministry of Education guidelines for ELLP assessment and teacher practice. Firstly, although the Ministry advocates using 'a wide range of assessment tools to rate learners against the ELLP, this would not appear to be the case (Table 4, above). Although this range would be extended if informal classroom assessments and observations were included, evidence from these is also informal, and difficult to account for. The term 'a wide range' seems to be problematic, as the Ministry has not given clear guidance as to what this means. Another gap appears to exist between Ministry provision of online resources to assist teachers assessing ELLs, and their use by teachers. These resources are unlikely to be known or used by mainstream teachers, as even the ESOL teachers in the study were not familiar with these resources. Further, there seems to be a gap between the Ministry advice that ELLP assessment should be "based on your school's usual age-appropriate assessment tools, activities, and observations" and teacher practice. As the National Literacy Standards do not require assessment of oral language skills, there may be very little, if any, evidence of these skills which can be derived from 'usual' mainstream assessment

activities. Finally, there seems to be a gap between some of the assessment activities reported by teachers and the descriptors in the ELLP. For example, running records provide information about reading behaviours, whereas the ELLP Reading descriptors describe features of texts. Similarly, there seem to be no direct links between informal observations of listening and speaking and the ELLP descriptors.

Recommendations

There are a number of actions which could be taken to ensure that both ESOL teachers and mainstream teachers are able to make sound OTJs about their ELLs' language skills. One recommendation is that the Ministry firstly clarifies how the term 'a wide range' of assessment activities should be interpreted, and then what constitutes an appropriate set of assessment activities that teachers can use to gather evidence for forming their OTJs for ELLP assessment. This would seem to be particularly important for oral language skills. A further recommendation is that the Ministry provide information and examples of how evidence obtained from either formal or informal assessment activities can be aligned to the ELLP descriptors. An additional recommendation is that schools find ways to provide time for both ESOL and mainstream teachers to source appropriate tasks for ELLP assessment.

Time is also needed for other aspects of ELLP assessment. Time is needed for teachers to complete some formal moderation, to ensure more reliable judgements. Time is also needed for both ESOL and mainstream teachers to become more familiar with the ELLP descriptors, so that teachers are clear about what the 'standards' are. In addition, it appears that ESOL teachers, who are expected to play a leading role in their school in regard to ELLP assessment, need more time to fulfill this role more effectively. This could take the form of inter-school professional development workshops - all three teachers mentioned that they did not know how other schools were carrying out ELLP assessment.

To summarise, the current study has revealed some of the gaps that currently appear to exist between recommended practice regarding ELLP assessment, and the actual practice of teachers in New Zealand primary schools, drawing on information reported by ESOL specialist teachers. However, although ESOL teachers play an important

supporting role in ELLP assessment, mainstream teachers now have the chief responsibility for judging ELLs' language skills for this assessment. Future research is therefore needed which focuses on the practices and attitudes of mainstream teachers as they carry out ELLP assessment.

List of References

- Bachman, L. & Palmer, A. (1996). *Language testing in practice*. Oxford, UK: Oxford University Press.
- Brown, H.D. (2001). *Teaching by principles: An interactive approach to language pedagogy*. Longman.
- Brown, H.D. & Abeywickrama, P. (2010) *Language assessment: Principles and classroom practice (2nd ed.)*. White Plains, NY: Pearson Education.
- Darr, C. (2011). Can test results help us make overall teacher judgements? *set: Research Information for Teachers, 1*, 61-63.
- Education Services Australia (n.d.). *Assessment for learning: Key questions*.
http://www.assessmentforlearning.edu.au/professional_learning/intro_to_afl/introduction_key_questions.html
- Fairtest: The national center for fair and open testing (2009). *Principles and Indicators for Student Assessment Systems*. Retrieved from <http://www.fairtest.org/principles-and-indicators-student-assessment-syste>
- Farhady, H. (2012). Principles of language assessment. In Coombe, C., Davidson, P., O'Sullivan, B. & Stoyhoff, S. (Eds). *The Cambridge guide to second language assessment*, pp. 133-139. New York, NY: Cambridge University Press.
- Klenowski, V. & Wyatt-Smith, C. (2010). Standards, teacher judgement and moderation in contexts of national curriculum and assessment reform.

Assessment Matters (2), 107-131.

McKay, P. (2006). *Assessing young language learners*. Cambridge: Cambridge University Press.

Ministry of Education. (n.d.a). *Overall teacher judgement*. Retrieved from <http://assessment.tki.org.nz/Overall-teacher-judgment>

Ministry of Education. (n.d.b). *Assessment online*. Retrieved from <http://assessment.tki.org.nz>

Ministry of Education. (n.d.c). *ESOL Online*. Retrieved from <http://esolonline.tki.org.nz>

Ministry of Education. (2008). *The English language learning progressions*. Retrieved from <http://esolonline.tki.org.nz/ESOL-Online/Student-needs/English-Language-Learning-Progressions>

Ministry of Education. (2010). *National Standards*. Retrieved from <http://nzcurriculum.tki.org.nz/National-Standards>

Ministry of Education (2014). ESOL funding and Ministry of Education support. Retrieved from <http://nzcurriculum.tki.org.nz/System-of-support-incl.-PLD/Learner-initiated-supports/ESOL-funding-and-MoE-support>

Ministry of Education. (2015). *ELLP professional support modules, Module 6: Using ELLP to support funding applications*. Retrieved from <http://esolonline.tki.org.nz/ESOL-Online/Student-needs/English-Language-Learning-Progressions/ELLP-professional-support-modules/Module-6-Using-ELLP-to-support-funding-applications>

Murray, D. & Christison, M. (2011). *What English language teachers need to know Volume II*. New York, NY: Routledge.

Poskitt, J.M., & Mitchell, K. (2012). New Zealand teachers' overall teacher judgments (OTJs): Equivocal or unequivocal? *Assessment Matters* (4), 53-75.

Richards, J. (2015). *Key issues in language teaching*. Cambridge, UK: Cambridge University Press.