

Sue Edwards

Waikato Institute of Technology,
Hamilton, New Zealand



LED
2015

Assessment of English language learners in New Zealand
primary schools: Purposes, principles and practices.

Overview

- **Background to the research**
 - Theory relating to assessment, and assessment of ELLs
 - Change in assessment of ELLs in NZ
- **The current study**
 - Research questions
 - Data gathering
 - Findings and analysis
- **Conclusions and further research**

One central idea



**THE PURPOSE
OF ASSESSMENT
IS TO IMPROVE
LEARNING.**

Back to Basics

Testing

Assessment

Evaluation

- “ A **test** is a method of measuring a person's ability, knowledge or performance in a given domain.”
 - (Brown & Abeywickrama, 2010, p. 3)
- “**Assessment** refers to the different ways of collecting information about a learner’s progress and achievement.”
 - (Thornbury, 2006, p. 18).
- “**Evaluation** is involved when the results of a test (or other assessment procedure) are used for decision making.”
 - (Bachman, 1990, as cited in Brown & Abeywickrama, 2010, p.5)

What do we assess?

From the
English
Language
Learning
Progressions

(NZ Ministry of
Education, 2008)

- ❑ **Skills?** (receptive, productive, literacy, communication, text-processing)
- ❑ **Proficiency?** (reading, speaking, English language, academic language)
- ❑ All of the above?

Approaches to assessment

Harmer (2007,
p. 379)

(Poskitt &
Mitchell (2012)
p. 55)

- **Formative** (ongoing, progress, monitoring) **vs** **Summative** (end of unit/period)
- **Formal** (e.g. standardised, large-scale, school-wide) **vs** **Informal** (e.g. teacher-made tests or tasks; classroom observation)
- “New Zealand values the central role of formative assessment in improving learning and teaching, and the professionalism of its teachers.”

Assessment principles

- Validity
 - Reliability
 - Practicality
 - Authenticity
 - Fairness
 - Sensitivity
 - Washback
 - Security
 - Usability
- “Perhaps the most important quality of any test is how **practical** it is to administer.”
(Baxter, 1997, p. 27).
 - “**Validity** is certainly the most important single characteristic of a test.”
(Farhady, 2012, p. 37)

Prioritising Principles

Brown
(2001,
p. 389)

- “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 test, then you are well on your way to making accurate judgements about the competence of the learners with whom you are working.”

The importance of accurate ELL assessment

**National
Association
for the
Education of
Young
Children
(2005, p. 5)**

- “Accurate assessment of language proficiency is important because these children may seem to be speaking English with ease when actually they are not fully capable of understanding or expressing themselves in complex ways and still lack vocabulary skills, auditory memory, ability to follow sequenced directions, and other markers of proficiency.”

The wider context: Standards and OTJs

Poskitt &
Mitchell (2012,
p. 54)

N.Z. Ministry of
Education (n.d)

“Critical to the implementation of National Standards in New Zealand is the notion of standards and the centrality of the OTJ.”

“An overall teacher judgment (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.”

Dependable teacher judgements

Klenowski &
Wyatt-Smith
(2010, p.
113)

(my
emphases)

” ...standards-referenced assessment relies on teacher judgement that can be made dependable if standards are promulgated in appropriate forms and teachers have the requisite conceptual tools and professional training.”

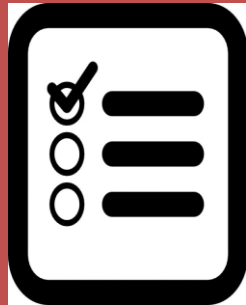
Concerns about standards

Poskitt &
Mitchell
(2013, p. 61)

OTJs can be problematic unless:

- Teachers are clear about what constitutes an OTJ
- They have common understandings of standards
- Such understandings are supported by clear criteria and exemplars of student work
- Teachers engage in moderation processes

Impact on assessment of ELLs



Assessment for ESOL funding eligibility

- **Prior to 2013:**
 - Comparison with native-speaker cohort (year group)

- **From 2103 (mandatory from 2015):**
 - Make an Overall Teacher Judgement (OTJ)
 - OTJ based on descriptors in the NZ [English Language Learning Progressions \(ELLP\)](#)

A change in assessment approach

THEN

- “A suggested approach is that teachers **view and assess each student’s strengths and needs in relation to those of the ‘typical’ students operating in the school’s mainstream classrooms and curriculum areas.**”

(Ministry of Education, 2004, p. 6)

NOW

- “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.**”

(Ministry of Education, 2013).

A change in scoring

THEN

- Teachers assigned a number (1-3) to each of 44 assessment criteria, to reflect comparison with cohort.
- Students were eligible for funding if their total score fell below a certain number (112).

NOW

- Teachers assign a number for each language skill (0-4), based on ELLP descriptors, to reflect achievement level.
- Students are eligible for funding if their total score falls below a certain number (which depends on their year group).

A change in assessment measures

THEN

Assessment criteria and suggested assessment measures were provided on the application form to be filled in.

NOW

- The assessment criteria are effectively the descriptors on the ELLP matrices, and there are no specific assessment measures provided.

Adopting the new system

- Workshops for ESOL Specialist teachers (ESTs)
- Online Professional Learning modules:
 - [Meeting the needs of English Language Learners](#)
 - [Using The English Language Learning Progressions](#)
- ESOL Online web site
- Online teacher forums

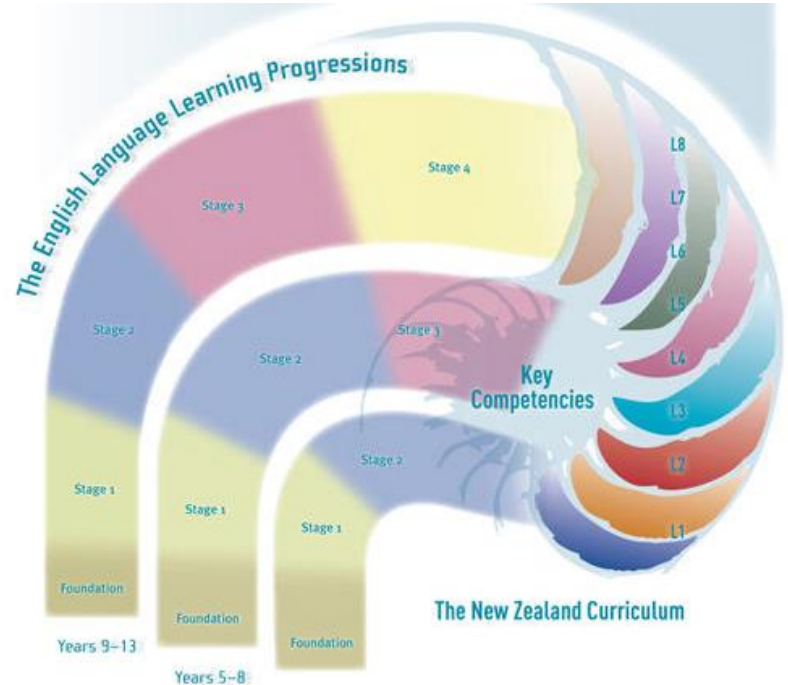


From Primary ESTs

- “A question – what moderation strategies do you have in place for the ELLP matrices?” (27/10/15)
- “At my school all the teachers meet in my room at 8 o’clock on a particular day usually around late July and mid December, and the teachers complete the matrices for their students.”(27/10/15)
- “I have found it useful to revise ELLP assessments during a staff meeting prior to when assessments are due.” (28/10/15)
- “At my school, we invited our ESOL expert at Canterbury University to run some PD with us.” (28/10/15)
- “the Oral Language Exemplar for the ELLP DVD is VERY good for practising moderating oral language during a PD session.”(28/10/15)

Purpose of the research

- How do ESOL specialist teachers (ESTs) decide which stage/s of the English Language Learning Progressions (ELLP) their ELLs have achieved, in order to make an OTJ for ESOL funding eligibility?



Research questions

- 1) How do the ESTs see the **purposes** of ELLP assessments?
- 2) Which **assessment principles** do ESTs take into account?
- 3) Which **assessment measures (tools/tasks)** do ESTs use (or recommend)?
- 4) What are the **positives** of using the ELLP for assessment?
- 5) What are the ESTs' **challenges** with using the new system?

Participants

Experience
with
assessing
ELLs:

A – 16 years

B – 20 years

C – 29 years

Role in the assessment of ELLs:

- Assess ELLs for other purposes - ALL
- Assessment for funding eligibility:
 - A: Assesses ELLs together with mainstream teachers
 - B: Proactively assists mainstream teachers
 - C: “Collaborates with mainstream teachers as required.”

Professional development of ESTs

For using the
new
assessment
system

- Ministry of Education ELLP workshops (for ELLP as a teaching and learning tool) (All)
- MOE workshop for new scoring system (All)
- ESOL Online PD materials (All)
- Participation in trial for ELLP assessment (A)
- ESOL Online discussion forum (B)

Research Contexts

**Refugee and migrant
ELLs:**

Maximum of 5 years

NZ-born ELLs:

**Maximum of 3 years
during first four years**

- School A – 30 ESOL-funded ELLs
- School B – 50 ESOL-funded ELLs
- School C – 67 ESOL-funded ELLs

Data gathering

“The interview is a social encounter, not merely a data collection exercise...”

(Cohen, Manion, & Morrison, 2011, p. 420)

- ❑ Qualitative study
- ❑ Semi-structured interviews
(Cohen, Manion, & Morrison, 2011, p. 236)
- ❑ Search for similarities and differences in practices and opinions

Findings 1: Purposes of ELLP assessment

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

Findings 2: Assessment principles

Planning assessment – Teacher A

“...principles of language assessment can and should be applied to...assessments of all kinds.”

(Brown & Abeywickrama, 2010, p. 25)

- Must be needs-based
- Use effective tools e.g. video
- Must tap into what teachers are already doing
- Timing – should be done after other assessments
- Speed – should be able to be done quickly

Planning assessment – Teacher B

“Language testing at any level ..must be based on theory as well as practice.”

Coombe, Folse, & Huble, 2007, p. xxii)

- Must be manageable for the classroom teacher, in terms of time
- Does it tell us what we’re wanting to know?
- Is it part of what teachers already do?
- Is it reliable – does it give the same results as other classroom-based assessment, relative to cohort?

Planning assessment – Teacher C

“All assessment tools and processes ... should be reviewed against three criteria – validity, reliability and usability.”

(Ministry of Education, 2005, p. 10)

- ❑ Start with prior knowledge about learners
- ❑ Select assessments known to be useful
- ❑ Eliminate assessments that are too difficult
- ❑ Ensure students won't be stressed by assessment
- ❑ Don't plan to assess too much in one session

Delivering assessment – Teachers A & B

Validity

- ❑ Ask the right questions (A)
- ❑ Focus on assessment of academic, not social, language (All)
- ❑ Strive for consistency by the assessor (B)
- ❑ Assess ELLs away from the mainstream class (B)

Reliability

Delivering assessment – Teacher C

Reliability

- ❑ Make the assessment within the reach of the learner
- ❑ Give encouragement and positive feedback for all attempts
- ❑ Provide a quiet, private environment
- ❑ Turn assessment into a game
- ❑ Not in front of their peers

Findings 3: Assessment tools/tasks:

	Tchr. A	Tchr. B	Tchr. 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 Nat. Stds.
Writing	Independent Writing samples, mainstream class	Unassisted writing done for English curriculum	Unassisted writing samples

Findings 4: Positives of ELLP assessment

Which system
do you prefer?

All 3 ESTs
prefer

ELLP
assessment

The previous system:

- ❑ “Too many criteria – a lot of reading and interpreting to be done” (C)
- ❑ Scoring errors – lots of adding points
- ❑ Variable interpretation of criteria
- ❑ “Teachers felt pressured for scores to improve each time.” (A and B)
- ❑ “It wasn’t helpful for our practice” (B)
- ❑ “It was a joke – think of a number” (A)
- ❑ “It didn’t match the current educational language” (C)

Positives of ELLP assessment

The new system – why the teachers prefer it

- “The new system gives us a model for pulling the teachers in.” (A)
- “I think this system helps you plan more for ‘where to next’” (C)
- “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.” (B)

Findings 5: Challenges with ELLP assessment

Arriving at an OTJ

- “The big struggle I think they’ve had is...it’s not a level of where they’re working at, it’s an achieved level.” (B)
- “One piece of evidence is not enough.” (C)
- “Well, It (one piece of evidence) has to be.” (A)
- “Reading’s a funny one – and it’s flaming wrong!” (A)
- “Mainstream teachers rate too highly, in general.”(B)
- “There might be different interpretations – it’s not standard across the school.” (C)
- “I don’t trust the teachers... they’d have everyone on Stage 2.” (A)

Challenges with ELLP assessment

The descriptors in the ELLP matrices

- “We had to learn how to interpret the descriptors” (C)
- Some of the descriptors are confusing for mainstream teachers e.g. minimal pairs (A and B)
- “Some things seem to be positive and some things seem to be negative – it doesn’t make sense to me.” (A)
- 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.” (B)
- “It’s curious – the Reading descriptors are not on what the child can do, but on the text type, and it seems to be out of step with the other three modes.” (C) (Also A and B)

Challenges with ELLP assessment

Sourcing assessment tasks

- “I remember this feeling of reinventing the wheel... why didn't they (MOE) trial tools and recommend tools?” (B)
- “It's the Oral [language assessment] - that's the tricky one.” (A)
- “It's the Listening and Speaking that's the hard one, really.” (B)
- What do we (schools) do for Listening and Speaking?” (B)

Challenges with ELLP assessment

Other areas of concern

- ‘There’s a great variation in the reliability of the mainstream teachers’ [judgements]’ (B)
- “There’s an element of resistance still by some [mainstream teachers]” (A)
- “For some of them this year, it’s at least a third of their class so we’re talking about quite time consuming” (C)
- “I don’t know that I’m happy with what we’re doing at the moment – it’ll be ongoing” (B)
- “I feel like it’s still in the learning stages” (B)

Conclusions

“Sometimes, it is just time pressures that lead us to use a quick and familiar assessment.”

(Spiller, 2009, p. 9)

- Practicality is the paramount principle for both ELL and mainstream teachers.
- Validity and particularly reliability may be being sacrificed.
- There does not seem to be “a wide range” of assessment measures being used.
- There are ongoing issues with using ELLP for assessment, although attitudes of ESTs are largely positive.

Recommendations to the MOE

- Provide **release time for mainstream teachers** to complete more valid and reliable assessment of their ELLs, including time for moderation.
- Provide **time and more support for ESOL specialists** who are expected to play a leading/guiding role in their school in regard to ELL assessment.
- **Draw teachers' attention to assessment tasks and tools** to use in making OTJs about ELLs (esp. Listening and Speaking).
- **Re-publish the ELLP (2008)** to remove references to cohort-level, and reflect the current emphasis on the achieved ELLP level.

Limitations of the research

- Small sample size
- Nature of semi-structured interviews
 - Different follow-up questions asked, depending on teachers' responses to the main question.
- ESTs are not the main assessors of ELLs for funding eligibility – the mainstream teachers are.

Further Research

- A larger study of both ESOL specialist and mainstream teacher thinking and practice in regard to the assessment of ELLs
- To investigate in detail the process of forming an OTJ about ELLs' language skills.
- To investigate the gap between Ministry recommendations and teacher practice

Questions/Discussion

- How many assessment tasks, activities or activities would constitute “a wide range”?
- Should the assessment tasks be directly related to the ELLP descriptors?
- What would constitute a useful blend of formal and informal assessment of ELLs?
- Should teachers be required to record the evidence on which the OTJs are based?

Thank you!

Sue Edwards
Waikato Institute of Technology
Hamilton
New Zealand

sue.edwards@wintec.ac.nz