

Parallel Session 7

Friday 08 September 17:00 - 18:30

Session Number 7A

Room: Humanities H2.03

Time: 17:00 - 18:30

BERA SIG: **Assessment**

TITLE: **Assessment**

SESSION TYPE: **Individual papers**

0220 Comparing apples with pears. Teacher knowledge about assessment and its impact on their practices.

Hill, M. University of Auckland

0210 Quality control of examination marking – some models and simulations

Bramley, T Cambridge Assessment
Bell, J F Cambridge Assessment
Claessen, M Cambridge Assessment
Raikes, N Cambridge Assessment

Session Number 7B

Room: Ramphal Lecture Theatre

Time: 17:00 - 18:30

BERA SIG: **Early Childhood Education and Care**

TITLE: **Enabling children's literacy development and early play experiences**

SESSION TYPE: **Symposium 8342**

Chair: **L. Miller**, Open University

Discussant: **L. Miller**, Open University

0598 Teaching assistants and support for literacy and language development

Cable, C. Open University
Eyres, I Open University
Hancock, R. Open University

0594 Young children's experience of agency in the literacy curriculum in four early years settings

Miller, L.K. Open University

0604 'First, fast and only': synthetic phonics for five year olds

Hall, K. Open University

0606 A study of the play experiences of six two-year-olds within the domestic spaces of their homes

Hancock, R. Open University
Gillen, J. Open University

Name	Session	Day	Time	SIG	Paper
Hanley, PM	2Z	Thursday	09.00 - 10.30	Teacher Education and Development	Chair
Hanley, PM*	2Z	Thursday	09.00 - 10.30	Teacher Education and Development	0310
Hansen, A*	7D	Friday	17.00 - 18.30	Educational Research and Educational Policy	0752
Harbinson, R	5R	Friday	09.00 - 10.30	Practitioner Research	0126
Hardy, G	6K	Friday	15.00 - 16.30	Mentoring and Coaching	0558
Hardy, G*	8O	Saturday	09.00 - 10.30	Science Education	0853
Hargreaves, L*	7D	Friday	17.00 - 18.30	Educational Research and Educational Policy	0746
Hargreaves, L	7D	Friday	17.00 - 18.30	Educational Research and Educational Policy	0749
Harold, B.*	7G	Friday	17.00 - 18.30	Learning in the professions	0041
Harper, H	3G	Thursday	14.30 - 16.00	Inclusive Education	0617
Harris, A	6I	Friday	15.00 - 16.30	Leading and Managing Schools and Colleges	0112
Harris, A	8B	Saturday	09.00 - 10.30	Early Childhood Education and Care	Discussant
Harris, B	Key 6	Friday	12.15 - 13.45	Social Justice	0055
Harris, D	8O	Saturday	09.00 - 10.30	Science Education	Chair
Harris, D P*	8O	Saturday	09.00 - 10.30	Science Education	0501
Harris, J	8K	Saturday	09.00 - 10.30	Post-compulsory and Lifelong Learning	0600
Harris, R.*	4F	Thursday	16.30 - 18.00	Educational Research and Educational Policy	0223
Harris, V*	3H	Thursday	14.30 - 16.00	Literacy and Language	0380
Harris, V	3H	Thursday	14.30 - 16.00	Literacy and Language	Chair
Harris, V	5K	Friday	09.00 - 10.30	Literacy and Language	0483
Harrison, C*	8A	Saturday	09.00 - 10.30	Assessment	0892
Harrison, C.	4Z	Thursday	16.30 - 18.00	Teacher Education and Development	0247
Harrison, R	7F	Friday	17.00 - 18.30	Leading and Managing Schools and Colleges	0871
Hartas, D*	7J	Friday	17.00 - 18.30	New Technologies in Education	0117
Haydn, T.A.	4F	Thursday	16.30 - 18.00	Educational Research and Educational Policy	0223
Hayes, R.*	2X	Thursday	09.00 - 10.30	Sexualities	0608
Hayes, S.	4Y	Thursday	16.30 - 18.00	Teacher Education and Development	0767
Hayes, S.*	4Y	Thursday	16.30 - 18.00	Teacher Education and Development	0766
Hayes, S.	4Y	Thursday	16.30 - 18.00	Teacher Education and Development	0763
Hayes, S.	4Y	Thursday	16.30 - 18.00	Teacher Education and Development	0764
Haynes, G.S.*	3Q	Thursday	14.30 - 16.00	Post-compulsory and Lifelong Learning	0019
Haynes, GS	3Q	Thursday	14.30 - 16.00	Post-compulsory and Lifelong Learning	Chair
Hayward, G	2Q	Thursday	09.00 - 10.30	Post-compulsory and Lifelong Learning	Discussant
Hayward, G*	6H	Friday	15.00 - 16.30	Higher Education	0447
Hayward, GF*	1R	Wednesday	15.15 - 16.45	Socio-cultural and cultural-historical activity	0492
Hearsum, A*	5F	Friday	09.00 - 10.30	Educational Research and Educational Policy	0889
Hempel-Jorgensen, A	5E	Friday	09.00 - 10.30	Educational Research and Educational Policy	0505
Henchy, D	6S	Friday	15.00 - 16.30	Practitioner Research	0660
Hennebry, ML*	3H	Thursday	14.30 - 16.00	Literacy and Language	0696
Hennessey, S*	7H	Friday	17.00 - 18.30	New Technologies in Education	0239
Hennessey, S.	7I	Friday	17.00 - 18.30	New Technologies in Education	0166
Herrington, N*	2F	Thursday	09.00 - 10.30	Educational Research and Educational Policy	0633
Herrington, N	2F	Thursday	09.00 - 10.30	Educational Research and Educational Policy	Chair
Herrington, N*	5Y	Friday	09.00 - 10.30	Teacher Education and Development	0348
Herrington, N*	P1	Friday	12.00 - 14.00	Poster Session	0487
Hewitt, D	1O	Wednesday	15.15 - 16.45	Race, Ethnicity and Education	0389
Hewlett, K	3G	Thursday	14.30 - 16.00	Inclusive Education	0617
Hextall, I	1N	Wednesday	15.15 - 16.45	Primary School Teachers' Work	0201
Hextall, I.	5Y	Friday	09.00 - 10.30	Teacher Education and Development	0665
Hey, V*	1G	Wednesday	15.15 - 16.45	Higher Education	0467
Hey, V	1G	Wednesday	15.15 - 16.45	Higher Education	Chair
Hey, V	1G	Wednesday	15.15 - 16.45	Higher Education	Discussant
Hey, V*	7T	Friday	17.00 - 18.30	Social Justice	0635
Hey, V	7T	Friday	17.00 - 18.30	Social Justice	Chair
Hey, V	7T	Friday	17.00 - 18.30	Social Justice	Discussant
Hibbett, M	4R	Thursday	16.30 - 18.00	Primary Education	0323
Hick, P*	8E	Saturday	09.00 - 10.30	Inclusive Education	0887
Higgins, S*	2S	Thursday	09.00 - 10.30	Practitioner Research	0229
Higgins, S*	7H	Friday	17.00 - 18.30	New Technologies in Education	0242
Higgins, S	7H	Friday	17.00 - 18.30	New Technologies in Education	Chair
Higgins, S	7H	Friday	17.00 - 18.30	New Technologies in Education	Discussant
Higham, J J S*	4P	Thursday	16.30 - 18.00	Post-compulsory and Lifelong Learning	0371
Hill, M	7A	Friday	17.00 - 18.30	Assessment	Chair
Hill, M.*	5R	Friday	09.00 - 10.30	Practitioner Research	0221
Hill, M.*	7A	Friday	17.00 - 18.30	Assessment	0220
Hines, C.L.*	P1	Friday	12.00 - 14.00	Poster Session	0628
Hixenbaugh, P	5G	Friday	09.00 - 10.30	Higher Education	0630
Hockings, C	3F	Thursday	14.30 - 16.00	Higher Education	0370
Hodgen, J*	2M	Thursday	09.00 - 10.30	Mathematics Education	0860
Hodgen, J	8A	Saturday	09.00 - 10.30	Assessment	0892

Comparing apples with pears. Teacher knowledge about assessment and its impact on their practices

Mary Hill

**Faculty of Education
The University of Auckland
New Zealand**

**Paper presented at the British Educational Research Association Annual Conference,
University of Warwick, 6-9 September 2006**

This paper reports research undertaken as part of the Great Expectations: strengthening teaching and enhancing learning in schools with diverse student populations through action research project funded by the Teaching and Learning Research Initiative (TLRI), a Ministry of Education initiative managed by the New Zealand Council for Educational Research (NZCER).

ABSTRACT

The *Teaching and Learning Research Initiative* project described in this paper found that, even in successful schools, a lack of teacher expertise in assessment caused roadblocks for evidence-based improvement of teaching and learning. This New Zealand Ministry of Education funded research project (Great expectations: strengthening teaching and enhancing learning in schools with diverse student populations through action research) aimed to build the capacity of teachers to reduce inequalities and maintain high student achievement in schools with diverse student populations. An associated aim was to increase the research capability of the teacher researchers within each of the participating schools.

There is no national testing of primary school children in New Zealand. Each self-managing school is responsible for setting its own achievement targets, monitoring achievement of its students against these targets and reporting achievement and any variance from the targets to the school community and the Ministry of Education annually. The teachers in the six primary schools involved in this project, in partnership with two university researchers, used a variety of standardised assessment tools for gathering this evidence. They used this assessment evidence to establish baseline data for their action research projects as well as for measuring student achievement for the reporting purposes explained above.

Data was collected from each of the six schools about the tools they used to collect assessment evidence, the use of these tools, how the information that they gathered informed teaching practice, and issues that they confronted in the area of assessment during their project. Although most of the schools were able to use the tools, the data indicated that a lack of technical understanding about

standardised assessment, issues related to the type of professional development and practical issues with some tools impeded assessment practices and led to some invalid and unreliable achievement results, particularly in the first year of this two-year project.

This paper briefly overviews the project, explains how New Zealand primary schools are required to monitor and report achievement and briefly introduces the tools that these schools used most often to gather achievement evidence. Data collection methods for the research are described and the findings, in terms of teacher knowledge and use of these tools, are presented. The paper then focuses upon difficulties that arose for particular schools and teachers in respect of assessment processes, how these related to their assessment knowledge/expertise (or lack of it) and how these teachers solved these problems. Discussion regarding the level of assessment knowledge and expertise primary teachers require in self-managing schools in order to effect evidence-based improvement of teaching and learning concludes the paper.

Introduction

This Teaching and Learning Research Initiative (TLRI) project investigated school-based action research for school improvement in 2004 and 2005. Researchers from the University of Waikato (and later, The University of Auckland) worked with teacher researchers from six primary schools to explore ways of changing the classroom practices of teachers.

It is now widely recognised that more detailed school and classroom research is needed to uncover the complexities of teaching and learning (Gorard, 2006; Ministry of Education, 2002). One of the greatest challenges in this kind of research is to describe what happens when teachers, students, and communities work together, in order to understand the relationship between teaching action, expectation, and student achievement (Nuthall, 1999). Within New Zealand, as elsewhere, investigations have looked into the impacts of assessment (both formative and for accountability purposes), the use of achievement evidence to inform teaching moves (Hill, 2000; Timperley & Parr, 2004), learning styles, leadership impacts, and teacher coaching (Robertson, 2005). Another factor known to be of considerable importance in teaching and learning is that of teacher expectations (Galton, Hargreaves, Comber, Wall, & Tell, 1999; Timperley & Phillips, 2003). While externally provided professional development has been shown to have an effect on teacher expectations, feelings of self-efficacy, and student achievement (for example, Timperley & Phillips, 2003), the rationale for this study was to investigate how schools themselves could draw on existing research and, through their own effort, initiate and sustain high expectations and increased student achievement.

The main aim of the project was to add to existing knowledge of teaching and learning in primary classrooms and use this new

knowledge to inform and help teachers and students. Together, the teacher researchers identified factors such as expectations that were linked to improving student achievement and teacher practices within their respective schools. Teaching strategies that led to improved achievement—especially literacy and numeracy—were identified, using a range of classroom research methods.

Building the capacity of teachers to reduce inequalities and maintain high student achievement in schools with diverse student populations was also an important aim of this project. An associated aim was to increase the research capability of the teacher researchers within each of the participating schools, in order to enhance their professional learning and research qualifications.

The six main research questions that guided the study were:

- How well are students achieving currently in each participating school?
- What expectations are held for students within each school?
- What teaching and school leadership practices strengthen learning in each context?
- How should teaching change to improve learning and achievement consistently throughout classrooms in each school?
- What capabilities do teachers and teacher researchers need to sustain constant improvement within their school?
- How can the findings of teacher research be best communicated to a wider audience of professionals, academics, and officials?

This paper spotlights some of the challenges regarding assessment that the teacher researchers encountered during the project.

Research design

The project began in 2004 by gathering baseline information from the schools about student learning and achievement, their expectations for their students, and the changes they were planning to implement and investigate. From late 2004 through to the end of 2005 each school used cycles of action research to plan interventions, implement them, observe and research their effects, and reflect on the results. Most schools moved through at least two full cycles of research over the two years.

Over the two years of the project (2004–2005), each of the six schools investigated the research questions through action research designed to raise both achievement and expectations. Case study methodology (Bassey, 1999; Hill, 2000) within an action research framework (Robertson, 2000) provided the overall research design. We generated an overarching conceptual framework (Robertson, Hill, & Earl, 2004) to refine, investigate and report on the research questions in terms of expectations, achievement, teaching, learning, assessment, professional learning communities, and leadership approaches.

Within this framework, each school designed and implemented its own action research project to investigate aspects relating to the research questions. In summary, each school

- identified the current achievement of students at each school, especially with respect to literacy and numeracy;
- undertook exploratory investigations of several innovative approaches to improve their teaching;

- analysed the data collected on the basis of categories and criteria developed by the research team;
- shared the findings with the other teachers in each school, between participating schools, and with other schools and interested parties as the project proceeded;
- developed further observational and analytic techniques that contributed to learning outcomes across these schools with diverse student populations; and
- communicated the findings to broader teacher, policy, and research audiences.

Symposia were held to plan, develop, discuss and support the research. The fifth symposium, in November 2005, was a more public event open to teacher researchers nationally, who were invited to share their experiences and findings. It was attended by over 100 teacher researchers and included an international keynote speaker (Dr. Lorna Earl, from the University of Toronto, Canada), 25 presentations, and workshops covering such aspects of research as ethical responsibilities, rigour and credibility, how to gather data effectively from children, and using theory to interpret qualitative data. University researchers and teacher researchers from each of the six schools in the Great Expectations project also led workshops and made presentations at this symposium.

The New Zealand assessment context

As explained above, each school began this project by sharing baseline information about student achievement in their schools. In New Zealand's self managing schools each individual school is responsible for setting its own achievement targets, monitoring achievement of its students against these

targets and reporting achievement (and any variance from the targets) to the school community and the Ministry of Education annually. Rather than using national tests as in the UK, New Zealand monitors student achievement nationally at Year 4 and Year 8 through sampling 3% of the school population across the entire curriculum in four-year cycles. This is known as the national Education Monitoring Project (NEMP). New Zealand also takes part in international surveys of performance such as PIRLS, PISA and TIMSS. In these ways, we are able to track and respond to macro trends in achievement for primary students nationally.

In order to provide comparative indicators of achievement New Zealand has invested in a range of assessment tools designed to be used by teachers within schools for a range of assessment purposes instead of using national cohort testing at the school level. For example, to measure reading progress there are standardised normed tests, national exemplars that indicate benchmarks for particular age groups and standardised procedures such as reading running records and the diagnostic survey (Clay, 1993).

Although some of these tools were developed progressively from the 1950s (for example, the Progressive Achievement Tests - PATs), most have more recent origins. The assessment tools for teaching and learning (asTTle), for instance, provide schools with the technology to choose the aspects of reading they wish to assess, construct valid and reliable tests, supply marking guides and then enable electronic analysis of the results. The analysis reports indicate children's individual, class and school-wide strengths and weaknesses allowing the information to be used diagnostically within a very short time of taking the tests; literally overnight in some schools. Teachers can also access a website that connects them to resources targeted to the specific needs of their learners.

Currently this tool is under development as a web based (rather than CD) tool.

There are also the assessment resource banks available on the web (assessment items categorised by curriculum area and level with marking guides and diagnostic information), new curriculum aligned PATs with electronic web based analysis tools, national exemplars of annotated students' work across the curriculum to guide teachers' expectations of achievement, and literacy and numeracy tools for diagnostic assessment to guide teaching. All of these assessment tools are provided free to schools and extensive professional development programmes have been provided nationally to assist teachers to learn how to use them productively for improvement as well as for monitoring achievement outcomes.

Findings related to assessment issues

In this context, the Great Expectations research team started out in 2004 to collect information about which tools each school was using to gather information about their school-wide achievement and the levels at which children in each of the schools were achieving. At our first meeting each school described the tools they were using to collect assessment data and the levels children in their schools were. We discovered that schools were collecting a broad range of information using most of the tools available in different combinations. Attempting to gather data on all the measures concurrently was far too complex a task so we restricted ourselves to collecting information about how the schools monitored progress in reading only.

Reading is an aspect of the English curriculum that all the schools had expectations for and collected student achievement information about. All of the schools used running records to ascertain reading achievement in the junior school area. However, while four used the “PM benchmarks” (levelled set texts with accompanying comprehension questions) the others had decided to use their own school-designed system for within-school consistency. Achievement in reading as measured by running records in these junior school departments was then aggregated to assess achievement within each year group. For example, one reported that in 2003 57% of Year One students were reading at “blue” level or above; 75% of Year Twos were reading at or above “turquoise”¹. In contrast, another reported in more general terms, that, using the PM benchmarks, 93% of their students (excluding their non English speaking background students) were reading at or above their chronological age.

But from here the similarities started to unravel and the complexity accumulated. Not only were issues of validity and reliability raised for further investigation (eg, were the texts suitable to provide a valid result? Did the teachers administer the running records consistently enough so that results could be meaningfully aggregated for this purpose? And so on), there was also a great deal of discussion about how these results were interpreted. For example, there was discussion regarding what “reading at or above turquoise” actually meant.

Teacher from School D: What is your age level for turquoise?

¹ A colour wheel that indicates increasing difficulty of texts is used to reflect progression in learning to read in New Zealand. Each colour can be related to a normative age level for reading.

Teacher from School F: Six and a half. End of year two.
Teacher from School D: Are you meaning that turquoise is at the six and a half year old level?
MH (Uni researcher): But wouldn’t children be at different ages at the end of year two if they begin on their fifth birthday? They could be between just under 6 and a half and just over seven by the end of year two.
Teacher from School F: That’s why we’re saying “at or above”

There followed some general discussion with several people talking at once and disagreeing about what level “turquoise” actually represented. Hence, although reading running records appeared to be regarded as “what works” for measuring achievement and setting expectations (particularly from a management point of view), there was a great deal of debate about how they should be interpreted, administered and used in general. This issue became a focus for improvement at each school throughout the 2 year project.

As well as raising issues of comparability, this discussion also alerted us to the fact that such statements had very little to do with driving teaching. The teachers explained that they reported in percentage terms in order to simplify student achievement measures for mandatory reporting to their boards of trustees, communities and the Ministry of Education. In fact,

because these measurements required a reasonable level of consistency in order for them to be aggregated for reporting, the running records were generally carried out as fast as possible in a short space of time and were not generally analysed in order to ascertain the strategies children were using (or not using). This issue of tension in the purposes of assessment was another that each school worked in different ways throughout the project to address.

There was even less consistency when it came to measuring the reading achievement of children above Year Three. Four of the schools used the Progressive Achievement Tests (PATs), but differed in whether they used the reading vocabulary or comprehension forms or both. Others used Supplementary Tests of Reading Achievement (STAR), Prose Reading Observation, Behaviour and Evaluation of Comprehension test (PROBE), asTTle (assessment tools for teaching and learning) tests, the NSW (Australian) reading tests, running records, or various combinations of each of these. One school designed their own reading comprehension test in addition to using the asTTle tools and the PAT comprehension test each year. When we discussed why this was, the school explained that they needed to triangulate the results in order to ensure they were getting an accurate picture of achievement. The school-designed tests are graded, use extracts from school journals and, in contrast with asTTle and the PATs, require a constructed response rather than a selected response from the students. The principal explained that this aspect alerted them to instances where students had simply randomly picked answers in the PAT but were working well below the level their PAT score had indicated.

Thus our initial findings were that student achievement was measured and reported idiosyncratically in each case. But

interestingly, as all of the schools were combining the “Planning and Reporting” requirements (Ministry of Education, 2002) with their school achievement monitoring, they had all set “targets” in terms of what they expected students in their school to achieve and were reporting against these. In all cases, these schools were attempting to use externally designed assessments, applied systematically within their school to enable them to measure and report progress and achievement over time. The targets schools had set were, in essence, their academic expectations of special focus for that year and were either phrased in terms of an expectation (for example, a certain percentage of children reading at or above their chronological age level) or as a stanine in the case where they used tests that provided these, or both.

It should be remembered that each of these schools is seen (and sees themselves) as successful. For reasons of manageability, these schools focus on one or two major targets, but there is complexity sitting behind the implied simplicity of annual reporting. In developing their annual targets, every one of these schools had made a decision to keep these manageable and had set one or very few (not more than 5) targets that they would measure and report on an annual basis as required by the Education Standards Act (2001).

In contrast to anecdotal information that suggests some schools in New Zealand at that time were setting a very large number of targets, these six schools targeted one priority area, such as literacy or numeracy for emphasis each year. The staff and board of trustees decided on these priority areas by collecting data on student achievement and comparing it with national (and other) norms in order to select the most relevant targets for

their school, for that year. One school, for example, which investigated the use of student learning meetings to improve expectation and achievement in 2004, targeted numeracy achievement across the school. This dovetailed with national and local professional development initiatives. Data collected on a regular basis throughout the two year project provided information that teachers analysed, discussed and used to plan interventions. Over the two years, using assessment information in this targeted manner, this school built a learning organization with regular scheduled learning meetings, a staff appraisal system, new ways of reporting to parents and raised numeracy achievement and expectations dramatically.

Although all of the schools had set academic expectations, it is important to emphasise that these were not the only expectations that drove their practice. Each school also indicated a broader set of expectations for their students that encompassed a set of values and goals but was articulated differently in each case. There is not room in a paper such as this to spell out all of these expectations, but one example is included here because it shows how expectations and achievement tend to flow from the particular values held within the school. One school, in a very low socio-economic rural area where about 80% of the students are Māori, had an emphasis on preparing students for life-long learning to succeed as citizens in New Zealand society. To achieve this the school constructed a graduate profile for students to be achieved by the end of Year Eight (the end of primary schooling in New Zealand).

Ideally, what would School student at our school “look like” when he/she leaves in Year 8?
 He/she would be:
 Socially confident

A confident learner
 Confident with personal relationships
 Confident about the future
 Personally confident
 Emotionally confident
 Confidently maintains personal safety
 Confidently skilful

Figure 1 Graduate profile summary

Under each of these headings, the school then described what each aspect meant in practice. For example, in being socially confident, a student would

Be comfortable greeting a variety of people in the community
 Be able to introduce themselves appropriately
 Be able to speak with confidence believing that what he/she has to say is of value and that people will listen.
 Be able to expresses him/herself freely though appropriately
 Be able to use non-verbal communication effectively
 Not shuffle when speaking to people
 Be able to respond appropriately in a variety of social and cultural settings
 Maintain personal integrity in all social situations
 Have a sense of fun
 Show leadership skills when appropriate
 See service to society as important

Figure 2 Socially confident behaviours indicated in the school’s Graduate Profile

Holding such an agreed set of expectations as a school community, this school set out to investigate how feedback, questioning and the development of a professional learning community could assist them to bring their profile to fruition. The indicators within the profile are clearly linked with characteristics of quality teaching (Harris, 2002; Alton-Lee, 2003) and provide a clear set of descriptors against which teachers can gather information, investigate progress, change teaching and report on progress and achievement. The academic expectations in this school (as in all six involved in the study) are clearly part of a broader conception of attributes needed to prepare for life in the 21st Century.

It is about creating a climate or a culture in the classroom – and in the school more widely- that systematically cultivates habits and attitudes that enable young people to face difficulty and uncertainty calmly, confidently and creatively. (Claxton, 2002, p.3)

Assessment challenges for teachers in self-managing schools

These examples from our findings related to assessment alerted us to the fact that several challenges still exist regarding the use of assessment in self-managing schools. In an earlier study, Hill (2000) reported that teachers experienced difficulty designing school systems of assessment in which assessment was used mainly for improving learning through focused teaching rather than where summative reporting purposes dominated. The findings of that study indicated that school leaders needed a level of theoretical assessment knowledge that enabled them to make school-based policy decisions that led to the balanced use of assessment for formative as well as summative purposes. In schools where assessment was used to

inform teaching and achievement levels were high, the teachers were also knowledgeable and skilled in assessment.

Our data from the project described in this paper also indicated a lack of technical assessment knowledge and skill but there was a clear intention by the leaders in these schools to increase their capability to use assessment to improve the educational outcomes for their students. Among the challenges confronting them in their desire to use assessment for learning in this way we discovered that: there was variation in the principals' expertise with assessment; teachers in some schools were not keen to use nationally normed tests in case the results were low; and, even after professional development, many teachers did not feel confident using the nationally provided assessment tools. Another challenge was that schools needed to organize data input of test results, the analysis of data and ways of reporting results that did not take teacher time away from other important teaching duties.

In order to address these issues, the teacher researchers used the notion of *distributed leadership* (Harris, 2005) to reconfigure leadership practice in their schools. In some there was a small attempt to do this using the existing hierarchy and class groupings. But in some emphasis was put on re organizing the way the entire school operated to effect change. In the second year, for example, one school released all the team leaders for one day a week so that they could co-ordinate valid and reliable assessments, hold student learning meetings with their team to analyse assessment data and make changes to their teaching, video teaching action and use peer coaching to facilitate improved teaching practice. This type of evidence-informed practice necessitated sophisticated assessment knowledge and

understanding but the rewards were spectacular improvements in student achievement.

One school, however, was not using the national assessment tools consistently or regularly in the first year of the project and thus was having trouble knowing just how well its students were performing. The short case study below describes how taking part in this TLRI project addressed some of the assessment challenges teachers at this school faced.

Case Study

Church School is a full primary school in one of the poorest communities in South Auckland. Virtually all of the students are from Pacific Islands backgrounds, the majority of which are Samoan. There are no European students at this school. Most of the students in this school do not speak English at home. Over the last few years this school has been involved in professional development projects designed to improve students achievement run by outside agencies. In 2003, when the TLRI was advertised, the principal and her two deputy principals decided that they would like to conduct a small research project to improve learning outcomes and expectations in reading comprehension. They approached the University of Waikato researchers in order to achieve this goal.

The main research question for their investigation was: How, if at all, does effective questioning improve reading comprehension? This question was derived from the teacher researchers' beliefs that questioning is important because it enhances and develops students' thinking ability. Church School decided to use Bloom's taxonomy of levels of thinking and questioning as a guide.

At the beginning of the project Church School mainly relied on running records and the diagnostic survey (Clay, 1993) in the junior school as well as reading comprehension tests designed by the school to measure reading progress. PATs and asTTle reading comprehension tests were also administered to the middle and senior classes at the beginning of 2004. The teacher researchers used these as a baseline for their research and shared these standardised results with the rest of the teachers in the school as well as the project. These results provided the basis upon which a set of expectations for reading comprehension for each year level was set. At the beginning of 2004 it was decided to use the asTTle reading comprehension tests as the main way of determining progress in reading comprehension.

Retesting in September 2004 produced really disappointing results. The teacher researcher explained to the university researcher that she didn't want to tell the teachers how badly the students had scored. After some discussions about how the retesting had taken place, it was discovered that the post-tests had tested different aspects of reading comprehension than had been assessed in February. Although the asTTle tools allow teachers to construct parallel tests on the same aspects of reading to measure progress over time, this school had chosen different aspects to test from those tested in February. They had therefore been comparing apples with pears! They explained that they had decided to choose the new aspects (finding information and recalling knowledge) because they were anxious that if they tested making inferences and evaluating (which had been the main aspects pre tested for) their students might not do well, even though these higher-level aspects of thinking on Bloom's taxonomy had been their focus for teaching during the year.

When new post-tests testing these higher-level aspects were administered in November, the results showed significant improvements.

In 2005 in order to increase the reliability of the results, the research team administered, collated and analysed the results. Although they took part in the decision making about planning and setting achievement expectations, the teachers were not involved in the administration of the tests. In 2005 the results were better again. In fact, through their more systematic approach to both assessment and teaching, significant improvements were made at all levels of the school. When compared against both the national norms and “schools like us” all cohorts had moved from below the average into or above the average performance nationally.

As a result of taking part in this project, Church School learnt a great deal about researching their own practice, using standardised assessment tools to gather evidence about learning and about working together as a professional learning community to achieve significant learning gains for their students.

Discussion and conclusions

The TLRI was set up in New Zealand with six principles set out under three value areas, namely: strategic value, research value and practice value. Under the *strategic value*, it was intended that these research projects address themes of strategic importance such as *reducing inequalities, addressing diversity, and understanding the processes of teaching and learning*. The findings from the Great expectations project clearly demonstrated that, in order to address these themes, self-managing schools require a range of well designed, standardised (and often norm-referenced) assessment tools suited to a

range of assessment purposes and, critically, school leaders and teachers skilled and knowledgeable in the use of such tools.

Additionally, although we found a wide range of assessment expertise within the schools in our project, the challenges faced in the case study school described above are not unusual. Experience teaching graduate courses for teachers in assessment also indicates that many teachers in New Zealand are not theoretically strong in aspects of assessment such as validity, reliability and the use of assessment *for* rather than *of* learning. The results of our project do indicate, however, that schools such as Church School can turn their knowledge and practices around, and effect very positive changes in, student achievement when they have the tools and the technical know how and financial resources to do so. Through their focus on raising particular important literacy and numeracy outcomes, Church school now sends its students to secondary school at or above the national norms for achievement in the curriculum areas they focused teaching and assessment of outcomes on. Through the use of the nationally provided tools they can measure progress reliably against the curriculum levels, against national norms and against the norms of schools similar to themselves in terms of rural/urban and socio-economic level of the community. This practitioner research project also indicated that it is possible for self-managing schools to access professional development specific to their needs in sustaining self-improvement. Each of the schools that took part in this project sought out the professional development they required to assist them to achieve their purposes of raising student achievement. Some attended national and international conferences. Others engaged curriculum and assessment consultants to assist them to learn teaching and assessment

theory and technique. As seen in the example in the case study, some needed multiple strategies to achieve their goals. Church school employed administrative staff to carry out technical assistance tasks, sent staff to cluster professional development meetings, employed consultants with expertise in teaching writing and called on university academics for assistance with the asTTle tools and research help.

What did become obvious from this project was that each school had needs that differed from the others. Putting them all through one structured professional development course in, for example, assessment or evidenced-based leadership would not have met their needs. Each school context, student and staff population was different. However, all needed leaders with sufficient skill in and understanding of assessment to guide teaching and establish school-wide policies and practices that led to improvement. Put another way, leading a learning organization/community requires levels of technical assessment skill and understanding that some principals and senior leaders have yet to acquire. In this TLRI project the teacher researchers and their colleagues in each school undertook a substantial amount of “just in time” professional learning that assisted them to raise achievement and expectations within each of their schools. This paper has highlighted some of the assessment learning that was required, but other areas for development included research capability, research ethics, mentoring and coaching, and organizational change.

References

- Alton-Lee, A. (2003). *Quality teaching for diverse students: Best evidence synthesis*. Wellington: Ministry of Education.
- Bassey, M. (1999). *Case study research in educational settings*. Buckingham: Open University Press.
- Claxton, G. (2002). *Building learning power*. Bristol: TLO Limited.
- Clay, M.M. (1993). *An observation survey of early literacy achievement*. Auckland: Heinemann.
- Galton, M., Hargreaves, L., Comber, C., Wall, D., & Tell, T. (1999). Changes in patterns of teacher interaction in primary schools: 1976-1996. *British Educational Research Journal*, 24 (1), 23-37.
- Gorard, S. (2006). Another interpretation of research capacity building: a response to Baron. *Research intelligence*, 96, August, 6-8.
- Harris, A. (2002). *School improvement: What's in it for schools?* London, Routledge-Falmer.
- Harris, A. (2005). Leading or misleading? Distributed leadership and school improvement. *Journal of Curriculum Studies*, 37(3), 255-265.

- Hill, M. (2000). *Remapping the assessment landscape: Primary teachers reconstructing the assessment landscape in self-managing primary schools*. Unpublished doctoral thesis. Hamilton: University of Waikato.
- Ministry of Education. (2002). Curriculum Stocktake Report. Wellington: Ministry of Education.
- Nuthall, G. (1999). Learning how to learn: the evolution of students' minds through the social processes and culture of the classroom. *International Journal of Educational Research*, 31(3), 141-256.
- Robertson, J.M. (1997). A programme of professional partnerships for leadership development. *Waikato Journal of Education*, 3, 137-152.
- Robertson, J.M., Hill, M., & Earl, L. (2004, November). *Conceptual frameworks in school-university action research communities*. A paper presented at the New Zealand Research in Education Conference, Wellington, 24-26 November.
- Timperley, H. and Parr, J. (2004). *Using evidence in teaching practice: Implications for professional learning*. Auckland: Hodder, Moe, Beckett.
- Timperley, H. & Phillips, G. (2003). Linking teacher and student learning to improve professional development in systemic reform. *Teaching and teacher education*, 19 (6), 643-658.