Literacy Professional Development Project

Evidence of Improved Student Outcomes

From the schools that participated in the Literacy Professional Development Project February 2006 – November 2007

Project Co-directors: Carolyn English, Lyn Bareta & Melanie Winthrop

Project Manager: Pam O'Connell

March 2008

Learning Media
Te Pou Taki Kōrero

Contents page

Executive Summary

The project outline

Literacy Professional Development Project is one part of the Ministry of Education's Literacy Strategy. The project is offered to schools with students in years 1–6 and/or years 7–8 and provides them with in-depth school-wide professional development in literacy. Schools can choose to focus on either reading comprehension or writing.

The project is led by a Leadership and Effectiveness team (LET) that consists of a project team, based at Learning Media; a team of researchers and consultants from the University of Auckland; and regional team leaders who each lead a cluster of literacy facilitators. The facilitators work with the literacy leaders, principals, and teachers of the participating schools, supporting them to take an inquiry and evidence-based approach to increasing the effectiveness of the literacy practices in their school. At all levels, the project works towards five outcomes:

- Evidence of improved student achievement
- Evidence of improved teacher content knowledge
- Evidence of improved transfer of understanding of literacy pedagogy to practice
- Evidence of effectively led professional learning communities
- Evidence of effective facilitation.

The project conducts an ongoing analysis into how well these outcomes are being met at the classroom, school, and project levels. This report focuses on the first outcome – "evidence of improved student achievement" – with regard to the third cohort of 127 schools that entered the project in February 2006. It is based on the national data collected for those schools from then until they completed their involvement nearly two years later in November 2007. 65 of these schools chose to focus on reading comprehension and 62 schools chose to focus on writing while on the project. 35% of the schools are full primary and 53% are contributing schools. The remaining are intermediates, and other schools with year 7 – 8 students. The decile ratings of the participating schools approximately match the decile ratings of New Zealand primary schools in general. 28% of the schools have only 1 – 5 teachers and 35% have 6 – 10 teachers.

Report structure

In the past, the Literacy Professional Development Project (LPDP) has followed a common structure for reporting on each of its outcomes. This structure seems to mirror the "Teacher inquiry and knowledge-building cycle to promote valued student outcomes" found in the *Teacher Professional Learning and Development: Best Evidence Synthesis Iteration* ("TPLD BES" – Timperley, Fung, Wilson, and Barrar, 2007). The project leaders have reflected carefully on this inquiry cycle, and reworked it to develop a "Project inquiry and knowledge-building cycle to promote valued student outcomes". This is presented in the first part of this report: "A collective responsibility: Developing project inquiry and knowledge-building to promote valued student outcomes".

This report follows the project's movement through one rotation of its inquiry and knowledge-building cycle as it worked with the schools that formed the February 2006 cohort. (These are the schools that joined the project in February 2006.) It focuses, in particular, on the third question of the inquiry: "What is the impact of our changed actions?" It also focuses on the first of the project's five outcomes: "Evidence of improved student achievement". A report on the other four will follow, following further analysis.

The second part of this report, "The project's improvement history" summarises the project's impact on students who attended schools that participated in the first two major cohorts. (These are the schools that joined in February 2004 and July 2004.) The LPDP was demonstrably successful in meeting its strategic goal of raising those students' overall achievement in reading and writing. It was especially successful in accelerating the rate of progress for the students who were at risk of underachieving. Nevertheless, analysis of sub-groups of students reveals room for improvement, especially for boys, Māori, and Pasifika. The report briefly describes the project's responses to those findings in terms of its learning and changed practices. In terms of learning, these responses have included further exploration of the classroom practices that make a difference for Māori and Pasifika students and the teaching and learning strengths and needs of new entrant students. In terms of practice, they include the decision to collect achievement data from all students from the age of five to year 8 and the selection of new tools with which to gather that data.

The third part of this report, "An evidence-based evaluation of LPDP's overall effectiveness" analyses the impact of the project's learning and changed practices on the literacy outcomes of the students who participated in the February 2006 cohort. This part of this report includes a discussion of the tools used to measure student achievement and asks questions about the effect of having tools that set two different kinds of "cohort expectations". One set of expectations is based on a set of standardised norms and another is based upon our aspirations for student improvement. LPDP's inquiry into the concept of "cohort expectations" is related to its consideration of the terms "at risk" and "of concern" and the specific support required for schools and teachers to analyse and respond to their students' strengths and needs.

The report concludes by reflecting back upon the project's learning from the February 2006 cohort and identifying its response to that learning as it plans for another cycle of learning and action. For example, the report notes that the project's understanding of the key concept of "sustainability" now includes ensuring that there is coherence between the "big ideas" about education that it is attempting to communicate and the activities it uses to transfer those ideas to teaching practice. As part of this shift, the project intends to support school leaders to use their planning and reporting processes as a mechanism for improving the literacy outcomes of all students. The planning and reporting documentation will also become a critical tool for monitoring the impact of the project on student outcomes, as will be seen in future reports.

Summary of key findings

The analysis of the impact of the project's learning and changed practices on the literacy outcomes of the students who participated in the February 2006 cohort was, once again, positive, with rates of progress for the majority of students being greater than those seen without project intervention and in accelerated rates of progress for those sub-groups of students traditionally over-represented in the lower bands of achievement. For example:

5 and 6 year olds, reading and writing focus:

The Observation Survey (Clay, 2002) was used to assess student achievement. There was a noticeable shift in stanine mean and decrease in

the proportion of students in stanine 1 – 3 for all tasks and for all sub-groups.

Years 4 – 8, reading and writing focus:

asTTle reading and asTTle writing was used to assess student achievement.

- In all year groups and for all demographic groups, the effect size is moderate to extremely high[1]. They range from 0.95 to 3.00 for students in schools with a writing focus and 0.53 to 2.11 for students in schools with a reading focus.
- In all year groups, both boys and girls are achieving at "cohort expectation" or better when compared to the national picture associated with the asTTle tool.
- In all year groups, Pasifika students are achieving at "cohort expectation" or better when compared to the national picture.
- The mean score for students in the lowest 20% for all year groups is now closer to each whole group's mean and is close to or better than the national picture.

Despite these positive findings, close analysis continues to reveal ongoing worries, puzzles, and inconsistencies. It is these concerns that continue to drive the project's quest for improved theories and practices. For example:

5 and 6 year olds, reading and writing focus:

- Although Māori students rate of progress is similar to or better than the whole cohort their mean score is one year behind their whole cohort for all tasks except Letter Identification for writing schools and Word Reading for reading focused schools.

Year 4 – 8, reading and writing focus:

- Māori students have in general made highly significant progress but the progress is not as great for students in years 6 and 7 reading schools when compared to all students.
- The range of asTTle scores increases over the years for both reading and writing schools and one of the outcomes of this is that the

[1] A moderate effect size is around 0.5, a large one is around 0.8 so an effect size of over 1.0 is extremely high.

proportion of students achieving below the curriculum expectation increased from year 4 to year 8.

A collective responsibility: Developing project inquiry and knowledge-building to promote valued student outcomes

Over the last three years, the Literacy Professional Development Project (LPDP) milestones have followed a common structure for reporting on each of its intended outcomes. First, the key findings were described and the improvements identified. Then, two actions were discussed: the areas that required further inquiry (those areas that we needed to understand more about before developing a practice response) and the areas that we were able to respond to straight away. This discussion included a description of the new or changed practices.

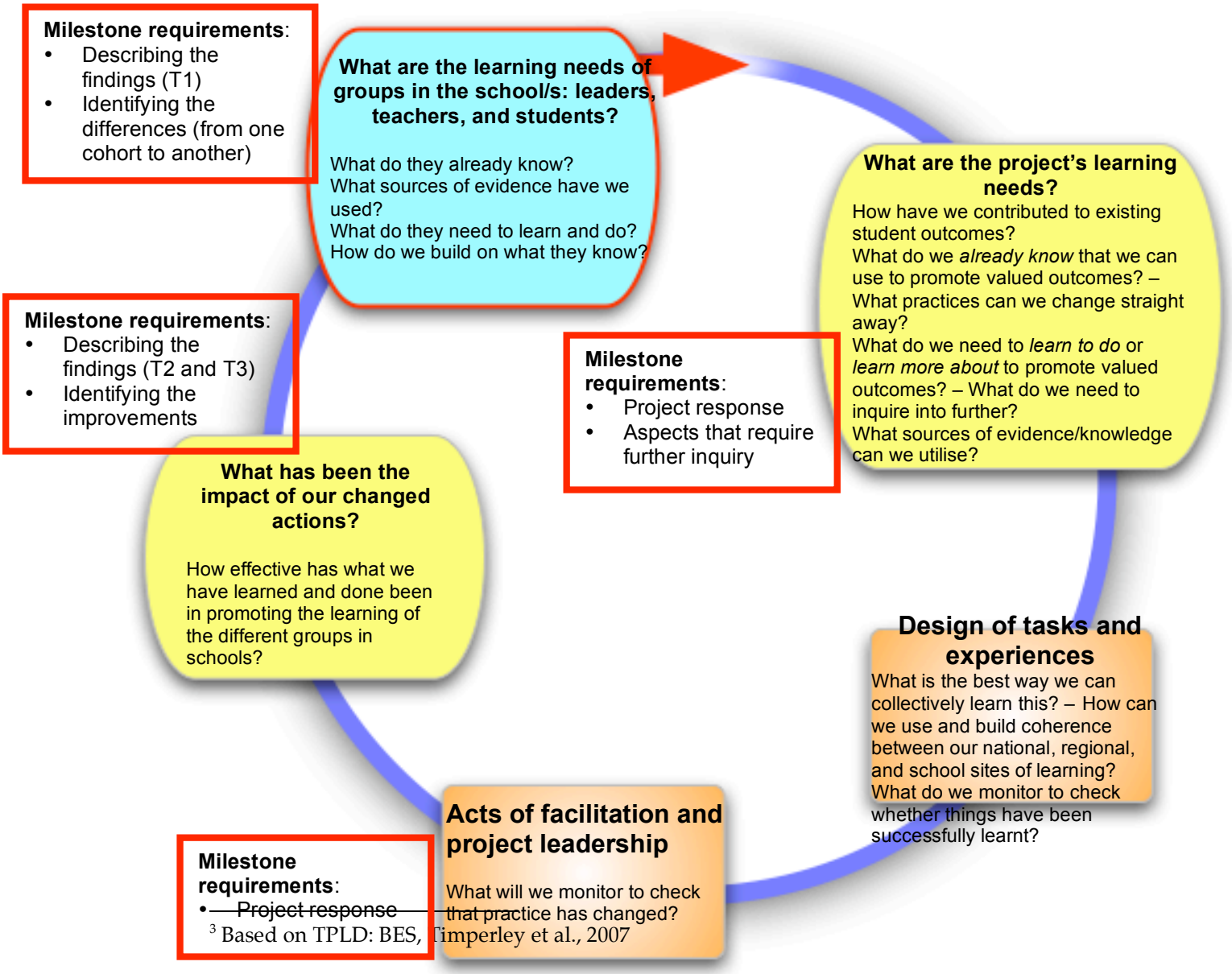
This reporting structure appears to mirror the "Teacher inquiry and knowledge-building cycle to promote valued student outcomes" found in the *Teacher Professional Learning and Development: Best Evidence Synthesis Iteration* ("TPLD BES" – Timperley, Fung, Wilson, and Barrar, 2007).[2] This similarity is shown in the diagram below. The wording has been re-worked to reflect two of the project's core values: it's adherence to the twin notions of "collective accountability" (that is, we believe that we are all responsible for promoting valued student outcomes) and of "collective good". (We believe that professional learning should be available for all so that we are all able to contribute to improving the system.) We hope that the use of this structure will help the reader visualise the project's processes of iterative learning, processes that are based on the assumption that all learning and practice changes must be in response to the analysis of student strengths and needs.

This report will follow the inquiry diagram by briefly describing the key findings from previous cohorts and the project's responses in terms of project learning and changed practices. The details of the project's inquiry over time are summarised in the paper: *Review of Learning, Change and Improvement in and through the Literacy Professional Development Project* (English, Bareta, and Dreaver). The major part of the report focuses on the third question of the inquiry: "What is the impact of our changed

[2] The project was first sat within this cycle in preparation for a paper that was presented to the 2007 Literacy Symposium (Bareta and English, 2007).

actions?" This section will also include changes made in response to the Time 1 and Time 2 findings, as related to each of the project's outcomes. (That is, there is another cycle of inquiry within the cycle, one that involves making comparisons over time.) This then leads us back to the project's responses (Question 2: "What are the project's learning needs?") and the start of another cycle of project learning and action.

Fig 1: Developing project inquiry and knowledge-building to promote valued student outcomes[3]



[3] Based on TPLD: BES, Timperley et al., 2007

Learning Media
Te Pou Taki Kōrero

The project's improvement history

The LPDP has been demonstrably successful in meeting its strategic goal of raising overall student achievement in writing and reading. There have been two recurring themes of improvement over the cohorts. The first is that, after taking into account expected growth and maturation, the gains in reading and writing achievement made by students from schools in the LPDP were equivalent to twice those that could be expected without the intervention. Even more importantly, the schools involved accelerated the rate of progress for many students who were at risk of underachieving. Collectively, their rates of improvement were four times the expected gains for each cohort as a whole.

Nevertheless, analysis of sub-groups of students reveals room for improvement. For example, when we looked at students who were in reading schools in the first two cohorts, we found that 85% of those who began in stanine one were still within the "at risk" band of stanines one–three after two years. Worse, over one third of the students who began in stanine one remained there. We were able to conduct a more detailed analysis of our small second cohort because we had collected data from three points in time. From this we learned that those students who shifted from stanine 1 for reading comprehension did so within the first year. We were left wondering what we needed to do differently to support those students whose reading comprehension had not improved or as the evaluation team wondered whether teachers needed effective specialist support to shift the other students. Māori and Pasifika students and boys have all been over-represented in the lowest 20% at the beginning of each cohort. There has not been a consistent pattern in their achievement patterns, either over the cohorts or between reading and writing. In some year groups and/or literacy foci, Māori students', Pasifika students', and boys' rate of achievement brought them to a cohort mean score. For other groups and/or literacy foci, there was a significant difference in the final mean scores when compared to the cohort as a whole or to the mean score for the lowest 20%. It is these inconsistencies that drive the project to improve all aspects of its theory and practice.

Since the analysis of the outcomes for the first two cohorts, the key foci for learning have been:

- Understanding classroom practices that make a difference for Māori and Pasifika students. The paper presented to the Literacy

Symposium (Bareta and English, 2007) was one of the outcomes of this learning.

- Exploring the usefulness of children's home language backgrounds for teachers to "know their students" and improve student outcomes. The Auckland regional team presented their findings on this issue at the February 2008 seminar.
- Understanding the teaching and learning strengths and needs of new entrant students. Lois Thompson presented a paper on this to the February 2008 seminar.
- Understanding the way teachers can make explicit links between reading and writing. Judy Parr and Melanie Winthrop have explored this at a number of seminars.
- Understanding the concept of "sustainability". Pam O'Connell has presented a number of sessions at seminars. This has been included with the researchers' work.

Major changed practices based on the analysis of findings from the first two cohorts have included:

- Collecting achievement data from all students from the age of five to year 8 at three points of time.
- Increasing the range of tools to collect this national data with the Observation Survey tools (Clay, 2002) used with five and six year olds, *Supplementary Tests of Achievement in Reading* (STAR) with year 3, and asTTle reading and writing with years 4–8.
- Altering outcome four to include the phrase "effectively led" and the addition of a new outcome about effective facilitation.
- Widening the research focus to involve the facilitators' case study schools (generally two per person). This has included research into the "chain of influence" and the impact of facilitator practice analysis conversations with teachers.

The purpose of this report is to evaluate the impact of these changed practices and our new learning (along with the strong project practices already in place) on promoting student outcomes.

An evidence-based evaluation of LPDP's overall effectiveness

The project has five contracted outcomes that collectively acknowledge that learning needs to occur at all levels and that certain conditions need to be created to promote the major strategic goal of improving student outcomes. Hence the outcomes influence each other. The project supports Elmore's (2002) contention that changed teacher, leader, and facilitator practice can only be described as improved if it leads to improved student outcomes. The outcomes are:

- Evidence of improved student achievement
- Evidence of improved teacher content knowledge
- Evidence of improved transfer of understanding of literacy pedagogy to practice
- Evidence of effectively led professional learning communities
- Evidence of effective facilitation.

The project follows an evidence-based inquiry and knowledge-building model that focuses on measurable improved student achievement, described both within a classroom setting and school-wide, as the primary goal.

Schools chose either a reading comprehension or a writing focus while in the project.

A detailed description of the project's improvement theory (theories of learning and theories of action) is found in the paper, *Review of Learning, Change, and Improvement in and through the Literacy Professional Development Project* (English, Bareta, and Dreaver).

A description of the schools involved

127 schools have completed two school years with the project. This particular cohort has been called "February 2006 cohort schools" in previous milestones. The following tables (Tables 1 – 3) give some background information about these schools.

Table 1: Number and percentage of schools with a reading comprehension or writing focus and number and proportion of types of schools within the February 2006 cohort

	Number of schools *Total = 127*	*Percentage of total*
Reading focus	65	51.2
Writing focus	62	48.8
All	**127**	**100**
Full primary	45	35.4
Contributing	67	52.8
Intermediate	11	8.7
Years 7–15	3	2.4
Composite Yrs 1-15	1	0.7
All	**127**	**100**

Table 2: Decile ratings for the schools within the February 2006 cohort

School decile	*Number of schools Total = 127*	*Percentage of total*
1	11	8.7
2	14	11.0
3	17	13.4
4	12	9.4
5	10	7.9
6	12	9.4
7	14	11.0
8	9	7.1
9	9	7.1
10	19	15.0
All	**127**	**100**

Table 3: Size of the reading and writing schools within the February 2004 cohort

Number of teachers in each schools	*Number of reading schools* *Total = 65*	*Number of writing schools* *Total = 62*	*Number of schools*	*Percentage of schools*
0-5	21	15	36	28.3
6-10	23	22	45	35.4
11-20	15	20	35	27.6
21-30	5	5	10	7.9
31+	1	0	1	0.8
All	**65**	**62**	**127**	**100**

The schools with 31+ teachers were all intermediates.

Identifying student outcomes: Evidence of improved student literacy achievement

Outcomes and expectations

The analysis of student data is based on two key assumptions. The first assumption is that all students can achieve at cohort. This means that all year groups as a whole should be within their expected range of achievement. Whatever the achievement pattern for each year group, it is expected to be found in all sub-groups.

Another assumption is that effective teaching can make a difference, and so it is reasonable to expect that teachers can accelerate the progress of students with achievement levels of concern. As described earlier, this has happened in previous cohorts.

The first set of data (Time 1) from the schools in this 2006 cohort made it clear that some year groups were not achieving at cohort and that some sub-groups within each year group had a wide range of patterns that did not reflect the group as a whole. For example, Māori and Pasifika students were over-represented in the lower band of achievement and under-represented in the higher band for each year group in schools with a reading focus. Māori students were over-represented in the lower band of achievement and under-represented in the higher band for each year group in schools with a writing focus.

Based upon a comparison of the mean score data with the mean for each of the tools that were used, it appeared that at the start this cohort had the following patterns:

- Five year olds from schools with either a reading focus or a writing focus and years 4–8 students from schools with a writing focus were performing below expectation
- Years 4–8 students from schools with a reading focus were performing at expectation
- Year 3 students with a reading focus were performing above expectation.

Over time, we have begun to describe what we mean by "achieving at cohort" but the descriptions are very reliant on the particular tool used. All tools used are standardised i.e. there is a norm population to compare to but the tasks have been referenced to different sources of expectation.

Normed and standardised tests means that each test has been designed so that the scores from students used in the norming exercise fit within a bell curve. It is assumed that the population used for norming is large enough to be representative of the total population that the tool is aimed for. Therefore this representation becomes the "expectation" of achievement for future cohorts.

The tasks within the Observation Survey and STAR tools have been referenced to generic reading ability expectations. It is possible to apply the aspirational goals of the project to this generic picture of achievement. Gwenneth Phillips described this with her phrase "breaking the bell" in her presentation for the Literacy Symposium 2007. The project's first goal is to improve the achievement of all students. This would be seen on the bell curve as a shift to the right and a mean greater than 5.0. The second goal is to increase the rate of achievement for the students most at risk of underachieving so they are no longer at risk (that is, to not have any student score within stanines 1–3 of the standardised picture).

There are a number of problems with the use of these tools when the goal is to improve student outcomes. One is that the small range of scores possible can lead a ceiling effect, as described in the March 2006 milestone (Learning Media, 2006) for STAR. Associated with this is the fact that it is difficult to see the true spread of achievement at both ends (that is, there is both a ceiling and a floor effect). This means that the range of scores

within the whole cohort appears to remain consistent over time (that is, an increased spread of achievement is not apparent within the analysis). Another problem is that it is difficult to know whether the aspirational goals are aligned with any of the other expectations that we have for students, such as those described in *The New Zealand Curriculum* (Ministry of Education, 2007).

The assessment tool asTTle is different in two ways pertinent to this discussion. It has normed data that enables people to compare their group of students to a similar group. As with the other tools, it is assumed that because the group of students used in the norming exercise is large enough, it is representative of the population for that particular year group. The first difference is that it has a large scoring range of 100–800, it is possible for students' scores to reflect their achievement of particular tasks. For years 4–8 there are more students within the 100–300 band than there are in the 500–800 band (particularly the earlier years). As a result, there are more students whose scores can only go up than there are whose scores can go down (a floor effect). The other difference is that the asTTle tasks are based on the knowledge and skills that content experts thought were a true reflection of a level, as described in the 2001 New Zealand curriculum document (now *The New Zealand Curriculum* Ministry of Education, 2007). This means the scores can then also be compared to the expectations that are described in the curriculum. However, this leads to another set of problems, as the tool's developers found, some year groups were not achieving at curriculum expectation. This is described in detail in *asTTle technical paper 22* (Auckland UniServices Ltd, 2003).

So for each tool we now have two sets of "cohort expectations". The first is based on the norming process and is a description of the typical picture as it is now and the second is a description of a picture based on improvement. This improvement can be inherent within the tool – such as the Observational Survey tools and STAR – as the reference is to some generic understanding about mastery and ability. Or it can be linked to other national expectations – such as the link asTTle has with the curriculum. Both these aspirational "cohort expectations" may be possible but the second set relies on the curriculum expectations being attainable for teachers and students. It is hoped that the analysis in this milestone will contribute to the discussions both within and outside the Ministry of Education about the following issues:

- Our expectation of achievement in both the short term and long term

- The questions we ask and the tools we use to evaluate the effectiveness of practice (teaching, leadership, and project)
- The language we use to describe the achievement and progress of particular groups of students.

A description of the tools used to identify achievement

The Observational Survey, STAR, and asTTle reading and writing were the tools chosen to collect national, school, and classroom-based student data. They were chosen because they enable people to ask questions about the impact of teaching/leadership/the project on student achievement in relation to "cohort expectations". There is an ongoing expectation that schools will use other assessment tools where appropriate to build a rich description of their students' strengths and needs.

Each assessment tool has been designed for a particular purpose. Furthermore, there are particular theories of literacy, learning, and assessment embedded within each tool. These different purposes and theories mean that there are limitations inherent within each tool. There are also limitations in the way people use each of the tools, often caused by misunderstandings about the theoretical rationale for the tool.

For this cohort, most facilitators were familiar with the STAR and asTTle writing tools as they had used them with all the schools in the earlier cohorts. They were also familiar with the Observation Survey tasks but may not have focused on the way teachers used them. Because of this, time at seminars and regional meetings was spent on understanding the inferences that could be made from individual students' results. The theory behind asTTle reading and about the inferences that can be made from the data were both completely new to facilitators so time at seminars and regional meetings was also spent exploring these aspects of the tool.

For schools to make the most of the literacy data they collect they need to take on an "inquiry way of being", to have deep literacy pedagogical knowledge, and sound data literacy. These knowledge and skills allow them to manipulate the data in ways that enable them to investigate their inquiry questions and respond in a pedagogically wise way. The project leaders try to model this for facilitators and schools in our reports and at national seminar presentations.

STAR

The limitations within this tool and in the way it is used were explored in an earlier milestone (March, 2006). One of the project's responses after analysing the Time 1 data was to inquire into the relationship between the assessment tools being used by schools, especially as students move through the years and only certain tools are available. It is hoped that the *Literacy Learning Progressions* (Ministry of Education, 2007) will be a guiding resource in understanding the information about student achievement and progression from the range of tools available.

asTTle writing

As with STAR, the limitations within this tool and its use were described in an earlier milestone (March, 2006). As facilitators have become more familiar with the tool, they have become critical of the inconsistencies in the language of the indicators. A new set of trial indicators, designed by a MOE funded team, will be used in 2008.

Facilitators and schools have found that students do not automatically transfer their learning about a particular literacy strategy from one writing purpose to another. This means that measuring progress by using a writing purpose that has not been taught may not provide a true reflection of the effectiveness of the teaching or of the students' writing strengths and needs.

Because of the facilitators' increased knowledge of the tool, the schools' knowledge of the tool, and the availability of the moderation guidelines, all student data was used in the national analysis. Within the project there has not been enough cross-moderation between teams to know whether we have any regional variations. The project has relied on each team leader identifying and supporting particular facilitators who are still struggling with the tool. This will be followed up at future seminars and regional meetings.

asTTle reading

Limitations to the tool

Some users appeared to develop a wariness of the tool as a mechanism for showing improvement because the results for whole cohorts did not match their expectations. For example, some schools found a significant number of students increasing by five sub-levels with no other classroom

evidence to support this increased achievement level. Others found that whole year groups decreased their scores over a year significantly more than other schools and/or groups of students with a decrease. One reason for this is that a test may have only one or two questions with a high SOLO taxonomy "rating". If a student gets these questions right, their aRs can be very high. Another reason may be that this is what you should expect to see when looking at longitudinal data with a fairly small cohort and that it is not until you have data over a three or four year period that a real trend will emerge. Often, though, there was no consistent reason within the tool or within the practices around the use of the tool that could explain these inconsistencies.

Limitations in the way the tool was used

Even though there are very clear guidelines and support for data entry, it is still possible to get it wrong. If people are not vigilant and do not compare their findings to what they know about the students from classroom observations, this mistake may not be picked up.

SOLO taxonomy is new to most teachers (and facilitators) and needs to be understood if the most is to be made from the reports at the classroom level. It appears that some schools did not spend enough time developing this knowledge.

Observation Survey tasks

These tasks have been designed to support teachers to notice what a child attends to as they read. It is expected that teachers will develop a profile of each child using the information gained from the whole range of tasks. Clay (2002) describes careful guidelines for use so that the information collected can be compared to a cohort of students at a similar age.

Limitations to the tool

The tasks are based on a theory of how young people come to master reading and writing continuous text. This means that there is a "ceiling" (this is the mastery level). The ceiling is particularly apparent in the following tasks: Letter Identification, Clay Word Reading, Hearing and Recording Sounds, as shown in the graphs in Clay (2002), page 153.

Each set of stanines is for a six-month period, for example, 5.00– 5.50 years. However, there is a huge amount of teaching time difference from 5.10 to 5.50 years. We suggested that schools use the same month for

collecting the data (for example, 5.10 and 6.10 years), assuming that students who have had one month's teaching will have mastered less than those who have had five months teaching.

Limitations in the way the tool was used

The project leaders assumed that the facilitators knew how to use the tool and knew the value of synthesising the data for a whole year group. Many facilitators assumed the schools knew how to administer the tool and knew how to respond to the findings. Both of these assumptions were often incorrect, as was very apparent once the national data for five year olds was collated. Quite a few schools had not collected information about their five year olds. Data was collected from 45 schools with a reading focus (of a possible 56 schools) and from 38 schools with a writing focus (of a possible 47 schools). We could see that many schools had not used the synthesised information, as it often came to us on pieces of handwritten paper. Many schools had used a range of tasks to assess five year olds – some of the tasks were related to the mastery of reading and writing continuous texts while others were not. Some schools had not used the tasks correctly or the scores to stanine conversion correctly. This was picked up by facilitators or by our final checking of the data before analysing it. More schools used the tasks with their six year olds and used them correctly.

Patterns of achievement

In this section, the patterns of achievement will be described according to the year groups and the particular tools used and only for the students who had data at all the national points of collection. The progress over the two years will be described in terms of the means score difference and the effect size. This will be compared to the national picture for the particular tool and to cohort one (February 2004) where possible. The achievement picture at Time 3 will also be described in relation to the "cohort expectation". This will be a comparison to a typical national picture and to the improved/aspirational national picture, as discussed earlier in this report. (See the section on "Outcomes and expectations".)

Five and six year olds

The data from five and six year olds was collected from all the schools but has been analysed according to the focus the school took (that is, all the reading comprehension schools' data was put together and all the writing

schools' data was put together). It was hoped that this would give us some insight into why the schools had chosen their particular focus and whether, at this year level, a professional learning emphasis on one focus or the other would lead to a greater rate of progress for the students. We found that there were a number of differences and some similarities in the two groups and all require further inquiry before making any statement about the professional learning emphasis. In this analysis only the data from students there at the two points in time have been used but it is clear that the use of the tasks improved over the two years.

Schools with a reading comprehension focus:

- Time 1 (as five year olds in 2006): 920 students with one or more task scores from 45 schools. (Letter Identification and Concepts about Print had the highest numbers of students.)
- Time 2 (as six year olds in 2007): 1254 students with one or more task scores from 56 schools. (Letter Identification and Writing Vocabulary had the highest number of students.)

Schools with a writing focus:

- Time 1 (as five year olds in 2006): 909 students with one or more task schools from 38 schools. (Letter Identification and Concepts about Print had the highest number of students.)
- Time 2 (as six year olds in 2007): 1407 students from 47 schools. (Concepts about Print and Writing Vocabulary had the highest number of students.)

What was interesting was that the reading tasks[4] were those used least frequently by both reading and writing focused schools and that the Writing Vocabulary task had the greatest increase in usage by both reading and writing focused schools.

The "cohort expectation" for each task is a bell curve with a mean stanine of five, as seen in Clay (2002). It was expected that the picture would look slightly to the left of this, as the students were assessed within the first month of their fifth or sixth birthday. At Time 1, the difference from the national mean was small for the reading task for writing schools, medium for the reading task for reading schools, and large for the other tasks no matter what the school focus was.

[4] Schools could use either the Duncan Word reading task or the Clay Word reading task.

For both reading and writing focused schools there was a noticeable shift to the right for all tasks and for all sub-groups (that is, the improvement was greater than expected). Even though the tasks have a positive correlation with each other (that is, the students who score high on one are likely to score highly in the others), there was huge variability in the size and shape of the cohort shift depending on:

- The school focus
- The actual task
- Whether we were looking at all students or just the Māori students
- The stanine at Time 1.

Examples of these are described in the sections below

Schools with a reading comprehension focus[5]

1. Variability in the proportion of students still at stanine 1 after a year of schooling.

Only 3% of students who were at stanine 1 for Writing Vocabulary and Word Reading at Time 1 were still in stanine 1 at Time 2 (as shown in Table 4).

Table 4: Reading focus; Percentage of students achieving a particular stanine level at Time 2 according to the stanine they achieved at Time 1 for Writing Vocabulary

% students at T1 stanine		*Stanine T2*								
		1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**
Stanine T1	**1**	**3.33**	20	28.33	20	20	3.33	3.33	1.67	0
	2	0	**13.46**	26.92	19.23	23.08	9.62	3.85	1.92	1.92
	3	0	11.84	**14.47**	32.89	25	9.21	3.95	0	2.63
	4	0	1.25	10	**26.25**	32.5	17.5	7.5	3.75	1.25
	5	0	0	5	8.33	**41.67**	23.33	16.67	5	0
	6	0	0	3.85	7.69	23.08	**19.23**	30.77	7.69	7.69
	7	0	0	0	0	8.33	33.33	**16.67**	33.33	8.33
	8	0	0	0	0	20	20	0	**20**	40
	9	0	0	0	0	0	0	0	0	**0**

[5] The analysis of the data for boys and girls from schools with a reading comprehension focus and the Letter Identification task does not appear to be accurate so has not been included in this report.

Yet 28% of students who were in stanine 1 for Concepts about Print at Time 1 were still in stanine 1 (see Table 5). Nearly 30% from stanines 6 and 7 appear to have slipped back to stanine 1 at Time 2. At this point, this is difficult to understand as it means shifting from a score of 8-15 to 0–7. To use the language of Literacy Learning Progressions, we assume that the development of concepts about print is a constrained skill. A constrained skill is unlikely to be lost, and so we surmise that this negative shift may be an outcome of the way in which the tool is administered.

Table 5: Reading focus; Percentage of students achieving a particular stanine level at Time 2 according to the stanine they achieved at Time 1 for Concepts about Print

% students at T1 stanine		*Stanine T2*								
		1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**
Stanine T1	**1**	**28**	30	8	4	18	6	6	0	0
	2	19.15	**21.28**	10.64	8.51	17.02	8.51	8.51	6.38	0
	3	12.77	13.83	**6.38**	13.83	17.02	11.7	12.77	8.51	3.19
	4	11.83	6.45	10.75	**7.53**	16.13	13.98	23.66	8.6	1.08
	5	2.7	5.41	2.7	10.81	**21.62**	16.22	32.43	5.41	2.7
	6	26.19	4.76	4.76	1.19	13.1	**9.52**	20.24	14.29	5.95
	7	30	0	5	5	20	5	**25**	0	10
	8	0	0	0	0	0	20	20	**0**	60
	9	0	0	0	0	14.29	14.29	14.29	0	**57.14**

2. Variability when comparing Māori student achievement with the whole cohort

Māori students had an accelerated rate of progress compared to the whole cohort for Concepts about Print and Word Reading. The two charts below (Figures 2 and 3) show what this looks like for Word Reading.

Figure 2: Reading focus; Word Reading bar graph presenting frequency of stanine achieved by students at time 1 and 2

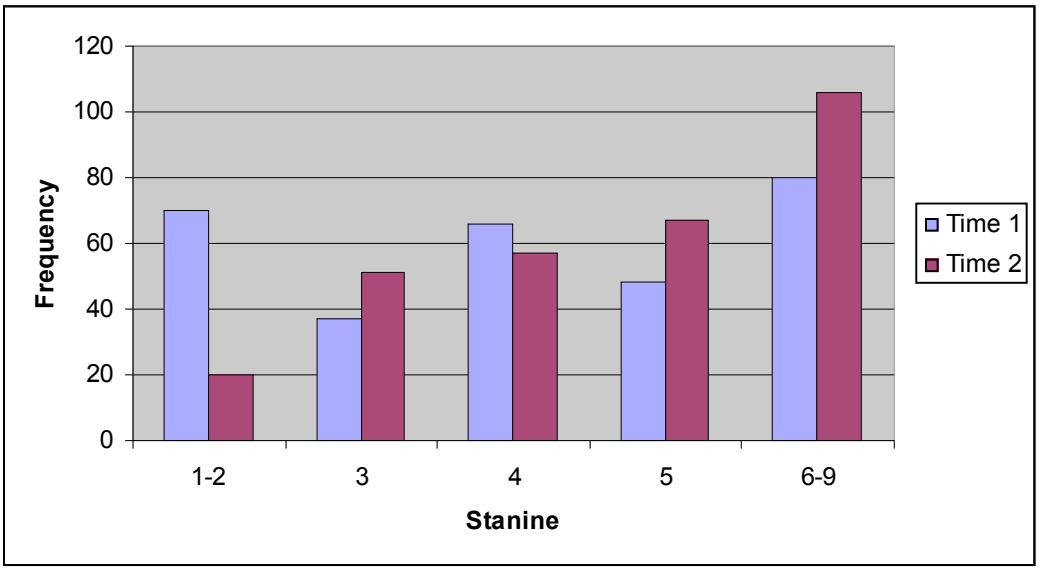
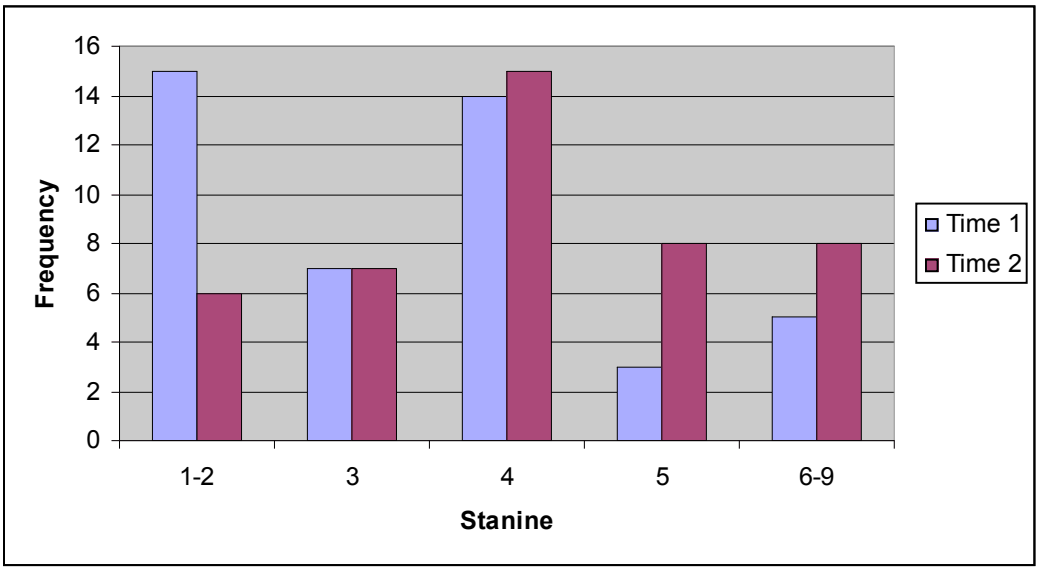


Figure 3: Reading Focus; Word Reading bar graph presenting frequency of stanine achieved by Māori students at time 1 and 2



Even with this accelerated progress for these two tasks, Māori students' mean stanine is still below the whole cohort's mean stanine at Time 2 for all tasks. Worryingly, it is one year behind for Word Reading (that is, their mean for Word Reading as six year olds is equivalent to that of the whole group as five year olds).

3. Variability in the different tasks when compared to a national picture and when comparing rates of progress.

Table 6 below shows the mean stanine and size of shift for each task for the whole group and for Māori students. The range for all students is 0.52–2.52. For Māori students it is 0.66–3.44. For both the whole cohort and the Māori students the relationship amongst the tasks has changed i.e. at Time 1 more students were scoring higher for Word Reading than the other two tasks and then by Time 2 Concepts about Print had more students scoring highly. The mean for Concepts About Print is well above the national mean by Time 2.

Table 6: Reading focus; Mean stanine scores and rate of progress

	All students			*Māori students*		
	Mean T1	*Mean T2*	*Difference*	*Mean T1*	*Mean T2*	*Difference*
Concepts about Print	3.87	6.39	2.52 (n=458)	2.77	6.25	3.44 (n=65)
Word Reading	4.10	4.62	0.52 (n=301)	3.45	4.11	0.66 (n=44)
Writing Vocabulary	3.44	4.99	1.55 (n=392)	2.75	4.28	1.53 (n=59)

Schools with a writing focus[6]

1. Variability in the proportion of students still at stanine 1 after a year of schooling.

11% of students for Writing Vocabulary and 17% for Word Reading at stanine 1 at Time 1 were still in stanine 1 at Time 2.

23% of students in stanine 1 for Letter Identification at Time 1 were still in stanine 1.

[6] The analysis of the boys and girls data from schools with a writing focus appears incomplete so has not been included in this report.

35% of students at stanine 1 for Concepts about Print at Time 1 were still in stanine 1 at Time 2.

These are much higher figures for stanine 1 after one year of schooling than the reading focus schools had.

Tables 7 and 8 show the shifts in Writing Vocabulary and Concepts about Print. These are presented so that the reader can compare the shifts in the writing focused schools with the shifts in the reading focused schools.

Table 7: Writing focus; Percentage of students achieving a particular stanine level at Time 2 according to the stanine they achieved at Time 1 for Writing Vocabulary

% students at T1 stanine		*Stanine T2*								
		1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**
Stanine T1	**1**	**10.91**	25.45	25.45	9.09	23.64	1.82	3.64	0.00	0.00
	2	8.62	**13.79**	17.24	22.41	18.97	8.62	3.45	3.45	3.45
	3	4.35	7.25	**30.43**	20.29	21.74	8.70	2.90	4.35	0.00
	4	3.85	5.77	15.38	**13.46**	32.69	17.31	5.77	3.85	1.92
	5	5.41	5.41	5.41	16.22	**18.92**	21.62	16.22	8.11	2.70
	6	5.00	0.00	10.00	10.00	27.50	**15.00**	12.50	17.50	2.50
	7	5.00	15.00	15.00	20.00	5.00	10.00	**20.00**	5.00	5.00
	8	0.00	0.00	0.00	25.00	50.00	25.00	0.00	**0.00**	0.00
	9	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	**0.00**

Table 8: Writing focus; Percentage of students achieving a particular stanine level at Time 2 according to the stanine they achieved at Time 1 for Concepts about Print

% students at T1 stanine		*Stanine T2*								
		1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**
Stanine T1	**1**	**35.35**	16.16	18.18	11.11	6.06	5.05	4.04	3.03	1.01
	2	22.08	**14.29**	9.09	14.29	22.08	6.49	9.09	2.60	0.00
	3	12.90	11.29	**19.35**	14.52	17.74	6.45	11.29	4.84	1.61
	4	12.12	4.55	10.61	**3.03**	25.76	10.61	12.12	15.15	6.06
	5	4.35	4.35	4.35	8.70	**26.09**	13.04	30.43	4.35	4.35
	6	9.38	1.56	1.56	1.56	14.06	**15.63**	43.75	7.81	4.69
	7	15.63	0.00	6.25	6.25	0.00	6.25	**31.25**	28.13	6.25
	8	11.11	11.11	0.00	0.00	0.00	0.00	22.22	**33.33**	22.22
	9	10.00	0.00	0.00	0.00	0.00	10.00	20.00	20.00	**40.00**

2. Variability when comparing Māori student achievement with the whole cohort

Māori students had an accelerated rate of progress compared to the whole cohort for Concepts about Print and Letter Identification.

The rate of progress for the cohort as a whole was five times that of Māori students for Word Reading. (See Figures 7 and 8) These charts have been included so that they can be compared to the same charts for the reading focused schools.

Māori students' mean stanine is still below the whole cohort's mean stanine at Time 2 for all tasks and, most worryingly, one year behind for all tasks except Letter Identification.

Figure 7: Writing focus; Frequency of all students at each stanine at Time 1 and Time 2 for Word Reading

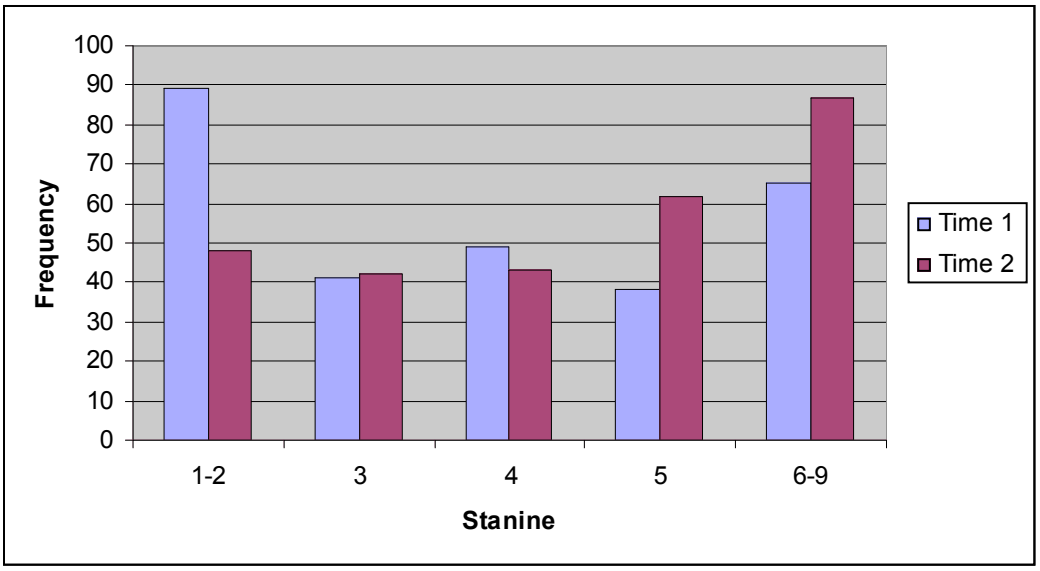
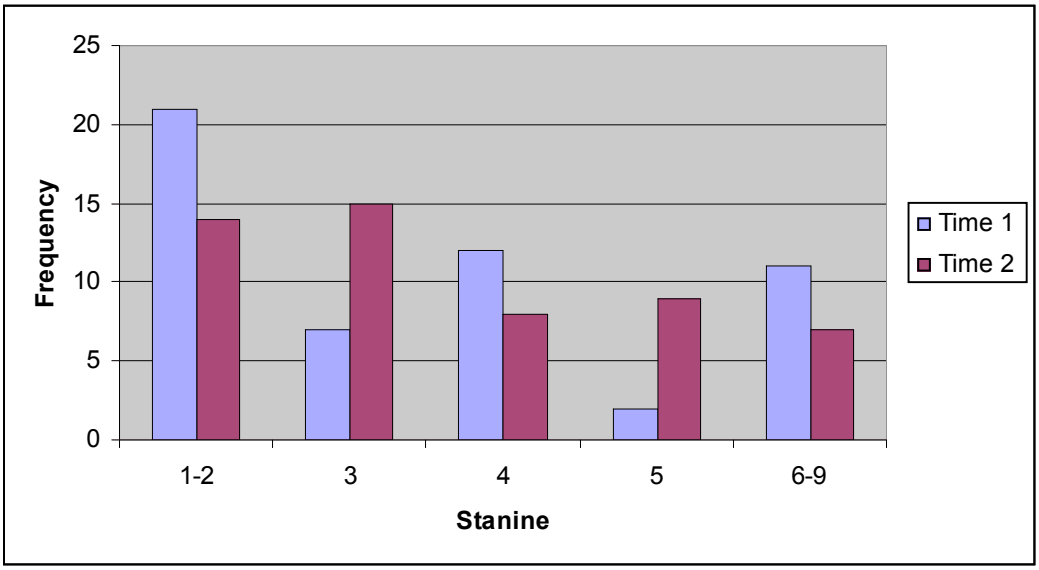


Figure 8: Writing focus; Frequency of Māori students at each stanine at Time 1 and Time 2 for Word Reading



3. Variability in the different tasks when compared to a national picture and when comparing rates of progress

The table below (Table 9) shows the mean stanine and size of the shift for each task for the whole group and for Māori students. The range for all students is 0.53–1.80. For Māori students the range is 0.09–2.21. Again, the mean score relationship among the tasks has changed from Time 1 to Time 2. For example, for Māori students more student had a higher score in Word Reading at Time 1 than in the other tasks and Letter Identification at Time 2.

Table 9: Writing focus; Mean stanine scores and rate of progress

	All students			*Māori students*		
	Mean T1	*Mean T2*	*Difference*	*Mean T1*	*Mean T2*	*Difference*
Concepts about Print	3.60	4.85	0.98 (n=435)	2.60	3.65	1.05 (n=91)
Word Reading	3.82	4.35	0.53 (n=282)	3.53	3.62	0.09 (n=53)
Writing Vocabulary	3.53	4.39	0.86 (n=335)	2.98	3.62	0.64 (n=58)
Letter Identification	3.67	5.47	1.80 (n=393)	2.77	4.98	2.21 (n=73)

Pictures across the different foci

There were three unexpected patterns that require further work to make sense of. These have not been described in the Executive Summary, as they need a lot more work done around them.

- The Time 1 mean scores were lower for the writing focused schools than the reading focused schools for all tasks except Writing Vocabulary. Unexpectedly, the rate of progress in Writing Vocabulary was less in writing schools than reading schools.
- Another conundrum is that the rate of progress in reading focused schools was greater than that in writing focused schools except for the whole groups' reading rate.
- And a third concern is that Māori students from writing focused schools had a mean stanine score 0.5 to 2.6 below that of their peers in reading focused schools, with both a lower Time 1 mean and a much lower rate of progress.

Year 3 students

In 2006, data on year 3 students was collected from schools with a reading focus (Time 1 and Time 2). This was discussed in the July 2007 milestone (Learning Media, 2007), along with the Time1–Time 2 shifts for years 4–8 and the analysis of the 2006 data on five year olds.

Year 4–7 students

The year identified describes the cohort at the beginning of the project. For example, year 4 means students in year 4 in 2006.

Schools with a reading focus:

- 3871 students with data at Time 1, Time 2, and Time 3
- The retention rates dropped for years 4, 5, and 7 in the following magnitude: Time 2: 92–81%, Time 3: 83-77%
- For year 6 the retention rate was Time 2: 74%, Time 3: 25%, as many of the schools were contributing schools
- Over the two years, the schools retained fewer Māori students and fewer from the lowest 20% (approximately 2% less than the cohort as a whole).

Schools with a writing focus:

- 3422 students with data at Time 1, Time 2, and Time 3
- The retention rates dropped for years 4, 5, and 7 in the following magnitude: Time 2: 90–87%, Time 3: 82–74%
- For year 6 the retention rate was Time 2: 75%, Time 3: 22%
- Over the two years, schools retained fewer Māori students (approximately 2% less than the cohort as a whole), and fewer from the lowest 20% (approximately 0.5 % less than the cohort as a whole).

A summary of the overall rates of progress shows:

- For all year groups, the effect size over the two years was greater than expected without an intervention. For years 4–6, the average increase in score was more than twice what would be generally expected. The average effect sizes were: reading 0.96, writing 1.05. The average expected effect sizes were: reading 0.69, writing 0.47. The writing mean effect size for the February 2004 cohort was 1.27.
- For all year groups, the rate of progress for those in the lowest 20% at Time 1 was double that of the cohort as a whole. The average effect sizes were: reading 1.79, writing 2.53. The writing mean effect size for the lowest 20% in the February 2004 cohort was 2.05.
- The rate of progress was above expectation for all ethnic groups (Pakeha, Pasifika, Asian, Other, and Other European students) in schools with a reading focus, except for Year 6 and 7 Māori students. The average effect size for Pasifika students was 1.30. The average effect size for Māori students was 0.73
- There was an accelerated rate of progress for Pasifika students for all year groups in schools with a writing focus. The average effect size was 1.15
- There was an accelerated rate of progress for Māori students for all year groups in schools with a writing focus. The average effect size was writing 1.05.
- There was an accelerated rate of progress for boys in all year groups, both for schools with a reading focus as well as schools with a writing focus. The average effect sizes were: reading 0.95, writing 1.09. These are in line with the overall effect size for the whole cohort.

- Students in a school in year 2 of the project achieve higher than students in year 1 of the project. For example the year 5 students in 2007 have a higher average mean score than the year 5 2006 students.
- The rate of progress is generally the same in both years for writing focused schools and around double that of the first year for reading focused schools in the second year of the project.

A summary of the overall improvement

Comparisons to the norm within the tool

- All year groups are achieving at or above "cohort expectation". The mean score for students in the lowest 20% for all year groups is now closer to each whole group's mean and is close to or better than the national picture.
- All ethnic groups at all year groups, except Māori in Years 6 and 7, are achieving at "cohort expectation" or better when compared to the national picture for reading focused schools.
- All ethnic groups at all year levels are achieving at "cohort expectation" or better when compared to the national picture for writing focused schools.
- In all year groups, both boys and girls are achieving at "cohort expectation" or better when compared to the national picture.

The following tables and graphs show this visually. The tables show the effect size. Hedge's corrected effect size is used as this corrects a bias introduced in small sample sizes (for the majority of these, it makes no change as the sample size is large enough). Descriptive statistics are also shown, along with a t-test to check for differences in the means from time 1 to time 3 (as an effect size does not always show the full picture).

The graphs use the mean asTTle score for each time to show the rate of progress for each cohort and then the sub-groups. Year 6 Pasifika student numbers are very low (12 from schools with a reading focus and 14 from schools with a writing focus) so have not been included in the graphs. Asian students' and Other European students' data have not been analysed this way as the groups are not really homogenous.

The first set of analysis is from the reading focused schools and the second set from the writing focused schools. Both sets have the whole group first then the gender analysis followed by the ethnicity analysis and the lowest 20% at Time 1.

Table 10: asTTle reading; descriptive statistics showing mean total score for students with data in all time points, according to year level.

Group	*Year*	*Mean T1*	*n*	*SD*	*Mean T3*	*SD*	*t-value for mean diff*	*Hedge's Corrected Effect Size*	*St. Error of Effect Size*	*Mean effect size*[7]
Overall	4	412.17	1059	77.77	490	76.4	0	**1.01**	0.05	0.96
Overall	5	442	932	78.68	512.36	76.7	0	**0.91**	0.05	
Overall	6	469.91	335	75.9	540.91	90.8	0	**0.85**	0.08	
Overall	7	513.1	1545	82.8	613.97	106.2	0	**1.06**	0.04	
Male	4	406.82	522	81	485.85	79.99	0	**0.98**	0.07	0.95
Male	5	432.11	488	81.8	503.72	76.7	0	**0.9**	0.07	
Male	6	459.35	185	75.8	534.24	94.6	0	**0.87**	0.11	
Male	7	502	832	82.8	600.16	106.4	0	**1.03**	0.05	
Female	4	417.5	534	74	494.41	72.6	0	**1.05**	0.07	0.98
Female	5	452.92	441	73.7	522	75.9	0	**0.92**	0.07	
Female	6	483.43	148	74.6	549.18	86.1	0	**0.81**	0.12	
Female	7	526.14	711	81.1	630.21	103.9	0	**1.12**	0.06	
NZ Euro	4	423.21	685	78.397	499.11	74.912	0	**0.99**	0.06	1.03
NZ Euro	5	450.81	583	75.435	520.83	75.699	0	**0.93**	0.06	
NZ Euro	6	480.06	214	74.782	559.41	85.865	0	**0.98**	0.1	
NZ Euro	7	531.57	962	78.008	640.37	99.05	0	**1.22**	0.05	
NZ Māori	4	377.54	178	65.433	444.34	65.297	0	**1.02**	0.11	0.73
NZ Māori	5	404.51	170	76.767	468.04	75.536	0	**0.83**	0.11	
NZ Māori	6	441.71	87	70.801	482.29	81.951	0.0006	**0.53**	0.15	
NZ Māori	7	480.23	334	77.934	531.92	104.7	0	**0.56**	0.08	
Pasifika	4	360.63	60	66.152	465.75	84.137	0	**1.38**	0.2	1.3
Pasifika	5	390.6	47	85.42	490.32	63.101	0	**1.32**	0.23	
Pasifika	6	410.58	12	92.051	532.11	66.83	0.0012	**1.46**	0.46	
Pasifika	7	451.24	83	68.231	531.5	85.99	0	**1.03**	0.17	
Asian	4	416.3	66	77.637	506.92	65.098	0	**1.26**	0.19	1.33
Asian	5	458.78	50	62.168	527.46	61.064	0	**1.11**	0.21	
Asian	6	506.33	9	31.698	587.44	59.307	0.0023	**1.62**	0.54	
Asian	7	501.52	85	80.519	606.15	73.6	0	**1.35**	0.17	
Lowest	4	302.4	205	41.457	420.58	67.329	0	**2.11**	0.12	1.79
Lowest	5	332.59	199	50.23	443.82	63.342	0	**1.94**	0.12	
Lowest	6	365.32	78	46.37	476.47	77.378	0	**1.73**	0.19	
Lowest	7	396.99	277	56.942	495.49	81.938	0	**1.39**	0.09	

[7] The asTTle national effect sizes are years 4 – 5 is 0.77, years 5 – 6 is 0.46, years 6-7 is 0.28 and years 7 – 8 is 1.26

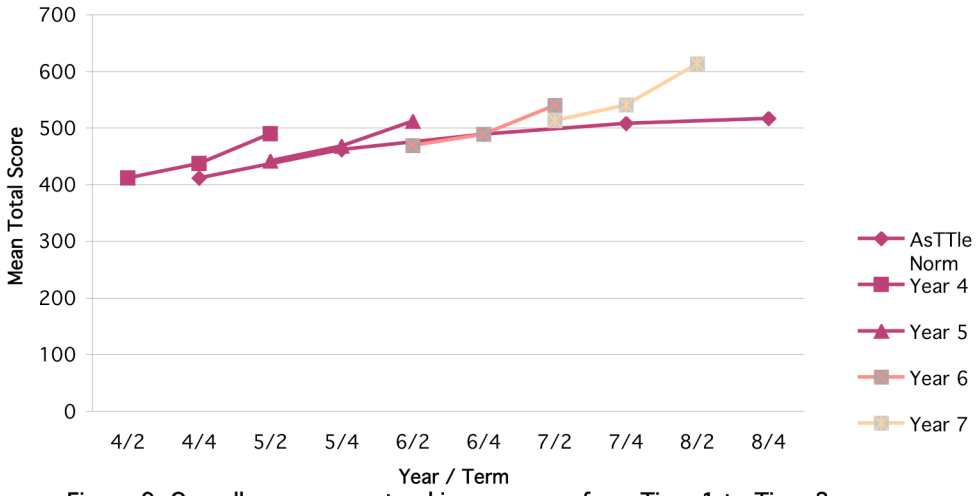


Figure 9. Overall mean score tracking progress from Time 1 to Time 3 for each year group compared with AsTTle norms.

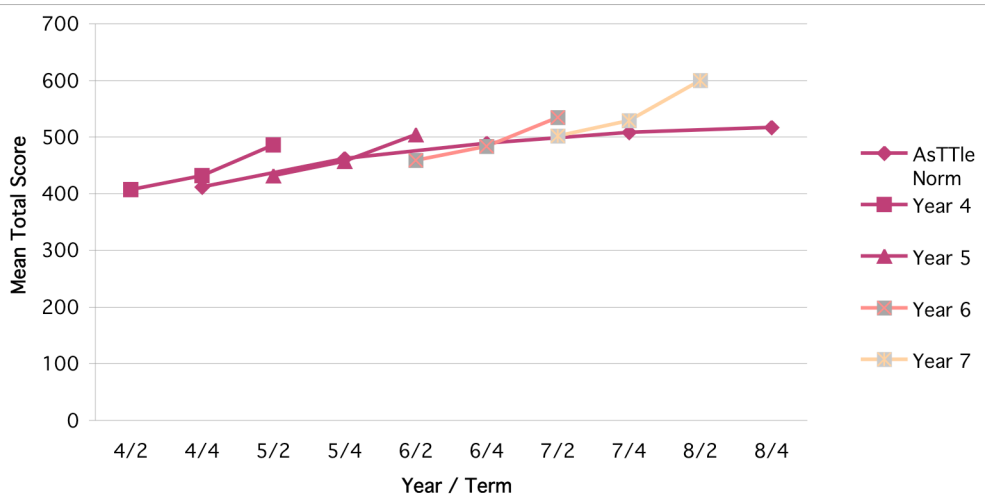


Figure 10. Overall mean score tracking progress from Time 1 to Time 3 for boys in years 4-7 compared with AsTTle norms.

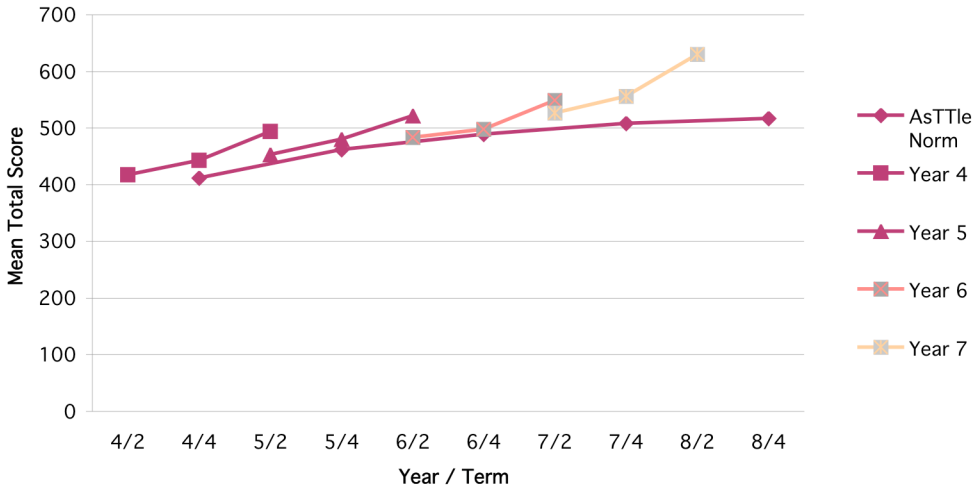


Figure 11. Overall mean score tracking progress from Time 1 to Time 3 for girls in years 4-7 compared with AsTTle norms.

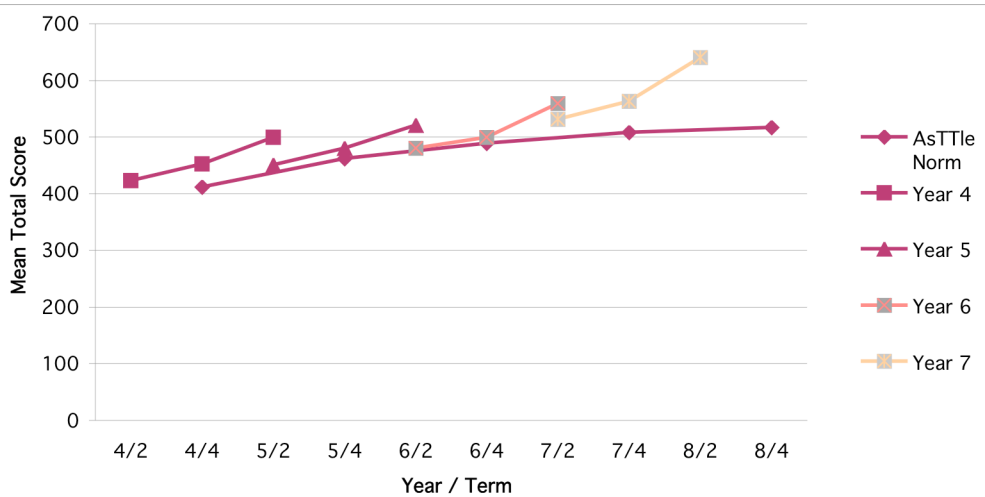


Figure 12. Overall mean score tracking progress from Time 1 to Time 3 for NZ European in years 4-7 compared with AsTTle norms.

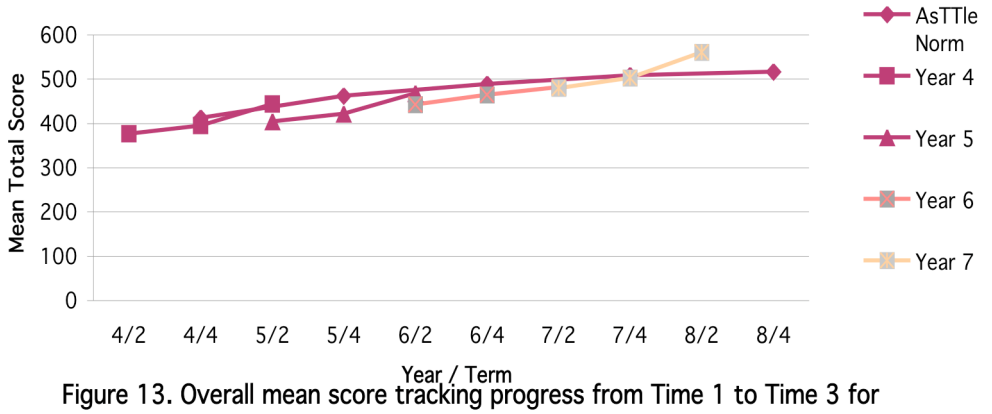


Figure 13. Overall mean score tracking progress from Time 1 to Time 3 for NZ Maori in years 4-7 compared with AsTTle norms.

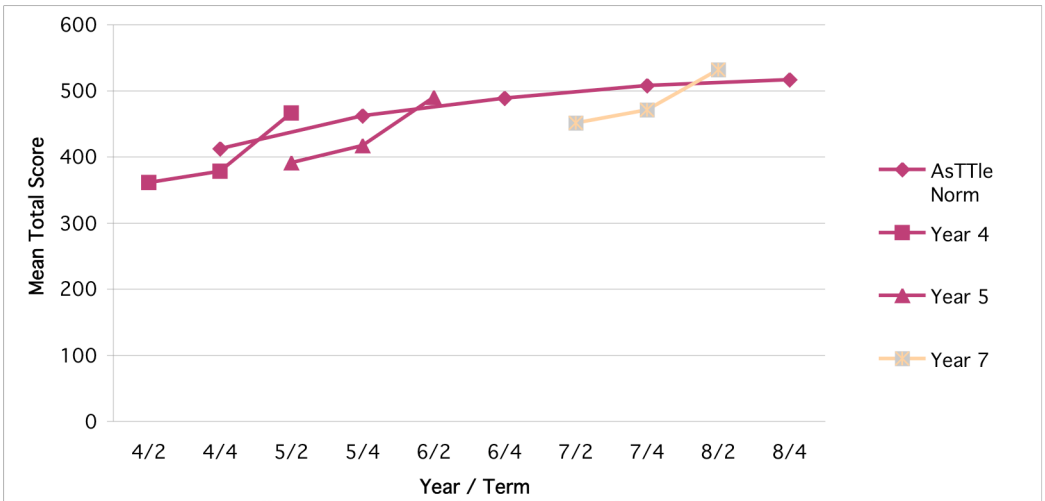


Figure 14. Overall mean score tracking progress from Time 1 to Time 3 for Pasifika in years 4-7 compared with AsTTle norms.

Year 6 analysis has not been shown, as there are only 12 students. The pattern does match the other year groups.

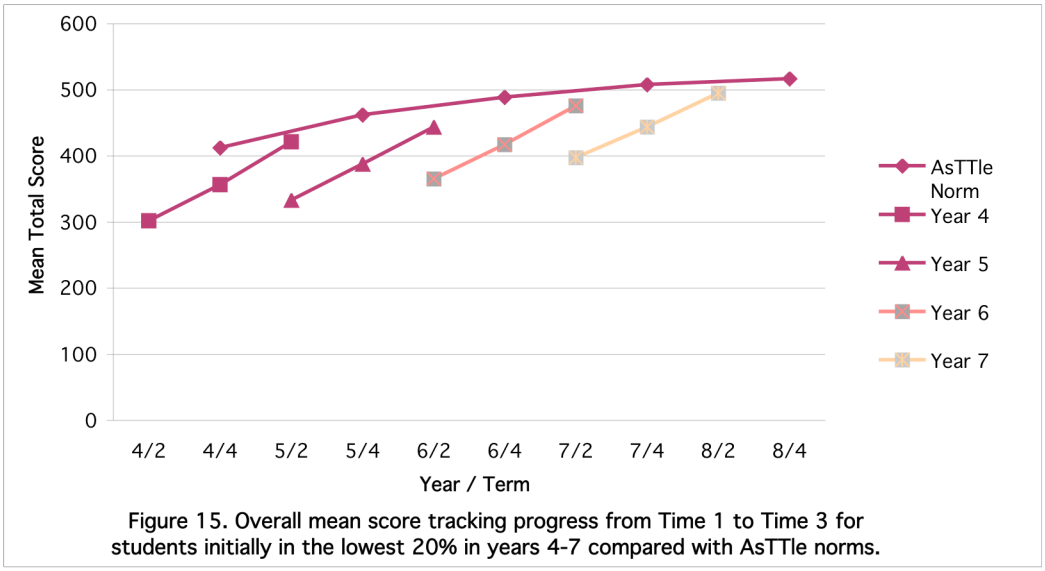


Figure 15. Overall mean score tracking progress from Time 1 to Time 3 for students initially in the lowest 20% in years 4-7 compared with AsTTle norms.

Writing focus

Table 11: Writing focus; descriptive statistics showing mean total score for students with data in all time points, according to year level.

DATA		Time 1			Time 3		p-value for mean diff (2-tailed T-test)	Hedges Corrected Effect Size	Standard Error of E.S. estimate	Average Effect Sizes[8]
Group	**Year**	Mean	n	SD	Mean	SD				
Overall	4	361.72	1318	123.85	493.42	110.98	0.0000	**1.12**	0.04	1.05
Overall	5	411.49	1331	120.44	527.91	102.46	0.0000	**1.04**	0.04	
Overall	6	461.51	348	116.84	577.72	110.11	0.0000	**1.02**	0.08	
Overall	7	467.85	445	140.04	595.28	103.03	0.0000	**1.04**	0.07	
Male	4	337.95	669	125.12	470.68	108.14	0.0000	**1.13**	0.06	1.09
Male	5	387.42	661	120.70	507.02	100.23	0.0000	**1.08**	0.06	
Male	6	441.83	162	115.55	562.26	110.83	0.0000	**1.06**	0.12	
Male	7	436.7	243	142.91	572.21	102.39	0.0000	**1.09**	0.10	
Female	4	386.23	649	117.68	516.85	107.23	0.0000	**1.16**	0.06	1.06
Female	5	435.24	670	115.44	548.51	100.52	0.0000	**1.05**	0.06	
Female	6	478.64	186	115.55	591.18	107.98	0.0000	**1.00**	0.11	
Female	7	505.32	202	127.07	623.03	97.01	0.0000	**1.04**	0.11	
NZ Euro	4	370.6	782	118.63	503.19	109.24	0.0000	**1.16**	0.05	1.06
NZ Euro	5	420.82	701	126.13	536.85	100.03	0.0000	**1.02**	0.06	
NZ Euro	6	480.95	197	103.44	596.88	108.33	0.0000	**1.09**	0.11	
NZ Euro	7	498.47	190	140.99	616.95	105.00	0.0000	**0.95**	0.11	
NZ Māori	4	329.83	248	130.44	455.43	120.05	0.0000	**1.00**	0.10	1.05
NZ Māori	5	378.53	306	109.13	491.29	107.28	0.0000	**1.04**	0.09	
NZ Māori	6	423.87	96	126.68	546.97	103.90	0.0000	**1.06**	0.15	
NZ Māori	7	452.89	146	128.35	574.12	92.03	0.0000	**1.08**	0.13	
Pasifika	4	325.73	124	128.91	473.07	93.44	0.0000	**1.30**	0.14	1.15
Pasifika	5	406.04	148	118.98	518.21	97.71	0.0000	**1.03**	0.12	
Pasifika	6	415.93	14	93.87	526.5	124.14	0.0133	**0.98**	0.40	
Pasifika	7	416.95	55	134.64	571.91	100.75	0.0000	**1.29**	0.21	
Asian	4	411.16	90	108.32	523.4	95.66	0.0000	**1.09**	0.16	1.16
Asian	5	448.69	111	103.12	574.32	91.41	0.0000	**1.28**	0.15	
Asian	6	454.81	16	111.85	564	111.29	0.0096	**0.95**	0.37	
Asian	7	454.15	34	133.64	616.79	112.19	0.0000	**1.30**	0.27	
Lowest	4	166.76	257	64.60	397.47	105.01	0.0000	**2.64**	0.12	2.53
Lowest	5	222.61	226	76.78	439.72	98.70	0.0000	**2.45**	0.12	
Lowest	6	292.93	71	69.92	480.07	109.62	0.0000	**2.02**	0.21	
Lowest	7	252.36	84	81.00	506.12	87.30	0.0000	**3.00**	0.22	

[8] asTTle national effect sizes; years 4 – 5 is 0.5, years 5 – 6 is 0.36, years 6 – 7 is 0.32 and years 7 – 8 is 0.72.

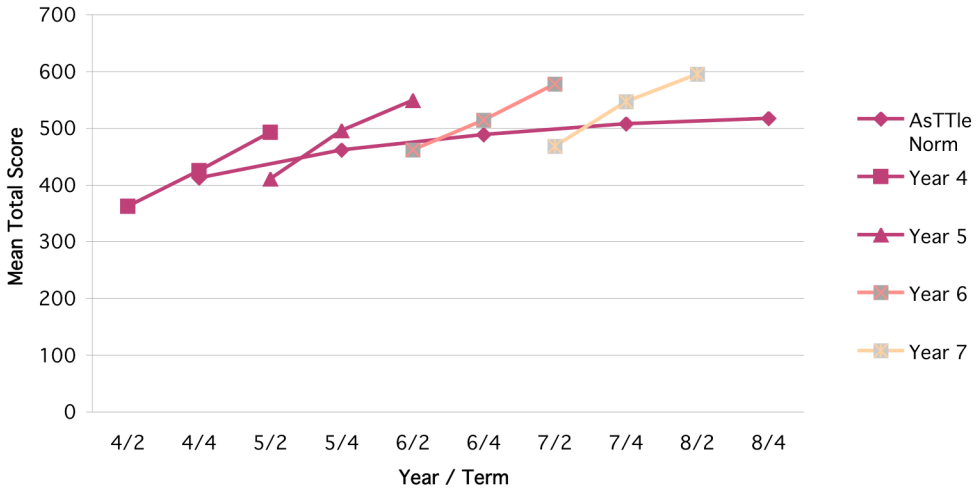


Figure 16. Overall mean score tracking progress from Time 1 to Time 3 for each year group compared with AsTTle norms.

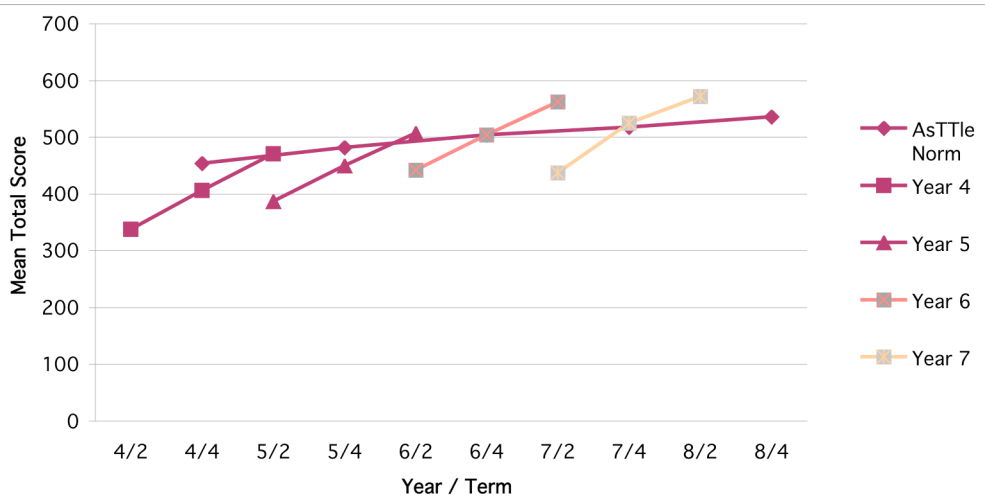


Figure 17. Overall mean score tracking progress from Time 1 to Time 3 for boys in years 4-7 compared with AsTTle norms.

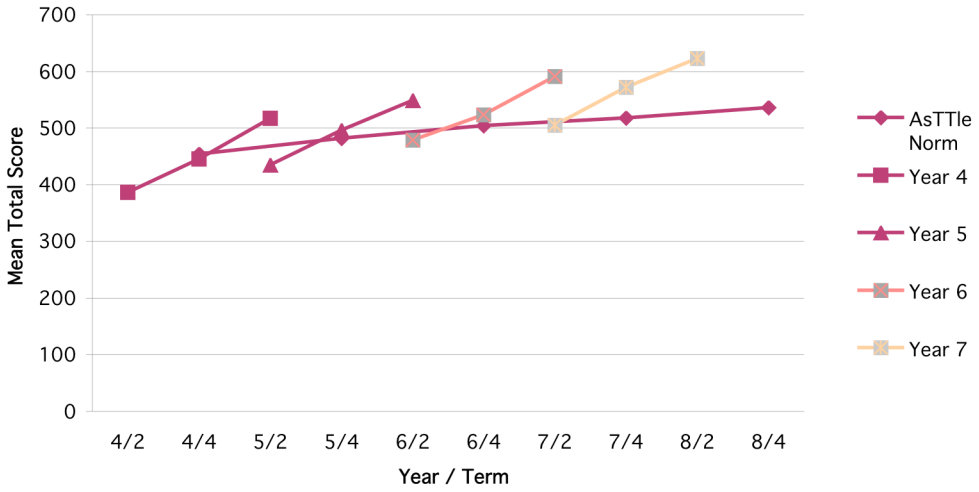


Figure 18. Overall mean score tracking progress from Time 1 to Time 3 for girls in years 4-7 compared with AsTTle norms.

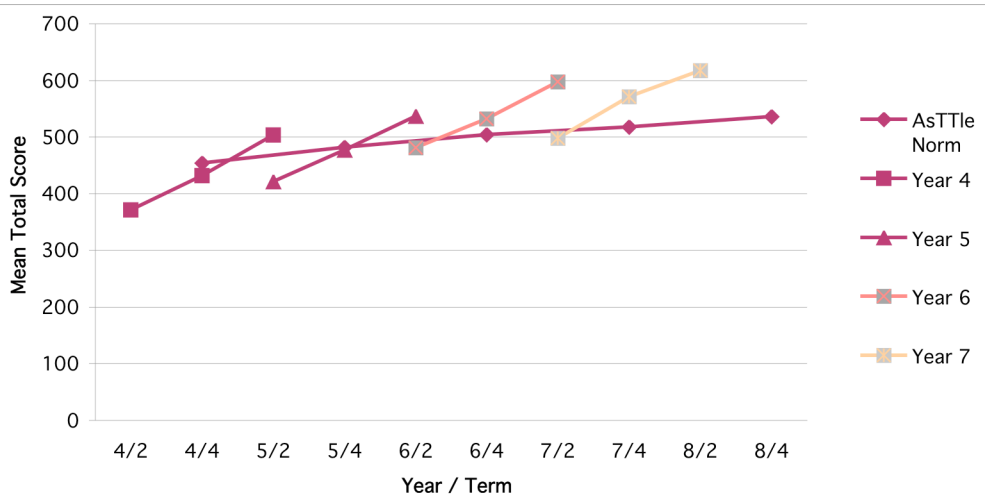


Figure 19. Overall mean score tracking progress from Time 1 to Time 3 for NZ European in years 4-7 compared with AsTTle norms.

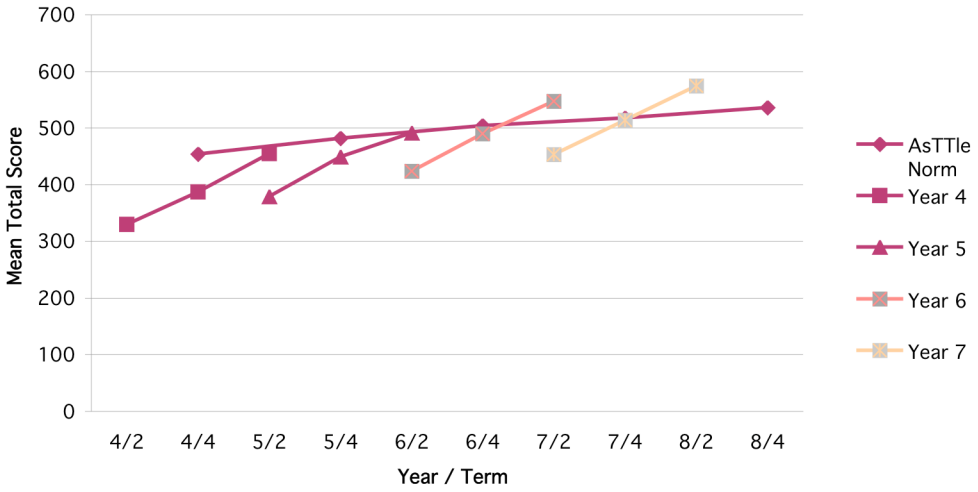


Figure 20. Overall mean score tracking progress from Time 1 to Time 3 for NZ Maori in years 4-7 compared with AsTTle norms.

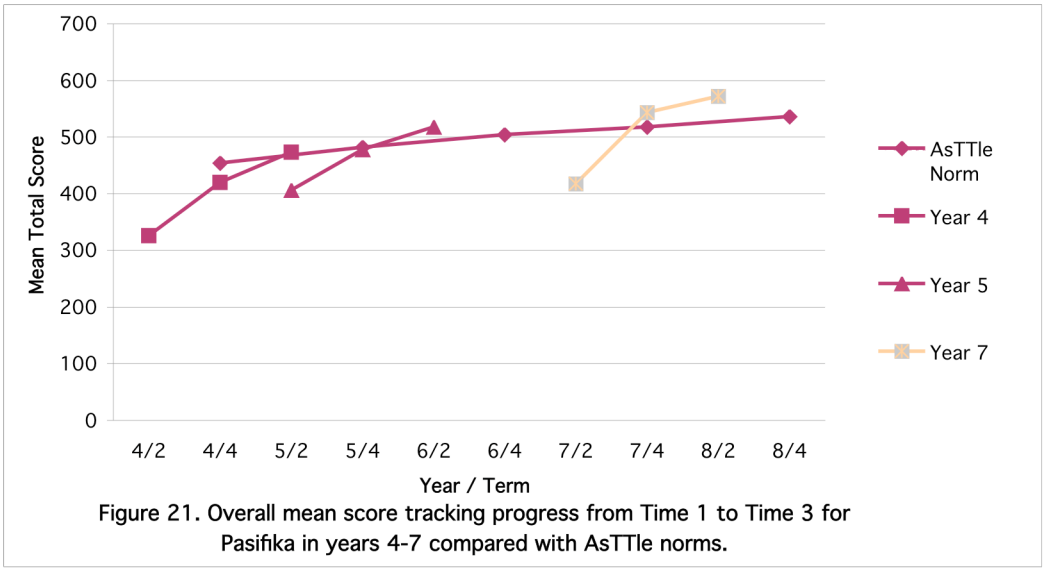


Figure 21. Overall mean score tracking progress from Time 1 to Time 3 for Pasifika in years 4-7 compared with AsTTle norms.

The year 6 data was removed from this analysis, as the numbers were too small to be reliably seen as representing a pattern.

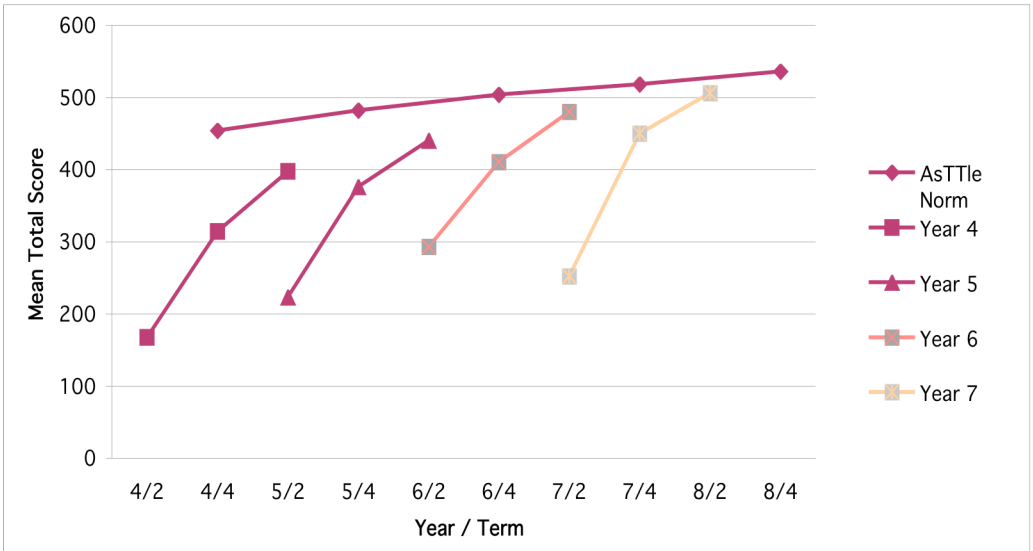


Figure 22. Overall mean score tracking progress from Time 1 to Time 3 for students initially in the lowest 20% in years 4-7 compared with AsTTle norms.

Comparisons to the New Zealand curriculum expectations

To undertake this comparison *The New Zealand Curriculum* "Years and Curriculum Levels" chart (MOE 2007 p 45) has been re-interpreted through the vertical year levels (rather than the curriculum levels). The proportion of students at each curriculum sublevel for Time 1, Time 2 and Time 3 has been mapped onto this interpretation of the year and curriculum levels. The findings are:

- The range of scores increases over the years for both reading and writing schools.
- The proportion of students achieving below the New Zealand curriculum "cohort expectation" increased from year 4 to year 8.
- The proportion of students achieving below New Zealand curriculum "cohort expectation" in the second year of the project is lower than the proportion below New Zealand curriculum "cohort expectation" in the first year of the project for the same year group. (For example, there were fewer year 7 students below expectation in 2007 than there were in 2006.) This suggests that participation in the project has enabled the schools to reduce the proportion of students achieving below New Zealand curriculum "cohort expectation".

The project's response

Our learning foci

To identify the effectiveness of an improvement based project there needs to be an understanding of the achievement levels the project is pursuing, as well as the increased rate of achievement certain groups need to make so as not to be left behind.

The inquiry is now more about the rate of achievement that is needed for a whole cohort if they achieving at cohort expectation. So the first question that follows is "What does 'achieving at cohort' mean?" This was discussed earlier in this milestone. In this analysis, it was decided to look very closely at the proportion of students at each curriculum sub-level (this was only possible with asTTle data) and compare this to the expectations set in *The New Zealand Curriculum*. This made us notice a couple of trends (as described in the section above) that we knew were there (see the Time 1 and Time 2 analyses for this cohort) but we did not have the detail of the trend and were uncertain of our contribution to it.

This analysis leads us to believe that the project should inquire into a number of related issues:

- From a project perspective we need to consider what we need to do if we are to continue in our quest to ask the hard questions and add depth to our understanding of the situation. This includes identifying who can help us in our learning and how that learning might take place.
- Secondly, from a project perspective we need to respond to the "cohort expectation" analysis and ensure that our work with schools and their self-review processes leads to a long-term increase in the rate of progress. We need to ensure that in the future our year 8 students are not as far below the New Zealand curriculum "cohort expectation" as this group is, who have only experienced a two-year focus on literacy. We also know that many Māori students are one year behind mean for a range of reading and writing behaviours after only one year at school. Taken together, we need to continue to support schools to think longitudinally and in a more sophisticated way about their specific targets.

Changed practices

The project has put many changed practices into place for the new cohort. This has been in response to learning from within the project such as the analysis of the Time 1 – Time 2 data (students, teachers, leaders, and facilitators), our research findings, our learning from external research, and our understandings of what is important in the wider policy context. Our understanding of the key concept of "sustainability" now includes ensuring that there is coherence between learning activities intended to communicate the "big ideas" of education and that the transfer of these ideas to teaching practice is explicit. This is evident in the following changes:

- The project is focusing on using each school's planning and reporting processes as a mechanism for helping to identify their school's strengths and needs so that they can use this knowledge as they plan what they will do to improve the literacy outcomes of all students. This includes supporting school leaders to see their role in this and to see that the process involves the ongoing monitoring of and response to the students who are most at risk.

- The project is supporting schools to think about the way they use effective literacy specialist support. The goal is to ensure that they that link specialist support to classroom practice so that the students can transfer their learning from outside the classroom to the classroom. We believe that this will lower the risk of slippage in achievement.
- The project is supporting schools and teachers to be explicit about the links between reading and writing so that students are helped to transfer the learning from one area to the other.

Identifying the effectiveness of the learning and changed practices

The planning and reporting documentation, including Education Review Office reports, will be a critical aspect of monitoring the impact of the project on student outcomes. This will be a key part of facilitator milestones from now on. It is hoped that we will see more emphasis on all student achieving at cohort expectation, a longitudinal view of progress and an action plan that reflects an improvement theory that shows continuous sophistication.

References

Auckland UniServices Ltd (2003). *asTTle technical paper 22: Validation Evidence of asTTle Reading Assessment Results: Norms and Criteria. Auckland: University of Auckland.* Available at www.tki.org.nz/r/asttle/pdf/technical-reports/techreport22.pdf

Bareta, L. and English, C. (2007). "Evidence-based Inquiry: A Collective Responsibility." Wellington: Learning Media. Paper presented at the August 2007 Literacy Symposium. "Lifting the Achievement of the Underachieving Reader and Writer in the New Zealand Mainstream Classroom."

Clay, M. M. (2002). *An Observation Survey of Early Literacy Achievement*. (2nd ed.) Portsmouth, NH: Heinemann.

Elmore, R.F. (2002a). "Bridging the Gap Between Standards and Achievement: The Imperative for Professional Development in Education". Published by the Albert Shanker Institute. Available at www.shankerinstitute.org/Downloads/Bridging_Gap.pdf Reprinted in Elmore, R. (2005). *School Reform from the Inside Out: Policy, Practice, and Performance*. Cambridge, MA: Harvard University Press.

English, Bareta, and Dreaver (2007). "Review of Learning, Change and Improvement in and through the Literacy Professional Development Project." Wellington: Learning Media.

Learning Media (March 2006). "Literacy Professional Development Project: Milestone 1". Wellington: Learning Media. Available at www.tki.org.nz/r/literacy_numeracy/pdf/literacy_professional_development_project.pdf

Learning Media (July 2007). "Literacy Professional Development Project: Milestone 2". Wellington: Learning Media

Ministry of Education (2007). *The New Zealand Curriculum.* Wellington: Learning Media.

Ministry of Education (2007). *Literacy Learning Progressions. Draft for Consultation.* Wellington: Learning Media. Available at www.literacyprogressions.org.nz/

Ministry of Education (ongoing). *asTTle: Assessment Tools for Teaching and Learning: He Pūnaha Aromatawai mō te Whakaako me te Ako.* Auckland: University of Auckland School of Education. See www.tki.org.nz/r/asttle/

NZCER (2001). *STAR: Supplementary Tests of Achievement in Reading: Years 4–9.* Wellington: NZCER.

O'Connell, P.,Timperley, H. and Parr, J. (2008). "Is sustainability of educational reform an article of faith or can it be deliberately crafted?" Paper presented at the International Congress for Schooling Improvement and Effectiveness Conference, 6–9 January 2008, Auckland, New Zealand.

Phillips, G. (2007). Paper presented at the August 2007 Literacy Symposium. "Lifting the Achievement of the Under-achieving Reader and Writer in the New Zealand Mainstream Classroom."

Timperley, H., Wilson, A., Barar, H., and Fung, I., (2007). *Teacher Professional Learning and Development: Best Evidence Synthesis Iteration [BES].* Ministry of Education: Wellington. Available at www.educationcounts.govt.nz/publications/series/ibes