

Sabbatical Report

To challenge the practice of intermediate schools extensively testing year seven children in term one in spite of having collected considerable data on the children from the contributing school in term four of the previous year.

Is the practice a valid use of teaching time?

Is there a better alternative?

Geoff Siave

Principal

Shirley Intermediate School

(Sabbatical taken term three 2011)

Rationale: In term one of every year our intermediate school teachers spend three to five weeks completing extensive assessment exercises on new children. This is in spite of having collected considerable data during the latter part of the previous year.

Is this practice simply doubling up the assessment process? Is this time which could have been better spent teaching?

<p><i>Excerpt from Shirley Intermediate School request to teachers of Year Six children at contributing schools (distributed early November each year)</i></p>	<p><i>Excerpt from Shirley Intermediate School Teachers' Brief:</i></p>
<p>Please complete the results for any of the following assessments which are available for your child:</p> <p>PAT reading comprehension Stanine</p> <p>PAT reading vocabulary Stanine</p> <p>Gloss</p> <p>IKAN</p> <p>Times Tables Testing in four operations</p> <p>Spellrite Essential Lists mastery</p> <p>Writing (Level)</p> <p>Please list any special learning needs:</p> <p>In your opinion is this child a suitable candidate for our extension class?</p>	<p>The following assessments are to be completed <u>during weeks one to four of term one.</u></p> <ul style="list-style-type: none"> • PAT reading comprehension • PAT reading vocabulary • Gloss • IKAN • Times Tables Testing in four operations • Spellrite Essential Lists • Writing samples <p>All tests are to be marked by class teachers and results entered into the computerised achievement data programme before the end of week five.</p>

Having been an intermediate principal for five years, I have worked with my senior staff in the latter part of every year to obtain a range of information about the children from our main contributing schools who have enrolled to begin with us the following year. We principally use this information to make our class placements in preparation for the following year.

Over the years we have frequently modified how we collect this information. For instance, five years ago just *one* senior staff member – usually a deputy principal - would visit all contributing schools during term four to collate academic and social information about the pupils enrolled to start with us the following year. We evolved this to the point where we now send *at least two* senior staff to interview each year six teacher at our main contributing schools. Another example of improving our transition system has been in the format of the summative information which we ask relevant year six teachers to complete. Our current compact (one A4 page) summative document is - we hope - concise and practical, conveying important figures and gradings, but also indicating for us which children require special attention – whether academically, socially or otherwise.

This information serves at least these four purposes:

- 1) Provides information for the relevant teacher(s) at our school
- 2) Identifies children for our classrooms which have a special character (extension class, high literacy needs class, Level one Maori Language class);
- 3) Allows us to manipulate – or manage – the balance of academic levels and behavioural concerns across the school.
- 4) Informs us – in good time – about children for whom we may need to make some special arrangements due to personal, social, or physical factors, in order that we can be well prepared for the new term.

Given the amount of effort which we put into this process before the child begins at our school, it has always struck me as ironic and wasteful that we then spend around four or five weeks of term one, the following year, thoroughly testing all of the children with our own extensive range of standardised tests. In spite of endeavouring to be very well informed about each child, we still do Progress and Achievement Tests (PAT) Mathematics, PAT Reading, PAT Listening, Star Reading, Gloss testing, spelling tests, writing samples, and spelling assessments. By the time we mark, analyse and collate these results school-wide, we are typically into week six before we are able to say with confidence that we know where children are “at” academically. Is this testing necessary?

Would the term one testing time be better used in *teaching*?

We should also acknowledge that the contributing schools have usually completed their own end of year assessments shortly before the children come to us. **Is their testing still valid in term one?**

Is there a more efficient way - a more seamless way – for the intermediate school to develop an early and accurate assessment of children?

Possible Outcomes

- A. That our current system is validated by the findings
- B. That a simpler system is devised in our own school - without compromising quality of information
- C. That contributing *and* intermediate schools are prompted to change their end of year or beginning of year assessment practices as a result of this focus.
- D. That *teaching time* is increased during term one of year seven.

Focus Questions

Before I began investigating this subject I had several focus questions:

- i. Is the term one testing a *perceived* need rather than a real one?
- ii. In whose eyes is the testing necessary?
- iii. Is the testing regime merely a historical practice which has persisted unquestioned?
- iv. Is there something wrong with the quality of information which we receive from contributing schools?
- v. Is there something wrong with the information which we *request* from those schools? (Are we asking the wrong questions?)
- vi. Are we over-inflating the so-called "summer effect" – the acknowledgement that *some* children's achievement *lowers* during the six week school holiday prior to starting a new school year?

Actual Methodology (as opposed to the original *proposed* methodology):

1. Interviewed Principals or senior staff from several Christchurch intermediate schools
2. Read materials from online sources such as NZCER, sabbatical reports, MOE, Masters papers.

Secondary Possible Outcome

I also anticipated that In the course of my research I would learn a lot about what schools do to smooth the overall transition of children. So in addition to my main purpose of investigating the transfer of academic information I hoped to learn - by osmosis, if you like - more about good practice in the overall transition process

Reporting Intention

- Board of Trustees
- Canterbury Association of Intermediate and Middle Schools (AIMS)
- South island AIMS conference
- Ministry of Education Sabbatical reports online

The Re-testing Issue – Are we wasting time by re-testing in term one?

Following my discussions with teachers and colleagues across the Christchurch intermediate school sector it appears to me that the reasons for re-testing can be linked to a *perceived* need for *reliable uniformity of data*. I believe that perception, reliability and (perceived need for) uniformity are central factors which lead a “doubling up” of assessment.

It appears to be common practice in New Zealand to extensively assess children in term one. While my report focuses on year seven at intermediate schools, this assessment focus appears to arise from any or all of the following:

1. The *perception* that children regress slightly over the summer break.
2. The desire of current teachers to get a uniform understanding of all children *in spite of whatever data has been passed on from previous teachers*.
3. The perception that standardised testing is the most reliable, valid and worthwhile method of assessment.
4. The prevalent school-wide assessment timetable which - across most schools - requires extensive testing in term one *and* term four.

5. A perception that the data from contributing schools might be unreliable. i.e. A lack of confidence in every teacher of every year six child assessing children with the same degree of accuracy, consistency, or diligence.
6. A perception that overall teacher judgements in any discipline are subject to a degree of variance.

Each of these perceptions and variables are consistent with a desire to ensure uniformity, and reliability of current data. It is useful to look at these two qualities a little more closely.

Uniformity of information from contributing schools

There are eleven intermediate schools in Christchurch. Each intermediate has at least four main contributing schools with a smaller number of pupils enrolling from outlying schools in any particular year. Schools also routinely get new enrolments from families that move into an area during the new school year. Consequently the number of schools from which intermediate schools actually draw pupils in any one year will usually exceed a dozen. Certainly that is the case at Shirley Intermediate School (roll: 230). The number of contributing schools to a larger (600+ pupils) intermediate school would likely be considerably more. Consider also that each main contributing school might have two or more year six classes. Because of these factors the actual number of *teachers* who taught (and therefore, *assessed*) the children is usually going to be even greater than the number of *schools* from which the pupils came. Ergo the effects of variations in methodology which have contributed to the assessments made in year six are potentially even greater than simply the number of contributing schools.

Each of these schools has its own consortium of assessment tools. There is currently no national requirement upon schools to use any particular assessment product (– although such a requirement may not be far away). There are certainly some assessment products which are common. The table following outlines some of the more common tests available in New Zealand.

Test or Assessment Strategy	Subject Area(s)	Outcome	Level of usage by local sampling
Progress and Achievement Tests (PAT) produced for many years by the <i>New Zealand Council for Educational Research</i> (NZCER)	Mathematics Reading Listening	Places students at a stanine level of achievement	6/6
Supplementary Test of Achievement in Reading (STAR) produced since 1999 by NZCER	Reading		6/6
Individual Knowledge Assessment of Number (IKAN) (Ministry of Education)	Numeracy	Places student at Knowledge Levels	3/6
Global Strategy Stagg Assessment (GLOSS)	Numeracy	Strategy stages across operational domains of addition/subtraction, multiplication/division, proportions/ratios.	5/6
E-asTTle Reading	Reading	Skills and curriculum level and	5/6

E-asTTle Writing	Writing	can show movement over time in areas specified by the school	1/6
E-asTTle Mathematics	Mathematics		2/6
Running Records	Reading	Identifies reading strategies used by pupil; approximates a reading "age"	6/6
Assessment Resource Bank (NZCER)	Various	Graduated assessment resources	1/6
Spell-Write (NZCER)	Spelling	Essential Lists Mastery	2/6

While this is by no means a comprehensive list, the variance between schools is already apparent and presents challenges for the year seven teachers at intermediate schools who might attempt to forego testing in term one. Such teachers would need to make a summative judgement on children from information which might be presented in many different formats. This is not impossible. The *Overall teacher Judgements* ("OTJs") currently required of teachers nationwide are examples of such judgements. In the meantime it is completely understandable that teachers and school managers would want to have an assessment made using a common source test.

Reliability of information from contributing schools

A further challenge exists around the reliability or authenticity of the data received from other teachers. Authenticity is a huge issue. Bearing in mind that "all tests have a margin of error and are only a snapshot of how students perform on a limited slice of curriculum on any one day" (NZCER 2010) how much credence can be placed on test results?

Can teachers be assured that all testing and assessment tasks carried out by other teachers are completed under similar conditions and judgements made using well moderated criteria? Standardised testing manuals for teachers typically include very precise instructions about the methods to be applied when conducting tests. To be valid, any "standardised" testing environment and procedure must be, well, *standard*. Wavering from these conditions puts the integrity of these tests in question. An obvious example of this would be in a PAT test situation where one teacher provides a quiet relaxed, studious class atmosphere, while another teacher has a higher threshold for noise levels during the test, but punctuates the test session with loud admonishments of some pupils. While this example is quite an obvious one I would suggest that teachers within and across schools create test conditions which vary in some ways as markedly as this.

Is it actually possible to emulate all the exact same conditions between tests? Simply put, the answer is "no". The best that teachers can hope for is that a very high degree of professional integrity and rigour has gone into the valuations which they receive from their peers at contributing schools. Teachers themselves do not routinely have faith in the reliability or appropriateness of standardised tests (Brown 2010). In short, such tests have been developed by academic and

education research bodies such as NZCER, and – bluntly – are perhaps as much as we can hope for as far as developing a standardised assessment tool.

Do we need new data in term one?

It seems logical that we need data which shows where each child is “at”, so that we can (a) deduce what the “next steps” should or could be, and (b) so that we can compare progress with past progress and future progress. Data of a formative or diagnostic nature would be most useful for this purpose. The focus of this paper, however, is on the summative assessments which are typically represented by standardised tests like the PTA. So if summative and diagnostic material will be of the most benefit for teaching purposes, why is so much time being spent in term one on summative assessment?

Simply put, the summative assessments in term one provide a starting point for school wide analysis which can be compared to data collected at other junctures in the school year. The school can thus measure academic progress of the student body as well as that of various cohorts of children (eg.girls, boys, Maori pupils).

The rationale behind this paper was that there is some doubling up of assessment, and I have been keen to explore whether this could be avoided? The issue to me has been that the testing process takes up valuable classroom time (plus the time used by the teacher in marking, and analysing each test) which could be better spent teaching.

From discussions with colleagues during the course of this sabbatical, there are several actions which teachers and principals could take which would counter the doubling up effect of the testing in term one, and which would minimise the loss of teaching time.

Alternatives and Safeguards 1: Validate the data collected in term four

One of the reasons cited for re-testing the children in term one has been a lack of complete faith of the intermediate teachers in the reliability of the data provided by contributing schools. This can be addressed in a number of ways, which include:

- a) Strengthen the professional development opportunities between the intermediate school and the contributing schools, particularly in the area of assessment. Ensure that this professional development involves teachers working with teachers from the other schools. In this way the teachers will develop common understandings and from these will come growing confidence between the teachers from different schools.
- b) Establish teaching interchanges between the schools so that teachers develop better understandings of how colleagues from other schools work.

Alternatives and Safeguards 2: Consider Authentic Assessment

Authentic testing offers the opportunity to collect a range of “real” information such as observations, samples of practice, photos, and recounts which can provide a depth of information about performance. An example of authentic assessment in writing is the letter written by a child to thank a parent helper who attended camp. Another would be a child’s journal about school camp. These pieces of writing are “authentic” if constructed completely by the child, for a genuine purpose (not for an assessment exercise). By comparison writing produced as part of an AsTTLe assessment are contrived and therefore un-authentic (although they can still be quite valid assessment tools).

Conventional testing	Authentic testing
Low treatment validity	Directly informs intervention
Scripted standardised procedures to sample skills and behaviours (to preserve the validity of normative measures)	Naturally occurring behaviours
Norm focused – levels, stanines	Evidenced –based
Dominated by Teacher input	Encourages collaboration of learner, team members and Whanau

Implications

An implication of not challenging the practice of extensively assessing children in term one at Year 7 is that the teachers accept that practice as “best” practice. Reflective teachers, engaged in teacher inquiry should be questioning practice at all levels. They should be investigating the thinking behind the practice. Indeed, practice should reflect researched methodology; theory in practice – praxis.

An implication from this inquiry is that no matter how much information is provided at the end of a school year on student achievement, the teacher beginning the new year will still be seeking “current” information. This implies that teachers accept that children’s understandings or skill levels are constantly changing – including during holiday periods. This applies equally to year seven.

Another implication is that at the beginning of the school year teachers are seeking *summative* data rather than formative information. They want a starting point, against which they can later measure progress. Some of the formal assessment tools do provide both summative and formative or diagnostic information. But principals felt that the diagnostic information – such as that found in the PAT tests – was largely overlooked by teachers. The (summative) stanine was the main information used from the PATs.

A further implication is that teachers consider the summative data to be more important than formative information. Why would this be so? The obvious reasons are that summative data matters when teachers and schools are held to be accountable. For simplicity sake, it is more convenient to sum up a child's abilities by locating the child on a scale eg. a "stanine 6", or "level 4b", or "reading age 11.6". The usefulness of such scales is undoubted. But such usefulness is restricted to when comparisons with cohorts are required. The summative assessments are of little use when analysing the learning needs of a child, which are typically more complex, and more individualised than can be represented by a single scale score.

The emphasis on the summative score implies that the beginning of year testing is not for the purposes of teaching and learning for the child. The benefit of the testing is for the development of pictures of assessment across cohorts within the school- year level, age, gender, ethnicity, etc. Politicians can use the summative data for any number of political wranglings (and they do!).

Conclusion

The apparent waste of teaching time in term one prompted my sabbatical topic. It has become very clear that in spite of extensive gathering of information about incoming pupils academic achievement levels, the beginning of year testing will continue for the purposes of uniformity, reliability, simplicity, and – most importantly - accountability. The tests are *not* required in order for teachers to teach well. One single test, in one curriculum area, at one particular time, provides a snapshot of learning for individuals, and for cohorts within and across schools, for the purpose of measuring difference. That is, for accountability purposes. Accountability matters at levels *above* the teacher-learner context. Accountability is for mainly for the benefit of accountable groups like managers and governments.

These standardised tests address the need for reliability and uniformity – although the level to which the assessments across classrooms meet a high degree of standardisation could be challenged, relying as it can upon the uniform monitoring of testing procedures of all teachers involved.

Is this collation of levels and grades important? It certainly is. Apart from showing the learner where he or she is, or was, or should be "at", the grades are not directly for the benefit of the learning of the learner. Grades show the teacher and school management progress over time.

However, the tests or assessments which will more likely actually *improve* learning are those which provide diagnostic or formative information, informing the teacher of what the learner is actually doing; his or her strengths and needs. They provide useful information for the teacher about the

learner's behaviours. They can be authentic in nature, highly specific, and can vary greatly in format – from formal and objective, to informal and subjective. Authentic assessment can be very beneficial for the learner as well as the teacher.

We will continue to have the summative testing, for the assessment of teachers, schools, and groupings of children with specific commonalities. But teachers and principals must maintain a focus on what is most important in these processes, and that is what we learn about each learner which helps the further learning of that child. Everything we do should have the aim of improving learning outcomes for the learner. Therefore we should be very clear that grades, and levels and stanines contribute little to a learner or to learning. They are a mere measuring rod by which we can make comparisons with others and with past and future progress. Apart from providing some incentive for improvement, they are tools for decision making at levels far from the learner. The usefulness remains mainly at those auspicious levels.

Teachers and principals must focus on authentic and diagnostic assessments while “rendering unto Caesar that which belongs to Caesar”. In other words, do the summative assessments but don't let them get in the way of actions which are of far more importance to improving learning. Make full use of the diagnostic aspects of the standardised tests – such as those found in AsTTle and the PATs. Do not let the *measuring* of improvement get in the way of actually facilitating improving. Minimise the summative testing, simplify the time taken analysing for summative purposes, maximise time spent on formative and diagnostic assessments - and on teaching!

Geoff Siave

October 2011

Addendum

In the course of my investigation I gleaned a lot of information from my colleagues across the intermediate schools in Christchurch around transition practices in general. While they sit outside the brief of this report, I have summarised them here. There were many great ideas and I have grouped them under three categories. Perhaps these will be a catalyst for some further research.

Initiatives which focus on academic achievements	Initiatives with holistic focus	Initiatives which focus upon smoothing the transition for the child(ren) or families
Senior intermediate staff visit the contributing school and interview each year six teacher about the children (academic and social focus)	Teacher from intermediate releases the Year six teacher to allow him or her meet with senior intermediate staff	After enrolment during year six for the following year, the family receives an acknowledgement letter from the intermediate – welcoming them etc. and providing details of any orientation day, etc.
Senior Intermediate staff compared notes and made decisions about placement of children.	All intermediates made class placements after considering the following: Recommendations from contributing schools, parental requests, academic levels, behavioural concerns, social needs, special character of classrooms, teacher-pupil compatibilities	The Orientation day includes children and parents. Most intermediates provided between an hour and two hours for this process. Usually involved address by principal, then tour of the school.
	Several intermediates held an interview day on Day one of the new year. Each child and parent met with the teacher for 15 or so minutes.	Orientation Day – some intermediates had the children meet their future teacher, However, most intermediates did not actually have classes confirmed until the first day of the new school year.
	Most schools held an informal meet the teacher session in first few weeks of the new term – The format varied considerably with formal and informal elements.	

References

- Ministry of Education Assessment in early childhood intervention . The authentic assessment alternative; The 4th annual educational psychological forum (November 2011).
<http://www.massey.ac.nz/massey/fms/Conference%20services/Logos/Woodward,%20Julia.pdf> Accessed December 4 2011
- Harris, Lois R. & Brown, Gavin T.L. (2010). Mixing interview and questionnaire methods: Practical problems in aligning data Practical Assessment, Research & Evaluation . <http://pareonline.net/getvn.asp?v=15&n=1>. Accessed May 31 2012
- New Zealand Council for Educational Research. NZCER *NZCER background paper on national standards for literacy and numeracy* paper 18, volume 2 (2008). www.nzcer.org.nz/pdfs/nationalstandards.pdf
Accessed September 18 2011

Acknowledgements

- Shirley Intermediate School Board of Trustees – for approving my sabbatical application
- Ministry of Education - for granting me the sabbatical
- Simon Thomas – for ensuring that I never had a concern about the school during my absence
- Lois Christmas – for gentle reminders to finish my report!
- My Canterbury AIMS Principal colleagues for providing insight into their transition processes