

asTTle – Assessment Tools
for Teaching and Learning



asTTle – Assessment Tools for Teaching and Learning

Contact details

John Hattie j.hattie@visiblelearning.co.nz
Debra Masters d.masters@visiblelearning.co.nz
Visible Learning Labs
University of Auckland
N452 Epsom
Private Bag 92019
Auckland. NZ.
<http://www.tki.org.nz/r/asittle/>

Brief details

An internet- based system for New Zealand teachers to select tests in Maths, Reading and Writing (English and Maori) in Yrs 4-12 (age 8-16), which is used formatively or summatively both for students and for teachers. Tests are compiled by teachers selecting test characteristics (eg balance of content and difficulty) from a large 10,000 item bank. The system uses dichotomous¹ and polychotomous² item response theory (IRT). The system is particularly strong in the way in which summary data are presented.

Applicability to HEIs

asTTle is a national schools system with all the economies of scale that that implies. However, with the increasing emphasis on the need for effective feedback to students in HE much can be learned from the way that performance data are presented for student and teacher. Also, the use of IRT to generate standardised tests 'on the fly' deserves careful investigation, particularly for student formative purposes, for example, by generating in-course self-assessments.

What was the problem?

The New Zealand Ministry of Education asked the University of Auckland to create an educational resource for assessing

literacy and numeracy (in both English and Māori) for levels 2 to 6. The resource was to fit in with the New Zealand educational philosophy and was not to fall into perceived traps of encouraging teaching to the test with its ensuing narrowing of the taught curriculum.

asTTle provides teachers, students, and parents with information about a student's level of achievement, relative to the curriculum outcomes as described by the NZ national curriculum (see links below).

The NZ philosophy

National assessment

- should not dictate but follow from the curriculum
- should not be wide and shallow but rather explore the understanding of the rich ideas underlying the curriculum
- must be low-stakes
- be part of a larger assessment strategy to reinforce desired educational values
- should provide copious comparative data but no league tables
- should be owned and valued by teachers
- should be controlled by teachers so that what is assessed matches their teaching
- should provide feedback to teachers, students, et al, to inform them where they are, where they are going and how they might get there.

The solution

Teachers use asTTle to create 12-60 minute paper and pencil or screen based tests (which can be adaptive) designed for their own students' learning needs. Once the tests are scored by the teacher using supplied marking schemes (or by the computer if the test is closed or objective), the asTTle tool generates interactive graphic reports that allow teachers to analyse student achievement against curriculum levels, curriculum objectives, and population norms.

¹ Meaning items whose answers have only two states, right or wrong

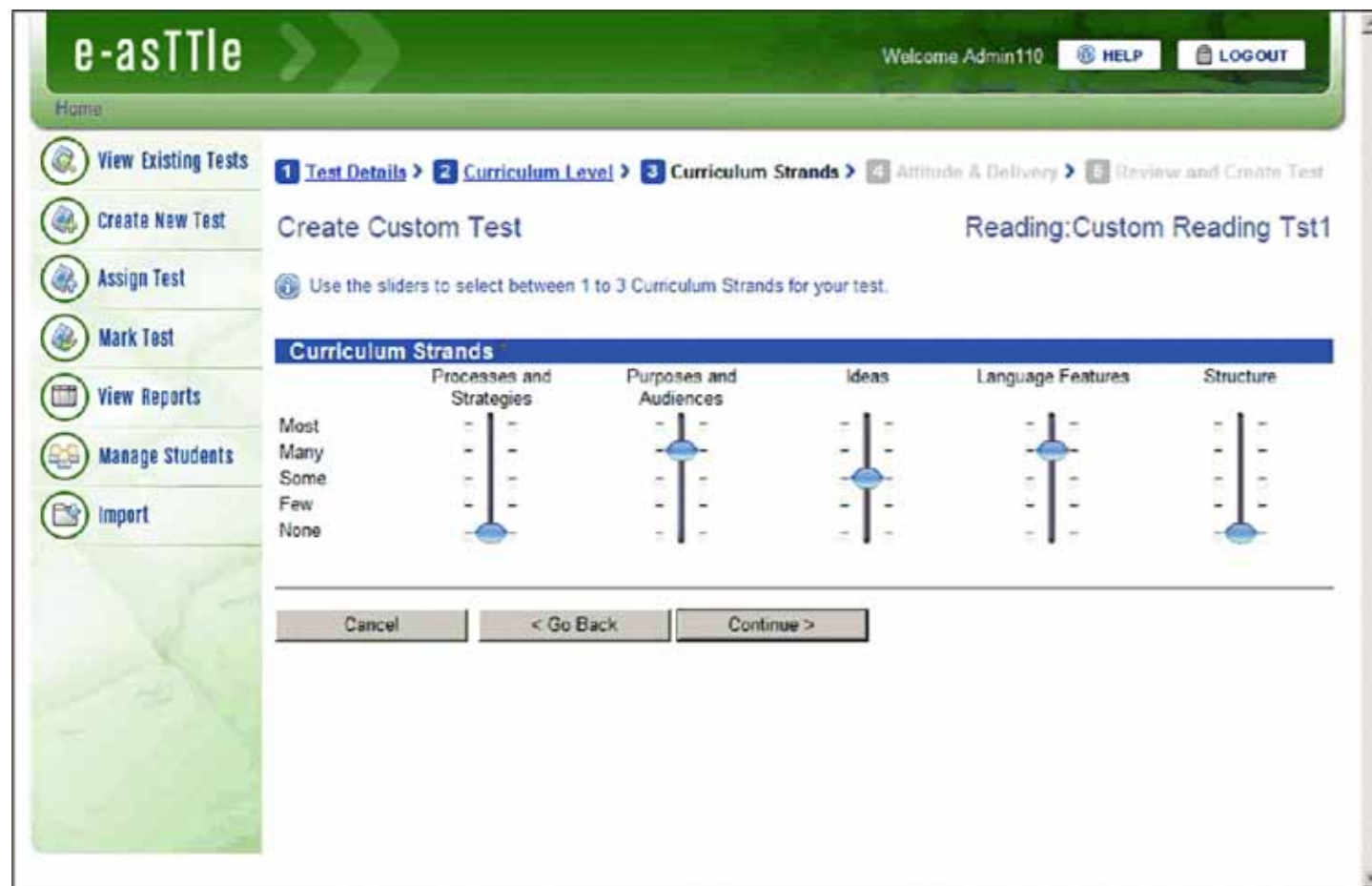
² Meaning items whose answers can be awarded partial credit

How does the solution work?

Rather than the balance of the test being determined centrally, control is passed to the teachers who can stipulate (within limits) the proportions of items addressing different rich ideas and the level of difficulty. Actual items are selected automatically according to the metadata stored in the system. The difficulty of any test created is calculated using IRT so that comparisons can

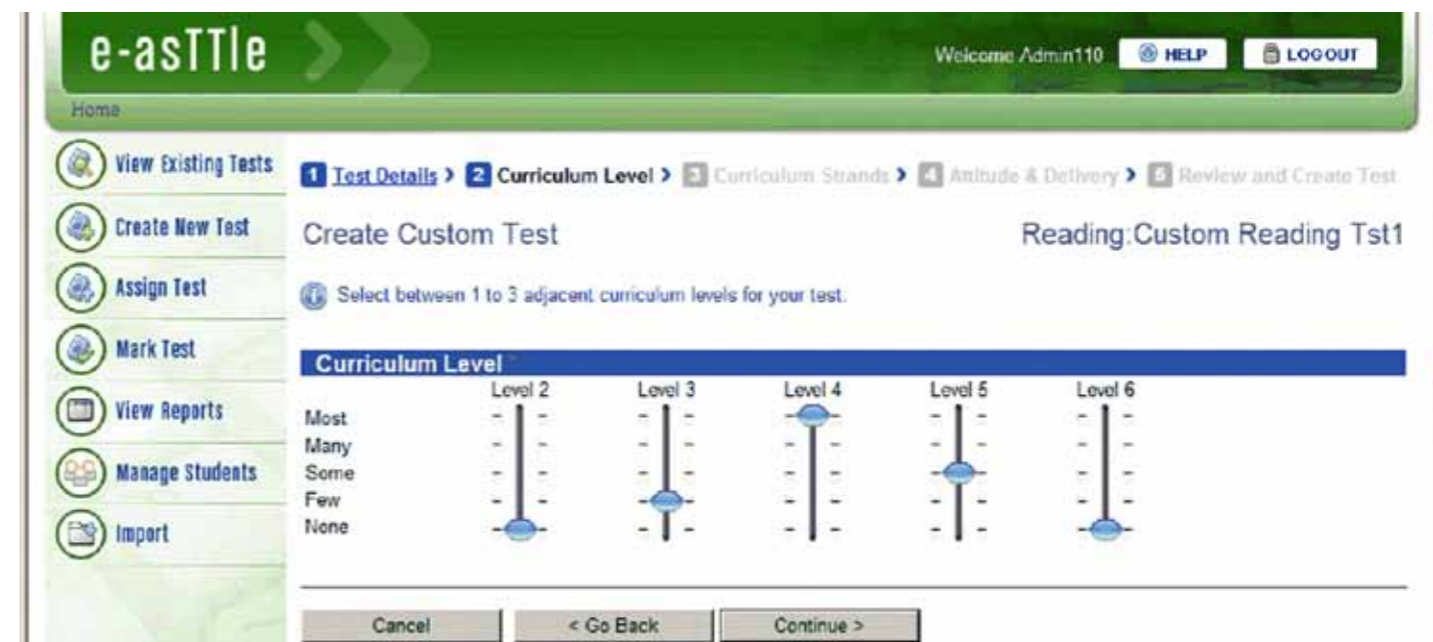
be made between different groups (who will have taken different tests), using a common scale that allows growth comparisons over the years of schooling. The items and their associated information (pictures, graphs, etc.) can be printed to any resolution and have high screen image resolution. The screen administration pays much attention to scrolling and positioning on the screen, and the paper versions maximise "white space" which has shown to improve student performance.

Customising the test - content

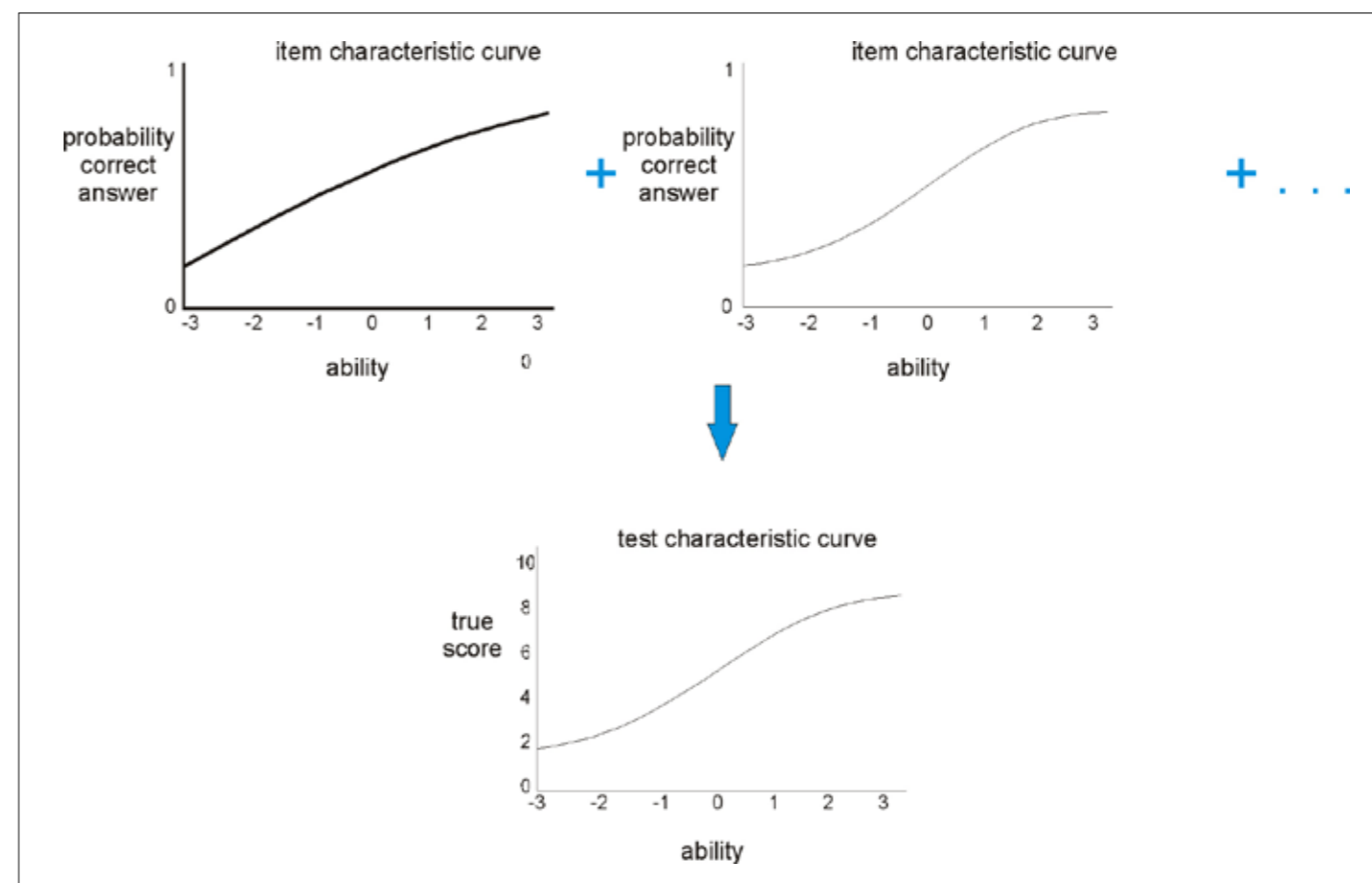


e-asTTle uses the curriculum strands from the 2008 New Zealand curriculum documents. After selecting the level, teachers make a choice among the strand options identified. The strands are subject-specific. Each strand area aggregates a number of achievement objectives as derived from the appropriate curriculum statement.

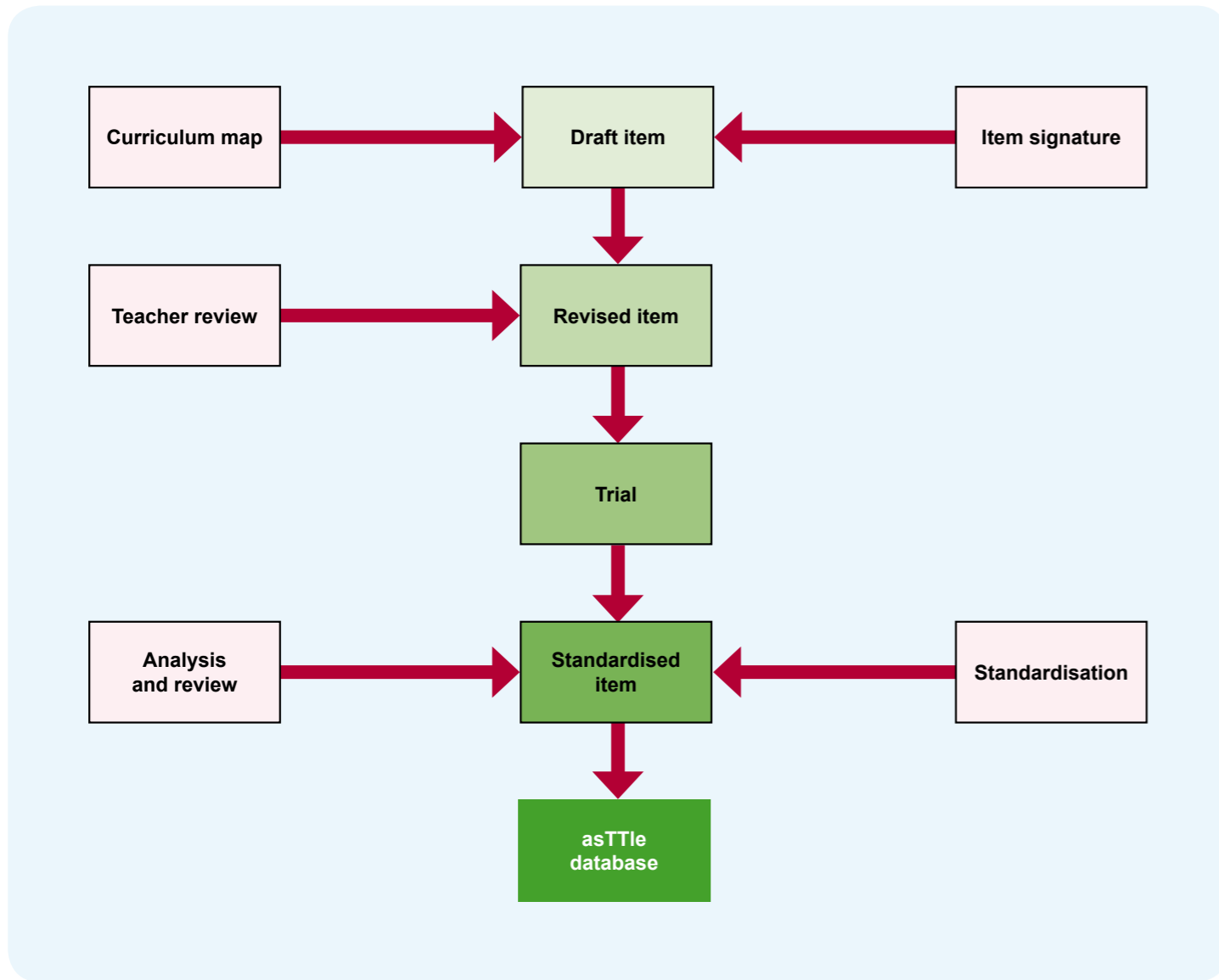
Customising the test – ability of group



Teachers adjust the sliders according to what they believe is the ability distribution of their class. Over time, this is a formative process, teachers' understanding of the difficulty of the levels improving with experience. Essentially they are generating a test characteristic curve (TCC). Each item in the itembank has measured psychometric characteristics. IRT generally postulates two variables, the difficulty of the item and the ability of the student. The item characteristic curve shows the relationship between the two, ie the probability of the item being answered correctly by candidates of different abilities. The test characteristic curve is simply the sum of those item characteristic curves and shows the relationship between underlying ability and the true test score.



The algorithm uses a linear programming heuristic (to maximise fit to the desired TCC) and then fits various constraints such as curriculum proportions, specific objectives, complexity, time, and reuse. The optimal test is delivered within 7-20 seconds.



Designing items and standards setting

It is notoriously hard to write items to assess specific sets of criteria. Here, item briefs are specified, an item generated and then the fit is critiqued.

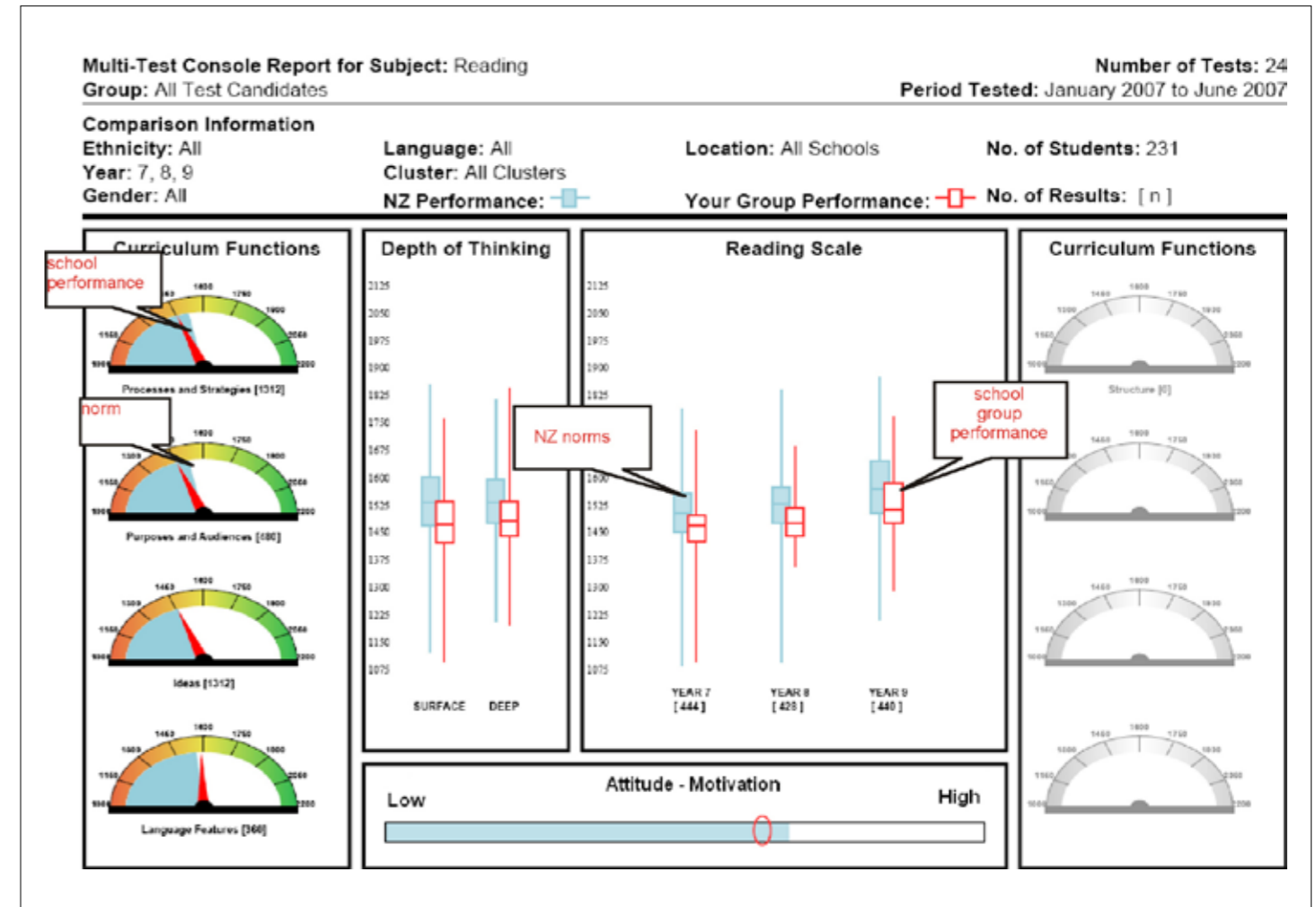
- 1 A detailed set of item specifications is designed to ensure coverage of the curricula, difficulty, complexity, etc.. This leads to control over the maximum number of items required and aims to ensure that there are sufficient items in all cells of the specification.
- 2 Professional item writers (who are/have been teachers) write or locate items to fit a specific set of item constraints – which can cover up to 20 attributes – their ‘signature’. Items may be right/wrong (dichotomous) or also yield partial credit (polychotomous). They include category sorting, close, insert text, complete or change diagrams, complete/solve equations or problems, draw shapes, explain reasoning or rule, formulate hypotheses, highlight, circle or underline, matching, multiple-choice, order or sequencing, proof or edit, short answers, complete an expression and click on hot spots.
- 3 Teachers’ workshops critique the match of the item to the specifications and, where necessary, change the item to ensure maximum fit.
- 4 Prior to inclusion, all items are trialled (sample 90,000 for first set) or a difficulty estimated by teacher workshops and then modified after trialling in the system.
- 5 The psychometric qualities of the items are then analysed using both dichotomous and polychotomous IRT processes (Rumm, Multilog). A three-parameter (difficulty, discrimination, guessing correction) model is first used to identify poor items due to high guess rate (>20%). Then a two-parameter model used to identify items with poor discriminators and finally a one-parameter model is used to do the scoring.
- 6 A series of standard setting workshops (using the Bookmark method http://www.education.ualberta.ca/educ/psych/crame/files/standard_setting.pdf) are conducted to determine the cut-points on the IRT asTTle score scale for the various levels in the curricula (each subdivided into three: Basic, Proficient, Advanced).

Reporting

There are a variety of reports which teachers can access, there is copious support material and, in addition, third party providers deliver professional development to schools about the interpretation of the reports and work with schools to integrate the methodology and reports into a whole-school evidence based set of practices. Some of these reports follow.

Console report

This compares the class or school performance (red) to NZ norms (blue).



Individual learning pathway

The individual learning pathway show a single student's performance on a single test., comparing the score on each objective to the overall estimated IRT score for that student.

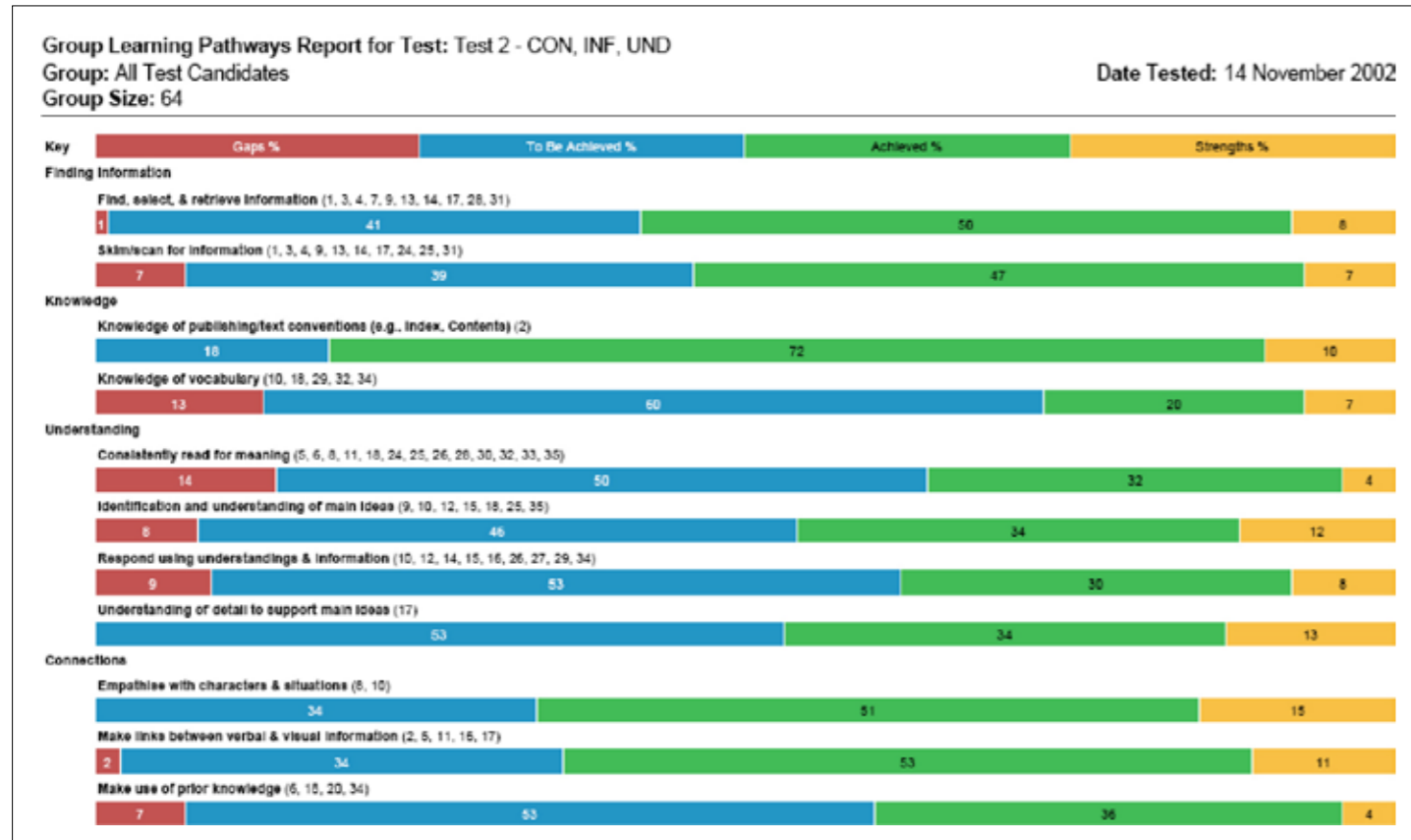
For the tests to provide maximum feedback, the student must be set as many questions that they pass as they do not pass. The

ILP shows how the student has performed (including objective and specific item information).

Students and teachers work together to interpret/act on these data and they are often transmitted to parents. The discussion centres around what the student knows, does not know, their strengths and gaps and aims then to discuss where the student should next place most attention.

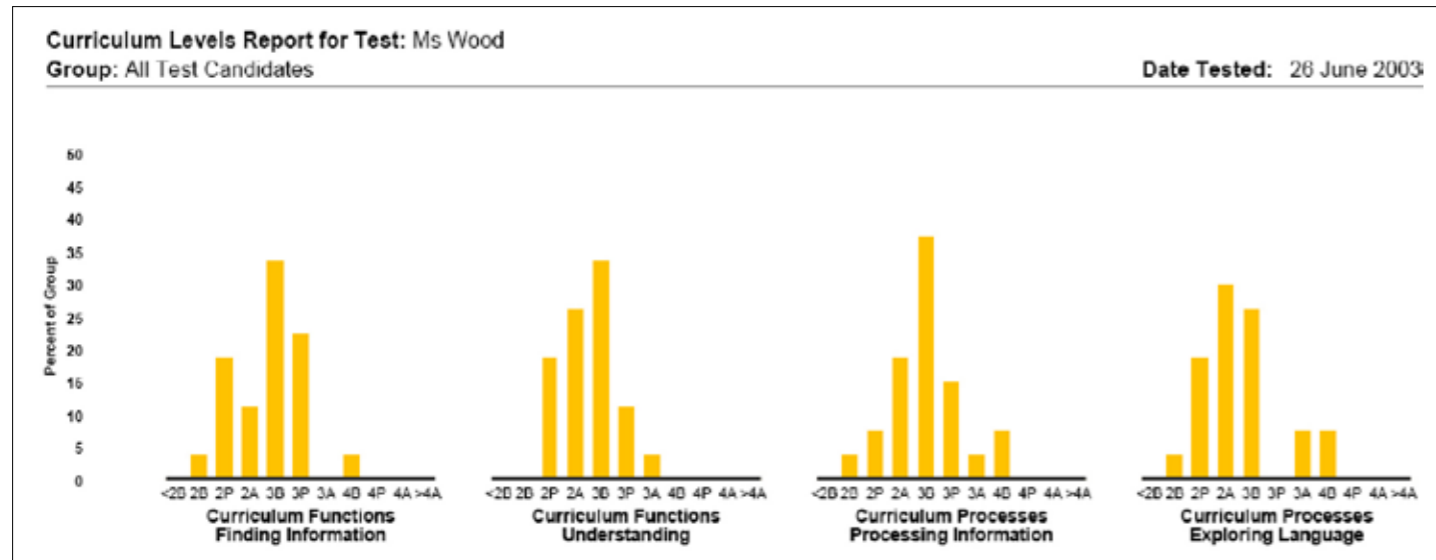
Group Learning pathways

The information in the individual learning pathway can be summed over a class, year, school or nationally.

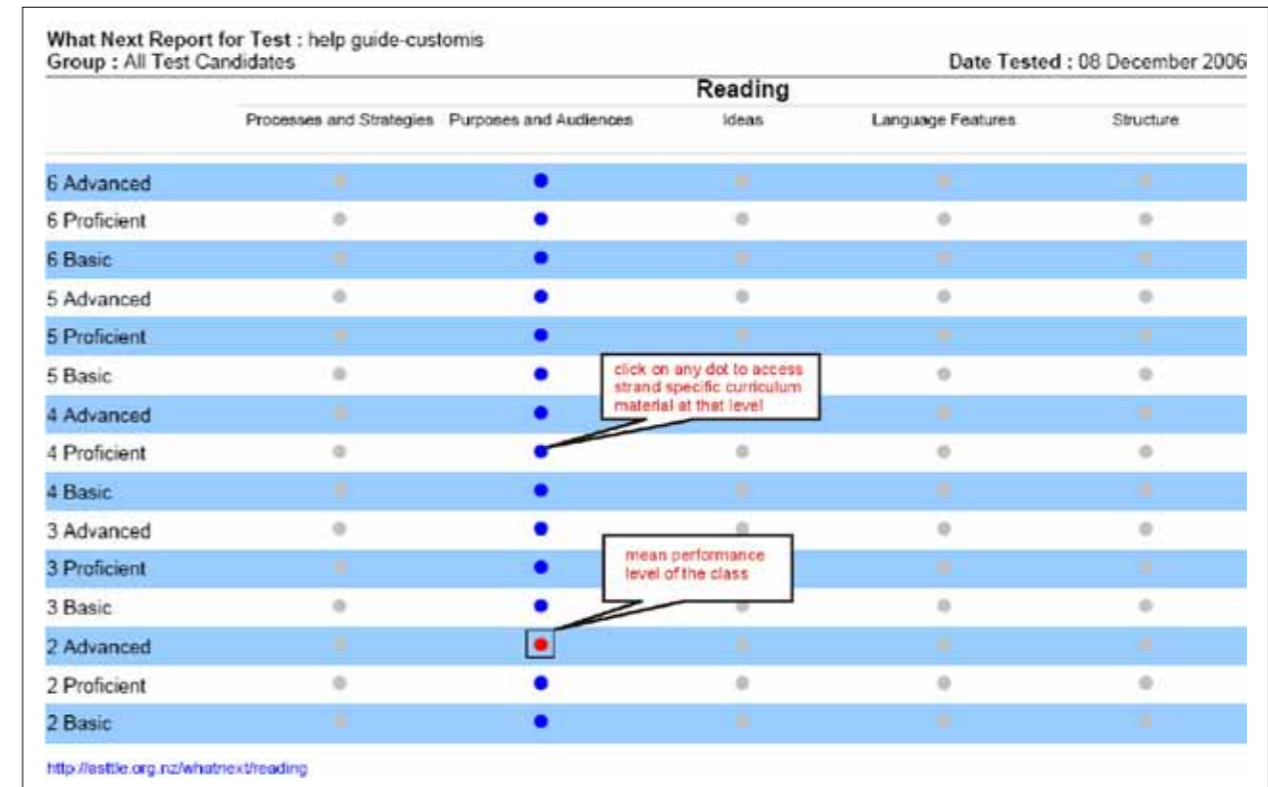


Curriculum level reports

Shows the distribution of pupils among the subject levels. If the user clicks on name of the curriculum strand, the names of students located in the various levels for that curriculum area will be presented.

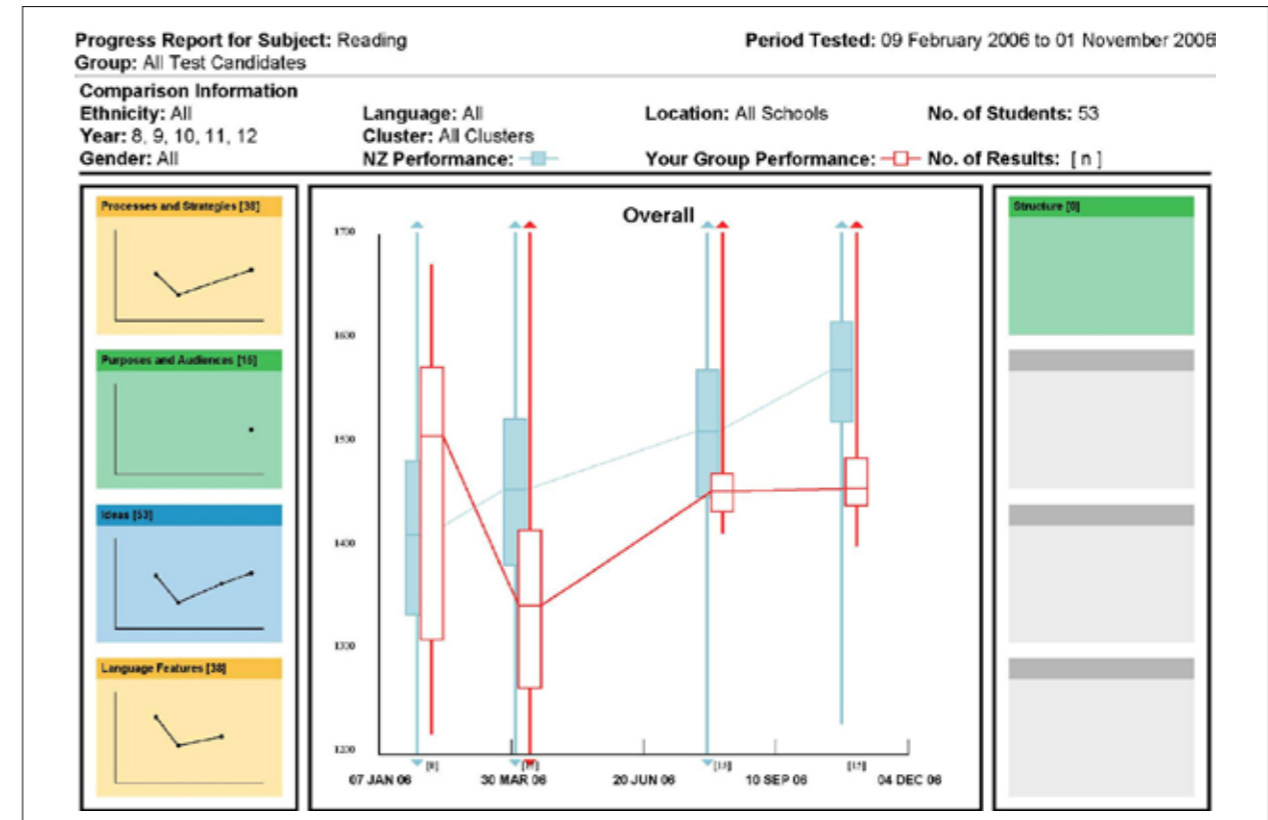


What next report



The What Next profile indicates the mean level of the chosen group of students in the curriculum areas assessed and assists teachers in locating appropriate teaching materials, generally at or above the indicated mean level of the class.

Student progress report



The Progress Report can be used to check that all students are making acceptable and substantial progress during the school year.

In sum, asTTle is a remarkably comprehensive support system for New Zealand teachers, empowering them and their students to take control of assessment in order to promote learning.

links and references

On the New Zealand national curriculum.

http://nzcurriculum.tki.org.nz/the_new_zealand_curriculum/learning_areas/english/english_curriculum_achievement_objectives__1

http://nzcurriculum.tki.org.nz/the_new_zealand_curriculum/learning_areas/mathematics_and_statistics/mathematics_and_statistics_curriculum_achievement_objectives

www.inca.org.uk/new-zealand-appendix-mainstream.html

On asTTle.

<http://www.tki.org.nz/r/asttle/>

http://www.tki.org.nz/r/asttle/e-asttle/faq_e.php#interpreting-reports6

Large-scale Assessment of Student Competencies; John Hattie; University of Auckland, New Zealand; Paper presented as part of the Symposium:

Working in Today's World of Testing and Measurement: Required Knowledge and Skills (Joint ITC/CPTA Symposium); 26th International Congress of Applied Psychology; July 16-21, 2006, Athens, Greece



asTTle – Assessment Tools for Teaching and Learning