Improving the quality of teacher-based assessment

Dr Peter J Keegan

13 September 2013

THE UNIVERSITY OF AUCKLAND
FACULTY OF EDUCATION
Te Kura Akoranga o Tāmaki Makaurau
Incorporating the Auckland College of Education

THE UNIVERSITY OF AUCKLAND
FACULTY OF EDUCATION
Te Kura Akoranga o Tāmaki Makaurau
Incorporating the Auckland College of Education

Aims of the Session

- Background/New Zealand context
- educating teachers for assessment for teaching and learning
- Basic & essential assessment concepts
- asTTle (assessment tool for teaching and learning)
- Improving the quality of teacher-based assessment

THE UNIVERSITY OF AUCKLAND
FACULTY OF EDUCATION
Te Kura Akoranga o Tāmaki Makaurau
Incorporating the Auckland College of Education

New Zealand Context

- 4 Million people, indigenous population, recent arrivals from the Pacific & wider Asian region
- Generally do well on international tests (TIMMS, PIRLs etc.) concerns about groups not doing well
- Assessments generally not compulsory, but recent compulsory reporting on National Standards (years 1 to 8)
- Educational for "knowledge economy"
- Can improve student achievement by improving teaching

THE UNIVERSITY OF AUCKLAND
FACULTY OF EDUCATION
Te Kura Akoranga o Tāmaki Makaurau
Incorporating the Auckland College of Education

Dr Peter J Keegan

- Teach university courses on assessment for teaching and learning
- Involved in the development of (standardized) assessment tools
- Provide inservice training and consultation on assessment
- Undertake educational research
- Parent

THE UNIVERSITY OF AUCKLAND
FACULTY OF EDUCATION
Te Kura Akoranga o Tāmaki Makaurau
Incorporating the Auckland College of Education

Key assessment concepts

- Conceptions of assessment
- Types of assessment (including standardized assessments)
- Reliability/Validity
- Measurement scales
- Measurement error
- SOLO taxonomy
- National Standards/Reporting of student results

THE UNIVERSITY OF AUCKLAND
FACULTY OF EDUCATION
Te Kura Akoranga o Tāmaki Makaurau
Incorporating the Auckland College of Education

Teacher conceptions of assessment

- Assessment to help both teachers and students improve their teaching and learning respectively
- Assessment to evaluate or certify student learning
- Assessment to evaluate or hold accountable schools and teachers
- Assessment has no meaningful purpose and so is ignored

THE UNIVERSITY OF AUCKLAND
FACULTY OF EDUCATION
Te Kura Akoranga o Tāmaki Makaurau
Incorporating the Auckland College of Education

Reliability

- *The consistency, stability, dependability, and accuracy of assessment results* (McMillan, J. H. 2001:65)
- An attribute of scores not tests
- Reliability is NOT the same as Validity
  - Something can be reliable but invalid
    - Inappropriate test scored accurately
  - Something can be valid but unreliable
    - Appropriate test scored inconsistently
  - We want both reliable and valid
    - Appropriate test scored accurately & consistently

Validity Defined

- Appropriateness of the inferences, uses, & consequences that result from assessment
- The soundness, trustworthiness, or legitimacy of the claims or inferences made on the basis of obtained scores
- Degree of soundness in the consequences of the inferences & decisions
- Not characteristic of a test; but a judgement

McMillan, p. 59

Validity Defined

- an integrated evaluative judgment of the degree to which empirical evidence and theoretical rationales support the *adequacy* and *appropriateness* of *inferences* and *actions* based on test scores or other modes of assessment
    - *Samuel Messick, 1989*
- What kind of evidence is needed to judge that the inferences and decisions are appropriate?

THE UNIVERSITY OF AUCKLAND
FACULTY OF EDUCATION
Te Kura Akoranga o Tāmaki Makaurau
Incorporating the Auckland College of Education

Two ways of looking at validity

- Types of Validity (traditional way)
- Messick's Validity Chain (everything done correct or chain breaks, i.e., becomes invalid)

Types of Validity (1)

- Face Validity – the degree to which a test does what it claims it can as judged by candidate or untrained observer
- Content Validity - is the content an appropriate coverage of skills, knowledge, abilities it is claiming to test ?
- Construct Validity – how test scores support the theoretical framework or construct being assessed

Types of Validity (2)

- Concurrent Validity – compared what is measured by test to a similar external test
- Predictive Validity – how well a test can predict "real world" behaviour.

Validity Chain

THE UNIVERSITY OF AUCKLAND
FACULTY OF EDUCATION
Te Kura Akoranga o Tāmaki Makaurau
Incorporating the Auckland College of Education

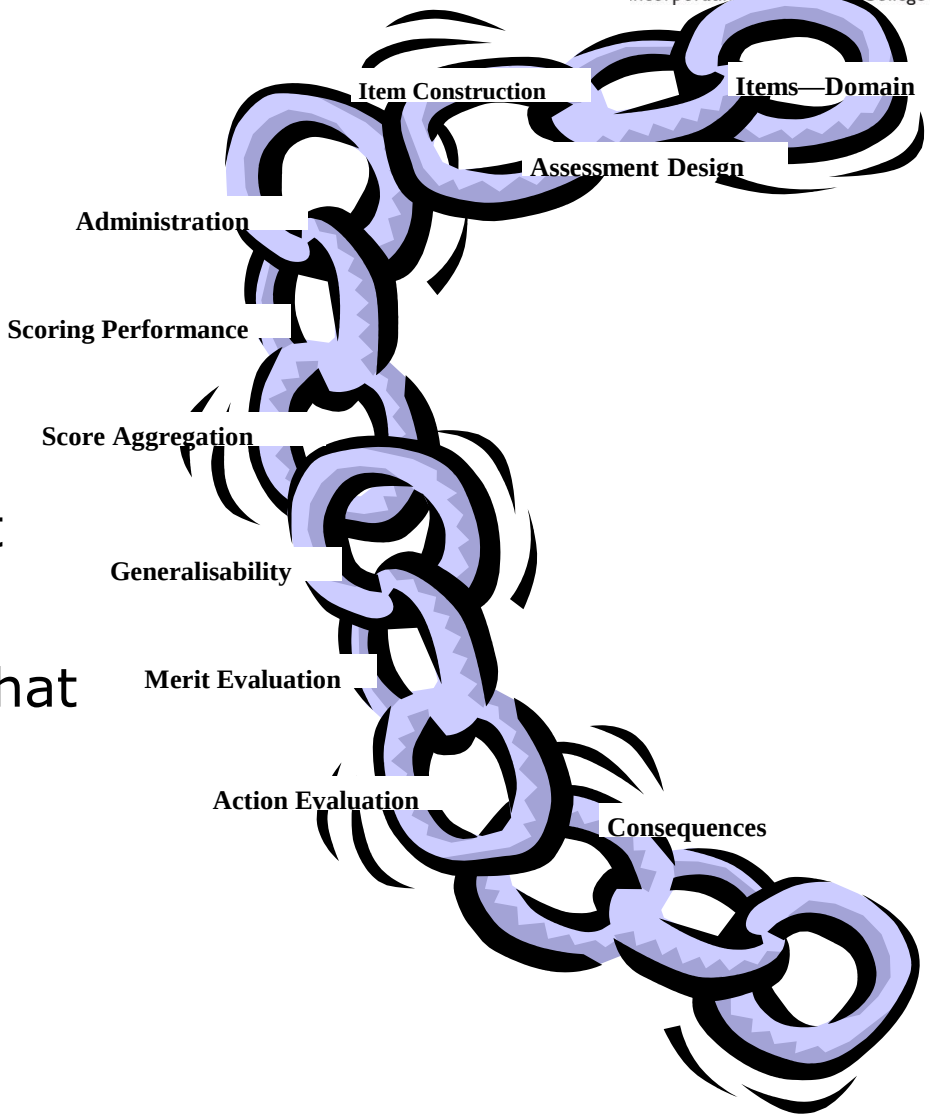
Chain as Metaphor[1]

All aspects are linked—weakness at any one point calls into question all inferences & decisions

No one link more important than any other

Links identify key aspects that must be evaluated ➔ Validation Evidence

[1]Chain from Crooks & Kane (1996)



THE UNIVERSITY OF AUCKLAND
FACULTY OF EDUCATION
Te Kura Akoranga o Tāmaki Makaurau
Incorporating the Auckland College of Education

Understanding Error

- Performance IS variable
- ALL educational assessment IS imperfect; 2 types of error exist
  - *Systematic*--can be controlled & identified; should be minimised
  - *Random*--not predictable as to size & direction; should be estimated

THE UNIVERSITY OF AUCKLAND
FACULTY OF EDUCATION
Te Kura Akoranga o Tāmaki Makaurau
Incorporating the Auckland College of Education

Sources of Error: Test Takers

- Health,
- motivation,
- mental efficiency,
- concentration,
- forgetfulness,
- carelessness,
- impulsiveness or subjectivity in responding,
- luck in random guessing
- And so on

THE UNIVERSITY OF AUCKLAND
FACULTY OF EDUCATION
Te Kura Akoranga o Tāmaki Makaurau
Incorporating the Auckland College of Education

Sources of Error: Situation

- Environmental factors (e.g., Heat & Light) in test room,
- level of learner preparedness,
- Prior knowledge of language of test
- Quality of previous teaching
- directions provided (significant source of error in school sssessment)

THE UNIVERSITY OF AUCKLAND
FACULTY OF EDUCATION
Te Kura Akoranga o Tāmaki Makaurau
Incorporating the Auckland College of Education

Sources of Error

- The MARKER (Evaluator/Assessor)
  - Idiosyncrasy or Subjectivity
  - Major source of error: look at essays & performance scoring
- Quality of Instrument
  - Major Source of error

THE UNIVERSITY OF AUCKLAND
FACULTY OF EDUCATION
Te Kura Akoranga o Tāmaki Makaurau
Incorporating the Auckland College of Education

Measurement scales, basic stats

- reporting scores, means, standard deviation
- distributions (normal etc.)
- scales, percentiles, stanines etc.
- conversions between scales
- displaying information/student scores visually
- comparisons between groups (effect sizes)
- longitudinal scores (over time)

THE UNIVERSITY OF AUCKLAND
FACULTY OF EDUCATION
Te Kura Akoranga o Tāmaki Makaurau
Incorporating the Auckland College of Education

Cognitive Processes
Surface & Deep Thinking

Structure of Observed Learning Outcomes (SOLO) Taxonomy

Analysis of the structure of student responses to assessment of given material by JB Biggs & K Collis, 1982

- SURFACE (increase in quantity)
  - Unistructural, Multistructural,
- DEEP (change of quality)
  - Relational, Extended Abstract

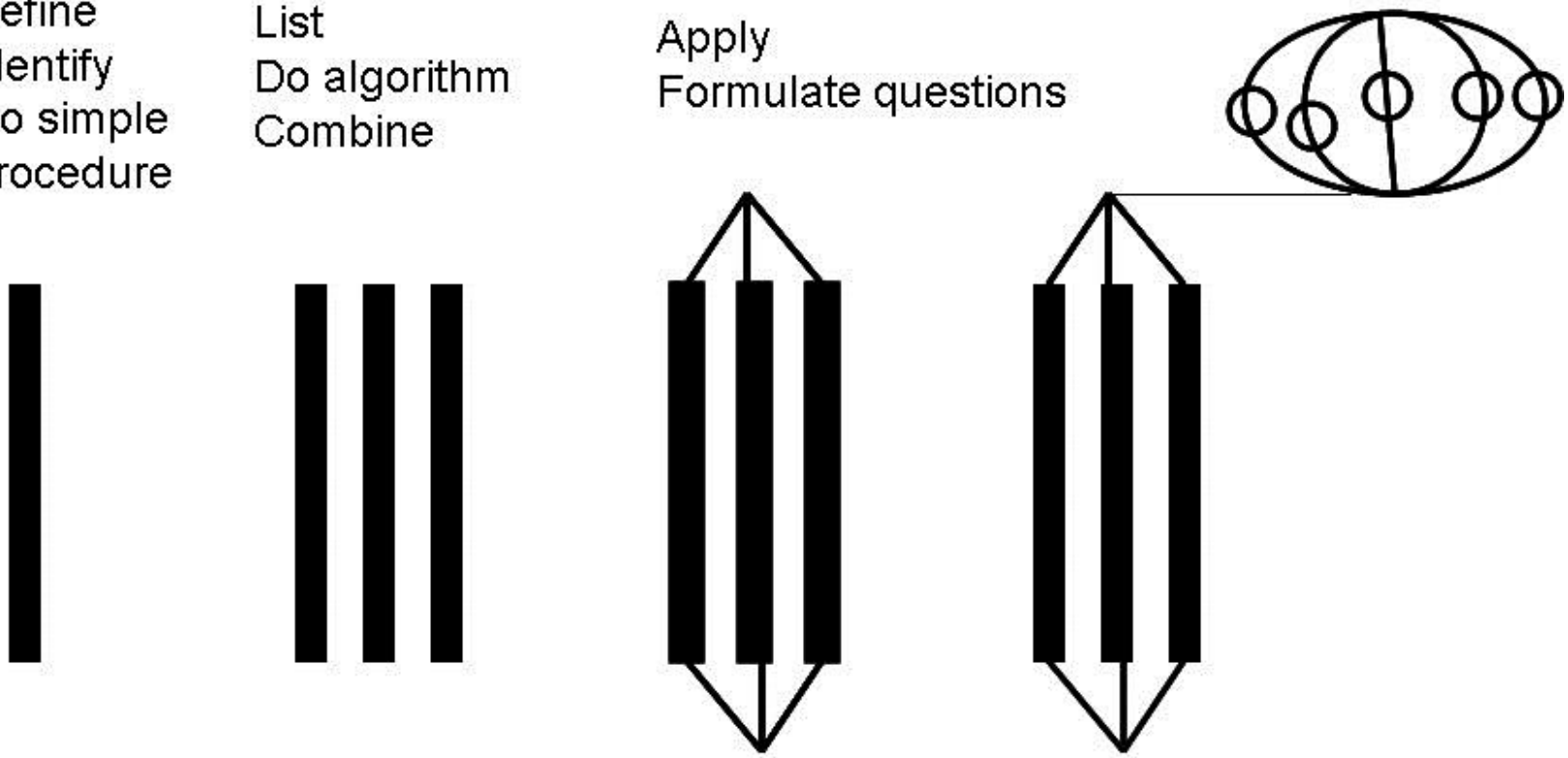
SOLO TAXONOMY (after Biggs and Collis)

Define
Identify
Do simple procedure

Define
Describe
List
Do algorithm
Combine

Compare/ contrast
Explain causes
Analyse
Relate
Apply
Formulate questions

Theorise
Generalise
Predict
Imagine
Hypothesise
Reflect



Unistructural **Multistructural** **Relational** **Extended abstract**

Students' perceptions of effective teaching

The concept of the caring teacher was particularly important at School A; clear explanation was more highly valued by students at School C; and School C student did not place as much importance on teacher humour. These variations may reflect the ethos of the school… another factor …might be the social background of the students. (Batten, Marland & Khamis, 1993, p. 16)

THE UNIVERSITY OF AUCKLAND
FACULTY OF EDUCATION
Te Kura Akoranga o Tāmaki Makaurau
Incorporating the Auckland College of Education

Surface Questions

Unistructural

What kind of teacher did School A students like?

_____

Multistructural

What two characteristics did School C students emphasise?

a) _____

b) _____

THE UNIVERSITY OF AUCKLAND
FACULTY OF EDUCATION
Te Kura Akoranga o Tāmaki Makaurau
Incorporating the Auckland College of Education

Relational

What might explain the differences between schools?

a) The schools had different ethical approaches
b) The teachers were of differing socioeconomic backgrounds
c) The teachers at one school were more caring
d) The schools had students from differing socioeconomic backgrounds

THE UNIVERSITY OF AUCKLAND
FACULTY OF EDUCATION
Te Kura Akoranga o Tāmaki Makaurau
Incorporating the Auckland College of Education

Extended Abstract

What do students look for in a teacher?

a) Friendliness, caring, and humour
b) An adult-figure not found at home
c) A person from a similar background
d) Whatever causes them to learn

THE UNIVERSITY OF AUCKLAND
FACULTY OF EDUCATION
Te Kura Akoranga o Tāmaki Makaurau
Incorporating the Auckland College of Education

asTTle (Assessment Tools for Teaching & Learning)

- Computer based online assessment tool
- Numeracy and Literacy (English and Māori)
- Curriculum based (year 4 & above)
- 2003-2005 CD-Rom, 2009 online (Ipad access under development)

THE UNIVERSITY OF AUCKLAND
FACULTY OF EDUCATION
Te Kura Akoranga o Tāmaki Makaurau
Incorporating the Auckland College of Education

asTTle Principles

- Free resource
- Voluntary (must be always be optional)
- Complements existing tests
- Open – no secrets
- Teacher driven, must be useful for teachers, loses purpose when required for external reporting

asTTle provides

- provides information about a student's level of achievement, relative to the curriculum achievement outcomes, for levels 2 to 6 and national norms of performance for students in years 4 to 12.
- 40-minute paper and pencil tests designed for their own students' learning needs. E-asTTle allows items to be completed online.

THE UNIVERSITY OF AUCKLAND
FACULTY OF EDUCATION
Te Kura Akoranga o Tāmaki Makaurau
Incorporating the Auckland College of Education

asTTle purpose

- To provide analysed assessment information to inform teaching and learning
- To provide externally referenced assessment information that will assist teachers to make valid, reliable, and nationally consistent judgements about the work and progress of their students

THE UNIVERSITY OF AUCKLAND
FACULTY OF EDUCATION
Te Kura Akoranga o Tāmaki Makaurau
Incorporating the Auckland College of Education

asTTle reports

The six major report formats provide 6 different ways of looking at the data from a single asTTle test.

1. Console Report

1. Tabular Output Report

2. Individual Learning Pathways Report

3. Group Learning Pathways Report

4. Curriculum Levels Report

5. What Next Report

THE UNIVERSITY OF AUCKLAND
FACULTY OF EDUCATION
Te Kura Akoranga o Tāmaki Makaurau
Incorporating the Auckland College of Education

At classroom level asTTle enables teachers to:

- Know at what level each learner is performing;
- Give learners focused feedback
- Personalize the learning to specific needs
- Develop and modify classroom programmes

THE UNIVERSITY OF AUCKLAND
FACULTY OF EDUCATION
Te Kura Akoranga o Tāmaki Makaurau
Incorporating the Auckland College of Education

At school level asTTle data can:

- be aggregated and used to evaluate teaching and learning and to inform strategic planning.
- Longitudinal data is an effective way of measuring school effectiveness.

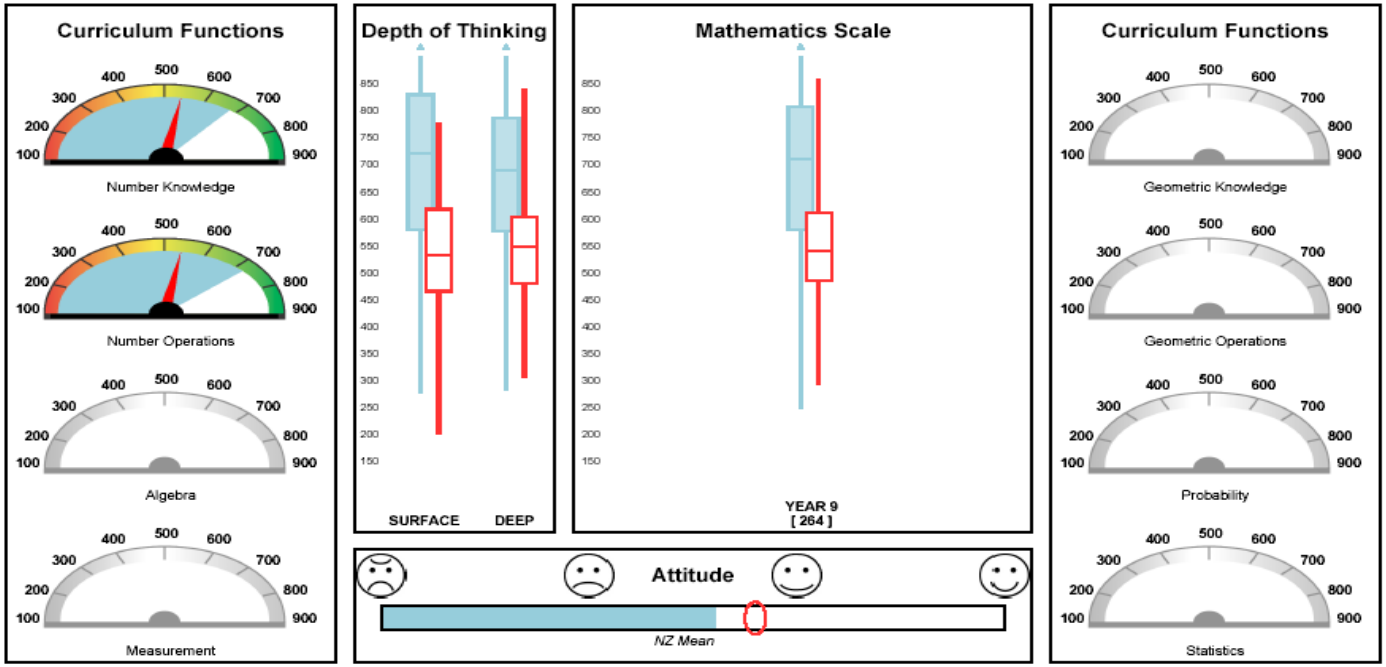
The Console Report

THE UNIVERSITY OF AUCKLAND
FACULTY OF EDUCATION
Te Kura Akoranga o Tāmaki Makaurau
Incorporating the Auckland College of Education

Console Report for Test: Entrance test Maths 2004
Group: All Test Candidates
Date Tested: 11 November 2003

Interaction Effects

Ethnicity: All
Year: 9
Gender: All
Language: All
Cluster: All Clusters
NZ Performance:
Location: All NZ Schools
Your Group Performance:
No. of Students: [*n*]



The Console Report in sections – the top

THE UNIVERSITY OF AUCKLAND
FACULTY OF EDUCATION
Te Kura Akoranga o Tāmaki Makaurau
Incorporating the Auckland College of Education

Console Report for Test: Entrance test Maths 2004
Group: All Test Candidates **Date Tested:** 11 November 2003

Interaction Effects

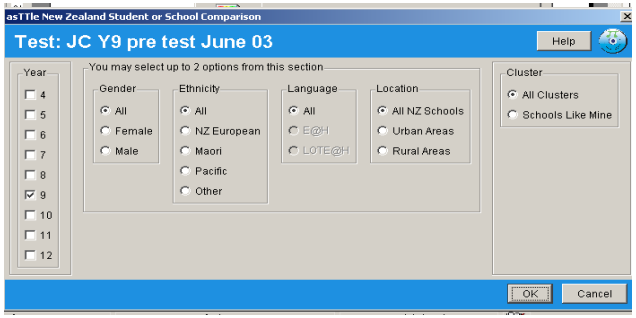
Ethnicity:	All	**Language:**	All	**Location:**	All NZ Schools
Year:	9	**Cluster:**	All Clusters		
Gender:	All	**NZ Performance:**		**Your Group Performance:**	**No. of Students: [*n*]**

General test information

The New Zealand comparisons you have chosen

For a multi-level class you can select one, two, or three year levels.

The default selection is for the year group with the most students in it and 'all' for every other category.



The Console Report in sections – the bottom

THE UNIVERSITY OF AUCKLAND
FACULTY OF EDUCATION
Te Kura Akoranga o Tāmaki Makaurau
Incorporating the Auckland College of Education

This shows the attitude your selected students have to the content tested on a scale shown by the smiley (or not) faces.

Your selected students' mean — remember some students will be outside the red circle.

The national mean for all students is shown by the green bar.

Remember that although attitude does not predict achievement it is still an important facet of children's learning.

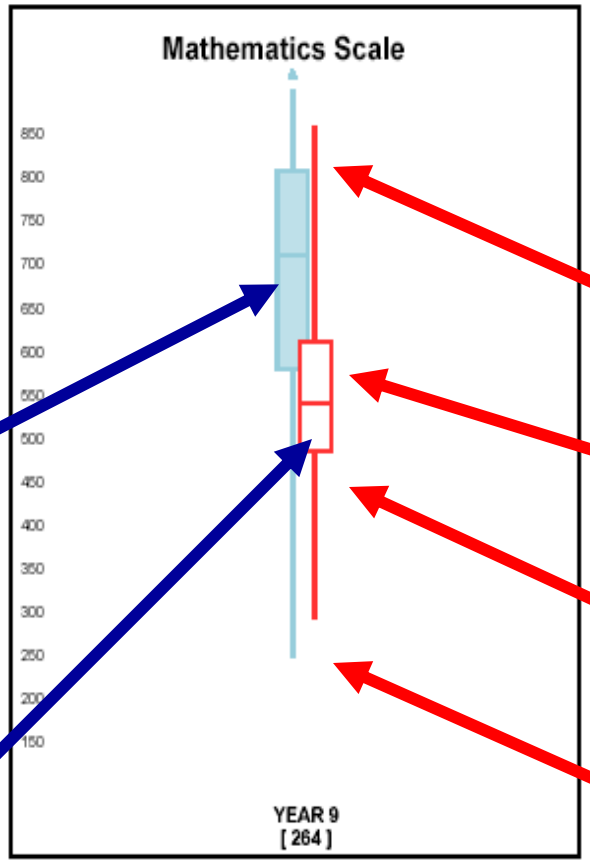
The Console Report in sections – the asTTle scales

THE UNIVERSITY OF AUCKLAND
FACULTY OF EDUCATION
Te Kura Akoranga o Tāmaki Makaurau
Incorporating the Auckland College of Education

This compares the distribution of scores for your class with the national distribution for reading, writing, or maths, based on the interaction effects you have chosen.

The national distribution is shown in blue

The **median** for your class is shown by the red line.



If you have chosen more than one year level in your class you will get a scale for each one.

Highest score

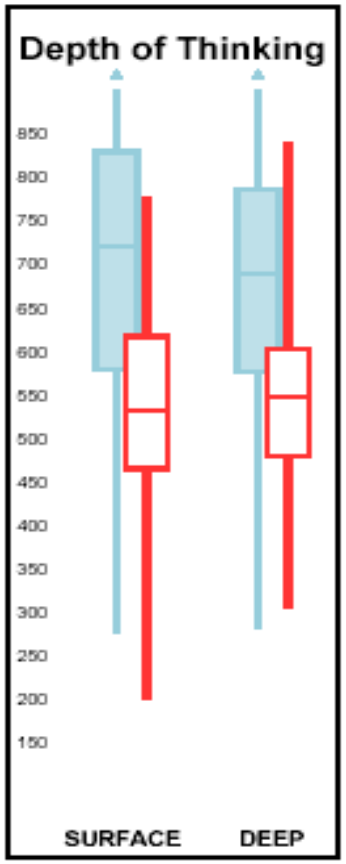
75th percentile

25th percentile

Lowest score

The Console Report in sections – Depth of Thinking

THE UNIVERSITY OF AUCKLAND
FACULTY OF EDUCATION
Te Kura Akoranga o Tāmaki Makaurau
Incorporating the Auckland College of Education



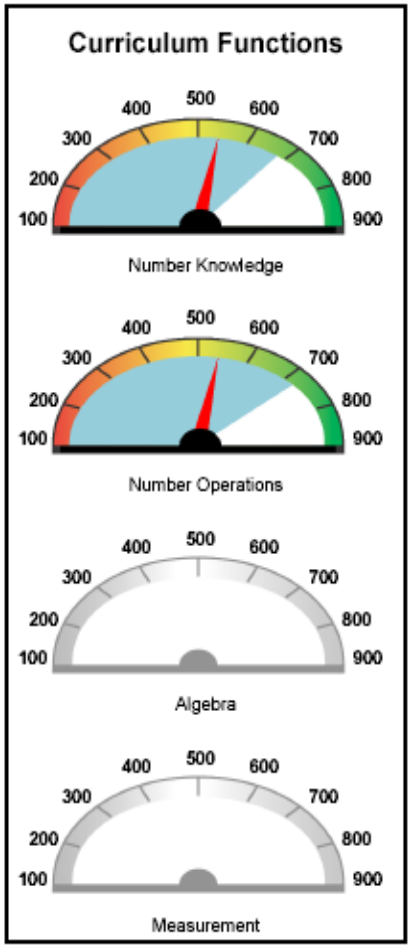
This shows the level of cognitive processing learners have used in the test. Both their surface thinking and their deep thinking is compared against the national mean for the comparison groups you chose.

Surface thinking is their ability to use one or unconnected lists of facts, information, or ideas to answer questions.

Deep thinking is their ability to relate the facts, ideas, or information to each other and to hypothesise about them in a more abstract manner.

The Console Report in sections – the sides

THE UNIVERSITY OF AUCKLAND
FACULTY OF EDUCATION
Te Kura Akoranga o Tāmaki Makaurau
Incorporating the Auckland College of Education



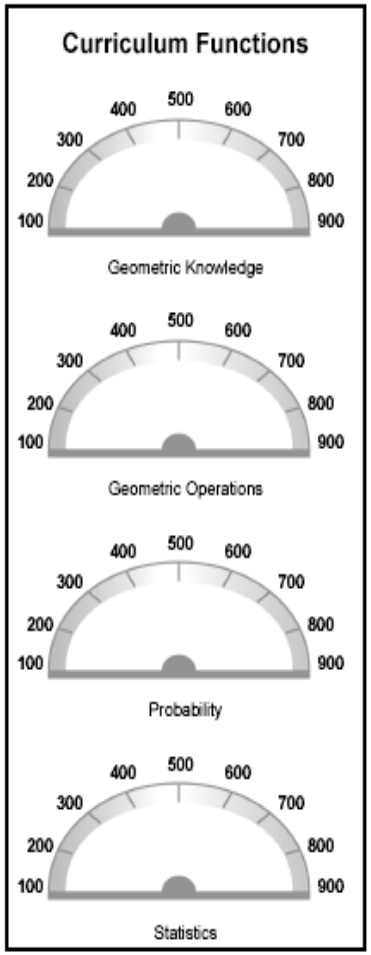
Information relating to the content areas you have focused your test on. Your class mean is compared to the national mean for the groups you have selected.

(For writing this would show all seven marking elements)

Note: Differences of more than 15 points (the standard error of measurement) are significant for teaching and learning.

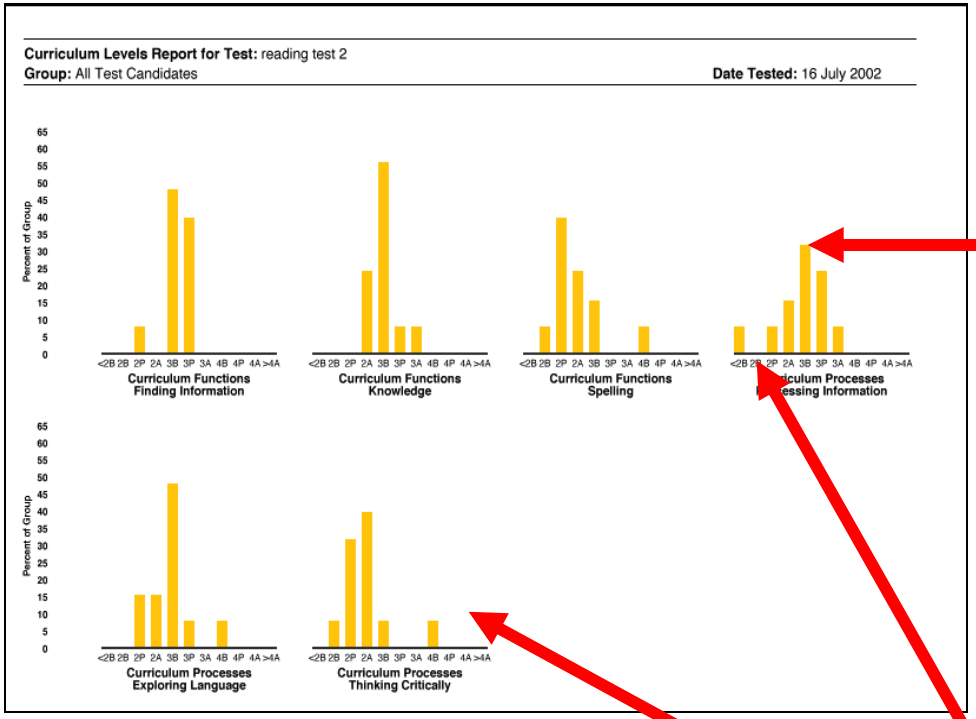
Your class mean is shown by the red arrow on the dial

The national mean for selected groups is shown by the blue shaded area



Curriculum Levels Report

THE UNIVERSITY OF AUCKLAND
FACULTY OF EDUCATION
Te Kura Akoranga o Tāmaki Makaurau
Incorporating the Auckland College of Education



This is the 'skyline' — showing you graphically the spread of your class over the curriculum levels.

Within each curriculum level there are three categories of ability to provide you with more precise information — basic (B), proficient (P), and advanced (A).

For reading – the curriculum functions you have tested are shown along with three curriculum processes.

For writing — the 'skyline'shows the seven elements the writing is marked on.

Curriculum Levels Report

THE UNIVERSITY OF AUCKLAND
FACULTY OF EDUCATION
Te Kura Akoranga o Tāmaki Makaurau
Incorporating the Auckland College of Education

Clicking on a graph will take you directly to a table showing which learners are at each level.

This report allows you to (a) group students appropriately and (b) monitor that learners are moving up levels throughout the year.

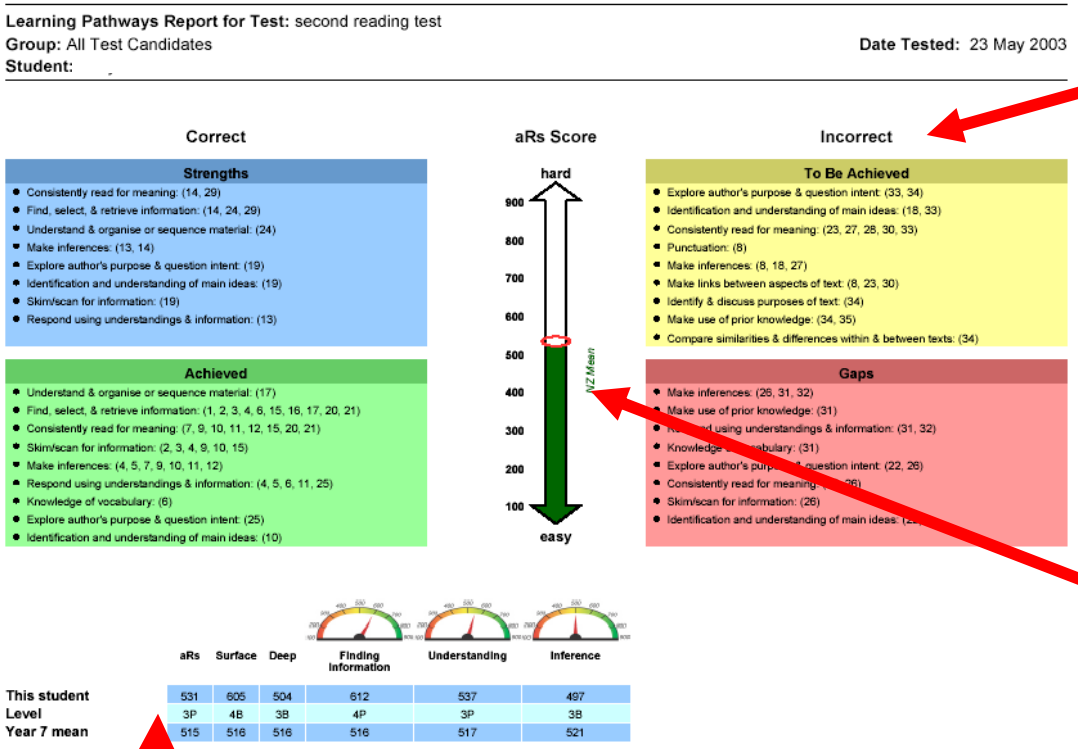
Curriculum Levels Report for Test: reading test 2
Group: All Test Candidates **Date Tested:** 16 July 2002

Finding Information (Click to Return to Graphs)

<2B	2B	2P	2A
		Sylvia Matthews	
3B	3P	3A	
Byron Elsworth Elanor Fatialofa Brian Johannsen Thomas Mirkwood Aroha Williams Jenna Wormswald	Susan Dillworthy Byron Emanual Samuel Freedman Anne Jacobs Louise Rose		
4B	4P	4A	>4A

Individual Learning Pathways Report

THE UNIVERSITY OF AUCKLAND
FACULTY OF EDUCATION
Te Kura Akoranga o Tāmaki Makaurau
Incorporating the Auckland College of Education



Learning Pathways Report for Test: second reading test
Group: All Test Candidates
Student:
Date Tested: 23 May 2003

Correct

Strengths
- Consistently read for meaning: (14, 29)
- Find, select, & retrieve information: (14, 24, 29)
- Understand & organise or sequence material: (24)
- Make inferences: (13, 14)
- Explore author's purpose & question intent: (19)
- Identification and understanding of main ideas: (19)
- Skim/scan for information: (19)
- Respond using understandings & information: (13)

Achieved
- Understand & organise or sequence material: (17)
- Find, select, & retrieve information: (1, 2, 3, 4, 6, 15, 16, 17, 20, 21)
- Consistently read for meaning: (7, 9, 10, 11, 12, 15, 20, 21)
- Skim/scan for information: (2, 3, 4, 9, 10, 15)
- Make inferences: (4, 5, 7, 9, 10, 11, 12)
- Respond using understandings & information: (4, 5, 6, 11, 25)
- Knowledge of vocabulary: (6)
- Explore author's purpose & question intent: (25)
- Identification and understanding of main ideas: (10)

aRs Score

hard
900
800
700
600
500
400
300
200
100
easy
NZ Mean

Incorrect

To Be Achieved
- Explore author's purpose & question intent: (33, 34)
- Identification and understanding of main ideas: (18, 33)
- Consistently read for meaning: (23, 27, 28, 30, 33)
- Punctuation: (8)
- Make inferences: (8, 18, 27)
- Make links between aspects of text: (8, 23, 30)
- Identify & discuss purposes of text: (34)
- Make use of prior knowledge: (34, 35)
- Compare similarities & differences within & between texts: (34)

Gaps
- Make inferences: (26, 31, 32)
- Make use of prior knowledge: (31)
- [illegible] using understandings & information: (31, 32)
- Knowledge [illegible]abulary: (31)
- Explore author's pur[illegible] question intent: (22, 26)
- Consistently read for meaning[illegible]26)
- Skim/scan for information: (26)
- Identification and understanding of main ideas: [illegible]

	aRs	Surface	Deep	Finding Information	Understanding	Inference
This student	531	605	504	612	537	497
Level	3P	4B	3B	4P	3P	3B
Year 7 mean	515	516	516	516	517	521

These reports are for individual learners to enable planning for specific needs. Each item in the test is placed in one of four quadrants.

The asTTle Reading scale (aRs) — this is the learner's overall mean score (shown by the red oval) compared to the national mean score (shown by the coloured bar).

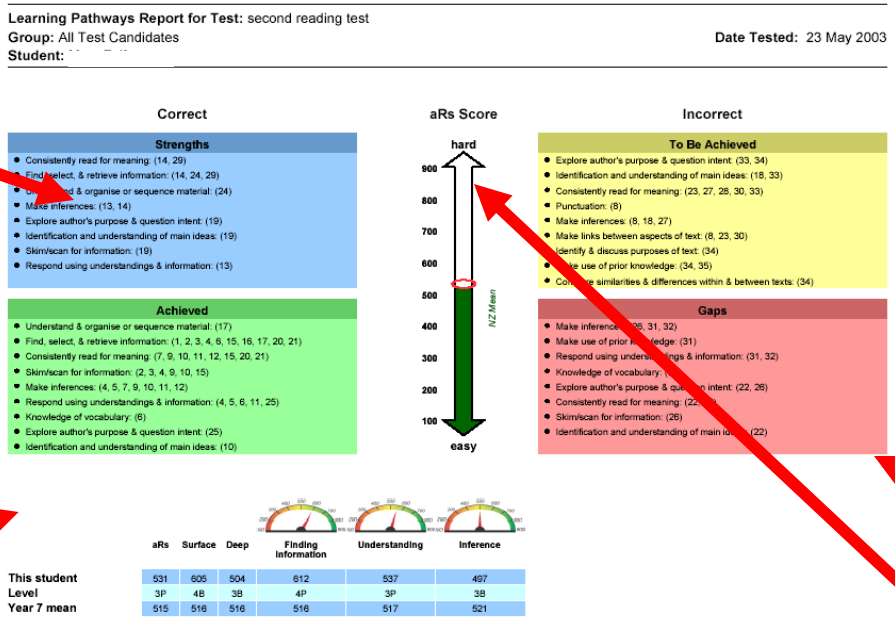
Console information for individual students gives scores and levels for: the content areas tested overall, surface and deep thinking, and the national mean for their year group.

Individual Learning Pathways Report

THE UNIVERSITY OF AUCKLAND
FACULTY OF EDUCATION
Te Kura Akoranga o Tāmaki Makaurau
Incorporating the Auckland College of Education

‘Hard’ questions the student got right

‘Easy’ questions the student got right



Learning Pathways Report for Test: second reading test
Group: All Test Candidates
Student:
Date Tested: 23 May 2003

Correct

Strengths
- Consistently read for meaning: (14, 29)
- Find, select, & retrieve information: (14, 24, 29)
- Understand & organise or sequence material: (24)
- Make inferences: (13, 14)
- Explore author's purpose & question intent: (19)
- Identification and understanding of main ideas: (19)
- Skim/scan for information: (19)
- Respond using understandings & information: (13)

Achieved
- Understand & organise or sequence material: (17)
- Find, select, & retrieve information: (1, 2, 3, 4, 6, 15, 16, 17, 20, 21)
- Consistently read for meaning: (7, 9, 10, 11, 12, 15, 20, 21)
- Skim/scan for information: (2, 3, 4, 9, 10, 15)
- Make inferences: (4, 5, 7, 9, 10, 11, 12)
- Respond using understandings & information: (4, 5, 6, 11, 25)
- Knowledge of vocabulary: (6)
- Explore author's purpose & question intent: (25)
- Identification and understanding of main ideas: (10)

aRs Score

hard
900
800
700
600
500
400
300
200
100
easy
NZ Mean

Incorrect

To Be Achieved
- Explore author's purpose & question intent: (33, 34)
- Identification and understanding of main ideas: (18, 33)
- Consistently read for meaning: (23, 27, 28, 30, 33)
- Punctuation: (8)
- Make inferences: (8, 18, 27)
- Make links between aspects of text: (8, 23, 30)
- Identify & discuss purposes of text: (34)
- Make use of prior knowledge: (34, 35)
- Compare similarities & differences within & between texts: (34)

Gaps
- Make inferences: (28, 31, 32)
- Make use of prior knowledge: (31)
- Respond using understandings & information: (31, 32)
- Knowledge of vocabulary: (
- Explore author's purpose & question intent: (22, 26)
- Consistently read for meaning: (22
- Skim/scan for information: (26)
- Identification and understanding of main ideas: (22)

	aRs	Surface	Deep	Finding Information	Understanding	Inference
This student	531	505	504	512	537	497
Level	3P	4B	3B	4P	3P	3B
Year 7 mean	515	516	516	516	517	521

‘Hard’ questions the student got wrong

‘Easy’ questions the student got wrong

The placement of the items in the four quadrants relates to the student’s level. ‘Hard’ items are those that would be difficult for this student, and ‘easy’ items are those that we would expect the student to get right – they are easy for this student.

aRs

Individual Learning Pathways Report – implications for teaching

THE UNIVERSITY OF AUCKLAND
FACULTY OF EDUCATION
Te Kura Akoranga o Tāmaki Makaurau
Incorporating the Auckland College of Education

Strengths
Take advantage by giving the student similar work at this level

To be achieved
Plan to teach these objectives at this level within the next term

Achieved
Stop teaching this type of material at this level to this student

Gaps
Investigate causes but don't 'skill & drill' teach these objectives – they are easy and the student will learn them quickly

Improving the quality of teacher based assessment (1)

- Teachers need to know fundament concepts of assessments
- Teachers need to be able to critique existing assessments
- Teachers may not always have time to create their own assessments, when doing so need to be aware of their limitations
- Teachers need assessment standardized tools that can provide high quality information on students

THE UNIVERSITY OF AUCKLAND
FACULTY OF EDUCATION
Te Kura Akoranga o Tāmaki Makaurau
Incorporating the Auckland College of Education

Improving the quality of teacher based assessment (2)

- Successful high quality tools need to have teacher input
- Tools need to revised on a regular basis
- Research needs to inform teacher practice in the classroom

References

THE UNIVERSITY OF AUCKLAND
FACULTY OF EDUCATION
Te Kura Akoranga o Tāmaki Makaurau
Incorporating the Auckland College of Education

- Biggs, J. B., & Collis, K. (1982). *Evaluating the Quality of Learning: The SOLO Taxonomy.* New York: Academic Press.
- Brown, Gavin T.L., Irving, S. Earl, & Keegan, Peter J. (2008). *An Introduction to educational assessment, measurement, and evaluation: Improving the quality of teacher-based assessment* (2nd ed.). Auckland: Pearson Education.
- Brown, G. T. L. (2012) Teachers' thinking about assessment: Juggling improvement and accountability. *Teacher: the International Education Magazine, 6*(2), 30-35.
- Collis, K., & Biggs, J. (1986). Using the SOLO Taxonomy. *set: Research information for teachers, 2*, item 4.
- Hattie, J., & Purdie, N. (1998). The Solo model: Addressing fundamental measurement issues. In B. Dart, & G. Boulton-Lewis. (Eds.). *Teaching and learning in higher education.* Melbourne: ACER.
- Keegan, Peter J., Brown, Gavin T.L., & Hattie, John A.C. (2013). A psychometric understanding of sociocultural factors in test validity: The development of standardised test materials for Māori medium schools in New Zealand. In S. Phillipson, K. Ku & N. Phillipson (Eds.), *Constructing Achievement: A Sociocultural Perspective* (pp. 42-54). London: Routledge.
- McMillan, J. H. (2001). *Classroom assessment: Principles and practice for effective instruction* (2 ed.). Boston, MA,: Allyn & Bacon