

# Report on Teacher Feedback from the First Calibration of Mathematics Assessments

Technical Report 24 Project asTTle, University of Auckland, 2002

Lyn Lavery & Gavin T. L. Brown  
University of Auckland

This report summarises feedback responses of the teachers ( $n=243$ ) involved in the first nationally representative calibration of asTTle mathematics papers in November 2001. Feedback was generally positive and from both teachers and students.

## Table of Contents

Question One .....	1
Question Two.....	2
Question Three.....	2
Question Four.....	2
Question Five.....	3
Question Six.....	3
Concluding comment .....	4

A total of 253 schools, a nationally representative sample, were invited to participate in the first calibration of the asTTle mathematics assessments conducted in November 2001. Of this population, 109 schools offered approximately 10,000 students, reasonably in proportion to the nation in regards to decile, school type, and school size, participated.

A total of 24 mathematics papers were calibrated. Eight mathematics papers for each year level (5, 6, and 7) were sent out, resulting in data from just over 9,000 students. The curriculum level of tasks in each paper was designed with the general year level of student in mind. In other words, Year 5 papers had less Level 4 materials and more Level 2 materials, while Year 7 papers had less Level 2 materials, and more Level 4 materials. It should be noted that this arrangement might nevertheless result in individual children receiving a paper that could have been too hard or easy depending on their own progress in mathematics.

Each teacher who administered the tests was asked to complete a teacher feedback form. The potential pool of respondents, estimated at about one teacher per 25 students, would be about 400. A total of 243 replies were received.

This response sample represented just over one half of all teachers potentially involved in administering the asTTle papers.

Responses were in the nature of comments to prepared questions. The comments were generally coded using a “Yes”, “No”, “Both yes and no”, or “No answer”. The category “Yes” indicates a favourable or positive response to the question, a “No” an unfavourable or negative response to the question, and “Both yes and no” indicates a response that contains both positive and negative comments. “No answer” includes comments that were incapable of meaningful interpretation.

## Question One

*Was the content appropriate for the age level and ability of the students?*

Overall, teachers were generally satisfied with the content in relation to the age level and ability of their students. Less than a fifth of responses were negative in nature. Approximately one quarter of responses fitted into the ‘both’ category, reflecting the difficulties in meeting the needs of every student.

Table 1  
*Appropriateness of content*

	Yes	No	Both	No answer	Total
N	140	40	58	5	243
Percentage	58	16	24	2	100

Favourable responses included: “Yes – the range allowed for all students to be able to answer some questions. Content was familiar to

*them and their understanding*” and *“I felt the content related very closely to the achievement objectives for Level 3 and 4”*. Responses from the ‘both’ category generally reflected mixed abilities in the classroom, for example, *“As for most class work - some struggled, and some found it easy”*. Of the unfavourable comments that were made, the majority either reflected that the assessment included content that they had not yet covered, or referred to the fact that ESOL and lower ability children struggled with the language used.

### Question Two

*Was the content interesting and engaging for the students?*

Teachers were overwhelmingly positive in regards to this question. Typical responses included: *“All the students found it interesting and wanted to keep doing it”* and *“The variety of methods for answering kept the children focused”*. While a tenth of responses were negative, teachers often qualified these as being due to environmental factors; for example, *“We have just done a lot of tests so they were not focused”*. Responses in the ‘both’ category included *“Although not all students found the test “interesting” they were all engaged”* and often mentioned that while the majority were interested and engaged a small proportion of the class were not.

Table 2

*Content interesting and engaging*

	Yes	No	Both	No answer	Total
N	188	23	26	6	243
Percentage	77	10	11	2	100

### Question Three

*Were the teacher instructions clear, easy to follow, and sufficient?*

The majority of teachers responded positively this question. While many simply gave a response of ‘Yes’, others were more specific. For example, *“Yes, the test was easy to follow, instructions were clear and it was simple to administer”*.

Table 3

*Teacher Instructions*

	Yes	No	Both	No answer	Total
N	189	25	24	5	243
Percentage	78	10	10	2	100

Negative responses often included suggestions for improvement. The four most commonly mentioned suggestions were: including an example ‘bubble’, clearer instructions for handing out, reducing wordiness and having a spare copy of the booklet for the teacher (particularly for when children are completing practice questions).

### Question Four

*Was the level of difficulty appropriate for all students?*

The issue of difficulty elicited a particularly mixed response from teachers, and due to the wide variation in responses this question was analysed slightly differently to the previous questions. While the four standard categories were used for more general responses, an additional three categories were created. These were ‘Difficult for a small group’, ‘Difficult due to language used’, and ‘Too hard/Too easy’. Overall the responses reflected the difficulties in designing an assessment suitable for all children, particularly given the wide range of abilities that so commonly occurs in classrooms. This was well reflected in one teacher’s comment, *“... there is an implication that all students on a particular year are at the same level – this is not the case. Therefore, no, it was not appropriate for all students”*.

Table 4  
*Level of difficulty*

Response	N	Percentage
Yes	81	33%
No	16	7%
Both	35	14%
Difficult for a small group	51	21%
Difficult due to language used	23	10%
Too hard/Too easy	24	10%
No answer/Un-interpretable	13	5%
Total	243	100%

Generally responses classified in the ‘Yes’ category were a positive response, however several teachers were more specific, for example, “A good mixture of easy to challenging questions presented”. Responses in the ‘both’ category generally either reflected a mixed class response, “Some struggled, some found easy”, or a mixture of positive and negative comments, for example “No, but it served as a good general overview of what a ‘typical’ student should understand at this level”. The ‘No’ category included responses that simply indicated ‘no’, or a general negative response. Comments about difficulty were included in the specifically created categories. One tenth of teachers felt that students found the test difficult due to the language used. As one teacher remarked, “The use of unfamiliar mathematics terminology in instructions made tasks more difficult”. A fifth of teachers indicated that the test was too difficult for a small group of their children. Over half the responses in this category made reference to ESOL, special needs, or ‘low-ability’ children. A typical comment from this category was as follows, “Children who are NESB or academically challenged were easily confused”. Another tenth of respondents felt that the test was either too difficult, “Overall the level of difficulty was too hard” or too easy, with one teacher commenting that they felt the questions were “ridiculously easy”.

Question Five

*What was the response of the students to the papers?*

All but four of the teachers commented on their students’ response to the papers. Only

fourteen percent reported negative responses, with the remainder reporting either a positive or neutral/mixed class response.

The wide variety of responses to this question can be seen in the example comments below:

- “Although they found the test hard they all said that they enjoyed doing it. The test led to a discussion of different maths concepts”
- “Children commented that it was ‘cool’ – the diagrams were interesting and kept them focused.”
- “Some children felt that the instructions were confusing. They said that there were some questions that were really easy but they found others quite difficult. Overall they thought it was okay.”
- “The students’ response to the paper was that it was too long and boring. When asked why it was boring the answer was they didn’t know how to do things, questions were hard and difficult to understand.”
- “They seemed to enjoy it and were happy to have a different challenge.”

Table 5  
*Student responses to paper*

Comment	N	Percentage
Positive	110	45%
Liked/enjoyed/fun	49	20%
OK/alright/no complaints	12	5%
Good/positive	45	18%
Liked assisting the University	4	2%
Negative	34	14%
General negative comment	12	5%
Too hard/too long/too easy	9	4%
Confusing/hard to understand	13	5%
Neutral/Mixed	88	36%
Neutral/ liked & disliked some	3	1%
Mixed general reaction	85	35%
Response unclear	7	3%
No answer	4	2%
Totals	243	100

Question Six

*Are there any general comments you would like to make?*

Just over half the teachers chose to make a general comment, and the responses were reasonably evenly mixed between positive, negative, neutral, or suggestions for improvement.

Table 6  
*General comments*

Response	N	Percentage
Positive	31	13%
Negative	36	15%
Both	6	2%
Neutral	29	12%
Suggestion for improvement	38	16%
No answer/Un-interpretable	103	42%
Total	243	100%

Positive comments often related to the project as a whole, for example: *“Can’t wait to get results and also use this tool for all my students”* and *“I am encouraged by the development and look forward to assisting in future”*. Other positive questions related more to the test itself, for example, *“Generally very impressed with the quality and relevance of questions”*. Negative comments generally related to issues such as the time limit, the difficulty, or the question wording. Suggestions for improvement were varied and often related to specific questions within the papers, setting-out of the booklets or to instructions (for either the teacher or the children). Several teachers also commented that the children would have preferred to receive the tests back. As one teacher commented, *“Students like feedback of their own. I think they would prefer it if the paper came back to them marked”*. Neutral comments usually related to some aspect of testing such as calculators not being available, student absences, etc.

#### Concluding Comment

Overall, the papers were well received by the teachers and students. Teachers commented particularly favourably on the interesting and engaging content as well as the clarity of teacher instructions. There were a few issues regarding level of difficulty although this generally applied to a small percentage of students. Several suggestions for improvement were made, often relating to the language used and general instructions given.

It should be noted, within the asTTle tool teachers will be able to prioritise the curriculum level of a test and will be able to inspect the test before administration. From this they can

identify whether certain students will require accommodations or exemptions, prior to running in-class assessments. It is expected that teachers will exercise professional judgment in determining the appropriateness of any test they create for their class of students.