PERFORMANCE OF PRIMARY 4 STUDENTS AT PERFORMANCE BASED ASSESSMENT TASKS

Ang Kok Cheng
CHIJ St Nicholas Girls' School
Singapore

Boo Hong Kwen
National Institute of Education
Nanyang Technological University
Singapore

Toh Kok Aun
National Institute of Education
Nanyang Technological University
Singapore

Abstract

To date, little is known about our pupils' ability in applying mathematical concepts to solving real-world hands-on problems – an essential individual survival skill in the 21st century. With the mode of assessment remaining largely unchanged over the last decade, a teacher's perception of his/her pupils' 'ability to do' relies much on the traditional pen-and-paper test which may not be holistic in assessing the application skills of the individual. This paper reports on a study which was aimed at finding out pupils' strengths and weaknesses in approaching hands-on problems in mathematics. Such information could be useful in helping classroom teachers to customise their lessons in a way that would enhance pupil learning and problem-solving. Five-hundred-and-four students (N=504) from 14 intact classes of five mainstream primary schools took part in the study. Five mathematics-related performance assessment tasks replicated from the TIMSS-95 Performance Assessment sub-study were used as the main instrument for data collection and the results analysed and reported.

Introduction

Locally, thinking is infused through the introduction of the Thinking Programme, which incorporates the eight core thinking skills from Marzano's Dimensions of Thinking into our Dimension of Learning Framework. The eight core skills include: generating, integrating, evaluating, focusing, information gathering, remembering, organising and analysing (Chua & Leong, 1997). Students are imbued with the dispositions for independent, lifelong learning, enabling them to apply these skills in real-life decision making and problem solving situations and to help them become critical, creative and self-regulated thinking learners. The present day role of the teacher has also shifted from a dispenser of knowledge to that of a resourceful facilitator skilled in questioning, flexible in applying different approaches, encouraging to those who need encouragement and humour at times to give the lesson some "kick".

However, infusion of thinking into the curriculum did not begin with these new initiatives. In the revised mathematics syllabus (MOE, 1990) for the New Education System (Primary), thinking skills are already manifested as the processes involved in mathematical problem solving and it comes in the form of metacognition, deductive and inductive reasoning. Schoenfeld (1989) commented that the move towards a problem solving based curriculum can be seen as going against rote, algorithmic mathematics, and more importantly, an attempt to introduce to students significant aspects of mathematical thinking.

Whilst mathematics education, in the local context, has received praises around the world, it may not be as holistic as it appears. Mathematical understanding is often measured within an examination setting revolving around single product problems. Much of assessment is still skewed towards standardized testing, while they served worthy pedagogical goals, it has denied our students the exposure to the larger aims and scope of a mathematical education. Students' ability to demonstrate the skill of reasoning, communicating about mathematics and expressing more complete conceptual understanding are also constantly being limited.

Rationale

While we took years to complete the implementation of the 1990 revise syllabus, almost a decade has since lapsed. With the reduction in syllabus content in 1999 and the new revision in 2000-2006 to replace the 1990 syllabus, we see ourselves at the crossroad again – what to adopt, what to discard and what to retain. With our vision of Thinking Schools Learning Nation (TSLN), it seems timely and necessary for us to know where, along the path, our students are in their "ability to do" and how best we can assess them more effectively for a knowledge-based ability-driven economy.

Significance

The primary aim of the study was to gain some insight into our students' "ability to do" by administering a hands-on performance-based mathematics assessment. Secondary to this, adopting the TIMSS-95 Performance Assessment sub-study tasks also serves as an indicative benchmark to find out whether our current Singapore sample is only as able as or better than the International community that participated in the TIMSS-95 sub-study (i.e. it tells a lot about our students' ability if they were placed below or at the same competency level as those of the 1995 international average).

Design and Data Collection

Administration procedures and performance tasks from the TIMSS-95 sub-study were adopted and used as the main source of data collection. Sample for the study was drawn from five-hundred-and-four (N=504) mainstream Primary 4 students from 14 intact classes of five Primary schools. An opportunity random sampling was used to select the schools from different school types (e.g. all boys, all girls, coed schools). Intact classes were used because the school set-up in Singapore did not facilitate a random pick of students from the Primary 4 cohort, as it is deemed to be too disruptive to the regular instructional programme. As the study adopted the rotation scheme (e.g. rotation 1 and 2) drawn up in the original sub-study, only three sets of data was collected for each task for every nine students sampled. A total of 168 sets of data were collected for each performance task. It was also not possible for the researcher to revisit the school to verify erroneous submissions; incomplete data sets collected were dropped from the analysis.

The 12 original tasks from the TIMSS-95 sub-study were reconstructed using information available from the IEA's TIMSS Performance Assessment Report (Harmon et al., 1997). For the study, seven of the twelve tasks, five mathematics-related tasks and two combined mathematics-and-science tasks were selected viz. : Mathematics Tasks (Dice, Calculator, Folding & Cutting, Around the Bend, Packaging) and Combined Mathematics and Science Tasks (Shadows, Plasticine). To ensure that data was collected in accordance to the administration procedures as reported in Harmon et al, 1997, data was jointly collected with another similar study examining the science performance aspects for Population 1. The reconstructed tasks were validated and examined for accuracy. The original scoring rubrics were also adopted (Harmon et al., 1997) and scoring standardized.

After the validation process was completed, a pilot test was carried out in two schools to ascertain how robust the reconstructed tasks were, e.g. durability and functionality, as well as to gather feedback from students with regards to the clarity of instructions given and whether they could relate to the problems presented to them. The twelve tasks were presented at nine different stations (Table 2A). This resulted in three stations with one "short" science and one "short" mathematics task each, two stations with one "long" science task each, two stations with one "long" mathematics task each and two stations with one combined mathematics / science task each. Each station required 30 minutes of working time. Each student was assigned to three stations, for a total testing time of 90 minutes. Because the complete circuit of nine stations enables nine students at one time, students were selected in sets of nine for the performance assessment.

Table 2A : Assignment of Tasks to Stations

Station	Task	
A	S1 M1	Pulse Dice
B	S2 M2	Magnets Calculator
C	SM1	Shadows
D	S3 M3	Batteries Folding & Cutting
E	S4	Rubber Band
F	M5	Packaging
G	S5	Containers
H	M4	Around the Bend
I	SM2	Plasticine

Five identical setups for each of the tasks were prepared, which allowed for 36 students to be sampled simultaneously; which translated to 12 data sets collected for each task. Should the class exceed 36, the fifth set of setup was then used to collect extra data sets. For an intact class, two sets of nine students were selected to follow the scheme for Rotation 1 and another two sets of nine students to follow the scheme for Rotation 2 (Table 2B).

Table 2B: Assignment of Students to Stations

Student Sequence Number	Rotation 1 Stations	Rotation 2 Stations
1	A, B, C	A, B, E
2	B, E, D	B, D, G
3	C, F, E	C, A, D
4	D, G, H	D, E, F
5	E, A, G	E, I , H
6	F, H, B	F, H, A
7	G, I , F	G, F, I
8	H, C, I	H, G, C
9	I, D, A	I, C, B

Should the class exceed 36 students, these students were assigned to follow the scheme for Rotation 1 and extra data sets were collected (Table 2C).

Table 2C: Station Assignment for a typical class of 40 students

Student	Station Assignment
1 to 9	Rotation 1 : Student Sequence 1 - 9
10 to 18	Rotation 2 : Student Sequence 1 - 9
19 to 27	Rotation 1 : Student Sequence 1 - 9
28 to 36	Rotation 2 : Student Sequence 1 - 9
37 to 40	Rotation 1 : Student Sequence 1 – 4

Prior to each administration, students were briefed on the rotation scheme assigned to them, a brief introduction of all the twelve tasks available as well as reminder to return the apparatus to their original position before moving on to the next station. A dry run was also conducted to familiarise students of the sequence of stations to go to. Students were also encouraged to provide detailed answers for open-ended questions.

During the administration of the data collection sessions, instructions perceived as unclear or ambiguous were clarified on the spot. The sessions were also closely monitored by the researchers to ensure that no exchange of views among students took place during the sessions. This was to ensure that students' personal mathematical understanding and their ability to apply conceptual knowledge were not under any external influence.

Scoring rubrics reproduced from the TIMSS-95 sub-study were used to score the students' responses. Four coders, including the researcher, were involved in the scoring process. All the coders were each given 10 identical scripts to score and the high values for the inter-rater reliability correlation give confidence concerning the reliability of the scoring rubrics for scoring the students' responses.

Assumptions

The study was based on the fact that the Primary 4 students had undergone basic mathematics education in schools. This is a fair assumption as mathematics is taught as a core subject from Primary 1 in Singapore. Hence, students were assumed to have an equal exposure to mathematics instruction.

The students selected were assumed to be able to handle all the tasks assigned to them even though the topics coverage at the TIMSS-95 sub-study included topics beyond the local syllabus coverage for Primary 4 Mathematics (Table 3).

Table 3 : Comparison of TIMSS-95 Content Coverage vs Local Syllabus Coverage

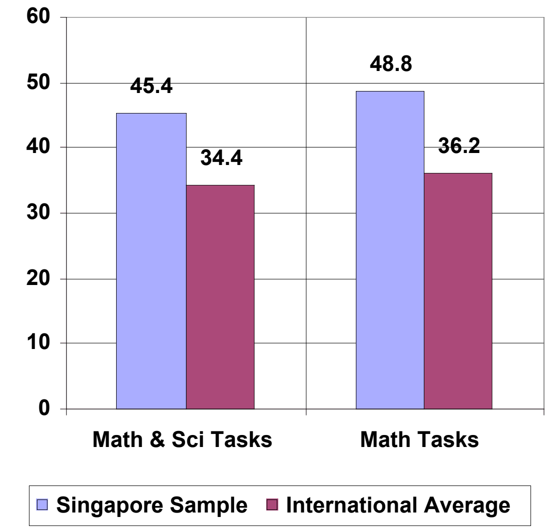
Task	**Content**	**Covered in the Mathematics Syllabus**
Dice (M1)	- Whole Number Operations - Data Representaiton and Analysis - Probability	Yes Yes No
Calculator (M2)	- Whole Numbers: Meaning and Operations - Data Representaiton and Analysis	Yes Yes
Folding & Cutting (M3)	- Geometry: Symmetry Transformation	No
Around the Bend (M4)	- Measurement and Units - Geometry: Position, Visualisation and Shape	Yes
	- 2D Polygons	No
	- 3D	No
	- Proportionality Problems	No
Packaging	- Measurement and Units	Yes
(M5)	- Geometry: Position, Visualisation and Shape	
Shadows (SM1)	- 3D - Measurement and Units - Geometry: Position, Visualisation and Shape - 2D Polygons - Geometry: Symmetry, Congruency and Similarity - Proportionality Problems	No Yes No Yes No
Plasticine (SM2)	- Measurement and Units - Proportionality Concepts and Problems	Yes No

Findings

Research Question 1
With the understanding that Singapore did not take part in the TIMSS-95 sub-study, how would Singapore Primary 4 students perform compared with the average student who made up the international community of the sub-study?

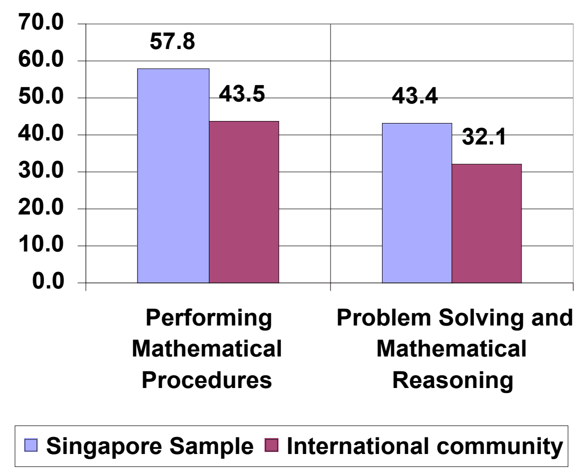
Our Singapore sample scored an overall performance average of 45.4% for the five mathematics-related and the two science-related tasks. This average increased to 48.8% when only the mathematics-related tasks were considered. Both scores were significantly better than the international performance average of 34.4% and 36.2% respectively at the 0.05 level of confidence. The difference in performance average between the two groups was found to be 11.0% (i.e. for mathematics and science tasks) and 12.6% (i.e for mathematics tasks only) which suggested that our Singapore sample found it easier with number manipulation than science hands-on activities.

Figure 1A: Comparison of Overall Performance Average (%) between Singapore Sample and international community



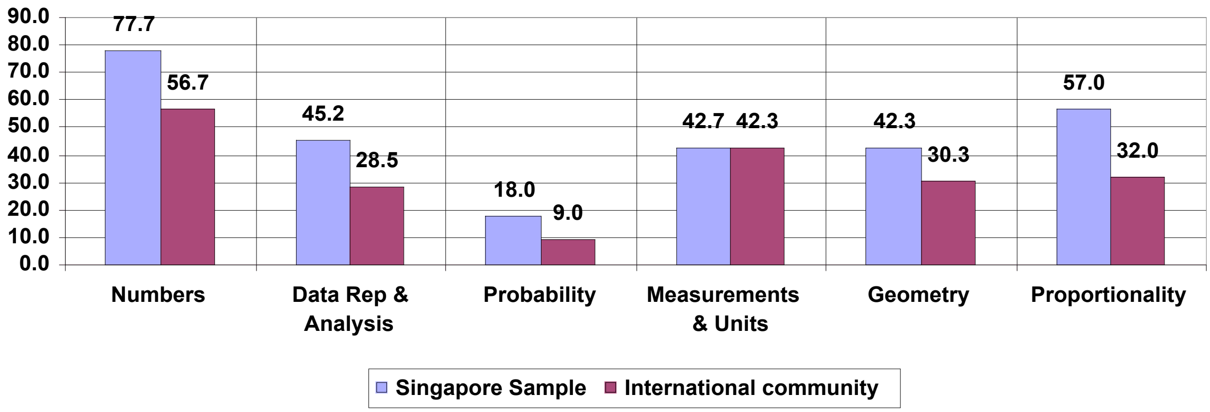
In the comparison of main performance expectations (i.e. performing mathematical procedures and problem solving and mathematical reasoning), our Singapore sample was also able to perform significantly better than the international community at the 0.05 level of confidence. The difference in performance average between the two groups was 14.3% (i.e. performing mathematical procedures) and 11.3% (i.e. problem solving and mathematical reasoning) which suggested that they were better in using equipment, performing routine procedures and using more complex procedures than with developing strategy, solving problems, predicting, generalising and conjecturing.

Figure 1B: Comparison of Overall Performance Expectation (%) between Singapore Sample and international community (Mathematics + Plasticine Tasks)



In the comparison of performance expectations sub-categories, our Singapore sample was also able to perform significantly better than the international community, at the 0.05 level of confidence, in all categories except in the category of measurements and units. Our Singapore sample encountered least difficulties with numbers (77.7%) and most difficulties in probability (18.0%). As the findings for performance expectation sub-category of probability and proportionality were each based on the results collected for one sub-task, the findings may not necessary reflect actual performance ability of the students.

Figure 1C : Comparison of Performance Expectation Sub-categories (%) between Singapore Sample and international community



These findings were found to be encouraging and it allowed the researcher to place our Singapore sample separate from the international community that took part in the TIMSS-95 sub-study. While the Singapore sample performed slightly better than the international community in the performance sub-category of measurement and units, the results were not significantly different. More work may be carried out to identify our students' weaknesses in this area so as to bring their ability to a higher level.

Research Question 2

How would Singapore Primary 4 students perform in each of the five TIMSS-95 Mathematics-based tasks and the two Combined Mathematics-and-Science tasks?

Our Singapore sample was able to obtain a higher score than the international community for all the tasks, except the Shadow task. The three greatest differences in scores between the two groups were noted for the Calculator task, followed by the Dice task and the Around-the-Bend task. The Around-the-Bend task was the easiest to complete successfully for our Singapore sample and the Shadow task hardest to complete. They were also able to perform significantly better than the international community, at the 0.05 confidence level, for the Dice, Calculator, Folding & Cutting, Around-the-Bend and Plasticine tasks but not for the Packaging and Shadow tasks.

Further examination revealed that the Packaging task tested mathematical concepts beyond the formal curriculum of our students. However, for the Shadow task, the science concept of light and shadow formation was already taught formally in the classroom. Their poor performance requires further investigation especially in the following areas:

a) Shadow of an object becomes smaller as the object is brought closer to the light source and vice versa (misconception).
b) Concept of light traveling in a straight line and spreading out from a source.
c) Hands-on practical ability especially in the category of measurement and units.
d) Written communications skills to make illustration of observations made and to provide explanation to such observations.
e) Students' lack of ability in making generalisations.

Research Question 3
What are the difficulties encountered by our Singapore Primary 4 students when performing each of the TIMSS-95 tasks adopted for the study?

a) The Dice Task (Performance Average : 59.0%)

Students had difficulties answering questions requiring them to describe or explain (e.g. describing pattern that is consistent with data presented; explaining why things happened in the manner observed). Carelessness on their part also attributed to partially correct or wrong answers (e.g. transfer error, inaccurate counting / recording of frequency). Students were capable of basic pattern recognition but most could not go beyond that of basic observation.

b) The Calculator Task (Performance Average : 59.0%)

Students were found to be lacking in communication skills and had difficulties articulating their answers in writing to others. Some students were also found to be careless in their work (e.g. inaccurate multiplication of a four-digit number by another four-digit number even with the use of the calculator). Students were capable of basic pattern recognition but most could not go beyond that of basic observation.

c) The Folding & Cutting Task (Performance Average : 46.0%)

It was found that students who took time planning and observing the cause-and-effect of an action were more successful than those who did not (e.g. having made an unsuccessful attempt, student took time off to examine the effects of the previous attempt before making a second attempt). Students who were unsuccessful either failed to follow the instructions given or did not know what to do.

d) The Around-the-Bend Task (Performance Average : 60.0%)

It was observed that quite a number of students had difficulties with the conversion of measurement involving half measurements (e.g. 1.5 m). There was also evidence to suggest that students had a higher success rate given hands-on concrete activities which helped them to visualise the answer.

e) The Packaging Task (Performance Average : 20.0%)

The study found that students had little or no prior concept of nets. Though a detailed explanation was given in the instructions sheet, those who were unsuccessful seemed to demonstrate a lacking in their ability to comprehend the instructions given. This inability may be interpreted either as :

1 **2**

i) an inability to response to non-routine open-ended problems, or
ii) the concept of nets may be too abstract for the Singapore sample for their age though it may have been included in the Population 1 mathematics curriculum of most countries that took part in the TIMSS-95 sub-study.

f) The Plasticine Task (Performance Average : 45.0%)

The study again surfaced the inadequacy of the students in communicating their answers to others. Students had little difficulties with basic direct measurement but encountered difficulties when the task involved higher order form of manipulation. Some students were also found to be inaccurate in their work which resulted in partially correct or totally wrong answers.

g) The Shadow Task (Performance Average : 29.0%)

A group of students were found to possess wrong conceptions about the related concepts of light and shadow (e.g. how the size of shadow varies with varying distance of light source). This came as a surprise to the researcher since the required content knowledge was covered formally in the curriculum. The study also found that students had difficulties making generalisations and had difficulties explaining their findings, giving superficial responses.

An overview of difficulties encountered by our Singapore sample of varying degrees is summarised in Table 4.

Table 4: Summary of difficulties encountered by some of the Singapore sample

Difficulty Encountered	Dice	Calculator	Folding & Cutting	Around-the-Bend	Packaging	Plasticine	Shadow
Answering questions requiring them to describe or explain. Answers could not go beyond that of basic observation.	X	X				X	X
Inaccuracy / Carelessness.	X	X		X		X	X
Lack in communications skills and ability to articulate answers clearly to others.	X	X				X	X
Lack ability to comprehend the instructions given or to follow instructions.	X		X		X		X
Solving problem without prior planning.			X				
Conversion of measurement involving half measurements.				X			
Misconceptions in the concept of light and shadow.							X
Making generalisations.							X
Lack prior concept of nets.					X		

Implications for Mathematics Education

The aims of this section are to highlight some areas of concern for educators engaged in the teaching of mathematics at the Primary level and suggest recommendations for them to incorporate into their daily teaching of the subject. These are:

a) Some students were found to be lacking, with varying degrees, in their ability to communicate their written answers to others. Classroom lessons could be structured to promote inter-personal communications, both verbally and written. The use of group work, cooperative learning strategies or concept mapping techniques may help expand students' ability in this area. Technical vocabulary specific to the subject should be emphasised as well.

b) The study found that students encountered difficulties with activities involving generalisations as well as recognising patterns beyond the observable (i.e. explaining the relationship behind the pattern). Classroom lessons could be planned to provide a structured progression from simple pattern recognition, progressing to a higher level and subsequently to one that involves generalisation. Ample opportunities could also be created for students to discuss and challenge each other's mental models through activities that will help them to simplify the process of pattern recognition and asking the 'why' questions.

c) Classroom teachers should also encourage and expect a high level of numerical accuracy in their students. Efforts should be made to impress upon the students the need to read the questions carefully and understand the instructions given before attempting the question. Cultivating good habits of checking and verifying final answers, be it through working backwards or estimating the reasonableness of the answers, should also be emphasised.

d) Students should be taught and provided with a structured approach towards problem solving. Given a generic structured approach, instructions should also advocate the need to plan the 'hows' of solving a problem before actually attempting it. This is meant to help students clarify and simplify the otherwise complex problem.

e) Comprehension and investigative work should be encouraged. Students should be constantly required to examine casual effects of factors related to a problem. By asking the "why" questions may, not only, help with finding root causes rather than symptomatic solutions but also a better understanding of the problem in general.

f) There was evidence to show that hands-on concrete activities helped students to visualise the answers and should be encouraged in the classroom.

g) Modern teaching based on constructivism should/does not refer to knowledge transfer from teacher to student. Instead, emphasis is on teacher as a facilitator, who helps learner construct/make sense of knowledge and integrate new knowledge into his/her cognitive structure.

Conclusion

The study has provided some insights into the performance of our Singapore Primary 4 students' ability in hands-on mathematical problem-solving through the use of a performance-based assessment. The Singapore sample was found to perform significantly better than many in the international community involved in the TIMSS-95 Performance Assessment sub-study. However, the study also revealed that the Primary 4 students do encounter varying degrees of difficulties and emphasis should be given to overcome these difficulties in the daily classroom teaching.

From the wealth of knowledge derived from conducting the study, it was evident that the use of performance-based assessment to complement traditional standardised testing methods to provide students with a broader and more holistic assessment of their ability as well as achievement in the content area should be explored further.

Abbreviated Bibliography

Chua, M.H. & Leong, H. (1997). The Thinking Programme in Singapore - An Overview. *Singapore paper presented at 7th International Conference on Thinking.* Retrieved Jan 2000, from http://eduweb.nie.edu.sg/sctt/

Goh, C.T. (1997). Shaping our future: Thinking Schools, Learning Nation. *Speech by Prime Minister Goh Chok Tong at the opening of the 7th International Conference on Thinking on Monday, 2 June 1997, at 9.00am at the Suntec City Convention Centre Ballroom.* Retrieved Jan 2000, from http://www1.moe.edu.sg/Speeches/020697.htm

Harmon, et al. (1997). *Performance Assessment: IEA's Third International Mathematics and Science Study (TIMSS)*. Chestnut Hill, MA: Boston College.

Kilpatrick, J. (1992). Problem solving in mathematics. In A. Lewd, (Ed.), *The international encyclopedias of curriculum.* Oxford, UK: Pergamon Press.

Krulik, S., & Rudnik, J.A. (1987). Problem Solving; In R. Shumway (Ed.), *Research in Mathematics Education* (pp.286-323). Reston VA: National Council of Teachers of Mathematics.
MOE (1990). Mathematics Syllabus 1990, *Ministry of Education, Singapore: Curriculum Planning Division.*

Nitko, A.J. (1996, 2nd ed). *Educational Assessment of Students.* University of Pittsburgh: Simon & Schuster Asia Pte Ltd.

Schoenfeld, A.H. (1989). Teaching mathematical thinking and problem solving. In L.B. Resnick, and L.E. Klopfer (Ed.) *Toward the thinking curriculum: Current cognitive research,* pp.83-103. New York: Macmillan. 1989 Yearbook of the Association for supervision

Schoenfeld, A.H. (1992). Learning to think mathematically: Problem solving, metacognition and sense making in mathematics. In D.A. Grouws, (Ed.) *Handbook of Research on Mathematics Teaching and Learning, Ch 15, 334-370.* New York: Macmillan.

Table 1A
Content Links and Performance Expectations of Performance Tasks

Task	Task Description	Content	Performar
Dice (M1)	Students investigate probability by repeatedly rolling a die, applying a computational algorithm, and recording the results. They observe patterns in the data and propose explanations in terms of probability for patterns that emerge.	- Whole number operations - Data Representation and Analysis - Probability	- Perform mathem - Conjectu
Calculator (M2)	Students perform a set of multiplications with a calculator and observe and record patterns of results. These data allow students to predict the results of further multiplications beyond the scope of the calculator.	- Whole Numbers: Meaning and Operations - Data Representation and Analysis	- Using ec - Recallin properti - Perform mathem - Develop - Predictir
Folding & Cutting (M3)	Students are shown pictures of rectangular shapes with pieces cut out of them. They try to make similar shapes by folding and cutting rectangles of paper. These are evaluated for accuracy of shapes and recognition of the axes of symmetry.	- Geometry: Symmetry transformation	- Problem - Predictir
Around the Bend (M4)	Students use a simulated section of hallway corridor made of cardboard, thin wood, or plastic to determine the dimensions of furniture that can be moved around a bend in the corridor. The furniture is represented by rectangles of varying dimensions cut out of cardboard. The students manipulate the rectangles in an attempt to determine rules about the maximum dimensions and the relationship between the length and width of the furniture that affects whether they will "go around the bend". The task involves understanding scale conversions and right-angled triangle relationships.	- Measurement and Units - Geometry: Position, Visualisation and Shape - 2D polygons - 3D - Proportionality Problems	- Perform procedu - Problem

Table 1B

Content Links and Performance Expectations of Performance Tasks (Continued)

Task	Task Description	Content	Perform
Packaging (M5)	Students design boxes for packaging four balls by experimenting with drawing boxes of various shapes and their nets. The students then construct the net of a box of actual size to hold the set of four balls.	- Measurement and Units - Geometry: Position, Visualisation and Shape - 3D	- Perform mathem - Problem
Shadows (SM1)	A flashlight is attached to the top of a box and directed toward a wall or projection screen from a distance of about 50cm. A 5 x 5 cm card is on a stand between the wall and the torch, perpendicular to the beam of light and parallel to the wall. Students experiment with the effect of distance on casting shadows by moving the card and measuring the different-sized shadows. They then find positions where the shadow is twice the size of the card and construct a general rule to predict when this will be true. The task sample science concepts of light and shadow formation and the mathematics concepts of similar triangles and proportion.	- Measurement and Units - Geometry: Position, Visualisation and Shape - 2D polygons - Geometry: Symmetry, Congruency and Similarity - Proportionality Problems	- Perform mathem - Problem - Conject - Generali
Plasticine (SM2)	Students are given a 20g standard weight and plasticine. Using a simple balance, they devise methods for measuring different amounts of plasticine, record their procedures, and save and label their plasticine samples so that their weight can be verified. In describing their strategies students may use concepts of proportionality or knowledge of alternative number combinations to achieve the desired masses.	- Measurement and Units - Proportionality Concepts and Problems	- Perform - Problem - Develop