Assessing Mathematics Learning Outcomes Through Performance Based Tasks

Ang Kok Cheng, *CHIJ St Nicholas Girls' School, Singapore*
Boo Hong Kwen, *National Institute of Education, Singapore*
Toh Kok Aun, *National Institute of Education, Singapore*

Abstract: Relatively little is known about pupils' ability in applying mathematical concepts to solving hands-on problems. With the mode of assessment remaining largely unchanged over the last decade, teacher's perceptions of their pupils' hands-on performance ability relies much on the traditional pen-and-paper tests which may not be sufficient in assessing the application skills of individual pupils. This paper reports on a study which was aimed at finding out pupils' strengths and weaknesses in approaching such hands-on problems in mathematics. Such information could be useful in helping classroom teachers customise their lessons in a way that would enhance pupil learning and problem-solving. Five-hundred-and-four students 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 analysis.

Introduction

In Singapore schools, thinking is infused through the introduction of the Thinking Programme incorporating core skills such as 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.

However, infusion of thinking into the curriculum did not begin with these new initiatives. In the revised 1990 mathematics syllabus (MOE, 1990), thinking skills were already manifested as processes involved in mathematical problem solving.

While mathematics teaching in Singapore may have received accolades from certain quarters, achievement 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, which may serve worthy pedagogical goals but has denied 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 thus being limited. Perhaps if performance-based assessment can be introduced to match the curriculum's learning targets, used appropriately and consistently by the teacher, performance assessments may improve the validity of classroom assessment result by complementing on existing assessment modes (Nitko, 1996, p. 258).

Rationale
While it took years to complete the implementation of the 1990 revise syllabus, almost a decade has since lapsed. With the newly proposed reduction in syllabus content in 1999 (MOE, 1997) and the revision to replace the 1990 syllabus by 2006 (MOE, 2000), schools are at crossroads again – what to adopt, what to discard, and what to retain. With the vision of Thinking Schools Learning Nation (TSLN), it seems timely and necessary to know where, along the path, students are in their hands-on performance ability and how best they can be assessed more effectively for a knowledge-based, ability-driven economy.

Significance
The primary aim of the study was to gain some insights into students' hands-on performance ability by administering a performance-based mathematics assessment. Secondary to this, adopting the TIMSS-95 Performance Assessment sub-study tasks also served as an indicative benchmark to compare the current Singapore sample with the international community that participated in the TIMSS-95 sub-study.

Research Questions
The study was aimed at addressing the following research questions:
- With the understanding that Singapore did not take part in TIMSS-95, how would Singapore Primary 4 students perform compared with the average student who made up the international community involved in TIMSS-95?
- 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?
- What are the difficulties encountered by our Primary 4 students when performing each of the TIMSS-95 tasks adopted for the study?

Design and Data Collection
Seven of the twelve performance tasks from the TIMSS-95 Performance Assessment sub-study were adopted and used as the main source for data collection (Appendix A). Administration procedures, rotation schemes, scoring framework and rubrics from the TIMSS-95 sub-study were also adopted (Harmon, et al., 1997).

The 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, given the limitation of resources, was used to select and represent schools from different school types (e.g., all boys, all girls, and coeducational schools). Intact classes were used because school set-ups did not facilitate a random selection 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 original rotation scheme, for every nine students sampled, only three sets of data were collected for each task. On average, 168 sets of data were collected for each task. As 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.

To ensure that data were collected in accordance with the administration procedures as reported in Harmon, et al. (1997), data were jointly collected with another similar study examining science performance for Population 1. The reconstructed tasks were validated and examined for accuracy. Thereafter, a pilot test was carried out in two schools to ascertain the robustness of the reconstructed tasks.

Scoring rubrics were also re-produced and 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. Inter-rater reliability correlation coefficients ranged from 0.95 to 0.99. The high values for the inter-rater reliability correlation thus gave confidence concerning the reliability of the scoring rubrics for scoring the students' responses.

Assumptions and limitation
One assumption of the study was 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 had equal exposure to mathematics instruction.

The study also assumed that the students selected were 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 1). This is a limitation of the study; however it might be of interest to compare the students with the average student who made up the international community involved in TIMSS-95.

Table 1

Comparison of TIMSS-95 Content Coverage vs Singapore Syllabus Coverage

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

Three broad-based comparisons were made to characterise student performance.

Comparison of Overall Performance Average

The Singapore sample scored an overall performance average of 45.4% for the five mathematics-related and the two science-related tasks (Figure 1A). This average increased to 48.8% when only the mathematics-related tasks were considered. Both scores were significantly higher than the international performance average of 34.4% and 36.2% respectively at the 0.05 level of confidence (Harmon, et al., 1997). 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 the Singapore sample found number manipulation tasks easier than science hands-on activities.
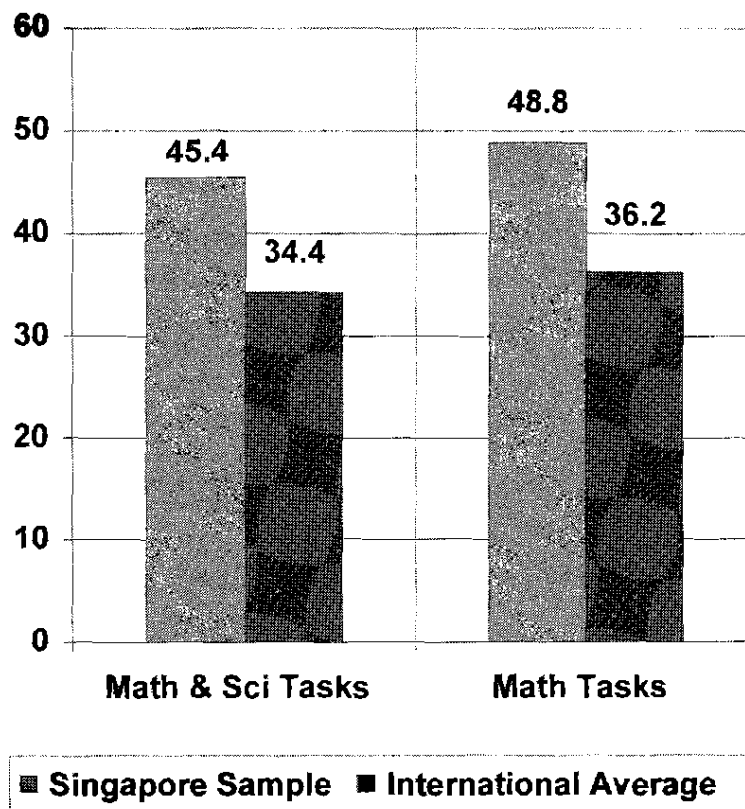

Figure 1A. Comparison of overall performance average (%) between Singapore sample and international community

Comparison of Overall Performance Expectation

Comparison of main performance expectations (i.e., performing mathematical procedures, and problem solving and mathematical reasoning) also suggested that the Singapore sample performed significantly better than the international community at the 0.05 level of confidence (Figure 1B). 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 .
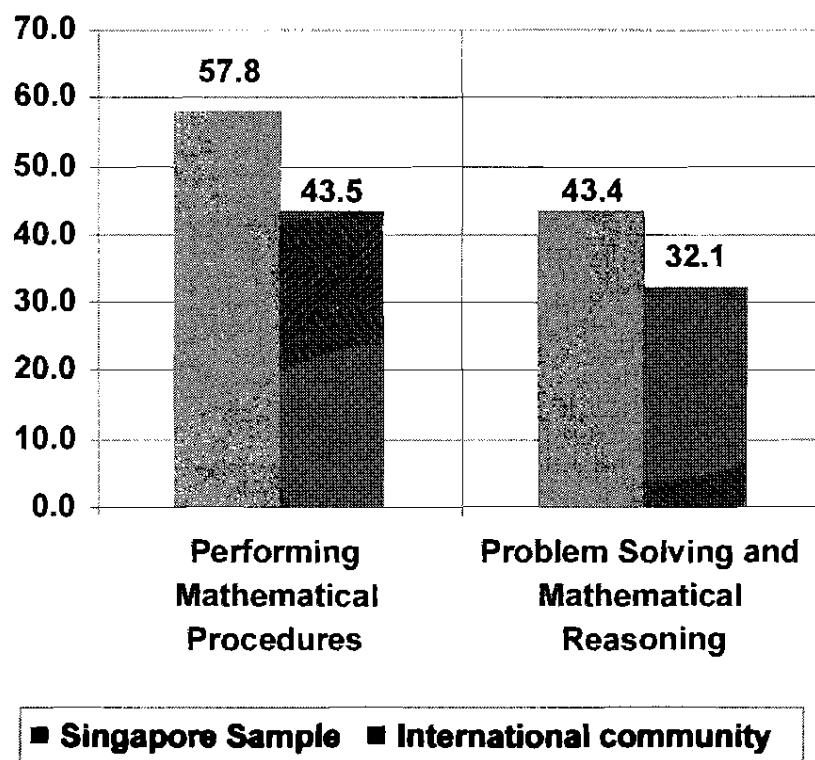

Figure 1B. Comparison of overall performance expectation (%) between
   Singapore sample and international community (Mathematics + Plasticine Tasks)

Comparison of Performance Expectation Sub-categories

In the comparison of performance expectations sub-categories, the Singapore sample also performed significantly better than the international community, at the 0.05 level of confidence, in all categories except in the category of measurements and units (Figure 1C). The Singapore sample encountered least difficulties with numbers (77.7%) and most difficulties with 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.
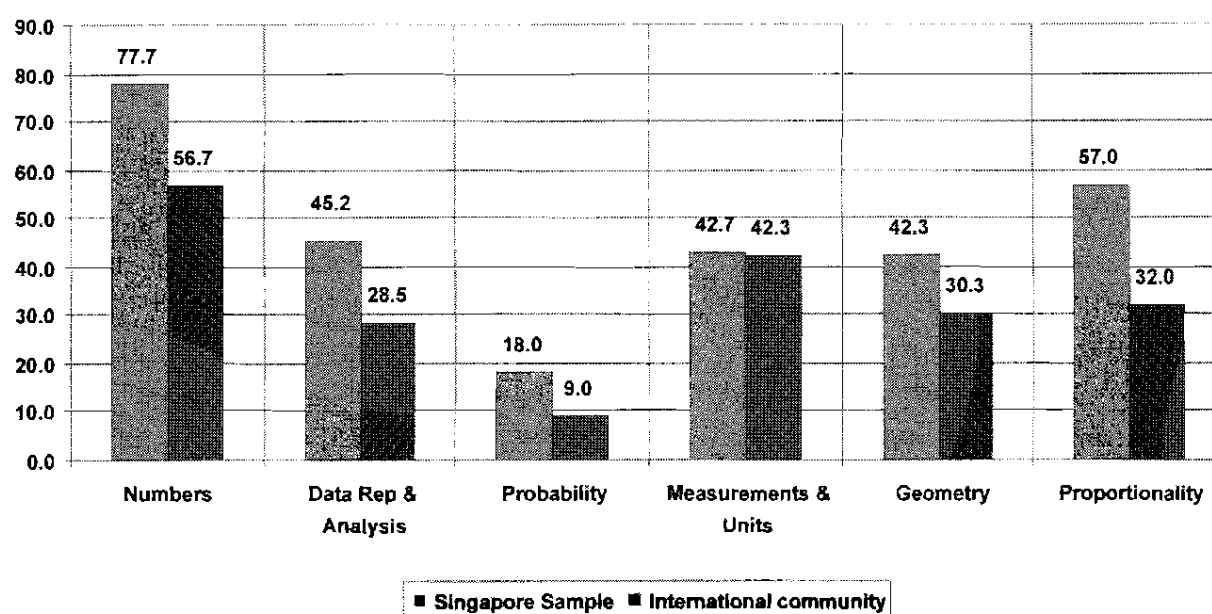
Figure 1C. Comparison of performance expectation sub-categories (%) between Singapore sample and international community

These findings allowed the researcher to distinguish the Singapore sample as a 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 research may be carried out to identify students' relative weaknesses in this area so as to raise their ability levels.

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?

The Singapore sample obtained 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 the Singapore sample and the Shadow task hardest. The students also performed 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 the students. However, for the Shadow task, the science

concept of light and shadow formation was already taught formally in the classroom which could be an indication of misconception at play.

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 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) and performed below expectation when no concrete materials were given.

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 lack the ability to comprehend the instructions given. This inability may be interpreted either as an inability to respond to non-routine open-ended problems, or to not understanding the concept of nets. The concept 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. Note however that nets is only taught to Singapore students as one of the topics that they have to master in Primary 6.

f) The Plasticine Task (Performance Average: 45.0%)
The study again brought out the inadequacy of the students in communicating their answers to others. Students had little difficulty with basic direct measurement but encountered difficulties when the task involved higher-order forms 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 possessed 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 is summarised in Table 2.

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 to suggest recommendations for them to incorporate into their daily teaching of the subject

a) Some students were found to be lacking, in 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.

Table 2

Summary of difficulties encountered by some of the Singapore sample

Difficulty Encountered	D[1]	C	F	A	Pk	Pl	S
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		

[1]D: Dice;   C: Calculator;   F: Folding and cutting;   A: Around-the-Bend
Pk: Packaging; Pl: Plasticine;   S: Shadow.

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 questions. Cultivating good habits of checking and verifying final answers, through either 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 an otherwise complex problem.

e) Comprehension and investigative work should be encouraged. Students should be constantly required to examine causal effects of factors related to a problem.

Asking the "why" questions may not only help with finding root causes rather than symptomatic solutions, but may also result in 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 hence should be encouraged in the classroom.

Conclusion

The study has provided some insights into the performance of Singapore Primary 4 students' ability in hands-on mathematical problem-solving through the use of a performance-based assessment. The 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.

From the wealth of knowledge derived from conducting the study, it was evident that the use of performance-based assessment could possibly complement traditional standardised testing in providing students with a broader and more holistic assessment of their ability (e.g., the skill of reasoning, communicating about mathematics and expressing more complete conceptual understanding) and should be explored further.

References

Ang, K. C (2002). *Performance of primary 4 students in performance-based assessment tasks (unpublished Master of Education thesis)*. National Institute of Education, Nanyang Technological University.

Chua, M. H., & Leong, H. (1997). The thinking programme in Singapore - An overview. *Singapore paper presented at 7th International Conference on Thinking.* Retrieved January 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, Singapore.* Retrieved January 2000, from http://www1.moe.edu.sg/Speeches/020697.htm

Harmon, M., Smith, T. A., Martin, M .O., Kelly, D. L., Beaton, A. E., Mullis, I. V. S, Gonzales, E. J., & Orpwood, G. (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.

Ministry of Education (MOE). 1990. Mathematics syllabus 1990. Singapore: Curriculum Planning & Development Division, Ministry of Education.

Ministry of Education (MOE). 1997. Towards thinking schools. Singapore: Ministry of Education .

Ministry of Education (MOE). 2000. Mathematics syllabus – Primary. Singapore: Curriculum Planning & Development Division, Ministry of Education.

Nitko, A. J. (1996). *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 & L. E. Klopfer (Eds.) *Toward the thinking curriculum: Current cognitive research* (1989 Yearbook of the Association for Supervision and Curriculum Development), pp. 83-103. New York: Macmillan.

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,* pp.334-370. New York: Macmillan.

Authors:
Ang Kok Cheng, CHIJ St. Nicholas Girls' School. caseyang@pmail.ntu.edu.sg
Boo Hong Kwen, Associate Professor, National Institute of Education, Nanyang Technological University. boowk@nie.edu.sg
Toh Kok Aun, Associate Professor, National Institute of Education, Nanyang Technological University. tohka@nie.edu.sg

Appendix A
Content Links and Performance Expectations of Performance Tasks

Task	Task Description	Content	Performance Expectations
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	- Performing routine and complex mathematical procedures -Conjecturing
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 equipment - Recalling mathematical objects and properties - Performing routine and complex mathematical procedures -Developing and describing strategy - Predicting
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 solving - Predicting

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	- Performing routine and complex procedures - Problem solving
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, visualization and shape - 3D	- Performing routine and complex mathematical procedures - Problem solving

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	- Performing routine and complex mathematical procedures - Problem solving - Conjecturing - Generalising
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	- Performing routine mathematical problem solving - Developing and describing strategy