# THIRD INTERNATIONAL MATHEMATICS AND SCIENCE STUDY
## Third National Report


Sue Harris with Wendy Keys and Cres Fernandes

nfer

# THIRD INTERNATIONAL MATHEMATICS AND SCIENCE STUDY
# Third National Report

## Performance Assessment

Sue Harris
with Wendy Keys and Cres Fernandes

nfer

Published in December 1997
by the National Foundation for Educational Research,
The Mere, Upton Park, Slough, Berkshire SL1 2DQ


Registered Charity No. 313392
ISBN 0 7005 1474 X

# CONTENTS


**Acknowledgements** i

**EXECUTIVE SUMMARY** 1

**CHAPTER 1 Introduction** 2

1.1 The context for Performance Assessment 2

1.2 The Performance Assessment component of TIMSS 3

1.3 The sample 4

1.4 The tasks 5

1.5 Administration 6

1.6 Marking and coding 6

1.7 National Curriculum links 7

**CHAPTER 2 Students' Achievement on the Mathematics and Science Performance Assessment Tasks** 8

2.1 Preface 8

2.2 Overall performance on mathematics and science tasks 10

2.3 Mathematics tasks 13

2.3.1 Overall performance on mathematics tasks 13

2.3.2 *Dice* 16

2.3.3 *Calculator* 20

2.3.4 *Folding and Cutting* 24

2.3.5 *Around the Bend* 28

2.3.6 *Packaging* 32

2.3.7 *Plasticine* 36

2.4 Science tasks 39

2.4.1 Overall performance on science tasks 39

2.4.2 *Pulse* 42

2.4.3 *Magnets* 46

2.4.4 *Batteries* 48

2.4.5 *Rubber Band* 50

2.4.6 *Solutions* 54

2.4.7 *Shadows* 58

2.5 Gender differences 61

2.5.1 Gender differences in overall performance 61

2.5.2 Gender differences in mathematics 61

2.5.3 Gender differences in science 62

2.6 Correlations between students' performance on the written tests and the Performance Assessment tasks 63

| | | |
|---|---|---|
| **CHAPTER 3** | **Students' Performance in the Mathematics and Science Skills Areas** | 67 |
| | 3.1 Preface | 67 |
| | 3.2 Mathematical skills | 68 |
| | 3.2.1 *PERFORMING MATHEMATICAL PROCEDURES* | 70 |
| | 3.2.2 *PROBLEM SOLVING AND MATHEMATICAL REASONING* | 71 |
| | 3.3 Science skills | 72 |
| | 3.3.1 *SCIENTIFIC PROBLEM SOLVING AND APPLYING CONCEPT KNOWLEDGE* | 73 |
| | 3.3.2 *USING SCIENTIFIC PROCEDURES* | 75 |
| | 3.3.3 *SCIENTIFIC INVESTIGATING* | 75 |
| **CHAPTER 4** | **Concluding Remarks** | 77 |
| **REFERENCES** | | 78 |
| **APPENDICES** | | 79 |
| | Appendix I National steering committee | 79 |
| | Appendix II Countries taking part in different components of TIMSS | 80 |
| **ANNEXES** | | 82 |
| | Annex 1 Assignment of tasks and students to workstations | 82 |
| | Annex 2 Mathematics skills covered by tasks, sorted by skills | 84 |
| | Annex 3 Mathematics skills covered by tasks, sorted by tasks | 85 |
| | Annex 4 Science skills covered by tasks, sorted by skills | 86 |
| | Annex 5 Science skills covered by tasks, sorted by tasks | 87 |

# ACKNOWLEDGEMENTS

This study could not have been completed without the assistance of a number of people and we would like to express our thanks to:

- Derek Foxman, the Project Consultant, for his constructive advice and suggestions
- Claudia Davis, who helped to coordinate the arrangements for administering the tasks in schools
- the team of administrators who visited schools to set up and administer the tasks
- the mathematics and science practitioners who marked students' responses to the tasks
- John Ashby and David Upton for their helpful comments on the draft report
- Pauline Pearce and Alison Bannerman for providing valuable secretarial support throughout the project
- Mary Hargreaves, who prepared the layout, Tim Wright, who designed the cover, and Enver Carim, who coordinated the publication of the report.

We would also like to acknowledge the support given by headteachers and teachers within schools, who responded warmly to our requests for the use of a science laboratory and provided the blackout facilities necessary for one of the tasks.

Finally, this study would not have been possible without the active involvement of Year 9 students, who approached the tasks with confidence and enthusiasm. Our special thanks are due to all those students.

# EXECUTIVE SUMMARY

Out of the 19 countries whose results are presented in the international report, the position of students in England was

- ranked second on their overall results on all 12 tasks
- ranked equal seventh on their results on the six mathematics tasks
- ranked second on their results on the six science tasks.

Only students in Singapore outperformed students in England on their overall score and on their score for the science tasks (by four percentage points and one percentage point respectively).

Overall, the mathematics results are better than the results in the written mathematics tests, and the science results are consistent with the results in the written science tests.

The mean scores of students in England were equal to or above the international means for five out of the six mathematics tasks. These results are better than the results in the TIMSS written tests in mathematics, in which the performance of students in England was lower than the international mean for five out of the six mathematics content areas in the written tests.

The mean scores of students in England were considerably higher than the international means for each of the six science tasks. These results are consistent with their performance on the written tests in science, in which the mean scores for students in England were above the international means in all of the five science content areas within the written tests.

In addition to determining students' performance on each of the tasks, further analysis focused on results in main performance categories (covering a range of mathematics and science skills respectively) across tasks. The results showed that students in England performed

- above the international means for the two main mathematics categories (*PERFORMING MATHEMATICAL PROCEDURES* and *PROBLEM SOLVING AND MATHEMATICAL REASONING*)
- above the international means for the three main science categories (*SCIENTIFIC PROBLEM SOLVING AND APPLYING CONCEPT KNOWLEDGE*, *USING SCIENTIFIC PROCEDURES* and *SCIENTIFIC INVESTIGATING*).

There were no statistically significant gender differences in the performance of girls and boys in England in either mathematics or science.

# CHAPTER 1
# Introduction

## 1.1 The context for Performance Assessment

Although cross-national studies have been conducted for more than 30 years by the International Association for the Evaluation of Educational Achievement (IEA), the data collection in mathematics and science studies has been predominantly by means of written achievement tests and questionnaires. A similar emphasis was evident in the first and second studies organised by the International Association for Evaluation of Educational Progress (IAEP), both of which focused on students' achievement in mathematics and science. However, some previous studies have included practical tasks for students as an optional component to supplement the main study, such as IEA's First International Science Study (FISS), their Second International Science Study (SISS) and IAEP2. It should be acknowledged, however, that relatively few countries conducted the practical components of previous studies.

In England, a noteworthy predecessor of the TIMSS Performance Assessment was the series of surveys initiated by the Assessment of Performance Unit (APU) which were conducted in England, Wales and Northern Ireland from 1977 to 1990. Mathematics and science surveys were both carried out with 11- and 15-year-olds, and, additionally, the science team surveyed 13-year-olds. The practical activities utilised in the APU mathematics and science surveys were mostly administered on a one-to-one basis, with a trained assessor presenting the task to the student and recording his/her responses. In some instances science activities were presented as a 'circus' of tasks which students attempted in turn. However, in the final mathematics survey, some students were expected to work collaboratively in groups of three. Collectively, the activities were designed to assess students in not only carrying out routine tasks, such as using measuring instruments, but also to identify the strategies they adopted in problem-solving situations and their reasons for choosing particular approaches.

With the introduction of the National Curriculum in 1989 and its associated assessment arrangements, the work of the APU drew to a close, although it is true to say '...some of the materials, particularly practical and oracy assessments, have influenced and extended good practice in their areas' (Foxman *et al.*, 1991).

With earlier work (such as that mentioned above) in mind, and sensitive to the growing movement to measure not only students' knowledge and understanding using written tests, but also practical, investigative, problem-solving and analytical skills by means of 'hands-on' activities, the TIMSS International Study Center set up the Performance Assessment component. In doing so, TIMSS has become the largest international study to include a practical assessment of students' performance, and, arguably therefore, the first to collect complementary sets of data that more accurately reflect the range of students' curricular experiences in mathematics and science.

## 1.2 The Performance Assessment component of TIMSS

The Performance Assessment constituted part of the Third International Mathematics and Science Study (TIMSS), a major international comparative study focusing on teaching and learning in these two subjects. This report presents information about the performance of students in England on the Performance Assessment component of TIMSS. Some summary information about the design and administration of the study as a whole is presented here for background purposes, but readers who seek more detailed information about the aims of the study, its design, and administrative and analytical processes are referred to other national[1] and international reports.[2]

England was one of more than 40 countries taking part in TIMSS. The study was designed to focus on the teaching and learning of mathematics and science in three target populations:

- upper primary students, aged about nine years (Population 1)
- lower secondary students, aged about 13 years (Population 2)
- upper secondary students, aged about 17/18 years (Population 3).

For each population there were two main aspects to the data collection:

- achievement tests in mathematics and science for the students
- questionnaires covering background information and other issues, such as attitudes to these subjects, for students, their mathematics and science teachers and their headteachers.

All countries taking part in TIMSS were required to survey lower secondary students (Population 2), and participation in the surveys of Populations 1 and 3 was optional. England was one of 26 countries that

1 See Keys *et al.*, 1996a; Keys *et al.*, 1996b; Harris *et al.*, 1997; Keys *et al.*, 1997a; and Keys *et al.*, 1997b for national information.

2 See Beaton *et al.*, 1996a; Beaton *et al.*, 1996b; Martin *et al.*, 1997; Mullis *et al.*, 1997; and Harmon *et al.*, 1997 for international information. Robitaille (1993) and Robitaille (1997) present useful background information about TIMSS and the countries that took part in the study.

took part in the survey of Population 1 and was one of more than 40 countries that surveyed Population 2. The funding for this work was provided by the Department for Education and Employment (DfEE) and carried out by the National Foundation for Educational Research (NFER).

The Performance Assessment was intended to supplement the data collected by means of the written tests in the main TIMSS survey. It consisted of a number of specific tasks that would allow students to demonstrate their practical, investigative, recording and analytical skills in mathematics and science in controlled situations. This element of the study was designed for Populations 1 and 2 only; participation in the Performance Assessment for Population 1 and/or Population 2 was optional. Twenty-one countries (including England) participated in this aspect of TIMSS with 13-year-old students[3], i.e. about half of the total number of countries that administered the written tests in mathematics and science to students (see Appendix II). The Performance Assessment in England was funded by the Office for Standards in Education (OFSTED) and carried out by the NFER.

## 1.3 The sample

The sample of schools that was invited to take part in the Performance Assessment was a sub-sample of the 127 schools that had participated in the main survey. In addition, the students who participated in the Performance Assessment formed a sub-sample of those that were involved in the main survey (i.e. had completed written tests and questionnaires). The study design involved the random selection of nine students from those in the upper grade only of the sample (i.e. nine Year 9 students and no students from Year 8). In addition to the first-selected nine students in each school, two replacement students were selected at the same time so that if one or two students were absent on the day of the administration of the tasks, the replacement student(s) could be substituted. Other than these replacement students identified at the time of drawing the sample, no other substitutions were permitted. More detailed information about the sampling is provided in Appendices III and V to the TIMSS national reports (Keys *et al.*, 1996b).

[3] In the international report (Harmon *et al.*, 1997) the results for two of the countries (Hong Kong and Israel) are not presented in the main tables and figures, but are given in an Appendix. This was because both of these countries had small samples.

## 1.4 The tasks

In total, 12 different tasks were used for the Performance Assessment:[4] five focusing on mathematics (M1-5), five focusing on science (S1-5), and two combination tasks with elements of both mathematics and science (SM1-2). (Within the international report, the information relating to students' performance on tasks SM1 and SM2 has been included within the main sections relating to science and mathematics respectively, and this approach will also be adopted within this report.) The time allowed for completion of particular tasks was either 15 minutes or 30 minutes. Typically, the 15-minute tasks required routine skills such as constructing tables to record results, and summarising observations, whereas the 30-minute tasks involved skills such as planning and carrying out an investigation, graphing results and drawing conclusions. All the tasks used in the Performance Assessment of 13-year-olds are reproduced within Chapter 2 of this report. The times allocated[5] for each of the activities are shown below:

| TASK | NAME | TIME |
|---|---|---|
| M1 | *Dice* | 15 mins |
| M2 | *Calculator* | 15 mins |
| M3 | *Folding and Cutting* | 15 mins |
| M4 | *Around the Bend* | 30 mins |
| M5 | *Packaging* | 30 mins |
| SM2 | *Plasticine* | 30 mins |
| S1 | *Pulse* | 15 mins |
| S2 | *Magnets* | 15 mins |
| S3 | *Batteries* | 15 mins |
| S4 | *Rubber Band* | 30 mins |
| S5 | *Solutions* | 30 mins |
| SM1 | *Shadows* | 30 mins |

Each task was presented to students in the form of a four-page booklet which provided details of the activity and asked questions related to the task. Most of the students' responses were written in the booklets, although some tasks required additional evidence of work, such as a graph of the measurements taken during the *Rubber Band* activity, scale drawings of pieces of furniture on squared paper for *Around the Bend*, or different masses of plasticine in the task of the same name.

[4] All except one of these tasks were also used with the nine-year-olds in a modified form. *Solutions* was replaced by another task (*Containers*) for the younger students.

[5] In some instances, students completed tasks in less time than was allocated; in others, students had not completed the entire task at the end of the time allowed.

## 1.5 Administration

The administration of the Performance Assessment in each of the 50 secondary schools that participated in this element of TIMSS was carried out by one of a team of administrators trained by the researchers at NFER. This approach was used for two reasons: firstly, to avoid over-burdening teachers within the schools; and secondly, to ensure that the administration of the tasks was as uniform as possible in each of the schools.

So as to ensure that the tasks were administered in similar conditions in all schools taking part in the Performance Assessment, the International Study Center provided detailed guidance about the types of equipment that should be used for each activity and a script for the administration.

The 12 tasks were organised to form activities at nine workstations: three of the workstations each presented two 15-minute tasks (one mathematics and one science task), and the remaining six workstations each presented one 30-minute task for students. During the administration, each student visited three different workstations and, depending on the combination of 15- and 30-minute tasks, attempted either three, four or five tasks altogether in a total working time of 90 minutes. At workstations with two tasks, students could choose which task to do first and were told when 15 minutes had elapsed so that they could move on to the second task if they had not already done so. The combinations of tasks ensured that all students attempted at least one mathematics and one science task. (Details of the combinations of tasks are given in Annex 1 at the end of this report; additional information about the administration is given in Appendix III to the TIMSS national reports (Keys *et al.*, 1996b).)

## 1.6 Marking and coding

The marking of students' responses to the Performance Assessment tasks was carried out under the supervision of the research team, using the coding guides supplied by the International Study Center. So as to reflect as accurately as possible the content and quality of responses, a two-digit coding system was utilised: using this approach, the first digit indicated the level of correctness (questions carried a maximum of one, two or three marks) and the second digit indicated the content of the response. (Further information about the marking is given in Appendix IV to the TIMSS national reports (Keys *et al.*, 1996b).) In addition to the marking based on students' written responses, some marks were allocated to additional material provided by students in accordance with the task requirements, such as the folded paper used in *Folding and Cutting*, and the nets produced in *Packaging*.

## 1.7 National Curriculum links

The written achievement tests completed by the students who participated in the TIMSS survey collected data about the performance of students in a range of content areas in the two subjects tested (for details of the curriculum frameworks see Robitaille, 1993). In terms of relevance to the National Curriculum Orders for mathematics and science, the content of the written tests was most related to those parts of the respective programmes of study concerned with knowledge and understanding of concepts:

Ma2: Number and Algebra[6]

Ma3: Shape, Space and Measures

Ma4: Handling Data

Sc2: Life Processes and Living Things[7]

Sc3: Materials and their Properties

Sc4: Physical Processes.

The Performance Assessment tasks, on the other hand, because of their practical nature, were more closely related to the process-oriented aspects of the National Curriculum Orders for mathematics and science respectively, i.e. Ma1: Using and Applying Mathematics and Sc1: Experimental and Investigative Science.[8] However, just as classroom work on these aspects of the programmes of study has to be set in a context (usually related to one of the content-focused attainment targets), so the Performance Assessment tasks similarly had their practical and investigative activities set in a context related to the content-focused attainment targets. Consequently, although these tasks were predominantly linked with Ma1 and Sc1, the context of each one had links with other attainment targets.

---

[6] At the time of testing (February(March 1995), the 1991 statutory Orders for mathematics were still in effect. The attainment targets at that time were as follows:

Ma2: Number

Ma3: Algebra

Ma4: Shape and space

Ma5: Handling data.

[7] At the time of testing, the attainment targets were as follows:

Sc2: Life and living processes

Sc3: Materials and their properties

Sc4: Physical processes.

[8] At the time of testing, the corresponding attainment target was Sc1: Scientific investigation.

# CHAPTER 2
# Students' Achievement on the Mathematics and Science Performance Assessment Tasks

## 2.1 Preface

Within this chapter, the data on 13-year-old students' achievement in the Performance Assessment tasks are presented for the 12 tasks. In total, 19 of the 21 countries that took part in the Performance Assessment are included in the tables within the international report[1]. It was decided that within this national report, comparisons would focus on the countries with which comparisons had been made in the national reports on the main written survey; this makes it possible to compare the performance of 13-year-old students in England with those of a selected number of other countries on both the written tests in mathematics and science, and also the practical activities in these subjects. Information about the performance of students in the following countries is included in both of the national reports relating to the main survey and also in this report on the Performance Assessment:

- Canada
- England
- Scotland
- Singapore
- Sweden
- Switzerland
- United States.

Four of the countries that were used for comparisons in the main reports did not take part in the Performance Assessment of 13-year-olds, so it is not possible to present data for these countries (France, Germany, Hungary and Japan). However, so as to increase the number of European countries with which comparisons are made, data relating to students' achievement in the Czech Republic and The Netherlands are also included in this report.

[1] In the international report, the results for England are shown as not satisfying the stringent international sampling requirements. These specified participation rates of at least 85 per cent of the sample of both schools and students, or a combined rate (for the product of school and student participation) of 75 per cent. After the exclusions (see Appendix V (Keys *et al.*, 1996b), Sections V.2 and V.4) permitted by the International Sampling Referee were taken into account, the combined participation rate in England was 73 per cent. Other countries that participated in the Performance Assessment of 13-year-olds and failed to meet the sampling requirements were: Australia; The Netherlands; and the United States.

This chapter is divided into five main sections, the first four of which cover the overall results on all 12 tasks; the results on the mathematics tasks; the results on the science tasks; and information relating to gender differences in performance in the mathematics and science tasks. The sections covering mathematics and science respectively present details of the overall performance of students in England on the six tasks for that subject, together with the following information for each of the tasks:

- the task and details of the international marking guides, together with an example of one student's responses to the task
- the maximum number of marks (one, two or three marks) allocated for each item within the task
- the mean percentage score[2] achieved by Year 9 students in England (and their counterparts in the selected other countries referred to above) for each item within the task, and for the task overall
- the international mean for each item within the task, and for the task overall.

The results on the six mathematics tasks are presented first (five tasks focused predominantly on mathematics and one with elements of both mathematics and science, but in which there is more emphasis on mathematics), followed by the results for the six science tasks (five tasks focused predominantly on science and one with elements of both mathematics and science, but in which there is more emphasis on science).

The final section of this chapter gives correlations between students' performance on the written achievement tests in mathematics and science and on the practical tasks in these subjects.

[2] Within the tables in this chapter, the mean percentage scores quoted for each item represent the percentage of the total possible marks for each item averaged over the students, some of whom would have given fully correct responses, some partially correct responses and some incorrect responses to specific items.

## 2.2 Overall performance on mathematics and science tasks

This section presents the mean percentage scores for mathematics and science, together with overall performance on all 12 tasks. In addition, details of the performance of students in all participating countries on each task are provided.

Table 2.2.1 shows the overall mean percentage scores for all 12 tasks, for the mathematics tasks and for the science tasks in each of the nine selected countries. Countries are ranked according to overall mean percentage score.

**Table 2.2.1 Overall performance on all tasks and in mathematics and science in nine countries**

| Country | Overall mean percentage score | Mathematics tasks | Science tasks |
|---|---|---|---|
| Singapore | 71 | 70 | 72 |
| **England** | **67** | **64** | **71** |
| Switzerland | 65 | 66 | 65 |
| Sweden | 64 | 65 | 63 |
| Scotland | 62 | 61 | 64 |
| Czech Republic | 61 | 62 | 60 |
| Canada | 60 | 62 | 59 |
| Netherlands | 60 | 62 | 58 |
| United States | 55 | 54 | 55 |
| **International mean** | **59** | **59** | **58** |

*Source: Table 2.3 (Harmon et al., 1997)*

Only students in Singapore outperformed students in England on the science tasks and in their overall performance on all 12 tasks. Students in three of the selected countries (Singapore, Switzerland and Sweden) exceeded the performance of students in England in mathematics, although the differences in performance amounted to six percentage points or less.

Table 2.1 is reproduced from the international report, and shows the performance of students in all the participating countries on each of the 12 tasks. It should be noted that although in the main TIMSS survey analysis was carried out to identify statistically significant differences in performance between students in different countries, this type of analysis was not carried out by the International Study Center on the Performance Assessment data. However, it is possible to examine the results for countries which took part in both aspects of TIMSS, based on their overall mean percentage scores for the mathematics and science tasks separately, and compare them with the data collected in the main survey.

**Table 2.1 Mean percentage scores overall and on Perfomance Assessment tasks**

| Country | Overall mean▼ | Mean percentage scores on tasks• | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | Mathematics tasks | | | | | Combination tasks | | Science tasks | | | | |
| | | *Dice* | *Calculator* | *Folding and Cutting* | *Around the Bend* | *Packaging* | *Plasticine* | *Shadows* | *Pulse* | *Magnets* | *Batteries* | *Rubber Band* | *Solutions* |
| Singapore | 71 (1.7) | 84 (1.6) | 60 (2.8) | 80 (2.6) | 63 (1.5) | 65 (2.5) | 66 (3.3) | 50 (3.5) | 60 (2.7) | 95 (0.9) | 79 (2.1) | 80 (1.5) | 68 (2.7) |
| †1 Switzerland | 65 (1.2) | 79 (1.4) | 61 (1.6) | 79 (1.9) | 54 (2.2) | 47 (3.3) | 73 (2.1) | 41 (2.1) | 51 (1.9) | 97 (1.2) | 75 (2.1) | 67 (1.9) | 57 (1.9) |
| Sweden | 64 (1.2) | 74 (2.4) | 51 (2.3) | 80 (2.5) | 65 (1.9) | 47 (2.3) | 72 (2.9) | 45 (1.9) | 45 (2.6) | 95 (1.6) | 71 (2.9) | 70 (2.4) | 50 (2.2) |
| † Scotland | 62 (1.7) | 76 (1.6) | 49 (3.1) | 71 (3.9) | 58 (2.1) | 51 (3.9) | 61 (2.5) | 36 (2.4) | 55 (2.9) | 98 (0.9) | 68 (2.4) | 75 (1.8) | 51 (2.3) |
| Norway | 62 (0.8) | 72 (1.9) | 59 (1.6) | 73 (2.1) | 62 (1.3) | 59 (2.4) | 67 (2.3) | 39 (2.0) | 48 (1.6) | 91 (2.0) | 67 (1.7) | 63 (1.9) | 42 (1.8) |
| Czech Republic | 61 (1.3) | 73 (2.5) | 54 (2.0) | 73 (3.2) | 58 (1.5) | 43 (4.6) | 68 (2.6) | 37 (1.9) | 46 (2.9) | 86 (2.3) | 66 (2.8) | 65 (3.6) | 59 (2.3) |
| Canada | 60 (1.3) | 77 (1.8) | 60 (1.5) | 59 (2.5) | 53 (2.0) | 57 (3.2) | 65 (1.9) | 35 (1.6) | 46 (2.4) | 92 (1.5) | 62 (2.1) | 71 (2.0) | 48 (2.1) |
| New Zealand | 60 (1.4) | 73 (1.2) | 55 (1.6) | 75 (2.3) | 60 (1.4) | 44 (2.5) | 63 (2.2) | 29 (2.0) | 44 (2.1) | 93 (1.6) | 68 (1.6) | 65 (1.8) | 48 (2.1) |
| Spain | 54 (0.8) | 73 (2.2) | 53 (2.1) | 61 (3.1) | 53 (1.9) | 28 (2.3) | 45 (2.5) | 36 (1.7) | 36 (2.1) | 96 (1.4) | 73 (1.7) | 51 (2.0) | 41 (2.3) |
| Iran, Islamic Rep. | 52 (2.0) | 58 (1.8) | 48 (3.7) | 58 (2.9) | 34 (3.2) | 43 (5.0) | 81 (2.6) | 43 (1.5) | 55 (4.5) | 45 (4.9) | 52 (4.0) | 56 (5.4) | 50 (3.5) |
| Portugal | 47 (1.1) | 76 (1.8) | 39 (2.1) | 58 (3.1) | 43 (1.8) | 31 (3.2) | 41 (2.5) | 25 (1.5) | 24 (2.5) | 94 (1.6) | 50 (2.2) | 51 (2.3) | 36 (2.4) |
| Cyprus | 46 (1.0) | 68 (2.2) | 40 (1.9) | 48 (2.4) | 42 (1.5) | 14 (2.1) | 52 (2.4) | 18 (1.5) | 33 (2.1) | 86 (2.3) | 66 (2.2) | 59 (2.3) | 29 (2.9) |
| ***Countries not satisfying guidelines for sample participation rates*** | | | | | | | | | | | | | |
| Australia | 65 (1.2) | 78 (2.4) | 59 (1.9) | 74 (3.3) | 58 (1.8) | 55 (2.8) | 73 (2.9) | 36 (1.9) | 54 (2.6) | 92 (1.4) | 71 (1.8) | 64 (2.4) | 59 (2.2) |
| 2 **England** | **67 (0.9)** | **79 (1.6)** | **62 (1.4)** | **69 (3.1)** | **63 (1.5)** | **53 (2.5)** | **55 (2.4)** | **46 (2.3)** | **59 (2.2)** | **99 (0.6)** | **77 (2.0)** | **79 (1.4)** | **68 (2.1)** |
| Netherlands | 60 (1.3) | 76 (2.2) | 59 (2.3) | 71 (2.4) | 67 (1.9) | 53 (2.9) | 44 (2.5) | 35 (2.8) | 45 (2.6) | 94 (2.1) | 63 (2.9) | 70 (1.9) | 43 (2.7) |
| United States | 55 (1.3) | 71 (2.1) | 56 (1.9) | 68 (2.0) | 48 (1.8) | 28 (2.5) | 53 (2.1) | 28 (1.9) | 50 (2.0) | 85 (2.5) | 56 (1.9) | 63 (2.4) | 48 (2.2) |
| ***Countries not meeting age/grade specifications*** | | | | | | | | | | | | | |
| 3 Colombia | 39 (1.8) | 49 (4.0) | 31 (1.6) | 43 (5.7) | 34 (4.4) | 20 (3.0) | 41 (2.7) | 22 (2.5) | 11 (1.0) | 96 (1.3) | 55 (2.2) | 40 (3.7) | 26 (2.3) |
| 3 Romania | 62 (1.9) | 76 (2.3) | 66 (2.6) | 84 (2.3) | 58 (3.1) | 51 (4.1) | 63 (4.1) | 36 (2.8) | 41 (3.6) | 83 (3.5) | 75 (2.2) | 45 (3.0) | 63 (2.6) |
| Slovenia | 61 (1.0) | 78 (1.4) | 58 (1.5) | 82 (2.0) | 55 (1.9) | 45 (3.8) | 63 (1.9) | 31 (1.8) | 40 (3.2) | 92 (1.9) | 71 (1.8) | 64 (1.7) | 49 (2.0) |
| **International mean** | **59 (0.3)** | **73 (0.5)** | **54 (0.5)** | **69 (0.7)** | **54 (0.5)** | **44 (0.7)** | **60 (0.6)** | **35 (0.5)** | **44 (0.6)** | **90 (0.5)** | **67 (0.5)** | **63 (0.6)** | **49 (0.5)** |

• *Mean of percentage scores across items in task: all items weighted equally (see overall task mean in Chapter 2).*
▼ *Mean of percentage scores across tasks: all tasks weighted equally.*
( ) *Standard errors appear in parentheses. Because results are rounded to the nearest whole number, some totals may appear inconsistent.*

† *Met guidlines for sample participation rates only after replacement schools were included.*
1 *National Desired Population does not cover all of International Desired Population – German-speaking cantons only.*
2 *National Defined Population covers less than 90% of National Desired Population for the main assessment.*
3 *School-level exclusions for performance assessment exceed 25% of the National Desired Population .*

*Source: Harmon et al., 1997*

Of the 20 countries whose students achieved significantly higher mean scores for mathematics in the TIMSS written tests for Year 9 (or equivalent) students, nine also took part in the Performance Assessment of 13-year-olds. However, when students' overall results on the mathematics tasks are compared, there were only six countries (out of the 19 for which results on the Performance Assessment are presented in the international report) in which students' mean percentage scores for mathematics were higher than those of students in England. Students in only three of the eight countries selected for comparison outperformed their counterparts in England (Singapore, by six percentage points; Switzerland, by two percentage points; and Sweden, by one percentage point).

In science, only students in the Czech Republic and Singapore had achieved significantly higher mean scores than students in England in the written tests. In the Performance Assessment, students in England outperformed their counterparts in the Czech Republic by 11 percentage points, and were only one percentage point below the mean percentage score achieved by students in Singapore.

# 2.3 Mathematics tasks

## 2.3.1 Overall performance on mathematics tasks

In this section, summary information about the performance of students in each of the nine selected countries is presented for the six mathematics tasks, as well as information about the performance of students in England on each task as compared with the international mean.

Table 2.3.1 shows the mean percentage scores achieved by students in each of the selected countries for the mathematics tasks, together with an overall percentage for all the mathematics tasks.

**Table 2.3.1 Mean percentage scores for mathematics tasks for nine countries**

| | | Mathematics tasks | | | | | |
|---|---|---|---|---|---|---|---|
| **Country** | Mean % scores for mathematics | *Dice* | *Calculator* | *Folding and Cutting* | *Around the Bend* | *Packaging* | *Plasticine* |
| Singapore | 70 | 84 | 60 | 80 | 63 | 65 | 66 |
| Switzerland | 66 | 79 | 61 | 79 | 54 | 47 | 73 |
| Sweden | 65 | 74 | 51 | 80 | 65 | 47 | 72 |
| **England** | **64** | **79** | **62** | **69** | **63** | **53** | **55** |
| Czech Republic | 62 | 73 | 54 | 73 | 58 | 43 | 68 |
| Canada | 62 | 77 | 60 | 59 | 53 | 57 | 65 |
| Netherlands | 62 | 76 | 59 | 71 | 67 | 53 | 44 |
| Scotland | 61 | 76 | 49 | 71 | 58 | 51 | 61 |
| United States | 54 | 71 | 56 | 68 | 48 | 28 | 53 |
| **International mean** | **59** | **73** | **54** | **69** | **54** | **44** | **60** |

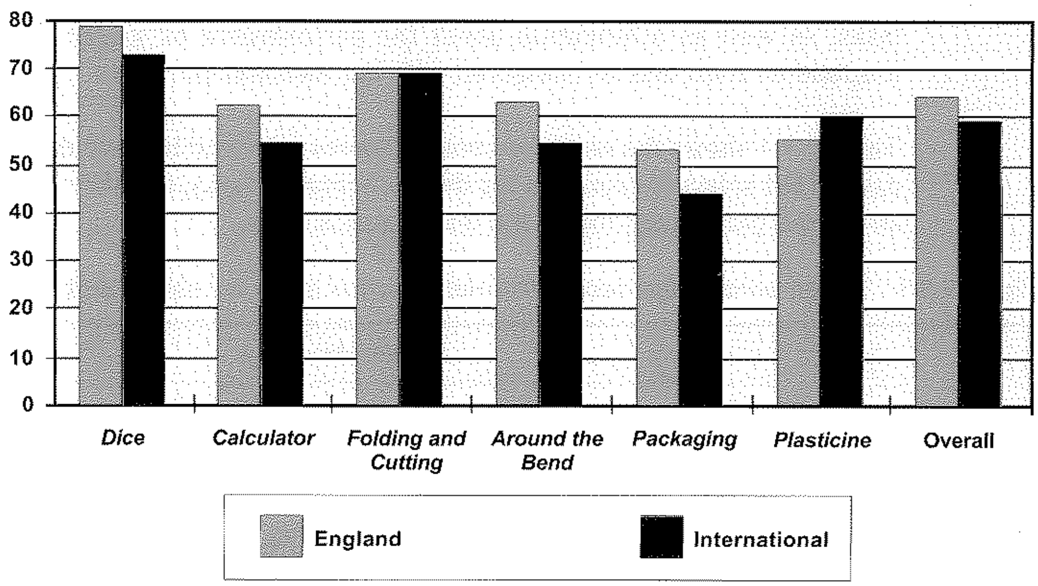
*Source: Tables 2.1 and 2.3 (Harmon et al., 1997)*

It is interesting to note the range of performance by students in the selected countries on the different mathematics tasks. For some tasks, such as *Calculator*, the performance of students in each of the selected countries was not substantially different from the international mean, and the difference between the highest and lowest scores amounted to only 13 percentage points. However, two of the tasks — *Plasticine* and *Packaging* — showed a greater range in performance, with differences between the highest and lowest scores of 29 percentage points and 37 percentage points respectively. It is possible that the skills addressed in some of the tasks were widely covered within the selected countries, whereas other skills were less widely taught.

Despite the considerable differences in performance on some tasks, when the results for the six mathematics tasks were aggregated to produce an overall percentage score for mathematics, there was relatively little variation between the performance of students in the selected countries, with differences of no more than five percentage points between seven of the nine countries.

Figure 2.3.1 shows the performance of students in England on each of the mathematics tasks, and for mathematics overall, compared with the international means.

**Figure 2.3.1 Performance on mathematics tasks**




*Source: Table 2.1 (Harmon et al., 1997)*

It can be seen from Figure 2.3.1 that the mean percentage scores achieved by students in England on the mathematics tasks were in line with or higher than the international mean for five out of the six tasks. Where students' performance exceeded the international means, the differences ranged from five to nine percentage points. However, for *Plasticine*, their performance was five percentage points below the international mean. Overall, these results are better than those for the TIMSS written achievement tests in mathematics, in which the performance of students in England was lower than the international means in five out of the six mathematics content areas defined for the written achievement tests, and above the international mean in only one area.

The following sections present more detailed information about the performance of 13-year-old students in England on each of the six mathematics tasks within the Performance Assessment, and draw comparisons with the performance of students in other countries.

## 2.3.2 *Dice*

**EXAMPLE OF FULL TASK WITH SCORING CRITERIA**

**DICE**

### DICE

**At this station you should have:**

One dice
A shaker

**Read ALL directions carefully.**

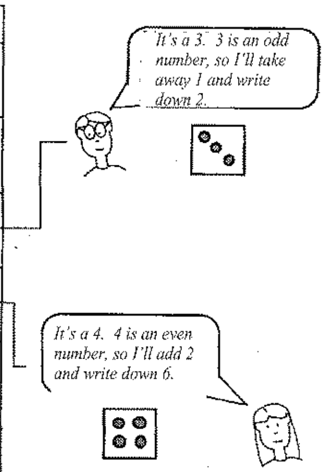
**Your task:**
Find out what happens when we use a rule to change the numbers that turn up when a dice is thrown.

***The rule for changing the numbers is:***

- If an ODD number turns up, take away 1 and write down the result.
- If an EVEN number turns up, add 2 and write down the result.

**1.** In the table below, two examples have already been recorded for you. Use the rule to find out what the other changed numbers will be. Complete the table.

| Number on dice | Changed numbers |
|---|---|
| 1 | 0 |
| 2 | 4 |
| 3 | 2 |
| 4 | 6 |
| 5 | 4 |
| 6 | 8 |




**ITEM 1**

**ITEMS 2 and 3**

**2.** What do you notice about the numbers you recorded?

The odd numbers start with 0 then go up in 2's to 4
The even numbers start from 4 and go up to 8 by adding 2's
There are also no odd numbers on the right.

**3.** Throw the dice 30 times. Each time you throw the dice, change the numbers that turn up using the rule. Each time record the number on the dice and the changed numbers. Write the numbers in the tables below.

| Number on dice | Changed number |
|---|---|
| 4 | 6 |
| 6 | 8 |
| 2 | 4 |
| 5 | 4 |
| 3 | 2 |
| 5 | 4 |
| 3 | 2 |
| 5 | 4 |
| 5 | 4 |
| 5 | 4 |
| 4 | 6 |
| 4 | 6 |
| 5 | 4 |
| 1 | 0 |
| 1 | 0 |

| Number on dice | Changed number |
|---|---|
| 6 | 8 |
| 4 | 6 |
| 6 | 8 |
| 5 | 4 |
| 4 | 6 |
| 5 | 4 |
| 4 | 6 |
| 6 | 8 |
| 6 | 8 |
| 6 | 8 |
| 4 | 6 |
| 6 | 8 |
| 4 | 6 |
| 1 | 0 |
| 2 | 4 |

4. Look again at the table you filled in for question number 3. How many times did you record each of the following numbers in the "Changed Number" column?

| Changed numbers | Number of Times Recorded |
|---|---|
| 0 | III |
| 1 | 0 |
| 2 | II |
| 3 | 0 |
| 4 | ~~IIII~~ ~~IIII~~ |
| 5 | 0 |
| 6 | ~~IIII~~ III |
| 7 | 0 |
| 8 | ~~IIII~~ II |

5a. What changed number did you record most?

The number 4

5b. Why did it happen this way?

because when you get an even number you add 2. So it is still an even one. With an odd number you take 1 away to make it an even number. 8 could not be the most because you can only get that by adding 2. 0 could not be the most because you only get that by taking one away which leaves 4 because that can be added to and taken away to get that number.

**PUT YOUR MATERIALS BACK THE WAY YOU FOUND THEM SO THAT SOMEONE ELSE CAN USE THIS STATION.**

**ITEMS 4 AND 5**

## CRITERIA FOR FULLY-CORRECT RESPONSE

**Item 1 – Change numbers according to algorithm to complete table.**
Applies algorithm correctly (0, 4, 2, 6, 4, 8).
*Total Possible Marks: 2*

**Item 2 – Identify and describe pattern in numbers.**
i) Describes pattern that is consistent with data. ii) Patterns and trends may be one or more of the following: all numbers are even; numbers range from 0 to 8; number 4 occurs twice; rule for obtaining sequential numbers, such as +4, -2, +4, -2.
*Total Possible Marks: 1*

**Item 3 – Apply algorithm to dice throws and record resulting numbers in table.**
i) Completes at least 25 throws of dice. ii) Applies algorithm correctly.
*Total Possible Marks: 2*

**Item 4 – Count frequency of each changed number recorded in table.**
Response consistent with data table.
*Total Possible Marks: 2*

**Item 5a – Identify most frequently recorded number in table.**
Response is consistent with data.
*Total Possible Marks: 1*

**Item 5b – Explain most frequently recorded number in table.**
Provides plausible explanation to account for the predominance of observed number.
*Total Possible Marks: 1*

Table 2.3.2 shows the mean percentage scores for each item within *Dice* for students in the selected countries.

**Table 2.3.2 Mean percentage scores for *Dice* in nine countries**

| Country | Overall mean for task | Mean percentage scores on items within task | | | | | |
|---|---|---|---|---|---|---|---|
| | | Q1 Complete table | Q2 Describe pattern | Q3 Apply algorithm | Q4 Count frequencies | Q5A Identify most frequent number | Q5B Explain findings |
| | | *2 marks* | *1 mark* | *2 marks* | *2 marks* | *1 mark* | *1 mark* |
| Singapore | 84 | 97 | 90 | 95 | 84 | 95 | 44 |
| **England** | **79** | **97** | **83** | **93** | **73** | **90** | **38** |
| Switzerland | 79 | 91 | 86 | 94 | 69 | 86 | 45 |
| Canada | 77 | 92 | 84 | 90 | 75 | 88 | 31 |
| Netherlands | 76 | 97 | 82 | 96 | 72 | 87 | 21 |
| Scotland | 76 | 93 | 73 | 93 | 70 | 87 | 41 |
| Sweden | 74 | 94 | 65 | 92 | 71 | 81 | 44 |
| Czech Republic | 73 | 93 | 75 | 83 | 73 | 78 | 39 |
| United States | 71 | 89 | 76 | 88 | 69 | 77 | 29 |
| **International mean** | **73** | **90** | **71** | **90** | **71** | **83** | **33** |

*Source: Table 1.15 (Harmon et al., 1997)*

The performance of students in England exceeded the international mean for *Dice* in each of the six parts of the task as well as for the task overall. Students in only one country (Singapore) achieved an overall mean percentage score for *Dice* which was higher than that achieved by their counterparts in England.

## 2.3.3 *Calculator*

**EXAMPLE OF FULL TASK WITH SCORING CRITERIA**

***CALCULATOR***

**ITEMS 1 and 2**

### CALCULATOR

**At this station you should have:**

A calculator

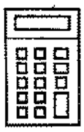
> **Your task:**
> Use a calculator to help you explore a number pattern, and to find missing numbers.

**Before answering the questions read these notes:**

When you use the calculator:

- Make sure that you press the correct keys.
- Make sure that you read the display carefully.

**1.** Use the calculator to find the answers to these multiplications.



34 x 34 = 1156

334 x 334 = 111556

3334 x 3334 = 11115556

**2.** What do you notice about the muliplications and the pattern of answers?

The number of 1's in the answer is one more than the number of 3's in one of the numbers multiplied

eg. 34 x 34 no. of 3's = 1 no. of 1's = 2

334 x 334 no. of 3's = 2 no. of 1's = 3

The number of 5's in the answer is the same as the number of 3's in one of the numbers multiplied

eg. 334 x 334 no. of 3's = 2 no. of 5's = 2

3334 x 3334 no. of 3's = 3 no. of 5's = 3

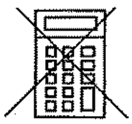
**ITEMS 3 to 6**

**3.** Now use the pattern to write down what you think the answer will be to the multiplication below WITHOUT using the calculator.

(4 x 3's) (5 x 1's, 4 x 5's)

33334 x 33334 = 1111155556

**4.** Now write down what you think the answer will be to the multiplication below WITHOUT using the calculator.



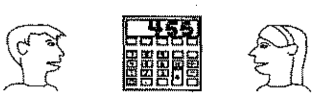
(7 x 1's, 6 x 5's)

3333334 x 3333334 = 11111115555556

(6 x 3's)

**5.** How did you find the answer to questions 3 and 4?

By counting the number of 3's in the number being multiplied and applying the rule I discovered previously, i.e. add 1 to the no. of 3's to get the no. of 1's ; have the same no. of 5's as no. of 3's.



**6.** Rachel tells Ahmed that she multiplied two whole numbers together using a calculator and the answer was 455, but she's forgotten the numbers. She can remember two things about them:

- both numbers had 2 digits
- both numbers were less than 50

**Please turn the page**

Ahmed tries several numbers. He began by putting 7 x 64 into the calculator. But Rachel said, "I can give you at least three reasons why those numbers can't be the ones I used. "What were Rachel's reasons?

a. Both numbers had 2-digits, and 7 only has one digit.

b. Both numbers were less than 50, and 64 is greater than 50.

c. The answer was 455, so both numbers must be odd, and 64 is even.

After thinking a bit about the problem, Ahmed made some more tries and found the two numbers.

- Now you try to find the numbers Ahmed found.

You may use any method you like. Write down each of your tries here.

21 × 21 = 441
21 × 23 = 483
19 × 23 = 437
19 × 25 = 475
17 × 25 = 425
17 × 27 = 459
15 × 31 = 465
13 × 35 = 455

**ITEM 6 (contd.)**

## CRITERIA FOR FULLY-CORRECT RESPONSE

**Item 1 – Use calculator to perform multiplications.**
All 3 calculations correct (1156, 111556, 11115556).
*Total Possible Marks: 3*

**Item 2 – Identify pattern in answers.**
i) Identifies a correct pattern. ii) Includes the repetitions of 1, 5, and may include 6. iii) Identifies a relationship between these and the increasing number of digits or the increasing numbers of 3 in the multipliers.
*Total Possible Marks: 2*

**Item 3 – Predict answer to first (routine) calculation.**
Predicts answer based on application of correct pattern (1111155556).
*Total Possible Marks: 2*

**Item 4 – Predict answer to second (non-routine) calculation.**
Predicts answer based on application of correct pattern (11111115555556).
*Total Possible Marks: 2*

**Item 5 – Describe strategy for predicting answers.**
Describes pattern and a correct method of application.
*Total Possible Marks: 2*

**Item 6 – Factors of 455.** Responses to two parts are scored separately.

**List three reasons why Alison's factors are incorrect.** Lists 3 of the following, or other correct reasons: 7 is not a two-digit number; 64 is more than 50; 64 is an even number so the product will be even; neither 7 nor 64 is a multiple of 5.
*Total Possible Marks: 3*

**Find correct factors.** i) Identifies correct factors (35 x 13). ii) Shows use of a systematic method.
*Total Possible Marks: 2*

Table 2.3.3 shows the mean percentage scores for each item within *Calculator* for students in the selected countries.

**Table 2.3.3 Mean percentage scores for *Calculator* in nine countries**

| Country | Overall mean for task | Mean percentage scores on items within task | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | | Q1 Perform calculations | Q2 Identify pattern | Q3 Predict: routine application | Q4 Predict: non-routine application | Q5 Explain predictions | Q6 Factors of 455 | |
| | | | | | | | Reasons factors incorrect | Find correct factors |
| | | *3 marks* | *2 marks* | *2 marks* | *2 marks* | *2 marks* | *3 marks* | *3 marks* |
| **England** | **62** | **98** | **50** | **85** | **59** | **61** | **53** | **29** |
| Switzerland | 61 | 99 | 51 | 85 | 64 | 55 | 40 | 33 |
| Singapore | 60 | 98 | 33 | 84 | 64 | 45 | 53 | 45 |
| Canada | 60 | 97 | 44 | 86 | 64 | 47 | 50 | 30 |
| Netherlands | 59 | 97 | 37 | 77 | 58 | 42 | 78 | 25 |
| United States | 56 | 97 | 44 | 79 | 51 | 44 | 54 | 20 |
| Czech Republic | 54 | 96 | 45 | 76 | 58 | 45 | 44 | 15 |
| Sweden | 51 | 95 | 40 | 69 | 52 | 49 | 39 | 10 |
| Scotland | 49 | 97 | 44 | 65 | 43 | 45 | 35 | 15 |
| **International mean** | **54** | **97** | **40** | **75** | **55** | **42** | **45** | **21** |

*Source: Table 1.17 (Harmon et al., 1997)*

Students in England outperformed students in all the other countries selected for comparisons in their overall mean percentage score for *Calculator*. Furthermore, their mean percentage score for Q5, in which students were asked to explain how they worked out their prediction for another long multiplication (without using a calculator) was higher than that of the students in any of the other countries that participated in the Performance Assessment. In this context, it is worth noting that the main survey found that calculators were used more frequently by 13- and nine-year-olds in England than in any of the other countries that took part in TIMSS (Keys *et al.*, 1997a; 1997b). It is therefore possible that previous work using calculators to explore number patterns meant that students in England were more confident in answering this item than their counterparts in other countries.

## 2.3.4 *Folding and Cutting*

**EXAMPLE OF FULL TASK WITH SCORING CRITERIA**

# *FOLDING AND CUTTING*

### FOLDING AND CUTTING

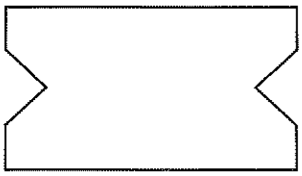
**At this station you should have:**

9 sheets of paper
Scissors
An envelope

**Your task:**
Fold and cut sheets of paper to make shapes which match the patterns given. For each shape you may fold the paper as often as you like, but ONLY ONE straight cut is allowed.

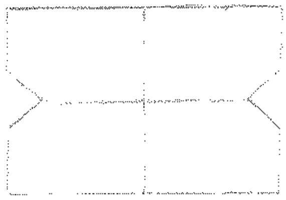
1. Look at Shape 1 below. Fold a sheet of paper as many times as necessary and make ONE STRAIGHT CUT so that when the paper is unfolded it has the same SHAPE as Shape 1. The SIZE of your paper and cutouts do not have to be the same as those shown here. If you are unsuccessful, you may try again with another sheet of paper. You may try this task THREE times.
   - Write number 1 on each sheet of paper you used for this task.
   - Write your full name on each sheet.

**Shape 1**



**ITEM 1**

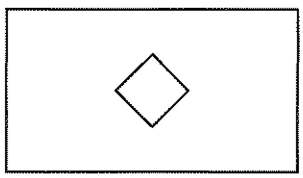
**Student's work (Item 1)**



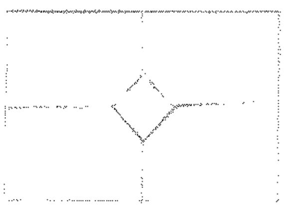
**ITEMS 2 and 3**

2. Do the same for Shape 2. Remember only ONE STRAIGHT CUT is allowed. You may try this task THREE times.
   - Write the number 2 on each sheet of paper you used for this task.
   - Write your full name on each sheet.

**Shape 2**

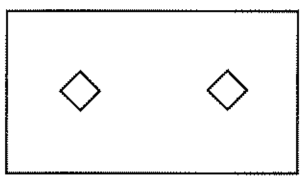


**Student's work (Item 2)**

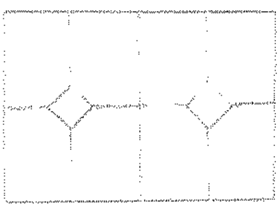


3. Do the same for Shape 3. Remember only ONE STRAIGHT CUT is allowed. You may try this task THREE times.
   - Write number 3 on each sheet of paper you used for this task.
   - Write your full name on each sheet.

**Shape 3**



**Student's work (Item 3)**



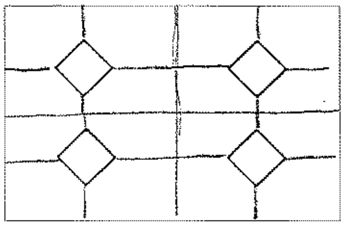
**Please turn the page.**

4. For this question, Shape 4 is drawn below. Instead of folding or cutting shape 4, you must THINK about how to get the pattern by folding a piece of paper and making one straight cut. DON'T FOLD OR CUT ANY PAPER FOR THIS QUESTION.

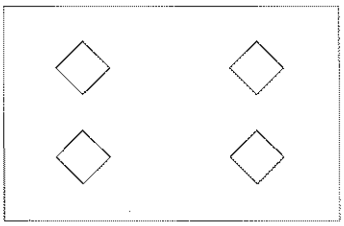
Instead draw on the diagram below the LINES you would see on a piece of paper that had been folded and cut.

Two copies of shape 4 are drawn here in case you are not satisfied with your first attempt and wish to try again. Remember, only draw lines to show where the paper should be folded.

**Shape 4**



**Shape 4**



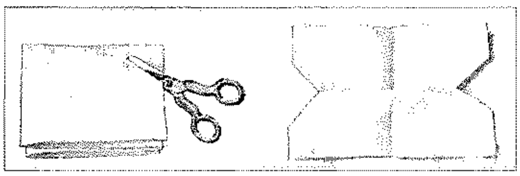
**PUT ALL YOUR SHEETS OF PAPER INTO YOUR ENVELOPE, INCLUDING YOUR UNSUCCESSFUL TRIES.**

**THROW AWAY ANY SCRAPS OF PAPER.**

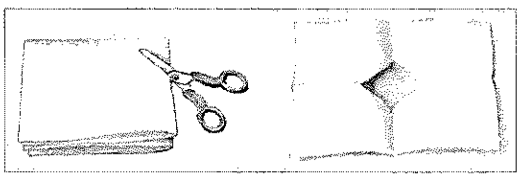
**ITEM 4**

## CRITERIA FOR FULLY-CORRECT RESPONSE

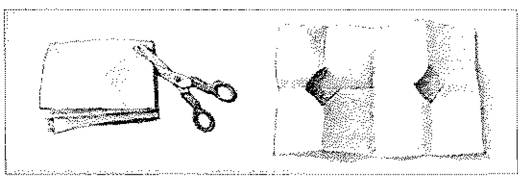
**Item 1 – Fold paper and cut out shape 1.**
i) Makes only one cut line.
ii) Places two fold lines correctly.
*Total Possible Marks: 2*



**Item 2 – Fold paper and cut out shape 2.**
i) Makes ony one cut line.
ii) Places two fold lines correctly.
*Total Possible Marks: 2*



**Item 3 – Fold paper and cut out shape 3.**
i) Makes only one cut line.
ii) Places four fold lines correctly.
*Total Possible Marks: 2*



**Item 4 – Predict and draw fold lines on shape 4.**
Shows six fold lines in the correct locations.
*Total Possible Marks: 3*

Table 2.3.4 shows the mean percentage scores for each item within *Folding and Cutting* for students in the selected countries.

**Table 2.3.4 Mean percentage scores for *Folding and Cutting* in nine countries**

| Country | Overall mean for task | Mean percentage scores on items within task | | | |
|---|---|---|---|---|---|
| | | Q1 Fold and cut Shape 1 | Q2 Fold and cut Shape 2 | Q3 Fold and cut Shape 3 | Q4 Predict and draw Shape 4 |
| | | *2 marks* | *2 marks* | *2 marks* | *3 marks* |
| Sweden | 80 | 84 | 88 | 86 | 62 |
| Singapore | 80 | 83 | 86 | 81 | 72 |
| Switzerland | 79 | 80 | 89 | 85 | 63 |
| Czech Republic | 73 | 78 | 84 | 75 | 55 |
| Scotland | 71 | 78 | 80 | 74 | 53 |
| Netherlands | 71 | 70 | 79 | 75 | 59 |
| **England** | **69** | **66** | **80** | **69** | **62** |
| United States | 68 | 72 | 82 | 75 | 45 |
| Canada | 59 | 60 | 72 | 63 | 42 |
| **International mean** | **69** | **71** | **79** | **72** | **53** |

*Source: Table 1.19 (Harmon et al., 1997)*

Although the overall percentage correct for students in England for *Folding and Cutting* was in line with the international mean (69 per cent), their performance was lower than the international means for two of the four questions, Q1 and Q3. However, for the final question, the mean percentage score for students in England exceeded the international mean by nine percentage points. Students in six of the other countries selected for comparisons achieved higher overall mean percentage scores for *Folding and Cutting* than their counterparts in England.

## 2.3.5 *Around the Bend*

**EXAMPLE OF FULL TASK WITH SCORING CRITERIA**

**AROUND THE BEND**

### AROUND THE BEND

**At this station you should have:**

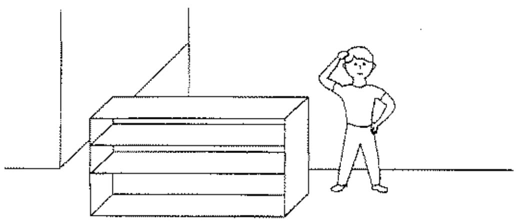
- Two rectangles of white card, A and B, which are models of pieces of furniture
- 1 cm squared graph paper to make different rectangles to models of other pieces of furniture (for question 5)
- Scissors
- A 30 cm ruler
- Plastic bag and labels
- Paper clips
- A model representing a corridor in a flat

**Your task:**
Find out what sizes of furniture can be moved around the bend in the corridor.

**Read this before answering the questions:**

Ray is moving into a flat which has the main rooms around a bend in the corridor leading from the front door.

What sizes of furniture will go around the bend in the corridor?

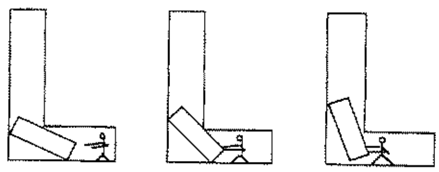


Ray wants to get some large pieces of furniture around the bend the right way up. He does not want to turn the pieces of furniture on their sides. He uses the models of the corridor and furniture to find out which pieces of furniture will go around the bend.

**Instructions**

**ITEMS 1 to 4**

Here are some pictures (not to scale) showing what could happen.



The models representing the furniture and the corridor in Ray's flat are to scale.
*Scale: 4 cm represents 1 m.*

1. Measure the lengths and widths of the two models of pieces of furniture in cm.

   A is 8 cm long and 4 cm wide.
   B is 4 cm long and 2 cm wide.

2. What are the lengths and widths of the two pieces of furniture in metrees?

   A is 2 m long and 1 m wide.
   B is 1 m long and 0.5 m wide.

3. Here is a list of furniture:

| | | | | |
|---|---|---|---|---|
| single bed | coffee table | 3-seater sofa | armchair | cot |
| double bed | dining table | 2-seater sofa | sideboard | |

Judging from their sizes:
What piece of furniture is A most likely to be? 3 - seater sofa
What piece of furniture is B most likely to be? coffee table

4. Which piece(s) of furniture (A or B or both) will go around the bend in Ray's flat and which will not?

   A will not go round the bend and B will go round the bend

**Please turn the page.**

5. Use the graph paper to make other models of pieces of furniture to the sizes listed in the table below. The sizes are all given in metres.

In the second column of the table suggest what the furniture could be.

In the third column find out if the piece of furniture will go round the bend, and check the answer.

| | Furniture Size Length (m) | Width (m) | What Furniture could be: | Goes around the bend? Yes, easily | Yes, just | No |
|---|---|---|---|---|---|---|
| C | 0.5 | 0.5 | chair | ✓ | | |
| D | 1.5 | 0.5 | ~~sideboard~~ bookcase | | ✓ | |
| E | 2 | 0.5 | sideboard | | ✓ | |
| F | 1 | 1 | a armchair | | ✓ | |
| G | 1.5 | 1 | dining table | | | ✓ |
| H | 2 | 1 | single bed | | | ✓ |

Easy = l + w = 1
Just l + w = 2/2.5
No = l + w 2.5/3

6. Whether or not a piece of furniture goes around the bend of Ray's corridor depends on its length and width. Look at the results you have for all the pieces of furniture A, B, C, D, E, F, G and H.

- Try to find a rule for working out from their lengths and widths whether or not a piece of furniture goes around the bend.

l = length w = width

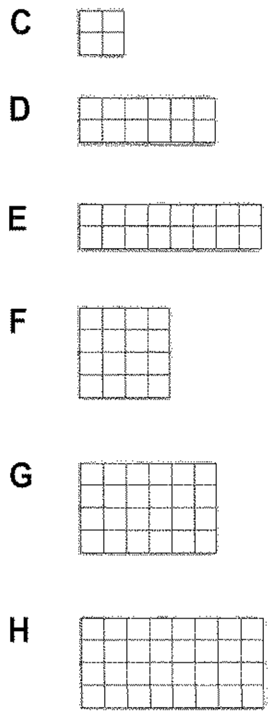
Yes, Easy = l + w = 1 (or less)

Yes. Just = l + w = 2 > 2.5

No = l + w = 2.5 > 3 >

**PUT THE PIECES OF FURNITURE YOU MADE IN THE PLASTIC BAG AND PUT YOUR NAME ON THE LABEL.**
**FASTEN THE BAG TO THIS PAGE WITH A PAPER CLIP.**
**LEAVE THE MODELS A AND B AT THE STATION.**

**Student's work (Item 5)**




**ITEMS 5 and 6**

## CRITERIA FOR FULLY-CORRECT RESPONSE

**Item 1 – Measure lengths and widths of two furniture models.**
Length and width measured correctly (in cm) for both pieces (A = 8 cm x 4 cm; B = 4 cm x 2 cm).
*Total Possible Marks: 2*

**Item 2 – Convert cm into meters using scale.**
Computes conversion of length and width accurately (A = 2 m x 1 m; B = 1 m x 0.5 m).
*Total Possible Marks: 2*

**Item 3 – Relate models to real-world furniture.**
Judgments are reasonable for both pieces.
*Total Possible Marks: 2*

**Item 4 – Solve problem: which pieces (s) of furniture will go around the bend.**
i) Response is consistent with measurements. ii) Response is correct (A will not; B will).
*Total Possible Marks: 2*

**Item 5 – Draw and make predictions about six models.**
Three aspects of responses were scored separately.

**Draw or cut models to scale.** i) Correctly applies scale. ii) Computations and drawings are accurate for all 6 pieces. (C = 2 cm x 2 cm); D = 6 cm x 2 cm; E = 8 cm x 2 cm; F = 4 cm x 4 cm; G = 6 cm x 4 cm; H = 8 cm x 4 cm).
*Total Possible Marks: 3*

**Relate models to real-world furniture.** Judgments are reasonable for all 6 pieces.
*Total Possible Marks: 2*

**Solve problem: which piece(s) of furniture will go around the bend.** Judges all six pieces of furniture correctly, based on drawings (C and D – Easily; E and F – Just; G and H – No).
*Total Possible Marks: 3*

**Item 6 – Find a general rule for whether furniture will or will not go around the bend.** Includes the correct relationship between length and width based on the corridor dimensions and scale: i.e., furniture will go around the bend if (1/2 x length + width) ≤1.5 m.
*Total Possible Marks: 3*

Table 2.3.5 shows the mean percentage scores for each item within *Around the Bend* for students in the selected countries.

**Table 2.3.5 Mean percentage scores for *Around the Bend* in nine countries**

| Country | Overall mean for task | Mean percentage scores on items within task | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | Q1 Measure models A and B | Q2 Convert using scale | Q3 Relate A and B to real furniture | Q4 Solve problem with A and B | Q5 Six models: Draw models to scale | Q5 Six models: Relate models to real furn. | Q5 Six models: Solve problem with models | Q6 Find general rule |
| | | *2 marks* | *2 marks* | *2 marks* | *2 marks* | *3 marks* | *2 marks* | *3 marks* | *3 marks* |
| Netherlands | 67 | 92 | 89 | 87 | 80 | 54 | 60 | 68 | 5 |
| Sweden | 65 | 89 | 95 | 81 | 46 | 70 | 57 | 73 | 12 |
| Singapore | 63 | 94 | 82 | 76 | 89 | 66 | 33 | 64 | 2 |
| **England** | **63** | **94** | **81** | **68** | **82** | **65** | **33** | **67** | **11** |
| Czech Republic | 58 | 95 | 83 | 61 | 79 | 51 | 44 | 51 | 4 |
| Scotland | 58 | 95 | 78 | 50 | 80 | 57 | 39 | 58 | 9 |
| Switzerland | 54 | 81 | 80 | 47 | 64 | 58 | 31 | 63 | 9 |
| Canada | 53 | 82 | 67 | 63 | 68 | 48 | 42 | 56 | 1 |
| United States | 48 | 68 | 53 | 66 | 62 | 33 | 45 | 52 | 3 |
| **International mean** | **54** | **84** | **69** | **66** | **69** | **47** | **42** | **53** | **5** |

*Source: Table 1.21 (Harmon et al., 1997)*

The overall mean percentage score for students in England was higher than the international mean for *Around the Bend.* Students' mean percentage scores were also above the international means for each part of the task, with the exception of that part of Q5 in which students were asked to suggest what each piece of furniture with the specified dimensions could be.

Very few students in any of the countries taking part in the Performance Assessment of 13-year-olds were able to suggest a general rule for determining whether or not a particular piece of furniture would go around the bend (Q6) which included the correct mathematical relationship (furniture will fit if ($\frac{1}{2} \times$ length + width) ($\leq 1.5$m).

## 2.3.6 *Packaging*

**EXAMPLE OF FULL TASK WITH SCORING CRITERIA**

**PACKAGING**

### PACKAGING

**At this station you should have:**

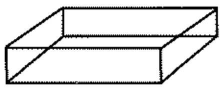
- 4 plastic balls packed in a square-shaped box
- Blu tac to stop the balls from rolling around
- Some thin card to make a package for the balls
- A pair of compasses
- A 30 cm ruler
- Two pieces of thick card to help measure the balls
- Scissors
- Paper clips

**Your task:**
Find different ways of putting 4 plastic balls in a box. Then make a box that the 4 balls will fit into.

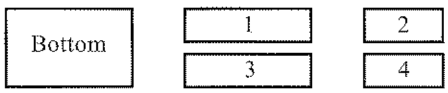
**Read this before answering the questions:**

The following shows what is meant by the net of a box.

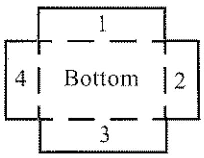
This box has a bottom and 4 sides.



The sides can be cut out separately:




Or the sides can be cut out in one piece and then fold along the dotted lines like this:

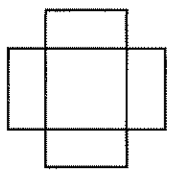



This is a net of a box.

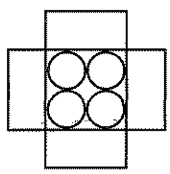
**Instructions**

**ITEM 1**

This is the shape of a net of a box like the one which holds the 4 balls. It is not drawn to size but if it were, you could fold up the sides and make the box.

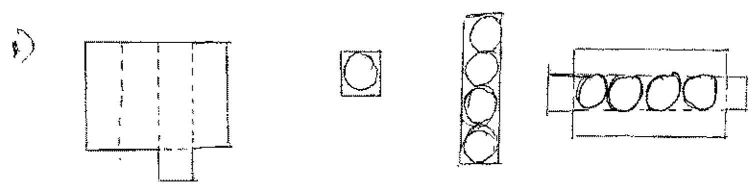


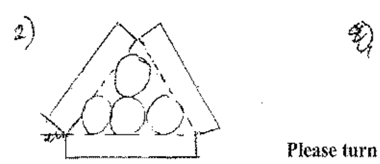
You have been given the box with the four balls just fitting in like this.



Other boxes with different shapes could be made so that the 4 balls would just fit in.

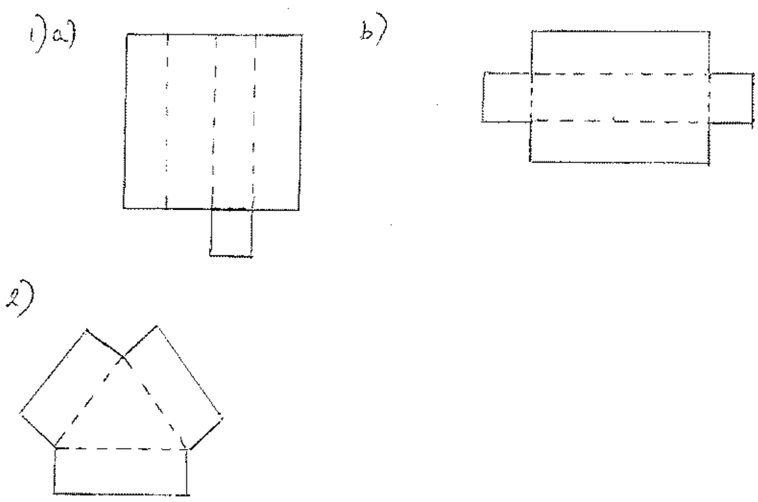
1. Use the balls to find 2 or 3 other shapes for boxes that the 4 balls will just fit into. Make a drawing of each box with the 4 balls in it.





**Please turn the page.**

**2.** Now make a drawing of the net for each box.




**3.** Choose ONE of the boxes you have drawn.

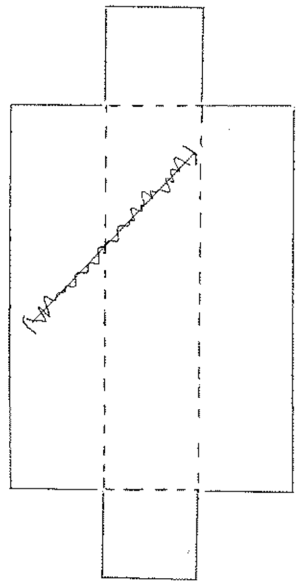
- Use a piece of plain card and make a NET of the box.
- Make the net the correct size to contain the 4 balls.

**ATTACH THE NET TO THIS PAGE WITH A PAPER CLIP.**

**LEAVE EVERYTHING ELSE AS YOU FOUND IT.**

**ITEMS 2 and 3**

**Student's work (Item 3)**



## CRITERIA FOR FULLY-CORRECT RESPONSE

**Item 1 – Draw three boxes that hold four balls in a "tightly packed" arrangement.**
i) Each box describes or shows all four balls. ii) Shows balls in "tightly packed" arrangements. iii) Draws at least two unique arrangements.
*Total Possible Marks: 2*

**Item 2 – Draw net for each box.**
i) Nets drawn are consistent with at least two of the ball arrangements. ii) Nets clearly show correct shape of base of box and side flaps required to constrain the balls in "tightly packed" arrangements. iii) Nets show side flaps and base of box in correct proportions (not necessarily in actual size).
*Total Possible Marks: 2*

**Item 3 – Construct net to scale.**
Constructs or draws a net for a box with the following requirements:
i) Net is consistent with one of the previous nets drawn. ii) Is constructed out of a single piece of cardboard or pieces are taped together and spread out into a net. iii) Includes base and side flaps that will constrain the ball in the "tightly packed" arrangement when folded up into a box.
iv) Dimensions of base and side flaps are within 4 mm of actual size required to hold the 4 balls.
*Total Possible Marks: 2*

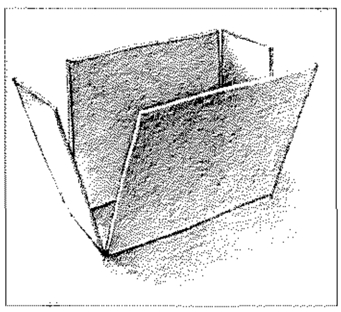


Table 2.3.6 shows the mean percentage scores for each item within *Packaging* for students in the selected countries.

**Table 2.3.6 Mean percentage scores for *Packaging* in nine countries**

| Country | Overall mean for task | Mean percentage scores on items within task | | |
|---|---|---|---|---|
| | | Q1 Draw boxes | Q2 Draw nets | Q3 Construct net to scale |
| | | *2 marks* | *2 marks* | *2 marks* |
| Singapore | 65 | 87 | 55 | 51 |
| Canada | 57 | 67 | 52 | 51 |
| **England** | **53** | **72** | **44** | **45** |
| Netherlands | 53 | 64 | 52 | 43 |
| Scotland | 51 | 59 | 41 | 54 |
| Sweden | 47 | 68 | 32 | 40 |
| Switzerland | 47 | 50 | 56 | 35 |
| Czech Republic | 43 | 53 | 39 | 38 |
| United States | 28 | 41 | 27 | 17 |
| **International mean** | **44** | **53** | **38** | **41** |

*Source: Table 1.23 (Harmon et al., 1997)*

The mean percentage scores for students in England were above the international means for overall performance on *Packaging*, and for each of the three parts of the task. Out of the countries selected for comparisons, only students in Singapore outperformed students in England on Q1, which asked students to draw two or three other boxes that the four balls would fit into. It is possible that students' activities in design technology provided relevant experiences for this task.

## 2.3.7 *Plasticine*

**EXAMPLE OF FULL TASK WITH SCORING CRITERIA**

**PLASTICINE**

### PLASTICINE

**At this station you should have:**

Some plasticine
A balance
Plastic bags
A 20 g and a 50 g mass (weight)
Coloured small circular sticky labels

**Read ALL directions carefully.**

**Your task:**
Use the balance to weigh different amounts of plasticine as carefully as you can. Then explain how you made them.

**Before starting the task:**

MAKE SURE THE PLANS ARE BALANCED WHEN EMPTY.

IF THEY ARE NOT, PUT YOUR HAND UP AND TELL THE TEACHER.

**1a.** Use the balance to make a lump of plasticine that weighs 20 g.

- When you have made the 20 g lump, write 20 g on a coloured label and stick it on the lump. Put the lump in a plastic bag.

**1b.** Write down how you made the 20 g lump.

I made the 20g by ~~be~~ putting a 20g balance in one of the pot on the scale and putting lumps of plastacine in the other till it balanced in the middle.

*Check:* 21g.

**ITEM 1**

**ITEMS 2 and 3**

**2a.** Use the balance to make a lump of plasticine that weighs 10 g.

- When you have made the 10 g lump, write 10 g on a coloured label and stick it on the lump. Put the lump in the plastic bag with the 20 g lump.

**2b.** Write down how you made the 10 g lump.

I put ~~X~~ the 20g balance in one pot and made a plastecine ball 20g in the other the same way ~~pe~~ by adding platacine then I took out the 20g balance and cut the 20g plastace ~~in half so it balance~~ in half put one on each end ~~at~~ of the scale till they balanced and used one as my 10g plastacine.

*Check:* 11g

**3a.** Use the balance to make a lump of plasticine that weighs 15 g.

- When you have made the 15 g lump, write 15 g on a coloured label and stick it on the lump. Place the 15 g lump in the plastic bag together with the other lumps.

**3b.** Write down how you made the 15 g lump.

I made a 20g lump like I had before then cut it and made a 10g lump using the scales. I put one 10g lump to one side then ~~mat~~ cut the other ten and made a 5g lump by using the scales and making the scales even. I then put the 5g lump and stuck it to ~~be~~ the ten ~~x~~ I had put aside before from the twenty.

*Check:* 16g

Please turn the page.

**4a.** Use the balance to make a lump of plasticine that weighs 35 g.

- When you have made the 35 g lump, write 35 g on a coloured label and stick it on the lump. Place the 35 g lump in the plastic bag with the other lumps.

**4b.** Write down how you made the 35 g lump.

I made the 35 by putting the 50 & 20 g balace in one side of the scales and making a 70g plastacine lump by previous methods. I cut the 70g in half and balanced on scales I took one even half of 70g which is 35g.

*Check:* 36g

**HAND IN THE BAG WITH THE LUMPS OF PLASTICINE YOU HAVE WEIGHED. MAKE SURE YOUR NAME IS ON THE BAG**

**LEAVE EVERYTHING ELSE AS YOU FOUND IT.**

**ITEM 4**

## CRITERIA FOR FULLY-CORRECT RESPONSE

**Item 1a – Weigh a 20 g lump of plasticine.**
Lump has correct mass (20 ±2 g). (Based on administrator measurement.)
*Total Possible Marks: 1*

**Item 1b – Describe strategy for making 20 g lump of plasticine.**
i) Method includes use of balance. ii) Method plausible for obtaining desired mass.
*Total Possible Marks: 2*

**Item 2a – Weigh a 10 g lump of plasticine.**
Lump has correct mass (10 ±2 g). (Based on administrator measurement.)
*Total Possible Marks: 2*

**Item 2b – Describe strategy for making 10 g lump of plasticine.**
i) Method includes use of balance. ii) Method plausible for obtaining desired mass.
*Total Possible Marks: 2*

**Item 3a – Weigh a 15 g lump of plasticine.**
Lump has correct mass (15 ±3 g). (Based on administrator measurement.)
*Total Possible Marks: 1*

**Item 3b – Describe strategy for making 15 g lump of plasticine.**
i) Method includes use of balance. ii) Method plausible for obtaining desired mass.
*Total Possible Marks: 2*

**Item 4a – Weigh a 35 g lump of plasticine.**
Lump has correct mass (35 ±3 g). (Based on administrator measurement.)
*Total Possible Marks: 2*

**Item 4b – Describe strategy for making 35 g lump of plasticine.**
i) Method includes use of balance. ii) Method plausible for obtaining desired mass.
*Total Possible Marks: 2*

This task was predominantly concerned with the application of mathematical skills (in relation to number combinations and proportion, for example); hence the results for this task are included with the other mathematics tasks, although it was originally classified as testing skills in both mathematics and science.

Table 2.3.7 shows the mean percentage scores for each item within *Plasticine* for students in the selected countries. Countries are ranked in order according to the overall percentage of correct responses for the task.

**Table 2.3.7 Mean percentage scores for *Plasticine* in nine countries**

| | Mean percentage scores on items within task | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | Q1A Weigh 20g lump | Q1B Descibe strategy 20g lump | Q2A Weigh 10g lump | Q2B Descibe strategy 10g lump | Q3A Weigh 15g lump | Q3B Descibe strategy 15g lump | Q4A Weigh 35g lump | Q4B Descibe strategy 35g lump |
| **Country** | **Overall mean for task** | *1 mark* | *2 marks* | *2 marks* | *2 marks* | *1 mark* | *2 marks* | *2 marks* | *2 marks* |
| Switzerland | 73 | 98 | 88 | 82 | 71 | 71 | 50 | 62 | 60 |
| Sweden | 72 | 88 | 97 | 80 | 69 | 73 | 51 | 57 | 58 |
| Czech Republic | 68 | 95 | 96 | 74 | 62 | 65 | 44 | 58 | 51 |
| Singapore | 66 | 99 | 82 | 65 | 60 | 64 | 44 | 60 | 53 |
| Canada | 65 | 93 | 86 | 68 | 58 | 71 | 40 | 54 | 49 |
| Scotland | 61 | 94 | 85 | 59 | 47 | 70 | 39 | 53 | 41 |
| **England** | **55** | **93** | **85** | **44** | **42** | **57** | **29** | **41** | **48** |
| United States | 53 | 91 | 65 | 50 | 34 | 76 | 24 | 46 | 40 |
| Netherlands | 44 | 95 | 80 | 35 | 29 | 31 | 17 | 29 | 38 |
| **International mean** | **60** | **93** | **86** | **63** | **52** | **61** | **37** | **47** | **44** |

*Source: Table 1.13 (Harmon et al., 1997)*

The overall performance of students in England on the *Plasticine* task was their weakest of any of the tasks in the Performance Assessment: it was the only task for which the students' overall mean percentage score was below the international mean. Furthermore, the mean percentage scores for students in England were below the international means for six out of the eight parts of *Plasticine*.

Students in six of the countries selected for comparison achieved a higher mean percentage score for the task overall than their counterparts in England. It is interesting to note that the only question for which the performance of students in England exceeded the international mean was Q4B, in which students were asked to describe their strategy for finding a 35g lump of plasticine using only the balance, the 20g and/or 50g masses provided and the lumps that they had already made. This performance is particularly surprising given that the performance of students in England was below the international mean for three similar items which asked students to describe their strategies for making 20g, 10g and 15g masses.

# 2.4 Science tasks

## 2.4.1 Overall performance on science tasks

In this section, summary information about the performance of students in each of the nine selected countries is presented for the six science tasks, as well as information about the performance of students in England on each task as compared with the international mean.

Table 2.4.1 shows the mean overall percentage achieved by students in each of the selected countries for all the science tasks, together with mean percentage scores for each of the science tasks.

**Table 2.4.1 Mean percentage scores for science tasks for nine countries**

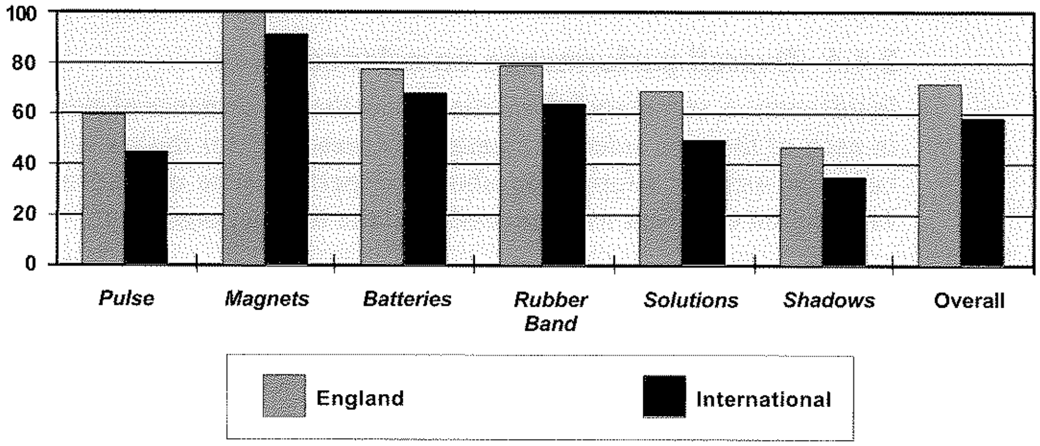
| | | Science tasks | | | | | |
|---|---|---|---|---|---|---|---|
| **Country** | **Mean % scores for science** | ***Pulse*** | ***Magnets*** | ***Batteries*** | ***Rubber Band*** | ***Solutions*** | ***Shadows*** |
| Singapore | 72 | 60 | 95 | 79 | 80 | 68 | 50 |
| **England** | **71** | **59** | **99** | **77** | **79** | **68** | **46** |
| Switzerland | 65 | 51 | 97 | 75 | 67 | 57 | 41 |
| Scotland | 64 | 55 | 98 | 68 | 75 | 51 | 36 |
| Sweden | 63 | 45 | 95 | 71 | 70 | 50 | 45 |
| Czech Republic | 60 | 46 | 86 | 66 | 65 | 59 | 37 |
| Canada | 59 | 46 | 92 | 62 | 71 | 48 | 35 |
| Netherlands | 58 | 45 | 94 | 63 | 70 | 43 | 35 |
| United States | 55 | 50 | 85 | 56 | 63 | 48 | 28 |
| **International mean** | **58** | **44** | **90** | **67** | **63** | **49** | **35** |

*Source: Tables 2.1 and 2.3 (Harmon et al., 1997)*

It is evident from Table 2.4.1 that, although students in eight out of the nine selected countries equalled or exceeded the international mean in their overall score for science, students in two countries (Singapore and England) outperformed their counterparts in the other selected countries by at least six percentage points.

As shown in Figure 2.4.1, for each of the science tasks, the performance of students in England was well above the international mean, with the differences ranging from nine to 19 percentage points. This shows a high level of achievement within science, and is consistent with the results for the written tests, in which the mean scores for students in England were above the international means in all of the five science content areas defined within the written achievement tests.

**Figure 2.4.1 Performance on science tasks**




*Source: Table 2.1 (Harmon et al., 1997)*

The following sections present more detailed information about the performance of 13-year-old students in England on each of the six science tasks within the Performance Assessment, and draw comparisons with the performance of students in other countries.

## 2.4.2 *Pulse*

**EXAMPLE OF FULL TASK WITH SCORING CRITERIA**

**PULSE**

### PULSE

**At this station you should have**

A watch with a second hand
A step on the floor to climb on

**Read ALL directions carefully.**

> **Your task:**
> Find out how your pulse changes when you climb up and down on a step for 5 minutes.

**This is what you should do:**

- Find your pulse and be sure you know how to count it. IF YOU CANNOT FIND YOUR PULSE ASK A TEACHER FOR HELP.
- Decide how often you will take measurements starting from when you are at rest.
- Climb the step for about 5 minutes and measure your pulse at regular intervals.

1. Make a table and write down the times at which you measured your pulse and the measurements you

| After: | beats per minute |
|---|---|
| 0 minutes | 108 |
| 1 minute | 144 |
| | 150 |
| 2 minutes | 162 |
| 3 minutes | 174 |
| 4 minutes | 180 |
| 5 minutes | 186 |

**ITEM 1**

**ITEMS 2 and 3**

2. How did your pulse change during this exercise?

It increased

3. Why do you think your pulse changed in this way?

As I did more exercise, my leg muscles needed more oxygen, so my heart beat faster to get the oxygen there, through the blood.

**PUT EVERYTHING BACK THE WAY YOU FOUND IT SO THAT SOMEONE ELSE CAN USE THE STATION.**

## CRITERIA FOR FULLY-CORRECT RESPONSE

**Item 1 – Measure pulse rates and record in table.**
Response is scored for both the quality of the presentation and the quality of data collection.

**Quality of presentation.**
i) Presents at least 2 sets of measurements in table. ii) Measurements are paired: time and number of pulse beats. iii) Labels table appropriately: data entries in columns identified by headings and/or units; units incorporated into headings or placed beside each measurement; headings or units for the number of pulse beats include the time interval.
*Total Possible Marks: 2*

**Quality of data.**
i) Makes at least 5 measurements (at rest, and 4 or more during exercise). ii) Pulse rates are plausible: 7 to 25 counts per 10 seconds (40-150 pulse beats per minute). iii) Pulse rate increases with exercise (may level off or slow near the end).
*Total Possible Marks: 3*

**Item 2 – Describe how pulse changes during exercise.**
i) Description consistent with data presented. ii) Description includes identification of the trend or pattern in the data.
*Total Possible Marks: 2*

**Item 3 – Explain why pulse changes.**
Includes the following three elements relating to physiological needs during exercise: i) role of muscle action (exercise results in need for more energy and oxygen in the muscles); ii) role of blood (more oxygen or food supplied by an increase in blood flow); iii) connection with heart action or pulse rate, (heart is pumping faster to supply more blood).
*Total Possible Marks: 3*

Table 2.4.2 shows the mean percentage scores for each item within *Pulse* for students in the selected countries. Countries are ranked in order according to the overall percentage score for the task.

**Table 2.4.2 Mean percentage scores for *Pulse* in nine countries**

| Country | Overall mean for task | Mean percentage scores on items within task: Q1 Measure pulse – Presentation | Q1 Measure pulse – Data quality | Q2 Describe trend | Q3 Explain results |
|---|---|---|---|---|---|
| | | *2 marks* | *3 marks* | *2 marks* | *3 marks* |
| Singapore | 60 | 59 | 56 | 82 | 42 |
| **England** | **59** | **65** | **59** | **75** | **39** |
| Scotland | 55 | 61 | 56 | 67 | 34 |
| Switzerland | 51 | 58 | 43 | 75 | 27 |
| United States | 50 | 54 | 43 | 72 | 33 |
| Canada | 46 | 53 | 44 | 60 | 26 |
| Czech Republic | 46 | 45 | 38 | 72 | 27 |
| Netherlands | 45 | 50 | 44 | 56 | 29 |
| Sweden | 45 | 45 | 50 | 62 | 22 |
| **International mean** | **44** | **49** | **41** | **60** | **27** |

*Source: Table 1.1 (Harmon et al., 1997)*

The mean percentages of correct responses achieved by students in England were substantially higher than the international mean, both for the task overall, and for each of the four parts of the task for which marks were allocated. For the two parts of Q1, the mean percentage scores achieved by students in England (65 per cent and 59 per cent respectively) were higher than any of the other countries selected for comparison. For the other two parts, only students in Singapore achieved higher mean scores than those of students in England.

## 2.4.3 *Magnets*

### MAGNETS

**At this station you should have:**

- 6 steel balls
- 10 hair pins or paper clips
- 6 plastic counters
- 2 steel bars
- 10 washers
- 2 magnets
- A 30 cm ruler

**Read ALL directions carefully.**

**Your task:**
Use the things in the bag to find out which magnet is stronger, A or B.

**This is what you should do:**

- Experiment with the things in the bag to complete the sentence below.

1. I found that magnet B is stronger.

**ITEM 1**

**EXAMPLE OF FULL TASK WITH SCORING CRITERIA**

**MAGNETS**

**ITEM 2**

2. Describe all the different ways you tried to find which magnet was stronger. You can draw pictures or diagrams as part of your answer if it helps you to explain.

<table>
<tr><th>What I did</th><th>What happened</th></tr>
<tr><td>For Magnet A and Maggnet B, I tested how many paperclips could be picked up using the magnets. I then repeated this with the washers, steel balls and steel bars.</td><td>I will draw a table of results to show what happened.
<table>
<tr><th></th><th>Magnet A</th><th>Magnet B</th></tr>
<tr><td>Paper clips</td><td>8 were lifted</td><td>10 were lifted</td></tr>
<tr><td>washers</td><td>2 were lifted</td><td>10 were lifted</td></tr>
<tr><td>Steel Balls</td><td>0 were lifted</td><td>6 were lifted</td></tr>
<tr><td>steel Bars</td><td>0 were lifted</td><td>2 were lifted</td></tr>
</table>
For all of the metal objects, magnet B was capable of lifting more than Magnet A. Therefore magnet B is stronger</td></tr>
</table>

**PUT ALL THE MATERIALS BACK IN THE BAG AND LEAVE THE STATION AS YOU FOUND IT.**

## CRITERIA FOR FULLY-CORRECT RESPONSE

**Item 1 – Identify stronger magnet.**
Correct magnet identified according to administrator's notes.
*Total Possible Marks: 1*

**Item 2 – Describe all tests used to identify stronger magnet.**
Includes at least one correct test that: i) includes description or clearly interpretable diagram; ii) shows how results of test were interpreted.
*Total Possible Marks: 1*

Table 2.4.3 shows the mean percentage scores for each item within *Magnets* for students in the selected countries.

**Table 2.4.3 Mean percentage scores for *Magnets* in nine countries**

| Country | Mean percentage scores on items within task | | |
|---|---|---|---|
| | Overall mean for task | Q1 Identify stronger magnet<br>*1 mark* | Q2 Describe strategy<br>*1mark* |
| **England** | **99** | **99** | **99** |
| Scotland | 98 | 99 | 96 |
| Switzerland | 97 | 98 | 97 |
| Singapore | 95 | 98 | 92 |
| Sweden | 95 | 95 | 95 |
| Netherlands | 94 | 96 | 93 |
| Canada | 92 | 95 | 89 |
| Czech Republic | 86 | 86 | 86 |
| United States | 85 | 90 | 81 |
| **International mean** | **90** | **92** | **88** |

*Source: Table 1.3 (Harmon et al., 1997)*

The performance of students in England on this task was higher than that of students in almost all other countries, both for the overall score and for each of the two parts of the task, although between the selected countries the range in the scores achieved was relatively small.

## 2.4.4 *Batteries*

**EXAMPLE OF FULL TASK WITH SCORING CRITERIA**

**BATTERIES**

### BATTERIES

**At this station you should have:**

A torch

Four batteries in a plastic bag: Batteries A, B, C, D

**Read ALL directions carefully.**

> **Your task:**
> Find out which of the batteries are good and which are worn-out.

**This is what you should do:**

- Think about how you could solve this problem.
- Then work out which batteries are good and which are worn-out.

1. Based on your investigation which of the batteries are good and which are worn-out? Write the letters of the batteries in the spaces below.

   Good batteries A, D

   Worn-out batteries B, C

2. Write down how you decided which batteries were worn-out.

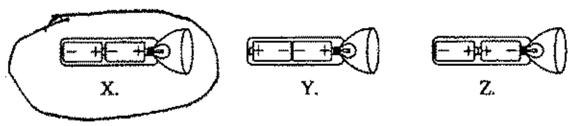
   I tried:

   A, D
   A, C
   A, B
   B, C
   B, D
   C, D

**ITEMS 1 and 2**

**ITEMS 3 and 4**

3. How should the batteries be put in the torch to give the brightest light? Here are 3 different ways of putting the batteries in the torch. Draw a circle around the picture that you think shows the correct way.




4. Why is the way you choose the best way to put in the batteries?

   Because its − then + then − then + so the energy is used in the correct way thus giving the best light.

## CRITERIA FOR FULLY-CORRECT RESPONSE

**Item 1 – Identify which batteries are good and which are worn out.**
All batteries correctly identified (per administrator notes).
*Total Possible Marks: 2*

**Item 2 – Describe how worn-out batteries were identified.**
i) Shows evidence of systematic and definitive testing of different combinations of batteries. ii) "Systematic" is evidenced by trying all combinations of batteries or trying selected combinations using reasoning and scientific knowledge to eliminate some trials.
*Total Possible Marks: 2*

Item 3 – **Identify which arrangement of batteries inside torch will produce the brightest light.**
Correct arrangement identified (X).
*Total Possible Marks: 1*

**Item 4 – Explain why chosen arrangements is the best one.**
i) Identifies correct arrangement. ii) Includes concepts of complete circuit and/or current flowing in one direction.
*Total Possible Marks: 2*

Table 2.4.4 shows the mean percentage scores for each item within *Batteries* for students in the selected countries.

**Table 2.4.4 Mean percentage scores for *Batteries* in nine countries**

| Country | Overall mean for task | Mean percentage scores on items within task: Q1 Identify good/bad batteries (2 marks) | Q2 Describe tests (2 marks) | Q3 Identify arrangement (1 mark) | Q4 Explain arrangement (2 marks) |
|---|---|---|---|---|---|
| Singapore | 79 | 83 | 72 | 98 | 63 |
| **England** | **77** | **89** | **71** | **91** | **56** |
| Switzerland | 75 | 87 | 77 | 94 | 41 |
| Sweden | 71 | 77 | 61 | 90 | 57 |
| Scotland | 68 | 72 | 59 | 94 | 47 |
| Czech Republic | 66 | 76 | 63 | 87 | 39 |
| Netherlands | 63 | 68 | 42 | 93 | 47 |
| Canada | 62 | 66 | 52 | 92 | 38 |
| United States | 56 | 59 | 35 | 97 | 34 |
| **International mean** | **67** | **74** | **59** | **91** | **42** |

*Source: Table 1.5 (Harmon et al., 1997)*

The mean percentage scores for students in England were ten percentage points or more higher than the international means for performance on the task overall, and for three of the four parts of the task. For one part of this task (Q1), the mean percentage score for students in England was higher than that achieved by students in any of the other selected countries.

## 2.4.5 *Rubber Band*

**EXAMPLE OF FULL TASK WITH SCORING CRITERIA**

**RUBBER BAND**

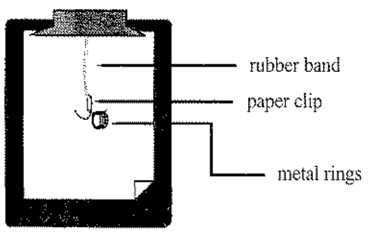
### RUBBER BAND

**At this station you should have:**

A clipboard with a rubber band
A large paper clip attached to one end of the rubber band
Metal rings to hang on the large paper clip
A 30 cm ruler
Some sheets of plain paper
2 sheets of graph or squared paper

**Read ALL directions carefully.**

**Your task:**
Find out how the length of the rubber band changes as more and more rings are hung on it.




**This is what you should do:**

- Hang the metal rings onto the paper clip one by one.
- Measure the length of the rubber band for each new ring.
- Record your measurements in the table on the next page.

**Instructions**

**ITEMS 2 to 4**

1. Write your measurements in the table. Remember to write a heading for each column.

| Number of metal rings | How long the elastic is ? |
|---|---|
| 0 | 3.8 cm |
| 1 | 4.0 cm |
| 2 | 4.2 cm |
| 3 | 4.4 cm |
| 4 | 4.7 cm |
| 5 | 5.2 cm |
| 6 | 5.6 cm |
| 7 | 6.2 cm |
| 8 | 7.5 cm |
| 9 | 8.3 cm |
| 10 | 9.0 cm |

2. Draw a graph to show your results on the paper provided. You may use a graph or a bar chart.

**ANSWER QUESTIONS 3 TO 6, USING YOUR TABLE, GRAPH, OR BAR CHART.**

3. When there are 2 rings on the paper clip and then 3 more are then added, how much longer does the rubber band become?

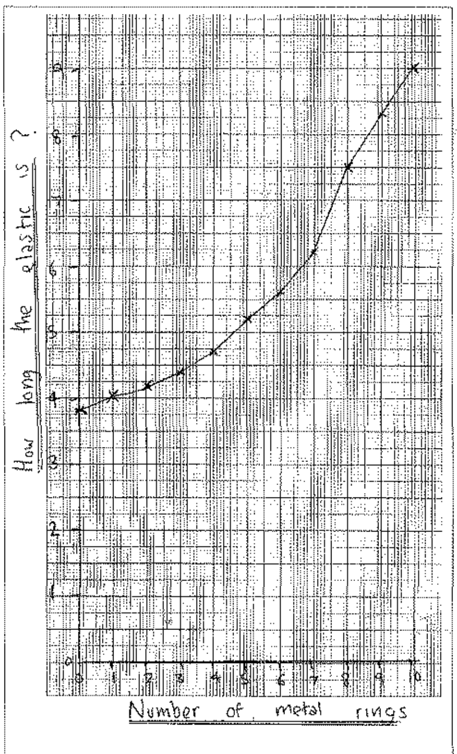
The rubber band becomes ___one___ cm longer.

4. Describe how the rubber band changed in length as more and more rings were added.

The rubber changed length every time another Ring was added. If one wring was added, the elastic stretched to 4.0 cm from 3.8 cm I found out that from no rings put on to 3 rings put on the elastic grew every 0.2 cm each time a ring was put on, Put the length grew after the 4th ring.

Please turn the page.

**Student's work (Item 2)**



5. What do you think would be the length of the rubber band if you could add 2 more rings than you have been given?

I think the total length of the rubber band might be 10·6 cm.

6. Why do you think this would happen?

I think this would happen because each time I put a ring on after the 3rd ring, the elastic grew bigger and bigger every time by at least 0·4 cm. When the 8th ring was put on the elastic grew by 0·7cm, when the 9th ring was put on the elastic grew by 0·8 cm and when the 10th ring was put on the elastic stretched by 0·7cm. So then I thought if the average was 0·8 cm I would put this as the wieght of each of the rings put on, So the end result of the elastic was 10·6 cm.

PUT EVERYTHING BACK THE WAY YOU FOUND IT SO THAT SOMEONE ELSE CAN USE THE STATION.

**ITEMS 5 and 6**

## CRITERIA FOR FULLY-CORRECT RESPONSE

**Item 1 – Record rubber band length as rings are added.**
Response is scored for both the quality of the presentation and the quality of data collection.

**Quality of presentation.** i) Presents at least 2 sets of measurements in table. ii) Measurements are paired: number of rings and length of rubber band. iii) Labels table appropriately: data entries in columns identified by headings and/or units; units incorporated into headings or placed beside each measurement.
*Total Possible Marks: 2*

**Quality of data.** i) Records length of rubber band for five or more different numbers of rings. ii) Shows reasonable trend in data: rubber band length increases with number of rings (at least for first few measurements); length *may* increase steadily at first and then stabilise or level off; elastic limit of rubber band *may* be exceeded and measurements toward the end show very large or erratic increases.
*Total Possible Marks: 3*

**Item 2 – Graph results (graph or bar chart).**
i) Axes correctly scaled. ii) Axes correctly labelled, including units where appropriate. iii) Measurement recorded in graph are consistent with data in table. iv) Graph reflects trend in data.
*Total Possible Marks: 1*

**Item 3 – Calculate increase in length of rubber band when rings are added.**
i) Records amounts consistent with data in table, graph, or bar chart. ii) Calculates increase correctly.
*Total Possible Marks: 2*

**Item 4 – Describe how rubber band length changes as more rings are added.**
i) Description corresponds to data in table or graph. ii) Identifies trend in data. Trend may show that rubber band length increases consistently with each added ring; initially rubber band length increases consistently, then begins to level off; increases become larger or erratic with more rings (elastic limit of band exceeded); no change in length occurs (rubber band too strong for weights, per administrator notes).
*Total Possible Marks: 2*

**Item 5 – Predict increase in length of rubber band.**
Makes reasonable prediction, based on the data presented in the table or graph.
*Total Possible Marks: 1*

**Item 6 – Explain reason for prediction.**
i) Refers to the increase in length as read from the table or extrapolated from graph. ii) Attempts to relate weight or number of rings to elasticity of the rubber band. iii) Response is consistent with data in table or graph.
*Total Possible Marks: 2*

Table 2.4.5 shows the mean percentage scores for each item within *Rubber Band* for students in the selected countries.

**Table 2.4.5 Mean percentage scores for *Rubber Band* in nine countries**

| Country | Mean percentage scores on items within task | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | Overall mean for task | Q1 Measure lengths | | Q2 Graph results | Q3 Calculate increase | Q4 Describe trend | Q5 Predict length | Q6 Explain prediction |
| | | Presen-tation | Data quality | | | | | |
| | | *2 marks* | *3 marks* | *3 marks* | *3 marks* | *2 marks* | *1 mark* | *2 marks* |
| Singapore | 80 | 95 | 99 | 67 | 67 | 87 | 84 | 61 |
| **England** | **79** | **95** | **98** | **76** | **55** | **84** | **80** | **68** |
| Scotland | 75 | 95 | 96 | 69 | 57 | 73 | 78 | 54 |
| Canada | 71 | 87 | 95 | 66 | 55 | 59 | 73 | 59 |
| Netherlands | 70 | 89 | 95 | 71 | 62 | 63 | 53 | 61 |
| Sweden | 70 | 83 | 93 | 55 | 64 | 65 | 72 | 59 |
| Switzerland | 67 | 93 | 93 | 31 | 58 | 73 | 68 | 53 |
| Czech Republic | 65 | 81 | 86 | 44 | 54 | 70 | 66 | 54 |
| United States | 63 | 83 | 88 | 54 | 45 | 68 | 59 | 41 |
| **International mean** | **63** | **85** | **88** | **50** | **47** | **64** | **59** | **49** |

*Source: Table 1.7 (Harmon et al., 1997)*

Again, the performance of students in England on *Rubber Band* was consistently above the international mean for each part of the task. For two of the questions (Q2: graphing the measurements taken during the activity; and Q6: explaining a prediction regarding the length of the rubber band if two additional rings were added), the mean percentage scores achieved by students in England were higher than those of the students in any of the other countries, for Q6 exceeding the international mean by 19 percentage points.

## 2.4.6 *Solutions*

**EXAMPLE OF FULL TASK WITH SCORING CRITERIA**

**SOLUTIONS**

### SOLUTIONS

At this station you should have:

- Hot and cold water
- Several beakers
- Some tablets
- A stirrer
- A clock or watch with a second hand
- A thermometer
- A 30 cm ruler

**Read ALL directions carefully.**

**Your task:**
Investigate what effect different temperatures have on how long it takes a tablet to dissolve.

**This is what you should do:**

- Plan an experiment to find out what effect different temperatures have on how long it takes a tablet to dissolve.

**1.** Write your plan here. Your plan should include

- what you will measure
- how many measurements you will make
- how you will present your measurements in a table.

I will measure how long it takes for a tablet to dissolve in a certain temperature.
I will make 4 measurements, 1 at 60 2 at 40 3 at 20 4 at 10°c
The water will be 200 ml deep. in °C
I will Stir very Slowly, in order to keep the Same ~~Speed~~ Speed
I know when the tablet has Fully dissolved, when there is know more tiny particles ~~so~~ on the Surface.

**ITEM 1**

**ITEMS 2 and 3**

**2.** Carry out your tests on the tablets. Make a table and record all your measurements.

| 1 | 2 | 3 | 4 |
|---|---|---|---|
| 18 Seconds 73 hundreds of a Second At 60°C | 36 Seconds 77 hundreds of a Second At 40°C | 45 Seconds 34 hundreds of a Second At 20°C | 57 Seconds 48 hundreds of a Second At 10°C |

**3.** According to your investigation, what effect do different temperatures have on how long it takes a tablet to dissolve.

If the temperature is higher, the tablet will dissolve more quickly. Between 60°C and 40°C there is a very big difference. So, If you wanted to dissolve these tablets quickly, 60°C is a good operating temperature. As the temperature gets lower, there are not very big gaps between the times.

**Please turn thepage.**

4. Explain why you think different temperatures have this effect.

I think that the hotter the temperature, then more the molecules are excited, they are probably moving faster when they are ~~less~~ hotter. So these excited molecules, try to break up the solid, and try and weave there way in to the gaps in the molecules

5. If you changed your plan while you were working, describe the changes you made and why you made them. If you did not change your plan, write "No change".

No change

**EMPTY YOUR BEAKERS INTO THE WASTE CONTAINER, DRY THEM, AND LEAVE EVERYTHING THE WAY YOU FOUND IT.**

**ITEM 4**

## CRITERIA FOR FULLY-CORRECT RESPONSE

**Item 1 – Plan investigation.**
i) Describes how the investigation will be conducted. ii) States what variables will be measured or observed; includes both solution time and temperature. iii) Provides control for other variables, or renders other variables irrelevant by design.
*Total Possible Marks: 2*

**Item 2 – Conduct investigation and record measurements in table.**
Response is scored for both the quality of the presentation and the quality of the data collection.

**Quality of presentation.**
i) Presents at least 2 sets of measurements in table. ii) Measurements are paired: dissolution time and temperature. iii) Labels table appropriately: data entries in columns identified by headings and/or units; units incorporated into headings or placed beside each measurement.
*Total Possible Marks: 2*

**Quality of data.**
i) Records solution time for at least three temperature points. ii) Measurements are plausible: time and temperature (10° to 100°C) iii) Records solution times that decline as temperature increases.
*Total Possible Marks: 3*

**Item 3 – Draw conclusions about effect of temperature.**
i) Conclusion is consistent with data table or other presentation of data (graph or text). ii) Describes relationship presented in the data.
*Total Possible Marks: 2*

**Item 4 – Explain conclusions.**
i) Relates higher temperature to greater energy or speed of particles (atoms, molecules, etc.). ii) Makes connection between greater speed or energy of water molecules and the effect on the tablet (may be implicit).
*Total Possible Marks: 2*

**Item 5 – Evaluate design and experiment; describe changes.**
i) Response is consistent with the way student recorded and described data ("no change" is acceptable if student plan was complete). ii) changes may be made in method, use of equipment, number of measurements taken, etc; reason for change must be included.
*Total Possible Marks: 2*

Table 2.4.6 shows the mean percentage scores for each item within *Solutions* for students in the selected countries.

**Table 2.4.6 Mean percentage scores for *Solutions* in nine countries**

| Country | Overall mean for task | Mean percentage scores on items within task | | | | | |
|---|---|---|---|---|---|---|---|
| | | Q1 Plan investi-gation | Q2 Conduct investigation | | Q3 Draw conclusion | Q4 Explain conclusion | Q5 Evaluate design |
| | | | Presen-tation | Data quality | | | |
| | | *2 marks* | *2 marks* | *3 marks* | *2 marks* | *2 marks* | *2 marks* |
| **England** | **68** | **66** | **82** | **75** | **89** | **36** | **59** |
| Singapore | 68 | 53 | 91 | 81 | 93 | 42 | 46 |
| Czech Republic | 59 | 60 | 71 | 63 | 86 | 28 | 48 |
| Switzerland | 57 | 55 | 79 | 79 | 85 | 11 | 36 |
| Scotland | 51 | 44 | 72 | 70 | 81 | 18 | 23 |
| Sweden | 50 | 51 | 55 | 65 | 77 | 18 | 34 |
| Canada | 48 | 39 | 64 | 59 | 76 | 26 | 27 |
| United States | 48 | 33 | 64 | 59 | 82 | 27 | 24 |
| Netherlands | 43 | 45 | 46 | 52 | 77 | 12 | 23 |
| **International mean** | **49** | **44** | **62** | **59** | **77** | **22** | **30** |

*Source: Table 1.9 (Harmon et al., 1997)*

The mean percentage score for students in both England and Singapore (the two highest-scoring countries on this task) was 68 per cent for *Solutions*, which was 19 percentage points above the international mean for this task. However, the students in these two countries had different strengths in terms of the questions in which they did particularly well. For example, the students in Singapore outperformed their counterparts in England in two parts of the task concerned with conducting the investigation, whereas the students in England achieved higher scores for planning the investigation and evaluating the design of the investigation.

The performance of the students in England exceeded the international means for individual parts of the task by 12–29 percentage points.

## 2.4.7 *Shadows*

**EXAMPLE OF FULL TASK WITH SCORING CRITERIA**

**SHADOWS**

**ITEMS 1 and 2**

### SHADOWS

**At this station you should have:**

A torch on a stand (this will be called "the light")
5 cm sq. card on a stand
Screen on which to form a shadow of the card
Metre ruler
30 cm ruler

**Read ALL directions carefully.**

> *When the card is between the light and the screen, the card makes a shadow on the screen.*
>
> **Your task:**
> Find out how the size of the shadow changes as you move the card and the light.

**This is what you should do:**

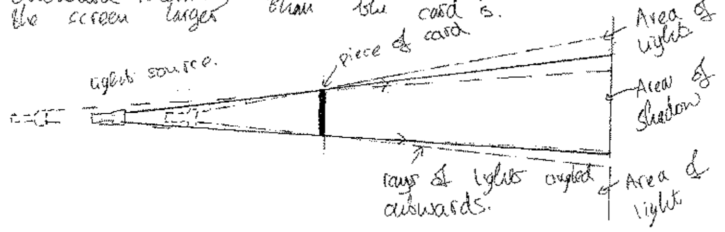
- Keep the card still and move the light closer to the card and further away from the card.

1. What happens to the size of the shadow?

As the light gets closer to the card the shadow size increases. As the light moves away from the carde the shadow size decreases

2. Why is the shadow always larger than the card? You may draw a picture or diagram as part of your answer.

Because the light source is ~~large~~ smaller than the card so therefore the rays of light which pass the card are always angled outward meaning that the shadow will be projected onto the screen larger than the card is.




**ITEMS 3 and 4**

3. Now find at least three positions where you can put the light and the card to make a shadow twice as wide as the card. Record the distance from the card to the screen and from the light to the card for these three positions.

| Position | Distance light ~~away~~ to card | Distance card ~~away~~ to screen |
|---|---|---|
| 1 | ~~50cm~~ 26cm | 2~~8~~4cm |
| 2 | ~~40cm~~ 20cm | 20cm |
| 3. | ~~60cm~~ 30cm | 30cm |

You are now going to do an investigation to try to find a general rule for how far away from the screen the card and the light should be placed to make the shadow twice as wide as the card.

**You will need to:**

- decide what to measure
- decide how to present your measurements clearly and simply
- draw what conclusions you can from your measurements

4. Describe what you did in your investigation. A picture may be useful.

~~The ligt~~ The apparatus were set up as shown in diagram below. The distance the light source was then moved, from 100 cm, nearer to the screen, at 10 cm intervals, and then the card was either moved nearer to ~~the li~~ ~~to~~ or farther from the light until the shadow was 10 cm². ~~This~~ ~~At~~ this point the distance from card to screen was measured and noted, and a table was made of the results The ratio between these two distances will the be worked out and a conclusion ~~drawn~~ drawn up from this.

**Please turn the page.**

**5.** Present your measurements as clearly as you can.

| Distance that light source was from ~~[illegible]~~ screen (cm) | Distance card was from screen when shadow was exactly 10 cm (cm) | Ratio between the (100. ~~[illegible]~~) (distance light : distance card) |
|---|---|---|
| 100 | 50 | = 2:1 |
| 90 | 44 | ≈ 2:1 |
| 80 | 40 | = 2:1 |
| 70 | 35 | = 2:1 |
| 60 | 29 | ≈ 2:1 |
| 50 | 25 | = 2:1 |
| 40 | 20 | = 2:1 |
| 30 | 16 | ≈ 2:1 |
| 20 | 10 | = 2:1 |
| 2 | | |

**6.** What general conclusion can you draw from these results? Try to write a rule that describes when the shadow will always be twice as wide as the card.

~~The light source has to be twice the distance~~

The distance from light source to screen has to be twice the distance from card to screen so as the shadow is twice the size of the card. Therefore the distance from light source to card must equal the distance from card to the screen.

**EMPTY YOUR MATERIALS BACK THE WAY YOU FOUND THEM SO THAT SOMEONE ELSE CAN USE THIS STATION.**

**ITEMS 5 and 6**

## CRITERIA FOR FULLY-CORRECT RESPONSE

**Item 1 – Describe how shadow size changes in response to distance of light.**
i) Comments appropriately on the size of shadow. ii) Comments on the relationship between the distance from light and size of shadow.
*Total Possible Marks: 2*

**Item 2- Explain why shadow is larger than card.**
i) Includes concept of light travelling in a straight line and spreading out from a source. ii) Explanation or diagram shows how the shadow is formed.
*Total Possible Marks: 2*

**Item 3 – Record distances for three positions where shadow is twice as large as card.**
i) Records at least 3 measurements where shadow is twice the size of card. ii) Measurements are paired: distance from light to card and distance from card to screen. iii) Measurements are plausible: the distance from card to screen and distance from light to card are equal (within ± 10%).
*Total Possible Marks: 2*

**Item 4 – Describe investigation.**
i) Includes description of how measurements were taken. ii) Includes taking measurements of both distances and shadow width.
*Total Possible Marks: 2*

**Item 5 – Present measurements.**
i) Measurements presented in a list, table or by graph. ii) Measurements are clearly and completely understandable with appropriate units, labels, and descriptors.
*Total Possible Marks: 2*

**Item 6 – Write a general rule to describe when shadow will always be twice as wide as card.**
i) Summarises data in sentences, formula, or diagram. ii) Indicates that shadow will be twice as wide as card when the distance from light to screen is twice the distance from light to card.
*Total Possible Marks: 2*

Although this task involved the application of some mathematical skills (in taking measurements, and, in the last question, determining a mathematical relationship), the main focus was a science investigation; hence the results for this task are included with the other science tasks, although the task was originally classified as covering both subjects.

Table 2.4.7 shows the mean percentage scores for each item within *Shadows* for students in the selected countries.

**Table 2.4.7 Mean percentage scores for *Shadows*[3] in nine countries**

| Country | Overall mean for task | Mean percentage scores on items within task | | | | | |
|---|---|---|---|---|---|---|---|
| | | Q1 Describe observ-ation | Q2 Explain observ-ation | Q3 Problem solve & record distances | Q4 Describe investig-ation | Q5 Present measure-ments | Q6 Conclude and generalise |
| | | *2 marks* | *2 marks* | *2 marks* | *2 marks* | *2 marks* | *2 marks* |
| Singapore | 50 | 90 | 55 | 41 | 39 | 46 | 29 |
| **England** | **46** | **77** | **33** | **23** | **47** | **71** | **23** |
| Sweden | 45 | 82 | 43 | 57 | 30 | 27 | 32 |
| Switzerland | 41 | 80 | 44 | 43 | 29 | 22 | 32 |
| Czech Republic | 37 | 87 | 48 | 32 | 27 | 8 | 19 |
| Scotland | 36 | 83 | 24 | 31 | 28 | 36 | 16 |
| Canada | 35 | 75 | 21 | 34 | 30 | 28 | 19 |
| Netherlands | 35 | 55 | 50 | 33 | 27 | 25 | 23 |
| United States | 28 | 64 | 20 | 13 | 27 | 34 | 11 |
| **International mean** | **35** | **75** | **33** | **30** | **27** | **25** | **21** |

*Source: Table 1.11 (Harmon et al., 1997)*

The performance of students in England was in line with or higher than the international mean for each part of *Shadows*, with the exception of Q3 (in which students were asked to find at least three positions where the shadow of the card was twice its actual width and to record the appropriate measurements). For Q5 (which asked students to present the measurements they had taken during the investigation), the mean percentage score for students in England (71 per cent) was higher than students in all other countries, and was almost three times as large as the international mean of 25 per cent.

[3] When considering the results for *Shadows*, it should be borne in mind that, in some schools, it was difficult for students to carry out the activity when there were inadequate blackout facilities. It is worth noting that students in all countries achieved higher scores for the first question than for any of the subsequent questions on this task, which might suggest that many students experienced difficulties with the taking (and, consequently, the interpretation) of accurate measurements relating to the size of shadows. One area of possible difficulty relating to this activity was whether or not students included the penumbra in their measurements.

# 2.5 Gender differences

## 2.5.1 Gender differences in overall performance

There were no significant gender differences in overall performance on the 12 tasks in England or in any of the other countries selected for comparisons, in common with the majority of countries that participated in the Performance Assessment.

## 2.5.2 Gender differences in mathematics

There were no significant gender differences in the performance of boys and girls in England on the six mathematics tasks, as in most other countries that took part in the Performance Assessment. However, there were statistically significant gender differences (i.e. statistically significant[4] at the 0.05 level) in two of the other countries selected for comparison:

- Singapore: a difference in favour of boys on *Around the Bend*
- Sweden: a difference in favour of girls on *Packaging*.

The cumulative effect of slight gender differences in two tasks (*Dice* and *Around the Bend*) resulted in a significant gender difference in the international means for these tasks, for *Dice* in favour of girls, and for *Around the Bend* in favour of boys.

Table 2.5.2 shows performance of boys and girls in England on the mathematics tasks together with the international means. Where there were significant gender differences, the numbers shown in bold type indicate the gender with the higher performance.

**Table 2.5.2 Overview of gender differences[5] in mean percentage scores for mathematics tasks for students in England**

| | Mean percentage scores on mathematics tasks | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | *Dice* | | *Calculator* | | *Folding and Cutting* | | *Around the Bend* | | *Packaging* | | *Plasticine* | |
| Country | Boys | Girls | Boys | Girls | Boys | Girls | Boys | Girls | Boys | Girls | Boys | Girls |
| England | 77 | 81 | 60 | 64 | 71 | 67 | 64 | 61 | 52 | 55 | 56 | 54 |
| International | 72 | **75** | 54 | 54 | 69 | 68 | **56** | 53 | 44 | 44 | 61 | 59 |

*Source: Table 2.5 (Harmon et al., 1997)*

[4] There is always a possibility that the given difference between two samples (in this case, boys and girls) could have arisen by chance. The level of statistical significance gives an indication of how likely this is. 0.05 is generally taken as the cut-off point between differences that are thought to be real and those which could have arisen by chance, i.e. a difference of this size is only likely to have arisen by chance five times out of one hundred.

[5] It should be noted that whilst a difference of, say, three or four percentage points in the international means may be statistically significant, a similar or larger difference in the performance of boys and girls in England may *not* be statistically significant. This is because the international mean is based on a much larger number of students, and, as a result, even small gender differences in performance may be statistically significant.

### 2.5.3 Gender differences in science

Although there were minor differences in the performance of boys and girls on the six science tasks, none of these was statistically significant either in England or in any of the selected countries, in common with the majority of countries that participated in the Performance Assessment. However, the cumulative effect of slight gender differences in one task (*Rubber Band*) produced a significant gender difference in the international mean for this task in favour of girls.

Table 2.5.3 shows the performance of boys and girls in England on the science tasks together with the international means. Where there were significant gender differences (i.e. statistically significant at the 0.05 level), numbers are shown in bold type.

**Table 2.5.3 Overview of gender differences[6] in mean percentage scores for science tasks for students in England**

| | Mean percentage scores on science tasks | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | *Pulse* | | *Magnets* | | *Batteries* | | *Rubber Band* | | *Solutions* | | *Shadows* | |
| Country | Boys | Girls | Boys | Girls | Boys | Girls | Boys | Girls | Boys | Girls | Boys | Girls |
| England | 58 | 60 | 99 | 99 | 77 | 77 | 78 | 81 | 64 | 71 | 45 | 47 |
| International | 44 | 44 | 91 | 89 | 68 | 65 | 61 | **65** | 48 | 51 | 36 | 34 |

*Source: Table 2.5 (Harmon et al., 1997)*

[6] It should be noted that whilst a difference of, say, three or four percentage points in the international means may be statistically significant, a similar or larger difference in the performance of boys and girls in England may *not* be statistically significant. This is because the international mean is based on a much larger number of students, and, as a result, even small gender differences in performance may be statistically significant.

## 2.6 Correlations between students' performance on the written tests and the Performance Assessment tasks

All the students that took part in the Performance Assessment component of TIMSS had already completed written tests in mathematics and science as part of the main survey. It is therefore of interest to determine to what extent students' results on the practical tasks are correlated with their results on the written tests. Correlation coefficients (Spearman) have been calculated between students' raw scores on each of the 12 tasks and the mathematics and science scores achieved by the students that attempted each of those tasks: these are shown in Table 2.6.1. The correlations between students' scores on the practical and the written components are all highly significant ($P<0.001$ i.e. the probability of these correlations occurring by chance is fewer than one in a thousand).

**Table 2.6.1 Correlation coefficients between each of the 12 tasks and mean mathematics and science scores on written tests**

| Task | Correlation coefficients | |
|---|---|---|
| | Mathematics | Science |
| *Dice* | 0.54 | 0.36 |
| *Calculator* | 0.55 | 0.40 |
| *Folding and Cutting* | 0.48 | 0.51 |
| *Around the Bend* | 0.53 | 0.51 |
| *Packaging* | 0.41 | 0.43 |
| *Plasticine* | 0.51 | 0.33 |
| *Pulse* | 0.45 | 0.47 |
| *Magnets* | 0.22 | 0.22 |
| *Batteries* | 0.52 | 0.42 |
| *Rubber Band* | 0.57 | 0.43 |
| *Solutions* | 0.47 | 0.48 |
| *Shadows* | 0.53 | 0.46 |

The correlation coefficients vary between tasks and also in the correlations between particular tasks and mathematics and science scores on the written tests. Four of the mathematics tasks (*Dice, Calculator, Around the Bend* and *Plasticine*) showed strong correlations (more than 0.5) with students' scores on the written mathematics tests; this finding is not unexpected, since one would expect students who do well on written mathematics tests also to do well in practical mathematics tasks. However, there were also strong correlations between three of the science tasks and students' scores on the written mathematics tests

(*Batteries*, *Rubber Band* and *Shadows*). Each of these tasks involved the application of mathematical skills in some items (e.g. in *Shadows* one item asked students to find at least three positions for the light and the card that would make the shadow of the card twice its actual size; and in *Batteries* students had to test all possible combinations of four batteries so as to find which two were good and which were flat). There was only a weak correlation between students' scores on *Magnets* and their scores on the written mathematics and science tests; this was because 67 per cent of the students who attempted the task achieved the maximum marks for the task (four marks).

These correlations should be set against the context of the correlation of 0.73 between the scores on the written mathematics and science tests achieved by all Year 9 students. In essence, this means that there is a stronger correlation between students' results on the written tests in both subjects than between their results on the Performance Assessment tasks and the written tests.

Additional analysis has been carried out to calculate the means achieved by students of different abilities on each of the 12 tasks. In doing this, the students have been grouped into three main ability bands according to their scores on the written mathematics tests, the science tests, and their overall results on the written tests in both subjects. These bands consist of the students whose scores place them in the lowest, middle and top thirds of performance on the aforementioned tests. The mean score on each task has been calculated for students in each of the ability bands for mathematics, science and overall results; these are shown in Table 2.6.2, together with the maximum possible score for each task.

The results shown in Table 2.6.2 are consistent with what one would expect, that the mean scores on each task are higher for the students in the top third ability band (according to results on the written tests) than for students in the middle band, who, in turn, achieved higher mean scores than those in the lowest ability band.

Table 2.6.2 also shows standard deviations; the standard deviation (SD) is a useful statistical measure for indicating the spread in a set of observations, in this instance the scores achieved by students in different ability bands on different tasks. About ninety-five per cent of cases in most datasets fall within the following range: the mean, plus or minus two standard deviations. The standard deviations shown for each task/ability band are an indication of the range of the scores achieved by students on the Performance Assessment tasks: for example, if we look at the results for *Calculator* for the lowest overall ability band, 95 per cent of students achieved the mean for that task plus or minus two SDs (i.e. 95 per cent scored between 2.2 and 14.2 out of a possible 16 marks). A larger standard deviation indicates that the scores were more spread out, whereas a smaller standard deviation indicates that the scores were more grouped together.

**Table 2.6.2 Mean scores on each of the 12 tasks achieved by students in different ability bands (according to their results on the written mathematics and science tests)**

| Task | Overall Ability Band | | | Maths Ability Band | | | Science Ability Band | | |
|---|---|---|---|---|---|---|---|---|---|
| | Bottom third | Middle third | Top third | Bottom third | Middle third | Top third | Bottom third | Middle third | Top third |
| ***Dice*** | | | | | | | | | |
| *(max. 9 marks)* Mean | 6.4 | 7.6 | 8.1 | 6.3 | 7.5 | 8.3 | 6.7 | 7.5 | 7.8 |
| Standard Deviation | 2.0 | 1.1 | 1.0 | 2.0 | 1.1 | 0.9 | 2.2 | 1.2 | 1.2 |
| *Unweighted Valid N* | *45* | *50* | *50* | *46* | *51* | *48* | *46* | *52* | *47* |
| ***Calculator*** | | | | | | | | | |
| *(max. 16 marks)* Mean | 8.2 | 10.5 | 12.0 | 8.1 | 10.3 | 12.1 | 8.1 | 10.5 | 11.8 |
| Standard Deviation | 3.0 | 2.7 | 2.4 | 2.7 | 3.1 | 2.3 | 3.1 | 2.9 | 2.3 |
| *Unweighted Valid N* | *47* | *53* | *47* | *45* | *52* | *50* | *42* | *59* | *46* |
| ***Folding and Cutting*** | | | | | | | | | |
| *(max. 9 marks)* Mean | 4.5 | 6.7 | 7.5 | 4.6 | 6.6 | 7.4 | 4.5 | 6.4 | 7.7 |
| Standard Deviation | 2.8 | 2.2 | 1.7 | 2.9 | 2.4 | 1.8 | 2.8 | 2.4 | 1.6 |
| *Unweighted Valid N* | *53* | *41* | *50* | *52* | *41* | *51* | *51* | *45* | *48* |
| ***Around the Bend*** | | | | | | | | | |
| *(max. 19 marks)* Mean | 8.9 | 12.1 | 13.3 | 9.0 | 11.9 | 13.1 | 9.4 | 11.3 | 13.4 |
| Standard Deviation | 3.6 | 2.8 | 2.7 | 3.6 | 2.7 | 2.9 | 3.8 | 3.3 | 2.2 |
| *Unweighted Valid N* | *47* | *48* | *51* | *48* | *42* | *56* | *43* | *51* | *52* |
| ***Packaging*** | | | | | | | | | |
| *(max. 7 marks)* Mean | 2.3 | 3.2 | 4.1 | 2.4 | 3.4 | 3.9 | 2.5 | 3.0 | 4.1 |
| Standard Deviation | 1.8 | 1.4 | 1.3 | 1.8 | 1.5 | 1.4 | 1.9 | 1.5 | 1.3 |
| *Unweighted Valid N* | *50* | *49* | *48* | *51* | *45* | *51* | *50* | *46* | *51* |
| ***Plasticine*** | | | | | | | | | |
| *(max. 14 marks)* Mean | 5.7 | 6.3 | 9.7 | 5.0 | 6.8 | 10.0 | 6.2 | 7.1 | 8.5 |
| Standard Deviation | 3.1 | 3.8 | 4.0 | 2.8 | 3.6 | 4.0 | 3.7 | 3.7 | 4.3 |
| *Unweighted Valid N* | *50* | *46* | *50* | *47* | *50* | *49* | *53* | *41* | *52* |
| ***Pulse*** | | | | | | | | | |
| *(max. 10 marks)* Mean | 4.4 | 5.9 | 6.8 | 4.3 | 6.0 | 6.9 | 4.7 | 5.5 | 7.0 |
| Standard Deviation | 2.5 | 2.4 | 2.6 | 2.3 | 2.5 | 2.7 | 2.6 | 2.5 | 2.5 |
| *Unweighted Valid N* | *45* | *50* | *50* | *46* | *51* | *48* | *46* | *52* | *47* |
| ***Magnets*** | | | | | | | | | |
| *(max. 4 marks)* Mean | 3.2 | 3.4 | 3.7 | 3.2 | 3.4 | 3.7 | 3.2 | 3.5 | 3.7 |
| Standard Deviation | 0.9 | 0.9 | 0.6 | 1.0 | 0.8 | 0.7 | 0.9 | 0.8 | 0.8 |
| *Unweighted Valid N* | *47* | *53* | *47* | *45* | *52* | *50* | *42* | *59* | *46* |
| ***Batteries*** | | | | | | | | | |
| *(max. 7 marks)* Mean | 4.4 | 5.5 | 6.0 | 4.3 | 5.4 | 6.2 | 4.5 | 5.3 | 5.9 |
| Standard Deviation | 1.9 | 1.2 | 1.2 | 1.9 | 1.3 | 1.0 | 1.9 | 1.5 | 1.2 |
| *Unweighted Valid N* | *55* | *42* | *50* | *55* | *41* | *51* | *53* | *45* | *49* |
| ***Rubber Band*** | | | | | | | | | |
| *(max. 15 marks)* Mean | 10.6 | 12.1 | 13.4 | 10.4 | 12.4 | 13.3 | 10.9 | 12.3 | 13.2 |
| Standard Deviation | 2.6 | 1.9 | 1.2 | 2.5 | 1.7 | 1.5 | 2.5 | 2.1 | 1.2 |
| *Unweighted Valid N* | *48* | *48* | *53* | *47* | *48* | *54* | *53* | *46* | *50* |
| ***Solutions*** | | | | | | | | | |
| *(max. 14 marks)* Mean | 7.6 | 10.3 | 11.2 | 8.0 | 10.2 | 11.4 | 7.9 | 9.9 | 11.2 |
| Standard Deviation | 2.6 | 2.9 | 2.5 | 2.8 | 3.0 | 2.3 | 3.1 | 2.9 | 2.3 |
| *Unweighted Valid N* | *39* | *55* | *53* | *50* | *45* | *52* | *38* | *51* | *58* |
| ***Shadows*** | | | | | | | | | |
| *(max. 12 marks)* Mean | 4.0 | 5.5 | 7.1 | 4.2 | 4.9 | 7.5 | 4.1 | 5.6 | 6.8 |
| Standard Deviation | 2.0 | 2.1 | 2.4 | 2.3 | 2.0 | 2.1 | 2.1 | 2.4 | 2.4 |
| *Unweighted Valid N* | *51* | *44* | *51* | *43* | *55* | *48* | *54* | *41* | *51* |

The standard deviations for *Magnets* indicate that the scores were not spread over a large range (the mean percentage score for students in England on this task was 99 per cent (see section 2.4.3), and, as noted above, 67 per cent of students achieved maximum marks on this task), whereas the largest standard deviations were to be found in the results of students in the top ability bands on *Plasticine*.

Most of the correlations between the students' results on the Performance Assessment tasks and on the written tests in mathematics and science were between 0.4 and 0.55, which suggests that although there was some overlap of skills tested, these two aspects of TIMSS also each tested additional skills and/or knowledge which the other component did not. Finally, it should be remembered that fewer students participated in the Performance Assessment than in the TIMSS written tests, and the numbers of students who attempted any one task were relatively small (approximately 50). Consequently, it would be unwise to place too much emphasis on the correlations identified in this analysis.

# CHAPTER 3
# Students' Performance in the Mathematics and Science Skills Areas

## 3.1 Preface

The previous chapter presented details of the performance of students in England on the 12 tasks that comprised the Performance Assessment component of TIMSS. However, in addition to collecting information about students' achievement on each task as a whole, a major part of the study design was concerned with determining students' performance on specific items within tasks which required particular skills. Consequently, as an alternative to presenting students' performance on the 12 tasks which included different skills within them, it is possible to consider students' performance in different skills areas (for example, performing mathematical procedures) across several tasks. The different skills areas were defined in the curriculum frameworks which apply to both the written and the practical components of TIMSS (Robitaille, 1993). The curriculum frameworks specify the range of mathematics and science skills (identified as Performance Expectations) addressed by the Performance Assessment component of TIMSS. The skills covered are quite similar to the process-focused parts of the National Curriculum Orders for mathematics and science (Ma1: Using and Applying Mathematics and Sc1: Experimental and Investigative Science). The particular skills addressed by items within the tasks are shown for mathematics in Annexes 2 and 3, sorted by skills areas and tasks respectively, and are shown similarly for science in Annexes 4 and 5.

Within the international report, the items covering different mathematics skills have been grouped into two main categories:

- *PERFORMING MATHEMATICAL PROCEDURES*
- *PROBLEM SOLVING AND MATHEMATICAL REASONING.*

The items covering different science skills have been grouped into three main categories:

- *SCIENTIFIC PROBLEM SOLVING AND APPLYING CONCEPT KNOWLEDGE*
- *USING SCIENTIFIC PROCEDURES*
- *SCIENTIFIC INVESTIGATING.*

Two main sections of this chapter cover the mathematics skills categories and the science skills categories respectively. More specifically, this chapter presents details of:

- the particular items within tasks that are classified under each of the skills categories
- the performance of students in England on each of the items within the two mathematics skills categories and the three science skills categories
- the international means for comparative purposes.

It should be noted that although the international classification allocates items within most tasks to either the mathematics skills categories or the science skills categories (reflecting the overall content of the task), items within *Plasticine* are included in both the mathematics and science skills categories.[1]

## 3.2 Mathematics skills

The range of skills addressed within the six mathematics tasks have been grouped into two main reporting categories in the international report (Harmon *et al.*, 1997). These categories, and the types of activity[2] that they represent, are shown below.

**MATHEMATICS**

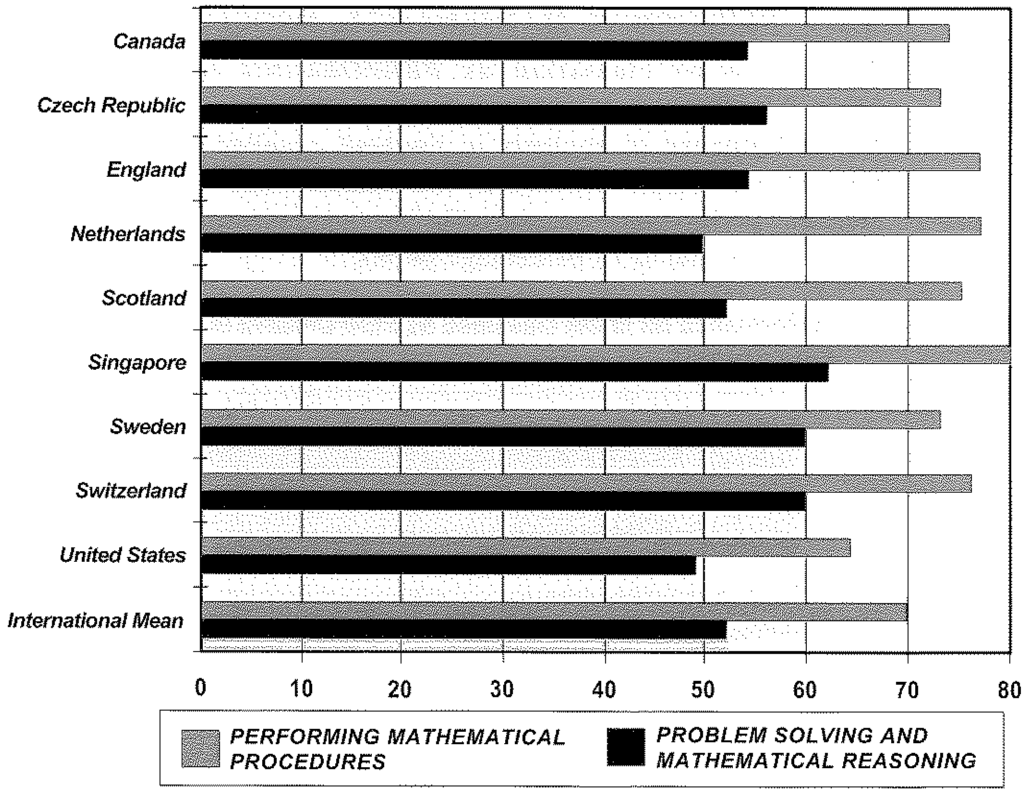
| ***PERFORMING MATHEMATICAL PROCEDURES*** (13 Items) | ***PROBLEM SOLVING AND MATHEMATICAL REASONING*** (21 Items) |
|---|---|
| • Using equipment | • Developing strategy |
| • Performing routine procedures | • Solving problems |
| • Using more complex procedures | • Predicting |
| | • Generalising |
| | • Conjecturing |

[1] The assignment of items to the science and mathematics skills categories is based on the primary performance expectation associated with each item, although each item has a number of associated sub-constructs. For example, in order to draw conclusions about which is the stronger of two magnets (a primary performance expectation classified under the category *SCIENTIFIC INVESTIGATING*), a student would have to demonstrate capability in two sub-constructs: (a) use equipment and (b) carry out experimental operations. Two items are not assigned to one of the main skills categories listed above: *Shadows* Q4 and *Plasticine* Q1B. In grouping items into the categories listed above, the International Study Center omitted these two items since their primary performance expectation category was communicating (i.e. not uniquely related to mathematics and/or science).

[2] The types of activity listed under the main reporting categories are the skills sub-categories identified in the international report (Harmon *et al.*, 1997). The complete classification of skills, as determined by the International Study Center, covering all the mathematics and science tasks, is shown in Annexes 2-5 to this report.

The relative performance of students in the nine selected countries on the two main mathematics reporting categories[3] is shown in Figure 3.2, together with the international means for comparative purposes.

**Figure 3.2: Mean percentage scores by mathematics skills categories**




*Source: Table 3.3 (Harmon et al., 1997).*

It is evident from Figure 3.2 that the performance of students in England, in common with their counterparts in all other countries that participated in the Performance Assessment, was higher in *PERFORMING MATHEMATICAL PROCEDURES* than in *PROBLEM SOLVING AND MATHEMATICAL REASONING*. Typically, students' scores in *PERFORMING MATHEMATICAL PROCEDURES* were approximately 20 percentage points higher than in *PROBLEM SOLVING AND MATHEMATICAL REASONING.* Within the selected countries, the differences ranged from 13 percentage points (Sweden) to 27 percentage points (The Netherlands), with the difference in England (23 percentage points) close to the international mean of 18 percentage points.

It is worth noting that the performance of students in England exceeded the international means for both mathematics skills categories, which may be unexpected given the rather disappointing results for mathematics in the TIMSS main survey (Keys *et al.*, 1996a), in which 13-year-olds scored below the international means for five out of six mathematics content areas.

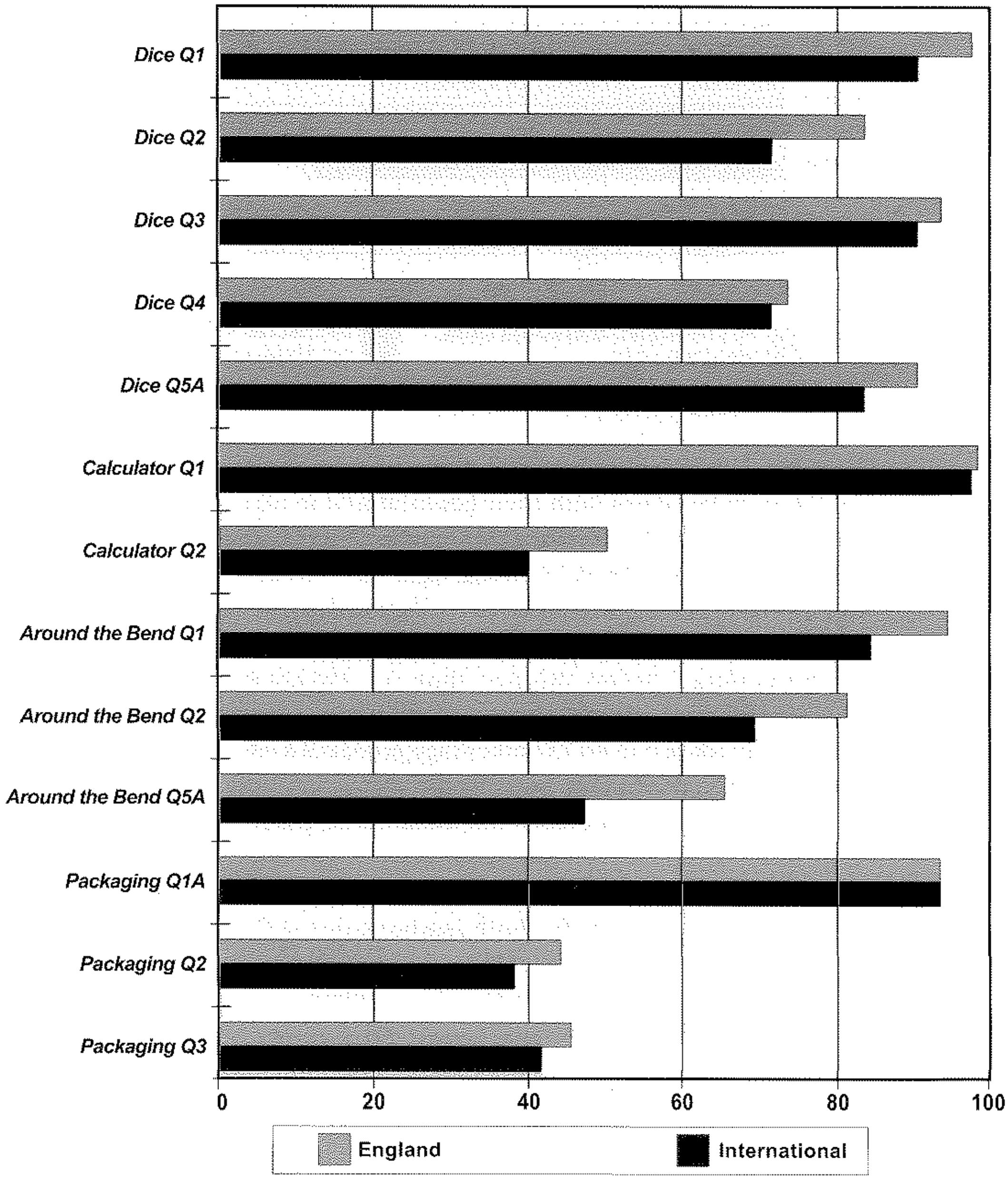
[3] Bars show the percentage scores averaged across items in each skills category (±2SE); items were weighted equally.

Within the next part of the chapter, figures show the performance of students in England on each of the two main mathematics categories listed above, together with international means for comparative purposes.

### 3.2.1 *PERFORMING MATHEMATICAL PROCEDURES*

The 13 items classified as *PERFORMING MATHEMATICAL PROCEDURES* were drawn from five tasks. As shown in Figure 3.2.1, the performance of students in England equalled or exceeded the international mean for each of the items in this category. Where students' performance was above the international mean, the differences ranged from one to 18 percentage points.

**Figure 3.2.1: Performance on items classified as *PERFORMING MATHEMATICAL PROCEDURES***

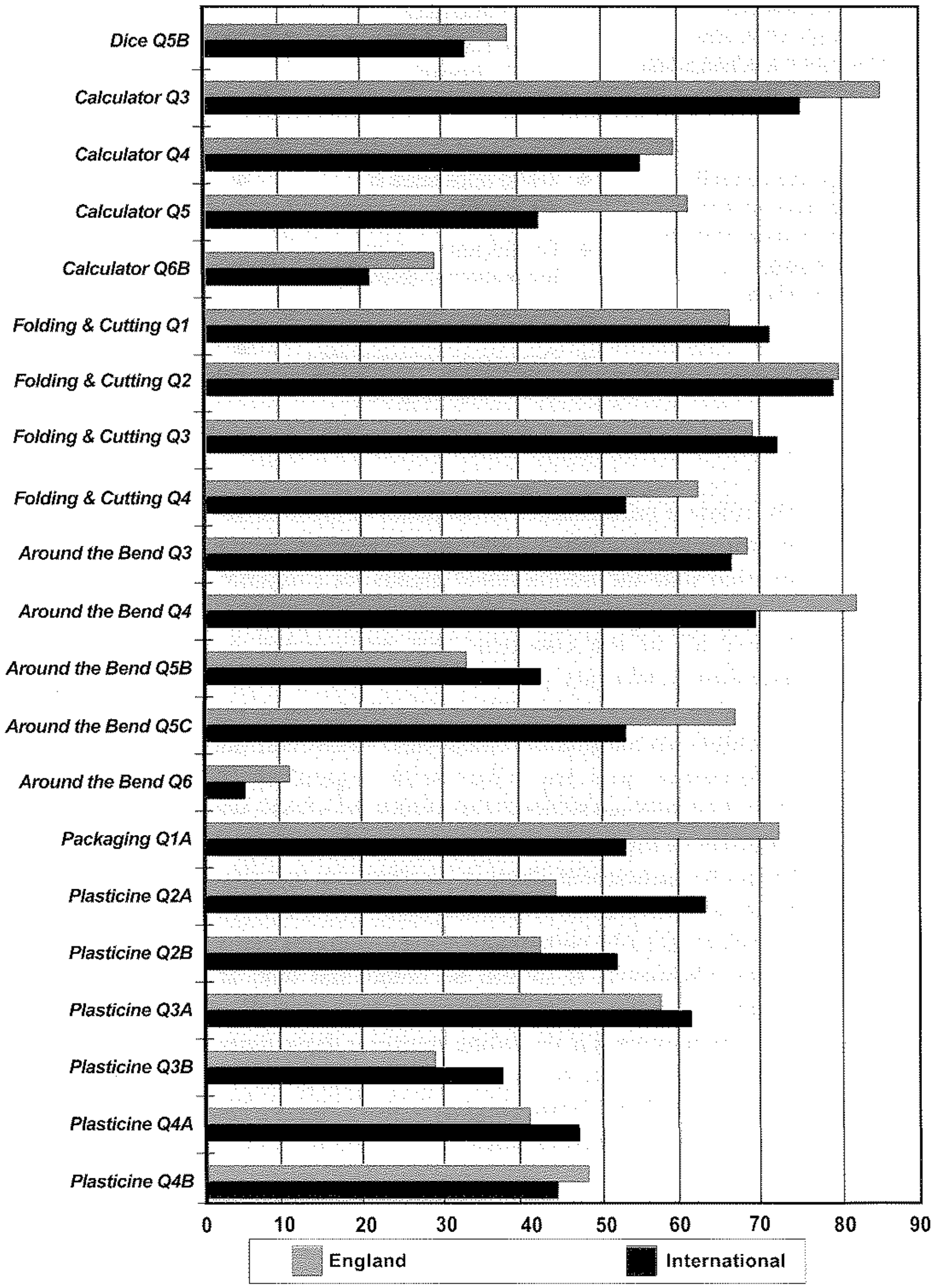



*Source: Tables 1.15, 1.17, 1.21, 1.23 and 1.13 (Harmon et al., 1997)*

### 3.2.2 *PROBLEM SOLVING AND MATHEMATICAL REASONING*

The 21 items classified as *PROBLEM SOLVING AND MATHEMATICAL REASONING* were drawn from all six mathematics tasks. The performance of students in England varied considerably on the items within this category, where performance ranged from 19 percentage points below the international mean (*Plasticine* Q2A) to 19 percentage points above the international mean (*Packaging* Q1 and *Calculator* Q5). Figure 3.2.2 shows students' performance on each of the items in this category.

**Figure 3.2.2: Performance on items classified as *PROBLEM SOLVING AND MATHEMATICAL REASONING***




*Source: Tables 1.15, 1.17, 1.19, 1.21, 1.23 and 1.13 (Harmon et al., 1997)*

## 3.3 Science skills

Like the mathematics items, the science items can be grouped into reporting categories representing different skills areas. The international report (Harmon *et al.*, 1997) shows three main reporting categories for science, each with a number of sub-categories as shown below.

**SCIENCE**

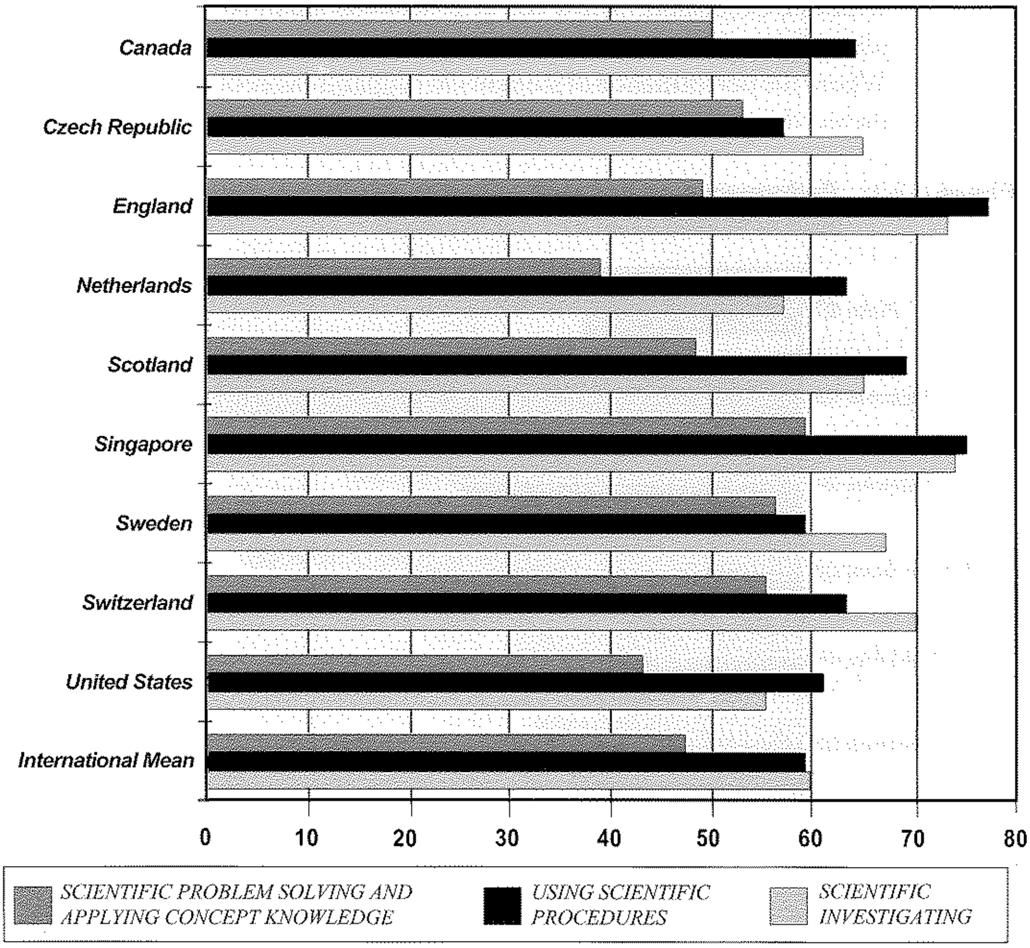
| *SCIENTIFIC PROBLEM SOLVING AND APPLYING CONCEPT KNOWLEDGE* (12 Items) | *USING SCIENTIFIC PROCEDURES* (7 Items) | *SCIENTIFIC INVESTIGATING* (16 Items) |
|---|---|---|
| • Applying scientific principles to solve quantitative problems or develop explanations | • Using apparatus or equipment<br>• Conducting routine experimental operations<br>• Gathering data<br>• Organising, representing and interpreting data | • Designing and conducting investigations<br>• Interpreting investigational data<br>• Formulating conclusions from investigational data |

The relative performance of students in the nine selected countries on the three main science reporting categories[4] is shown in Figure 3.3, together with the international means for comparative purposes.

Across all of the countries that participated in the Performance Assessment, students' performance in *SCIENTIFIC PROBLEM SOLVING AND APPLYING CONCEPT KNOWLEDGE* was at a lower level than for the other two main skills categories. However, the overall pattern of performance was less clear with regard to *USING SCIENTIFIC PROCEDURES* and *SCIENTIFIC INVESTIGATING*, with students in some countries doing better in the former and in others, better in the latter. The performance of students in England was just above the international mean in *SCIENTIFIC PROBLEM SOLVING AND APPLYING CONCEPT KNOWLEDGE* (49 percentage points as compared with 47 percentage points); however, it should be noted that half of the items in this category were derived from the mathematics task *Plasticine*. Irrespective of the arguments for including/omitting these items from this science category, it must be acknowledged that the performance of students in England would have been higher in this category had the items derived from *Plasticine* been omitted.

[4] Bars show the percentage scores averaged across items in each skills category (±2SE); items were weighted equally.

**Figure 3.3: Mean percentage scores by science skills categories**



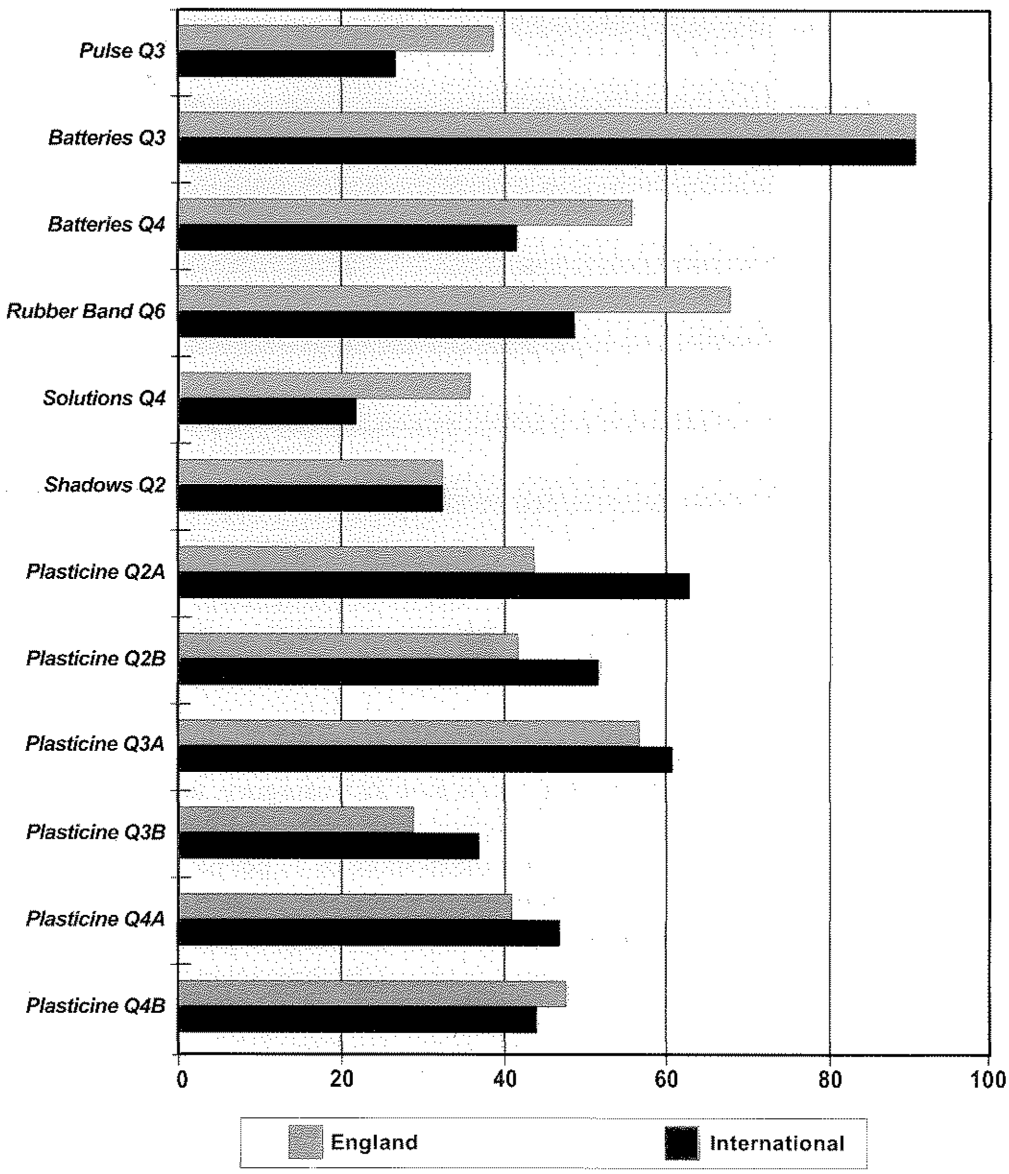

*Source: Table 3.1 (Harmon et al., 1997)*

In the *USING SCIENTIFIC PROCEDURES* category, the performance of students in England was higher than that of students in any other country that took part, with a mean score of 77 per cent. In *SCIENTIFIC INVESTIGATING*, only students in Singapore outperformed students in England, with mean scores of 74 per cent and 73 per cent respectively. This exceptional performance in two of the three science skills categories suggests that practical and investigative science is an area of the curriculum in which students in England are doing particularly well.

### 3.3.1 *SCIENTIFIC PROBLEM SOLVING AND APPLYING CONCEPT KNOWLEDGE*

The 12 items classified as *SCIENTIFIC PROBLEM SOLVING AND APPLYING CONCEPT KNOWLEDGE* were drawn from six tasks. It is worth remembering that six of the 12 items are from *Plasticine*, a task in which the main focus is on mathematics rather than science. The performance of students in England was equal to, or above, the international mean for the six items that form part of science tasks, but, with one exception, below the international mean for each of the items identified within the *Plasticine* mathematics task (see Figure 3.3.1). Where the performance of students in England exceeded the international mean, the difference ranged from four to 19 percentage points.

**Figure 3.3.1: Performance on items classified as *SCIENTIFIC PROBLEM SOLVING AND APPLYING CONCEPT KNOWLEDGE***

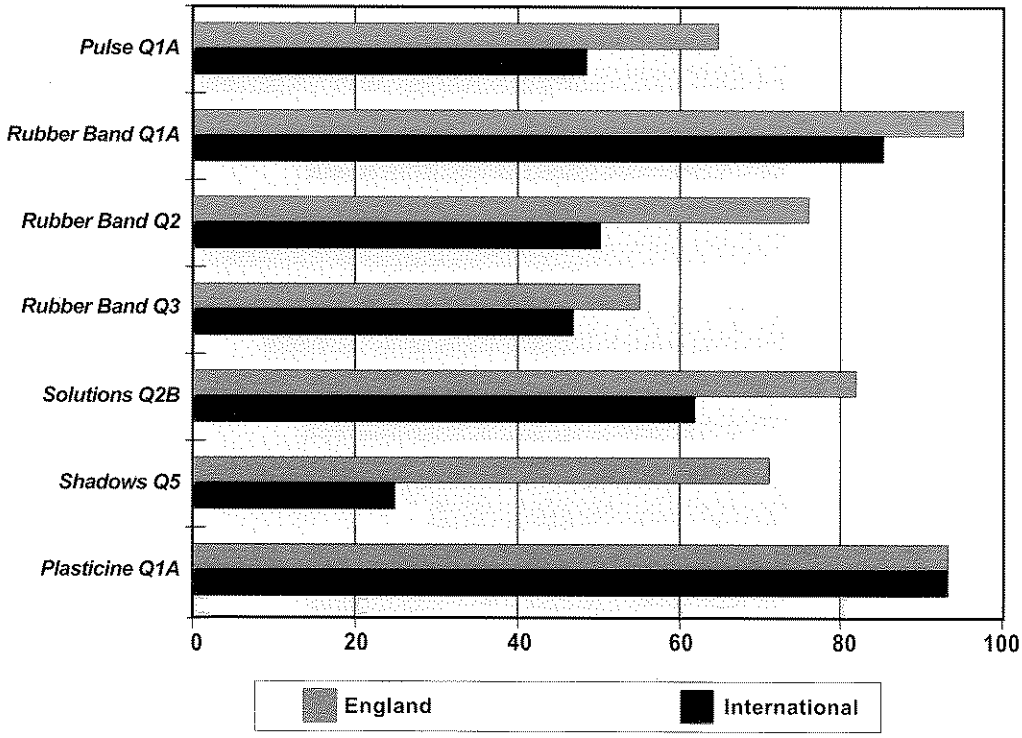



*Source: Tables 1.1, 1.5, 1.7, 1.9, 1.11 and 1.13 (Harmon et al., 1997)*

### 3.3.2 *USING SCIENTIFIC PROCEDURES*

The seven items classified as *USING SCIENTIFIC PROCEDURES* were drawn from five tasks. For each of these items, the performance of students in England either equalled or exceeded the international mean (see Figure 3.3.2). Where the students' performance was above the international mean, the difference ranged from ten to 46 percentage points.

**Figure 3.3.2: Performance on items classified as *USING SCIENTIFIC PROCEDURES***

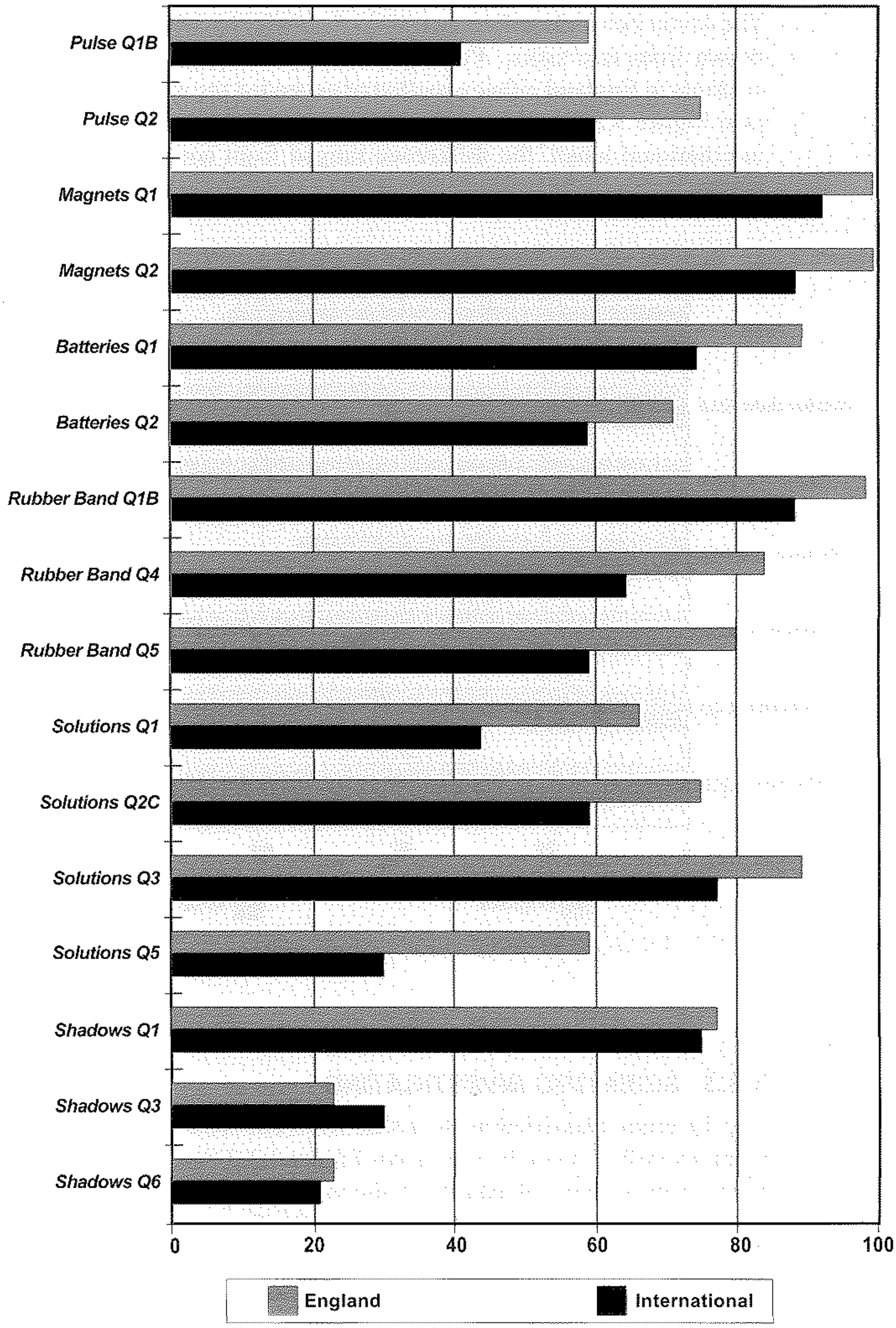



*Source: Tables 1.1, 1.7, 1.9, 1.11 and 1.13 (Harmon et al., 1997)*

### 3.3.3 *SCIENTIFIC INVESTIGATING*

The 16 items classified as *SCIENTIFIC INVESTIGATING* were drawn from six tasks. It can be seen from Figure 3.3.3 that the performance of students in England was above the international mean for all except one item in this category (*Shadows* Q3). Students' performance on 11 items exceeded the international mean by at least ten percentage points, and for one item (*Solutions* Q5) by 29 percentage points.

**Figure 3.3.3: Performance on items classified as *SCIENTIFIC INVESTIGATIING***




*Source: Tables 1.1, 1.7, 1.9, 1.11 and 1.13 (Harmon et al., 1997)*

# CHAPTER 4
# Concluding Remarks

The introduction of the National Curriculum in England in 1989 formally established the importance of process skills in mathematics and science as well as knowledge and understanding of concepts in these subjects at both primary and secondary school level. Prior to that date, the role of practical work in developing students' process skills had been highlighted by the Cockcroft Report (Cockcroft, 1982), the work of the APU in both mathematics and science, and the CSE Mode 3 examinations. However, since the implementation of the National Curriculum, TIMSS is the largest international study to supplement data collected by means of written achievement tests in mathematics and science with additional data relating to students' skills in practical activities in both subjects. Whilst the results of National Curriculum assessments in Year 9 present national data relating to both process skills and knowledge and understanding, TIMSS has collected data internationally and enables comparisons to be drawn between the performance of students in England and in other countries.

The results on the written achievement tests in mathematics and science and the results on the practical tasks complement each other, reflecting both subject knowledge and practical skills. The results of either component on their own present an incomplete (and therefore misleading) picture of the performance of 13-year-old students in England compared with their counterparts in other countries. Together, the data collected in both aspects of the study provide the most accurate reflection of the standards achieved by students in the full breadth of the National Curriculum in England in relation to other students world-wide.

Whilst many would argue that some of the results in the written mathematics tests showed cause for concern, students' performance in the practical mathematics tasks was better, both in comparison with the international means and in relation to the performance of students in the other countries selected for comparison. With this in mind, it seems appropriate to conclude that in addition to identifying consistent high performance in all aspects of science, TIMSS has demonstrated that there are aspects of mathematics in which students in England are also doing well.

# REFERENCES


BEATON, A.E., MARTIN, M.O., MULLIS, I.V.S., GONZALEZ, E.J., SMITH, T.A. and KELLY, D.L. (1996a). *Science Achievement in the Middle School Years: IEA's Third International Mathematics and Science Study (TIMSS).* Chestnut Hill, MA: Boston College.

BEATON, A.E., MULLIS, I.V.S., MARTIN, M.O., GONZALEZ, E.J., KELLY, D.L. and SMITH, T.A. (1996b). *Mathematics Achievement in the Middle School Years: IEA's Third International Mathematics and Science Study (TIMSS).* Chestnut Hill, MA: Boston College.

COCKCROFT REPORT. GREAT BRITAIN. DEPARTMENT OF EDUCATION AND SCIENCE. COMMITTEE OF ENQUIRY INTO THE TEACHING OF MATHEMATICS IN SCHOOLS (1982). *Mathematics Counts.* London: HMSO.

FOXMAN, D., HUTCHISON, D. and BLOOMFIELD, B. (1991). *The APU Experience 1977–1990.* London: School Examinations and Assessment Council, Evaluation and Monitoring Unit.

HARMON, M., SMITH, T.A., MARTIN, M.O., KELLY, D.L., BEATON, A.E., MULLIS, I.V.S., GONZALEZ, E.J. and ORPWOOD, G. (1997). *Performance Assessment in IEA's Third International Mathematics and Science Study (TIMSS).* Chestnut Hill, MA: Boston College.

HARRIS, S., KEYS, W. and FERNANDES, C. (1997). *Third International Mathematics and Science Study, Second National Report. Part 1: Achievement in Mathematics and Science at Age 9 in England.* Slough: NFER.

KEYS, W., HARRIS, S. and FERNANDES, C. (1996a). *Third International Mathematics and Science Study, First National Report. Part 1: Achievement in Mathematics and Science at Age 13 in England.* Slough: NFER.

KEYS, W., HARRIS, S. and FERNANDES, C. (1996b). *Third International Mathematics and Science Study, National Reports. Appendices.* Slough: NFER.

KEYS, W., HARRIS, S. and FERNANDES, C. (1997a). *Third International Mathematics and Science Study, First National Report. Part 2: Patterns of Mathematics and Science Teaching in Lower Secondary Schools in England and Ten Other Countries.* Slough: NFER.

KEYS, W., HARRIS, S. and FERNANDES, C. (1997b). *Third International Mathematics and Science Study, Second National Report. Part 2: Patterns of Mathematics and Science Teaching in Upper Primary Schools in England and Eight Other Countries.* Slough: NFER.

MARTIN, M.O., MULLIS, I.V.S., BEATON, A.E., GONZALEZ, E.J., SMITH, T.A. and KELLY, D.L. (1997). *Science Achievement in the Primary School Years: IEA's Third International Mathematics and Science Study (TIMSS).* Chestnut Hill, MA: Boston College.

MULLIS, I.V.S., MARTIN, M.O., BEATON, A.E., GONZALEZ, E.J., KELLY, D.L. and SMITH, T.A. (1997). *Mathematics Achievement in the Primary School Years: IEA's Third International Mathematics and Science Study (TIMSS).* Chestnut Hill, MA: Boston College.

ROBITAILLE, D.F. (Ed) (1993). *Curriculum Frameworks for Mathematics and Science* (TIMSS Monograph No. 1). Vancouver: Pacific University Press.

ROBITAILLE, D.F. (Ed) (1997). *National Contexts for Mathematics and Science Education. An Encyclopaedia of the Education Systems Participating in TIMSS.* Vancouver: Pacific Educational Press.

# APPENDIX I
# National steering committee

| Member | Association |
|---|---|
| Mark Neale (Chairman, 1993–1995) | Department for Education and Employment |
| Michael Richardson (Chairman 1995–1997) | Department for Education and Employment |
| Robert Wood | Department for Education and Employment |
| John Gardner | Department for Education and Employment |
| Dr Seamus Hegarty | Director, National Foundation for Educational Research |
| Dr Wendy Keys | National Foundation for Educational Research |
| Sue Harris | National Foundation for Educational Research |
| Cres Fernandes | National Foundation for Educational Research |
| George Smith (1993–1996) | OFSTED |
| Christine Agambar (1996–1997) | OFSTED |
| Richard Browne | School Curriculum and Assessment Authority |
| Miranda Simond (1993–1995) | School Curriculum and Assessment Authority |
| Carolyn Swain | School Curriculum and Assessment Authority |
| Dr John Marks | Consultant |
| Dr Hilary Steedman | Centre for Economic Performance, LSE |
| Prof. Geoffrey Howson | University of Southampton |
| Dr Barbara Jaworski (1993–1995) | Mathematical Association (secondary) |
| Roy Ashley (1995–1997) | Mathematical Association (secondary) |
| Susan Sanders | Mathematical Association (primary) |
| Dr Alan Eales | Association of Teachers of Mathematics (secondary) |
| Marjorie Gorman | Association of Teachers of Mathematics (primary) |
| Judith Lee | Association for Science Education (secondary) |
| Mick Revell | Association for Science Education (primary) |
| Brian Semple | Scottish Office Education Department |
| Hywel Jones | Welsh Office Education Department |
| Mike Richards (1993–1996) | Welsh Office Education Department |
| Andrew George (1996–1997) | Welsh Office Education Department |

# APPENDIX II
## Countries taking part in different components of TIMSS

### Continental Western Europe

| | | | PERFORMANCE ASSESSMENT | | |
|---|---|---|---|---|---|
| | *Population 1* | *Population 2* | *Population 1* | *Population 2* | *Population 3* |
| Austria | ● | ● | | | ● |
| Belgium *(Flemish)* | | ● | | | |
| Belgium *(French)* | | ● | | | |
| Cyprus | ● | ● | ● | ● | ● |
| Denmark | | ● | | ● | ● |
| France | | ● | | | ● |
| Germany | | ● | | | ● |
| Greece | ● | ● | | | ● |
| Iceland | ● | ● | | | ● |
| [1]Italy | | ● | | | |
| Netherlands | ● | ● | | | ● |
| Norway | ● | ● | ● | ● | ● |
| Portugal | ● | ● | ● | ● | |
| Spain | | ● | | ● | |
| Sweden | | ● | | | ● |
| Switzerland | | ● | | ● | ● |

### English-speaking

| | | | | | |
|---|---|---|---|---|---|
| Australia | ● | ● | ● | ● | ● |
| Canada | ● | ● | ● | ● | ● |
| England | ● | ● | | ● | |
| Ireland | ● | ● | | | |
| New Zealand | ● | ● | ● | ● | ● |
| Scotland | ● | ● | | ● | |
| United States | ● | ● | ● | ● | ● |

[1] *Argentina, Italy and Indonesia were unable to complete the steps necessary for their data to appear in the international reports. Because the characteristics of its school sample are not completely known, achievement results for the Philippines are not included in the main tables of the international reports. Mexico chose not to release its results for the international reports.*

## Eastern Europe

| | | | PERFORMANCE ASSESSMENT | | |
|---|---|---|---|---|---|
| | *Population 1* | *Population 2* | *Population 1* | *Population 2* | *Population 3* |
| Bulgaria | | ● | | | |
| Czech Republic | ● | ● | ● | ● | ● |
| Hungary | ● | ● | ● | ● | ● |
| Latvia | ● | ● | | | ● |
| Lithuania | | ● | | | ● |
| Romania | | ● | | ● | |
| Russian Federation | | ● | | | ● |
| Slovak Republic | | ● | | | |
| Slovenia | ● | ● | | ● | ● |
| Ukraine | | ● | | | |

## Asia and Pacific Region

| | | | | | |
|---|---|---|---|---|---|
| Hong Kong | ● | ● | ● | ● | |
| [1]Indonesia | ● | ● | | | |
| Japan | ● | ● | | | |
| Korea | ● | ● | | | |
| [1]Philippines | | ● | | | |
| Singapore | ● | ● | ● | ● | |
| Thailand | ● | ● | | | |

## Other countries

| | | | | | |
|---|---|---|---|---|---|
| [1]Argentina | | ● | | | |
| Colombia | | ● | | ● | |
| Iran | ● | ● | ● | ● | |
| Israel | ● | ● | ● | ● | ● |
| Kuwait | ● | ● | | | |
| [1]Mexico | ● | ● | | | ● |
| South Africa | | ● | | | ● |

# ANNEX 1

## Assignment of tasks and students to workstations

The total of 12 tasks were allocated to nine workstations for the administration: six workstations had one 30-minute task and three workstations had two 15-minute tasks. The nine selected students in each school visited three workstations and, depending on the combination, attempted three, four or five tasks during the 90-minute working time.

Two different sets of combinations of tasks were devised by the International Study Center. Students in any one school were assigned tasks from either one combination or the other (rotation 1 or 2) according to procedures laid down by the Study Center. The tasks allocated to particular workstations are shown in Table A1.1, and the assignment of students to workstations is shown in Table A1.2.

**Table A1.1: Assignment of tasks to workstations**

| Workstation | Task(s) | |
|---|---|---|
| A | S1<br>M1 | *Pulse*<br>*Dice* |
| B | S2<br>M2 | *Magnets*<br>*Calculator* |
| C | SM1 | *Shadows* |
| D | S3<br>M3 | *Batteries*<br>*Folding and Cutting* |
| E | S4 | *Rubber Band* |
| F | M5 | *Packaging* |
| G | S5 | *Solutions* |
| H | M4 | *Around the Bend* |
| I | SM2 | *Plasticine* |

**Table A1.2: Assignment of students to workstations**

| Student No. | Rotation 1 Workstations | Rotation 1 Workstations |
|---|---|---|
| 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 |

# ANNEXES 2–5

## Mathematics and science skills covered by tasks

The two main mathematics skills categories and the three science skills categories (each identified as Performance Expectations by the International Study Center) represent aggregations of the sub-categories which were linked to particular items within tasks. The specific skills areas linked to each of the items within the six mathematics tasks are shown sorted by skills areas and by tasks respectively in Annex 2 and in Annex 3, and the specific skills areas linked to each of the items within the six science tasks are shown sorted by skills areas and by tasks respectively in Annexes 4 and 5. It should be noted that whilst these Annexes list all the items within the 12 tasks, the main skills categories (Performance Expectations) identified by the International Study Center do not include all items, for example two items that were assigned by the International Study Center to the category Communicating (*Shadows* Q4 and *Plasticine* Q1B) were not included in any of the main mathematics or science skills categories.

Within the tables shown in these Annexes, the Reporting Category column shows the Performance Expectation category as listed in the Curriculum Frameworks (Robitaille, 1993).

## Annex 2 Mathematics skills covered by tasks, sorted by skills

| Task | Question number | Reporting Category | Skill |
|---|---|---|---|
| *Calculator* | 6a | 2.1.3 | Recalling mathematical objects and properties |
| *Calculator* | 1 | 2.2.1 | Use of equipment |
| *Plasticine* | 1a | 2.2.2 | Performing routine procedures |
| *Dice* | 1 | 2.2.2 | Performing routine procedures |
| *Dice* | 3 | 2.2.2 | Performing routine procedures |
| *Dice* | 4 | 2.2.2 | Performing routine procedures |
| *Around the Bend* | 1 | 2.2.2 | Performing routine procedures |
| *Around the Bend* | 2 | 2.2.2 | Performing routine procedures |
| *Packaging* | 2 | 2.2.2 | Performing routine procedures |
| *Shadows* | 1 | 2.2.3 | Performing complex procedures |
| *Shadows* | 5 | 2.2.3 | Performing complex procedures |
| *Dice* | 2 | 2.2.3 | Performing more complex procedures |
| *Dice* | 5a | 2.2.3 | Performing more complex procedures |
| *Calculator* | 2 | 2.2.3 | Using complex procedures |
| *Around the Bend* | 5a | 2.2.3 | Using complex procedures |
| *Packaging* | 3 | 2.2.3 | Performing complex procedures |
| *Plasticine* | 2b | 2.3.2 | Developing strategy |
| *Plasticine* | 3b | 2.3.2 | Developing strategy |
| *Plasticine* | 4b | 2.3.2 | Developing strategy |
| *Calculator* | 5 | 2.3.2 | Developing and describing a strategy |
| *Shadows* | 3 | 2.3.3 | Problem solving |
| *Plasticine* | 2a | 2.3.3 | Problem solving |
| *Plasticine* | 3a | 2.3.3 | Problem solving |
| *Plasticine* | 4a | 2.3.3 | Problem solving |
| *Calculator* | 6b | 2.3.3 | Problem solving |
| *Folding & Cutting* | 1 | 2.3.3 | Problem solving |
| *Folding & Cutting* | 2 | 2.3.3 | Problem solving |
| *Folding & Cutting* | 3 | 2.3.3 | Problem solving |
| *Around the Bend* | 4 | 2.3.3 | Problem solving |
| *Packaging* | 1 | 2.3.3 | Problem solving |
| *Calculator* | 3 | 2.3.4 | Predicting |
| *Calculator* | 4 | 2.3.4 | Predicting |
| *Folding & Cutting* | 4 | 2.3.4 | Predicting |
| *Shadows* | 6 | 2.4.3 | Generalizing |
| *Around the Bend* | 6 | 2.4.3 | Generalizing |
| *Shadows* | 2 | 2.4.4 | Conjecturing |
| *Dice* | 5b | 2.4.4 | Conjecturing |
| *Around the Bend* | 3 | 2.4.4 | Conjecturing |
| *Around the Bend* | 5b | 2.4.4 | Conjecturing |
| *Shadows* | 4 | 2.5.3 | Describing/discussing |
| *Plasticine* | 1b | 2.5.3 | Describing and discussing |

## Annex 3 Mathematics skills covered by tasks, sorted by tasks

| Task | Question number | Reporting Category | Skill |
|---|---|---|---|
| *Dice* | 1 | 2.2.2 | Performing routine procedures |
| *Dice* | 2 | 2.2.3 | Performing more complex procedures |
| *Dice* | 3 | 2.2.2 | Performing routine procedures |
| *Dice* | 4 | 2.2.2 | Performing routine procedures |
| *Dice* | 5a | 2.2.3 | Performing more complex procedures |
| *Dice* | 5b | 2.4.4 | Conjecturing |
| *Calculator* | 1 | 2.2.1 | Use of equipment |
| *Calculator* | 2 | 2.2.3 | Using complex procedures |
| *Calculator* | 3 | 2.3.4 | Predicting |
| *Calculator* | 4 | 2.3.4 | Predicting |
| *Calculator* | 5 | 2.3.2 | Developing and describing a strategy |
| *Calculator* | 6a | 2.1.3 | Recalling mathematical objects and properties |
| *Calculator* | 6b | 2.3.3 | Problem solving |
| *Folding & Cutting* | 1 | 2.3.3 | Problem solving |
| *Folding & Cutting* | 2 | 2.3.3 | Problem solving |
| *Folding & Cutting* | 3 | 2.3.3 | Problem solving |
| *Folding & Cutting* | 4 | 2.3.4 | Predicting |
| *Around the Bend* | 1 | 2.2.2 | Performing routine procedures |
| *Around the Bend* | 2 | 2.2.2 | Performing routine procedures |
| *Around the Bend* | 3 | 2.4.4 | Conjecturing |
| *Around the Bend* | 4 | 2.3.3 | Problem solving |
| *Around the Bend* | 5a | 2.2.3 | Using complex procedures |
| *Around the Bend* | 5b | 2.4.4 | Conjecturing |
| *Around the Bend* | 6 | 2.4.3 | Generalizing |
| *Packaging* | 1 | 2.3.3 | Problem solving |
| *Packaging* | 2 | 2.2.2 | Performing routine procedures |
| *Packaging* | 3 | 2.2.3 | Performing complex procedures |
| *Plasticine* | 1a | 2.2.2 | Performing routine procedures |
| *Plasticine* | 1b | 2.5.3 | Describing and discussing |
| *Plasticine* | 2a | 2.3.3 | Problem solving |
| *Plasticine* | 2b | 2.3.2 | Developing strategy |
| *Plasticine* | 3a | 2.3.3 | Problem solving |
| *Plasticine* | 3b | 2.3.2 | Developing strategy |
| *Plasticine* | 4a | 2.3.3 | Problem solving |
| *Plasticine* | 4b | 2.3.2 | Developing strategy |
| *Shadows* | 1 | 2.2.3 | Performing complex procedures |
| *Shadows* | 2 | 2.4.4 | Conjecturing |
| *Shadows* | 3 | 2.3.3 | Problem solving |
| *Shadows* | 4 | 2.5.3 | Describing/discussing |
| *Shadows* | 5 | 2.2.3 | Performing complex procedures |
| *Shadows* | 6 | 2.4.3 | Generalizing |

## Annex 4 Science skills covered by tasks, sorted by skills

| Task | Question number | Reporting Category | Skill |
|---|---|---|---|
| *Batteries* | 3 | 2.2.2 | Applying scientific principles to solve problems |
| *Plasticine* | 2a | 2.2.2 | Applying scientific/mathematical principles to solve quantitative problems |
| *Plasticine* | 2b | 2.2.2 | Applying scientific/mathematical principles to solve quantitative problems |
| *Plasticine* | 3a | 2.2.2 | Applying scientific/mathematical principles to solve quantitative problems |
| *Plasticine* | 3b | 2.2.2 | Applying scientific/mathematical principles to solve quantitative problems |
| *Plasticine* | 4a | 2.2.2 | Applying scientific/mathematical principles to solve quantitative problems |
| *Plasticine* | 4b | 2.2.2 | Applying scientific/mathematical principles to solve quantitative problems |
| *Pulse* | 3 | 2.2.3 | Applying scientific principles to develop explanations |
| *Batteries* | 4 | 2.2.3 | Applying scientific principles to develop explanations |
| *Rubber Band* | 6 | 2.2.3 | Applying scientific principles to develop explanations |
| *Solutions* | 4 | 2.2.3 | Applying scientific principles to develop explanations |
| *Shadows* | 2 | 2.2.3 | Applying scientific [or mathematical] principles to develop explanations |
| *Solutions* | 2a* | 2.3.1 | Using equipment |
| *Rubber Band* | 3 | 2.3.2 | Performing routine operations |
| *Plasticine* | 1a | 2.3.2 | Conducting routine experimental operations procedures |
| *Pulse* | 1a | 2.3.4 | Organising and representing data |
| *Rubber Band* | 1a | 2.3.4 | Organising and representing data |
| *Rubber Band* | 2 | 2.3.4 | Organising and representing data |
| *Solutions* | 2b | 2.3.4 | Organising and representing data |
| *Shadows* | 5 | 2.3.4 | Organising and presenting data |
| *Solutions* | 1 | 2.4.2 | Designing investigations |
| *Solutions* | 5 | 2.4.2 | Designing investigations |
| *Pulse* | 1b | 2.4.3 | Conducting an investigation |
| *Rubber Band* | 1b | 2.4.3 | Conducting an investigation |
| *Solutions* | 2c | 2.4.3 | Conducting an investigation |
| *Shadows* | 3 | 2.4.3 | Conducting an investigation |
| *Pulse* | 2 | 2.4.4 | Interpreting investigational data |
| *Magnets* | 2 | 2.4.4 | Interpreting investigational data |
| *Batteries* | 2 | 2.4.4 | Interpreting investigational data |
| *Rubber Band* | 4 | 2.4.4 | Interpreting investigational data |
| *Rubber Band* | 5 | 2.4.4 | Interpreting investigational data |
| *Shadows* | 1 | 2.4.4 | Interpreting investigational data |
| *Magnets* | 1 | 2.4.5 | Drawing conclusions from investigational data |
| *Batteries* | 1 | 2.4.5 | Drawing conclusions from investigational data |
| *Solutions* | 3 | 2.4.5 | Formulating conclusions from investigational data |
| *Shadows* | 6 | 2.4.5 | Formulating conclusions from investigational data |
| *Shadows* | 4 | 2.5.2 | Sharing scientific information |
| *Plasticine* | 1b | 2.5.2 | Sharing information |

* Due to problems of misinterpretation of this item in some countries (administrators of the tasks in schools were asked to check that students were using the thermometer correctly), this item was omitted by the International Study Center in the aggregation of items to science skills areas.

## Annex 5 Science skills covered by tasks, sorted by tasks

| Task | Question number | Reporting Category | Skill |
|---|---|---|---|
| *Pulse* | 1a | 2.3.4 | Organising and representing data |
| *Pulse* | 1b | 2.4.3 | Conducting an investigation |
| *Pulse* | 2 | 2.4.4 | Interpreting investigational data |
| *Pulse* | 3 | 2.2.3 | Applying scientific principles to develop explanations |
| *Magnets* | 1 | 2.4.5 | Drawing conclusions from investigational data |
| *Magnets* | 2 | 2.4.4 | Interpreting investigational data |
| *Batteries* | 1 | 2.4.5 | Drawing conclusions from investigational data |
| *Batteries* | 2 | 2.4.4 | Interpreting investigational data |
| *Batteries* | 3 | 2.2.2 | Applying scientific principles to solve problems |
| *Batteries* | 4 | 2.2.3 | Applying scientific principles to develop explanations |
| *Rubber Band* | 1a | 2.3.4 | Organising and representing data |
| *Rubber Band* | 1b | 2.4.3 | Conducting an investigation |
| *Rubber Band* | 2 | 2.3.4 | Organising and representing data |
| *Rubber Band* | 3 | 2.3.2 | Performing routine operations |
| *Rubber Band* | 4 | 2.4.4 | Interpreting investigational data |
| *Rubber Band* | 5 | 2.4.4 | Interpreting investigational data |
| *Rubber Band* | 6 | 2.2.3 | Applying scientific principles to develop explanations |
| *Solutions* | 1 | 2.4.2 | Designing investigations |
| *Solutions* | 2a* | 2.3.1 | Using equipment |
| *Solutions* | 2b | 2.3.4 | Organising and representing data |
| *Solutions* | 2c | 2.4.3 | Conducting an investigation |
| *Solutions* | 3 | 2.4.5 | Formulating conclusions from investigational data |
| *Solutions* | 4 | 2.2.3 | Applying scientific principles to develop explanations |
| *Solutions* | 5 | 2.4.2 | Designing investigations |
| *Shadows* | 1 | 2.4.4 | Interpreting investigational data |
| *Shadows* | 2 | 2.2.3 | Applying scientific [or mathematical] principles to develop explanations |
| *Shadows* | 3 | 2.4.3 | Conducting investigations |
| *Shadows* | 4 | 2.5.2 | Sharing scientific information |
| *Shadows* | 5 | 2.3.4 | Organising and presenting data |
| *Shadows* | 6 | 2.4.5 | Formulating conclusions from investigational data |
| *Plasticine* | 1a | 2.3.2 | Conducting routine experimental operations procedures |
| *Plasticine* | 1b | 2.5.2 | Sharing information |
| *Plasticine* | 2a | 2.2.2 | Applying scientific/mathematical principles to solve quantitative problems |
| *Plasticine* | 2b | 2.2.2 | Applying scientific/mathematical principles to solve quantitative problems |
| *Plasticine* | 3a | 2.2.2 | Applying scientific/mathematical principles to solve quantitative problems |
| *Plasticine* | 3b | 2.2.2 | Applying scientific/mathematical principles to solve quantitative problems |
| *Plasticine* | 4a | 2.2.2 | Applying scientific/mathematical principles to solve quantitative problems |
| *Plasticine* | 4b | 2.2.2 | Applying scientific/mathematical principles to solve quantitative problems |

* Due to problems of misinterpretation of this item in some countries (administrators of the tasks in schools were asked to check that students were using the thermometer correctly), this item was omitted by the International Study Center in the aggregation of items to science skills areas.

nfer

# THIRD INTERNATIONAL MATHEMATICS AND SCIENCE STUDY
## Third National Report

This report presents the national results for secondary school students in Year 9 on an optional component of the Third International Mathematics and Science Study (TIMSS). This component, known as the *Performance Assessment*, measured students' practical, investigative, problem-solving and analytical skills by means of 'hands-on' tasks in mathematics and science. The results on this aspect of TIMSS supplement the findings from the written tests completed by 13-year-old students, and the questionnaires completed by the students, their mathematics and science teachers and their headteachers which were the subject of two earlier national reports.

The report shows the six mathematics tasks and the six science tasks that were attempted by students, and gives the results for students in England together with comparisons with other countries and the international averages. In addition to presenting information about students' performance on each part of the 12 tasks, the report shows the results of additional analyses which were carried out to determine performance on the following skills areas across different tasks:

- performing mathematical procedures
- problem-solving and mathematical reasoning
- scientific problem-solving and applying concept knowledge
- using scientific procedures
- scientific investigating.

The results from the practical activities presented in this report show that the performance of students in England was equal to or above the international average for 11 out of the 12 tasks. This is consistent with the results on the written tests in science attempted by the same students, but contrasts with their results on the written tests in mathematics.

Based on a national sample of 450 students in 50 schools which was part of a world-wide sample of 15,000 students in 1,500 schools in 21 countries, the TIMSS *Performance Assessment* represents the largest international study of students' skills in practical mathematics and science activities carried out to date.

This report will be of interest to all readers of the previous national reports on TIMSS, including teachers, school governors, LEA advisory teams, policy makers and researchers.

ISBN 0 7005 1474 X

£15.00