Learning from TIMSS

Implications for Teaching and Learning Science

at the Junior Secondary Level

Edited by
Benny Hin Wai Yung

Published by the Education and Manpower Bureau
Printed by the Government Logistics Department,
The Government of the Hong Kong Special Administrative Region

Contents

Acknowledgements

The National Research Coordinator for the TIMSS Hong Kong Study would like to acknowledge the financial support provided by the Hong Kong Research Grants Council for carrying out the main study of TIMSS 2003. The Study Centre is very grateful to the support given by the principals and teachers in the participating schools. Without such committed professional input from the large numbers of mathematics and science teachers who contributed to the various stages of the main study of TIMSS 2003, it could not have been so successfully completed. The Study is also indebted to Miss Ada S.F. Tso for her dedicated undertaking of the main study to provide such a comprehensive database for analysis.

We would also like to acknowledge the financial support provided by the Education and Manpower Bureau (EMB) of the Government of the Hong Kong Special Administrative Region for carrying out the secondary analysis of science achievement data at the junior secondary level in Hong Kong. Input and advice from EMB colleagues at various stages of this secondary analysis study are also very much appreciated. Nonetheless, the opinions expressed are those of the authors and do not necessarily reflect the views of the supporting agency.

We would also like to express our gratitude to Mr. Arthur K.Y. Lau for his contribution in undertaking the secondary analysis of the TIMSS data for Hong Kong and Miss Vanessa, L.Y. Ho for organizing the teacher workshops and compiling teachers' ideas as well as providing editorial support for the publication of this book.

The secondary analysis study is also indebted to three practicing science teachers, Mr. Chan Sheung Yee, Miss Wong Wai Lan and Mr. Leo Chi Shing, for their contributions in preliminary analyses of item statistics, identifying suitable TIMSS items for further discussion during the teacher workshops, and their professional input in facilitating a lively and invigorating discussion atmosphere during the workshops.

Lastly, we would like to express our sincerest thanks to the teachers who attended the three repeated whole-day workshops, during which they contributed much of their experiences and expertise in science teaching. We strongly believe that the production of this book would never have come to fruition without the collective wisdom of this group of enthusiastic teachers. To them, we salute again.

Foreword

For years, I have been indifferent to TIMSS, PISA and all the other large-scale international tests of scientific knowledge and understanding. I considered them to be, at best, a waste of time and resources. More seriously, I regarded them as potentially highly divisive in their tendency to invoke dangerously misleading comparisons. My major concern was that they provide politicians, journalists and others who might wish to engage in 'teacher bashing' with some powerful ammunition, but ammunition of a highly dubious nature because of the often superficial way in which the data are analyzed. What, I wondered have science teachers and their students to gain from such an enterprise? Benny Yung, Siu Ling Wong and Man Wai Cheng have provided some convincing answers to my question and have allayed many of my anxieties. They have shown, very clearly and persuasively, how teachers can utilize assessment data to shed light on issues of curriculum and pedagogy. Through a painstaking analysis of the performance of Hong Kong students in TIMSS 1995, 1999 and 2003, ably supported by some enthusiastic science teachers in a series of workshops, they have produced an easy-to-follow guide to the ways in which TIMSS data can be used productively to benefit teachers, students and curriculum decision-makers.

The authors have shown that, when deployed sensitively and with the insight of experienced science educators, data from TIMSS can identify serious shortcomings in our approach to the teaching and learning of science and can identify some priorities for change. By extending the analysis to encompass all three TIMSS studies, much useful information has been collected on the ways in which Hong Kong's recent curriculum reforms have impacted the performance of students. Judging by the substantially better performance of Hong Kong students in 2003 (4^{th} in overall performance, compared with 14^{th} and 15^{th}, respectively, in 1995 and 1999) that impact has been generally positive, at least in the aspects of science addressed by the TIMSS tests. However, this does not mean that there is room for complacency. Indeed, the authors point out several aspects of the curriculum where change is needed - most notably, the need to take account of students' prior misconceptions in designing teaching/learning activities, to provide more opportunities for students to design their own scientific investigations, to establish better links between different topics in the curriculum and with everyday life experiences, and to utilize a wider range of learning activities and associated assessment methods (especially those related to higher order thinking skills). By involving teachers (more than 250 teachers participated in the workshops), the authors have also contributed much to local professional development needs. I believe that this is likely to be the most significant and long-lasting impact of the enterprise. The authors have ably demonstrated to these teachers, and to those teachers who will read this book, how careful and systematic analysis of assessment data, whether it be large-scale international projects like TIMSS or tests and examinations given to one class in one school, can identify learning problems. They have provided a model for teachers to use in analyzing assessment data of all kinds to shed light on issues of curriculum and pedagogy and to generate powerful insight into the ways in which we might improve science curriculum, teaching and learning. The authors are to be commended for this achievement and I recommend this informative, persuasive and eminently readable text to all science teachers eager to engage in educational research as a means to reflect on their day-to-day practice.

Derek Hodson
Professor of Science Education,
Ontario Institute for Studies in Education of the University of Toronto, and
Distinguished Visiting Professor in the University of Hong Kong
April, 2006

Chapter 1

Learning from TIMSS

Benny Hin Wai Yung

The aim of TIMSS, Trends in International Mathematics and Science Study, is to improve the teaching and learning of mathematics and science by providing data about students' achievement in relation to different types of curricula, instructional practices, and school environment. It is a project of the International Association for the Evaluation of Educational Achievement (IEA), an independent international cooperation of national research institutions and government agencies that has been conducting studies of cross-national achievement since 1959. Conducted first in 1995 and then in 1999 and 2003, the regular four-year cycle of TIMSS studies provides countries with an opportunity to obtain comparative information about their students' achievement in mathematics and science on a longitudinal basis. More importantly, TIMSS also asks students, their teachers, and their school principals to complete questionnaires about the curriculum, schools, classrooms, and instruction. This data gives policy makers, curriculum specialists, and teachers a dynamic picture of implementation of educational policies and practices around the world, providing an invaluable perspective from which to consider educational reform and improvement.

TIMSS 2003 was administered at the eighth and fourth grades (i.e. Secondary 2 and Primary 4). With funding from the Education and Manpower Bureau (EMB), HKSAR, the present project is undertaken to perform secondary analysis of the TIMSS 2003 data with a focus on the Secondary 2 students' science achievement only. It is hoped that the findings of this project will generate useful information for improving science education in Hong Kong (HK).

Our task began with a preliminary secondary analysis of the TIMSS 2003 HK data set. The findings, together with the set of draft research questions that guided the preliminary analysis, were then shared with a group of enthusiastic science teachers, educators and EMB officers in a brainstorming session during which the research questions were further refined and additional research questions were identified. A second round of analysis of the data was then undertaken to yield answers for the final set of research questions.

Following the second round of data analysis, three identical workshops, each of 6 hours duration, were conducted for enthusiastic science teachers. The aims were: (1) to disseminate some aspects of the findings that are related to classroom teaching in general (e.g. gender-biased science instruction), and (2) to inform teachers of students' weaknesses and/or their common misconceptions on specific topics. Using this information, participating teachers suggested ways to revise the curriculum, improve teaching and enhance learning.

Because of the confidentiality of the TIMSS test items, our discussions with the teachers during the workshops had to be restricted to only those items that were released to the public (which were about one-third of the total items used in a single study). It was decided that we should widen our scope of analysis to include the released items of TIMSS 1999 and 1995. By doing so, we were able to examine our students' performance across time on certain topics that were tested on more than one occasion. Such a decision proved to be fruitful and has yielded useful information on how curricular changes have had an impact on our students' science achievement.

With reference to item statistics of the released test items, teachers discussed in groups during the workshops, identified students' misconceptions and/or weaknesses, and shared with one another their own experiences of tackling learning problems. The present report is a compilation of the recommendations made by this group of teachers during the workshops. Their recommendations range from revisions of the curriculum to suggestions of classroom practices to tackle specific misconceptions and/or areas of weakness in students' performance.

More than 250 teachers attended the workshops. In general, they rated the workshops highly in contributing to their professional development. We sincerely hope that through this book - a compilation of the collective wisdom of this group of teachers - more teachers in HK can learn from TIMSS.

1.1 Purpose of this book

The purpose of this book is to give teachers a number of suggestions and ideas for teaching more effectively the topics that HK students found difficult in the TIMSS tests. It also provides curriculum developers with useful information for modifying the curriculum. The emphasis of the book is on suggestions for further improvement. For these reasons, the many questions on which HK students did well are not discussed here. In some cases, the TIMSS test items do not correspond to a specific topic in the HK Junior Secondary Science Curriculum; nevertheless they all represent significant information that demands attention.

1.2 Organization of the book

The book is divided into six chapters. Chapter One introduces the background and origin of the project. Chapters Two to Four focus on students' common misconceptions in biology, physics and chemistry respectively. Chapter Five examines students' understanding about the nature of science. Chapter Six looks at gender differences in performance and related issues. Chapter Seven summarizes the main recommendations derived from this project.

Wherever appropriate, a common format is adopted throughout the book when the discussion is making reference to a specific TIMSS test item. It comprises four parts:

- **The Question (2003)** - shows the actual item used in the study, the number in bracket indicates the year of study; wherever appropriate, similar items from the 1995 and/or 1999 TIMSS tests are included to facilitate comparison of students' performance across time
- **Result** - includes the item statistics of HK students and the international average score
- **Noteworthy Points on Students' Performance** - suggests possible explanations for the observed results
- **Implications / Suggestions for Improvement** - suggests modifications to the curriculum or ideas for teaching the topic more effectively

1.3 Types of item and item statistics

Two main types of items were used in TIMSS - multiple-choice questions and constructed-response questions. A discussion of information for each type of item follows.

Multiple-Choice Questions

In order to differentiate performance of students of different abilities, the distractor options developed for multiple-choice questions were usually designed to resemble students' commonly committed mistakes. In many cases, these mistakes reflect the common misconceptions students hold, their lack of knowledge, etc. Below is an example to illustrate what can be learned through the analysis of students' responses to the different options.

The Question (1995)

Which BEST explain why green marine algae are most often restricted to the top 100 metres of the ocean?

A. They have no roots to anchor them to the ocean floor.
B. They can live only where there is light.
C. The pressure is too great for them to survive below 100 metres.
D. If the algae lived below 100 metres they would be eaten by animals.

To answer this 1995 TIMSS question, students needed to first pinpoint that there is no/little light at great depths of the ocean. Hence algae cannot carry out photosynthesis, and would therefore not be able to survive. Thus the correct answer to this question is B. The table below shows the percentage of students opting for each answer, together with the possible type(s) of error corresponding to each option.

Option [% of students]	Nature of associated error
A – [12.6]	Students were lack of knowledge about green marine algae.
(B)* – [49.3]	The correct answer. *In subsequent chapters, the correct answer will be identified by being bracketed as shown.

C – [29.8]	Students only considered pressure as the crucial factor, but failed to consider the necessity of light for photosynthesis by green algae. They failed to appreciate that light is the determining factor even if the algae were to be able to survive the great pressure at the ocean floor.
D – [7.7]	Students just blindly recalled that algae are producers and will be fed upon by animal consumers. They failed to appreciate that while this feeding relationship hold true, algae will continue to exist to sustain a stable ecosystem as long as there is enough sunlight for the algae. In sum, like those who opted for C, this group of students also failed to appreciate that light is the limiting factor down at the sea floor. In addition, they seemed to lack a clear understanding of the dynamic relationship among organisms in a food chain/food web.

As indicated in the table above, we can infer the kind of misconceptions that students might hold from the option they chose. The percentages of students opting for each answer will also tell teachers which is/are the area(s) that need(s) most attention for improvement.

Constructed-Response Questions

There were two types of constructed-response questions in TIMSS - short-answer and extended-response. Short-answer questions comprised mostly single-step questions, each assigned 1 mark. Extended-response questions (also known as structured questions) were multi-step questions worth 2 or more marks. There is an increasing trend towards more extended-response questions in recent TIMSS studies.

The constructed-response questions in TIMSS were designed to collect information on the processes and types of thinking that students use in answering the questions. A careful analysis of students' answers makes it possible to examine some of the complex and multi-stage thinking that takes place when students are attempting the questions. To allow such analysis to be undertaken, each of the students' answers had to be encoded using two-digit codes:

- the first digit indicates the correctness level of the answer;
- the second digit identifies the method or approach used in the problem or the misconception or type of error demonstrated.

Thus, a combination of the two digits represents a diagnostic code for identifying students' way of thinking, common errors and misconceptions when answering the questions. As an example, the coding guide for the question shown below allows for analysis of students' understanding of the advantage of having two eyes to see with, rather than one eye.

The Question (1995)

What is the advantage of having two eyes to see with rather than one eye?

The Marking Scheme / Coding Guide

Code	Response
Correct Response (any of the following four answers worth 1 mark)	
10	Mentions that 2 eyes allow depth perception or better perception of distance.
11	Mentions that 2 eyes allow seeing more or a wider field of vision. *Example: Seeing at a wider angle.*
12	Mentions that with two eyes, one is still working even if one eye is damaged.
19	Other correct answers.
Incorrect Response	
70	Mentions seeing twice as much. *Examples: We see twice as much with two eyes.* *You only see half as much with one eye.*
71	Refers to energy or effort. *Example: There is more energy in two eyes.* *A single eye will more easily get tired.*
79	Other incorrect answers.
Nonresponse	
90	Crossed out/erased, illegible, or impossible to interpret.
99	Blank

As mentioned above, the first digit in each code corresponded to the correctness level. If the answer worth 1 mark (as in this case), it received a code with a first digit of 1. If the answer worth 2 marks (as maybe in other questions), the code for the correct answers would then begin with a 2. In addition, some of the questions were given partial credit or full score. In those cases, the first digit would then be 1 or 2 depending on the degree of correctness of the answer. Codes for incorrect answers began with the digit of 7; and 9 represented blank or crossed-out answers. The second digit corresponded to the method or approach used if correct, or to the misconceptions or type of error if it was wrong.

Chapter 2

Learning from Students' Performance in Biology and Environmental Science-related Questions

Benny Hin Wai Yung

2.1 Performance in General

In TIMSS 2003, there were five science content domains: Earth Science, Environmental Science, Life Science, Physics and Chemistry. Table 2.1 shows the achievements of Hong Kong (HK) students in the five content domains. The international averages are also presented for comparison.

Table 2.1 Students' achievements in the five science content domains in TIMSS 2003

Science Content Domain	**HK**		**International Average**		**HK - Int**
	Mean*	S.E.**	Mean*	S.E.**	Mean*
Earth Science (31 items)	549	2.9	474	0.5	+75
Environmental Science (27 items)	555	2.6	474	0.5	+81
Life Science (54 items)	551	2.9	474	0.5	+77
Physics (46 items)	555	2.8	474	0.5	+81
Chemistry (31 items)	542	2.6	474	0.5	+68

* TIMSS used item response theory methods to summarize the achievement results on a scale with a mean of 500 and a standard deviation of 100.

** SE stands for Standard Error

HK students scored above the international average in all five domains, by 68 to 81 scale-points. In consequence, HK came fourth amongst the 46 participating countries – a great improvement on the 15th and 14th position in TIMSS 1999 and 1995, respectively.

2.2 Performance in Individual Items

Though HK students did particularly well in Environmental Science, close examination of their responses to some of the items revealed room for further improvement. In particular, more attention needs to be paid to environmental issues like global warming, ozone depletion and acid rain. However, we shall begin our discussion with the biology topics. These include classification of living things, photosynthesis, sense organs and ecology.

Wherever appropriate, for each of the TIMSS items discussed below (and in subsequent chapters), attempts will first be made to identify the specific **content or concept(s) that is/are being tested** in the item. Efforts will also be made to identify the **relevant Science (S1-3) topic(s)** – that is, relevant topic(s) in our Science Curriculum (Secondary 1-3). The rubric of **the question** is then presented, followed by **results** of students' performance. **Noteworthy points** on students' performance are then highlighted and followed by a discussion of their **implications** and possible ways of improving the situation. For the sake of consistency, boys' and girls' scores are reported in every item irrespective of whether there is any significant gender difference.

2.2.1 Classification of living things

Content / Concept assessed: Characteristics of mammals

Relevant Science (S1-3) topic: Looking at living things (Unit 2)

The Question (2003)

Cats are most closely related to which of the following animals?

A. Crocodiles
B. Whales
C. Frogs
D. Penguins

Result

	HK (%)
A	23.2
(B)	27.6
C	19.7
D	28.4

% correct			
HK			**International**
Boys	**Girls**	**Overall**	
29.4	26.0	27.6	26.6

Noteworthy Points on Students' Performance

There was a relatively even distribution in the percentage of students choosing each option. Several reasons could have contributed to this outcome:

1. ***Question rubric –*** Ambiguity of the term "closely related" perhaps confused students, and led them to interpret the question in various incorrect ways. Students might have classified the animals in terms of 'cats walking on land' and 'penguin walking on ice' as "closely related", or 'both cats and penguins like to eat fish' as "closely related".
2. ***Common misconceptions –*** Many students do not know whales/dolphins are mammals and have mammary glands. Likewise, students have great difficulty classifying crocodiles as reptile rather than amphibian. Students tend to use general impressions/superficial characteristics to classify the animals rather than specific characteristic features. Some hold the misconception that penguins have hair instead of feathers. When penguins' feathers are wet they look more like hairs, consequently students believe they are mammals rather than birds.
3. ***Everyday life experience –*** HK students lack contact with nature and different kinds of animals.

Implications / Suggestions for Improvement

1. Use models/specimens/toys of animals to show students the similarities and differences between different animals. Remind students to pay attention to the characteristic/internal features of each animal. Focus on the characteristic features of mammal classification – mammary gland and birth/viviparity and ask students to give examples.
2. Show documentary of whales/dolphins' way of life to help explain their mammalian characteristics:
   i. Show video clips of whales/dolphins suckling milk (which is a semi-solid substance) from their mother. (Use the keyword "suckling" to search on the Internet for such video clips.) Discovery Channel Video has a documentary entitled *The Ultimate Guide: Dolphin*, which contains footage of a dolphin mother feeding semi-solid milk to a baby dolphin. The program focuses on dolphins' evolution and discusses scientific investigation issues, such as why dolphins can hold their breath for 4 to 5 minutes underwater. To understand dolphins' behaviours, scientists attached video cameras onto their tails to measure the frequency of its movement. The greater the frequency of tail movement, the greater is the consumption of oxygen. Scientists discovered that dolphins only need to make a few strokes to swim a considerable distance.

This low level of activity leads to a smaller need for oxygen, and explains why dolphins can stay under water for 4 to 5 minutes after each breath. Introducing students to this kind of research technique could stimulate them to link creativity and scientific investigation. For example, since dolphins cannot speak, what innovative methods could scientists use to confirm their speculations related to dolphins' behaviour? The above example can also be used to illustrate how technological advancement (video technology in this case) can contribute to scientific advancement.

ii. Show video of whales/dolphins using their blowholes to breathe (and blowing water out) above the water surface. This enables students to appreciate that dolphins, as mammals, are lung-breathers – unlike fishes, which breathe via gills.

iii. Show video of whales/dolphins giving birth.

3. Bring students to Ocean Park. Observe the caretaker measuring the dolphin's body temperature when it is sick. Only mammals and birds are warm blooded, while fish are cold-blooded. This would provide information to aid students when trying to remember the characteristics of dolphins.
4. Bring students to the science museum to examine whale/dolphin skeleton.
5. For more able students, teachers can talk about how the classification system was established and how it is evolving. Explain how scientists change the classification system according to each new discovery they make. Where appropriate, the dynamic nature of science that scientific knowledge is constantly changing can be introduced.

2.2.2 Photosynthesis

Content / Concept assessed: Conditions necessary for photosynthesis

Relevant Science (S1-3) topic(s): Living things and air (Unit 7)

The Question (1995)

Which BEST explain why green marine algae are most often restricted to the top 100 metres of the ocean?

A. They have no roots to anchor them to the ocean floor.

B. They can live only where there is light.

C. The pressure is too great for them to survive below 100 metres.

D. If the algae lived below 100 metres they would be eaten by animals.

Result

	HK (%)
A	12.6
(B)	49.3
C	29.8
D	7.7

% correct			
HK			International
Boys	Girls	Overall	
53.9	44.2	49.3	52.8

Noteworthy Points on Students' Performance

The relatively poor performance of HK students (49%) compared with the International Average (53%) might be due to the following reasons:

1. ***Everyday life experience*** – HK students rarely come across algae except during field trips. Of all countries, Japan performed the best (77%). It might be because Japanese eat a lot of algae and hence know more about algae including their habitat.
2. ***Gender difference*** **–** HK boys (54%) performed better than girls (44%). It might be that boys tend to explore more on the beaches than girls do. For example, boys like to explore at the ends of beaches, searching in rock pools, turning over stones to look for organisms, etc. while girls tend to stay on the beaches. This lends support to the above argument that everyday life experiences matter.
3. ***Fragmented teaching*** – Algae are introduced to students in S1 during the topic on Classification, which focuses on external features rather than looking at the organism as a whole (including its mode of nutrition). In S2, terrestrial plants are often used by teachers in teaching photosynthesis. They often miss the chance to bring algae back into the picture. Actually, the more able students may be challenged to find out why there are different kinds of algae (namely, green, brown and red) found at different depths of the sea, and how this distribution is related to photosynthesis.
4. ***Higher order thinking skills*** – Students could not systematically analyze the complex information contained in the question. They did not know how to apply learned knowledge to answer the question. In particular, students could not apply the concept that light is essential for photosynthesis and that *light intensity decreases as depth of water increases*.

Implications / Suggestions for Improvement

1. In planning their lessons, teachers should take into consideration students' everyday life experiences, including any gender differences. Dried green algae, being a favourite food of students can be used by the teacher as a starting point to

talk a little more about how algae adapt to their mode of life. This can be supplemented by, for example, showing video clips and discussing why algae found at different depths of water are of different colours.

2. Teachers should help students to link their present learning with their prior knowledge, consolidate and extend it further. This helps them to form a holistic picture of what they are learning in relation to what they have learned.
3. To clarify the relationship between light intensity and depth of water, teachers can recall students' memory of watching any documentary video of deep sea exploration (or actually showing them the video), and ask them why the explorers have to bring along illuminating devices. Then ask students what would happen to the plants if there is no/little light down there. Students need to understand that even though plants can stand the high pressure at great depth, there is no/or little light to allow for photosynthesis.

2.2.3 Sense organs

Content / Concept assessed: Advantages of having two eyes and two ears

Relevant Science (S1-3) topic: Sensing the environment (Unit 11)

The Question (1995)

What is the advantage of having two eyes to see with rather than one eye?

Result

Percentage		**Response**
Overall	**Breakdown**	
		Correct Answer **(Boys: 45.4%; Girls: 39.6%)**
43 (62.7)*	5.2	Mentions that 2 eyes allow depth perception or better perception of distance.
	35.8	Mentions that 2 eyes allow seeing more or a wider field of vision. *E.g. Seeing at a wider angle.*
	0.6	Mentions that with two eyes, one is still working even if one eye is damaged.
	1.4	Other correct answers.
		Incorrect Answer **(Boys: 44.6%; Girls: 49.4%)**
46.7	0.5	Mentions seeing twice as much. *E.g. We see twice as much with two eyes.* *You only see half as much with one eye.*
	2.4	Refers to energy or effort. *E.g. A single eye will more easily get tired.* *There is more energy in two eyes.*
	43.8	Other incorrect answers.
		No Answer **(Boys: 9.9%; Girls: 11%)**
10.5	10.5	Blank, crossed out/erased, illegible, or impossible to interpret.

*International average

The Question (2003)

What is the advantage of having two ears to hear with rather than one ear?

Result

Percentage**		Response
Overall	Breakdown	
		Correct Answer (Boys: 31.4%; Girls: 32.1%)
32 (34.8)*	10.2	Mentions being able to locate the position, direction and/or distance of the source of sound. *E.g.* *By having two ears, you can actually tell where a sound came from.* *With two ears you can hear which direction a noise is coming from.* *With two ears you can judge the distance the sound is away from you.* *With two ears you can tell if the sound is near or far.*
	19.7	Mentions hearing sounds from both sides (direction) with no mention of locating the source. *E.g.* *You can hear both sides of you.* *You can hear sounds from all around, not just one side.* *With two ears you can hear from more than one way.*
	0.7	Mentions that if hearing is lost in one ear, the other may still function. *E.g.* *In the result of being deaf in one ear, you have another one that is used.* *If you lost the hearing in one ear, the other one might still work.*
	1.1	Other correct answers.
		Incorrect Answer (Boys: 61.8%; Girls: 63.3%)
62.4	51.8	Gives only a general or vague response relating to how well one can hear. *E.g.* *You hear better.* *You can hear half as much with one ear.* *Two ears lets you hear a lot more.* *The volume is greater.*
	3.6	Mentions only that hearing is uneven/unbalanced with one ear. *E.g.* *If you had one ear, the sound would be uneven.* *Your hearing would be out of balance.* *Your hearing gets balanced better with two ears.*
	7.1	Other incorrect answers.
		No Answer (Boys: 6.7%; Girls: 4.6%)
5.6	5.6	Blank

*International average

** In some cases that follow (in this and subsequent chapters), the percentage total may not add up to 100 because of rounding up of figures to one decimal place.

Noteworthy Points on Students' Performance

The two questions test students on very similar concepts, using eyes and ears as the respective contexts. HK students did poorly in both the 1995 question on eye (43%) and the 2003 question on ear (32%). Students' poor performance might be attributed to:

1. ***Curriculum / Teaching emphasis*** – Our curriculum (and hence teaching) emphasizes structural details of the organs concerned and pays insufficient

attention to their functions in relation to the survival of an organism as a whole.

2. ***Communication skills –*** Students were unable to communicate ideas in a concise and coherent manner in written form – as indicated by the high non-response rates (11% and 6% for the 1995 and 2003 question, respectively). Some teachers attributed the slight improvement in 2003 to the switching of the medium of instruction to Chinese in most of the schools.

Implications / Suggestions for Improvement

1. Teaching (and curriculum revision) needs to look at the organism as a whole rather than placing too much emphasis on structural details of individual organs. Below are some activities for achieving this shift of emphasis:
   i. Show students the ETV on eye. The program talks about the advanatages of having two eyes.
   ii. Engage students in an activity on depth perception using two eyes compared with using one eye only.
   iii. Use examples from a wide variety of animals to demonstrate that different animals have eyes in different positions relative to their body structure. Ask students how this is related to the survival of the organisms concerned.
   iv. Take students to the life science related exhibits in the science museum. Instruct them to try the computer simulation programs that allow them to perceive the world through the "eyes" of different animals, such as a mosquito or a fish. Discuss with students the differences in vision between humans and other animals.
   v. Ask a student to stand in the middle of the room. Blindfold the student. Produce sounds from several locations. Ask the student to identify the origins of the sounds. Repeat this demonstration, but this time asks the student to cover one ear. Ask if he/she could locate the origins of the sounds.
2. Provide more opportunities for students to practice their written communicative skills – for example, require them to complete worksheets in complete sentences.

2.2.4 Ecology

Content / Concept assessed: Unwanted consequences of introducing a new species

Relevant Science (S1-3) topics: Looking at living things (Unit 2)

Living things and air (Unit 7)

The Question (1995)

What could be the unwanted consequences of introducing a new species to a certain area? Give an example.

Result

Percentage		Response
Overall	Breakdown	
		Correct Answer (Boys: 2.4%; Girls: 1.6%)
2.1 (13)*	1.4	States that the natural (ecological) balance will be upset. A realistic example of a species is given.
	0.5	States that the new species may take over and gives examples. *E.g. It may kill everything. If you put an alligator in a fish and duck pond, it will eat the fish and ducks.*
	0.2	Other correct responses with examples.
		Partial Answer (Boys: 37.4%; Girls: 45.1%)
40.5	11.5	Adequate explanation, but no concrete and realistic example is given. *E.g. There will be too many.*
	27.3	States the new species cannot live here.
	1.4	Other partially correct.
		Incorrect Answer (Boys: 36.4%; Girls: 30.5%)
34.1	0.7	Only an unrealistic example is given. *E.g. Introducing polar bears into the Sahara.*
	33.4	Other incorrect answers.
		No Answer (Boys: 23.9%; Girls: 22.3%)
23.2	23.2	Blank, crossed out/erased, illegible, or impossible to interpret.

*International average

The Question (1999)

A new species of fish was released into a lake. State two unwanted outcomes that could arise from the introduction of this new species.

Result

Response 1 (i.e. the first unwanted outcome stated in the answer)

Percentage		Response
Overall	Breakdown	
		Correct Answer (Boys: 53.3%; Girls: 59.5%)
56.3 (59.9)*	3.2	Mentions competition with native species (e.g. overpopulation, eating the limited food supply). E.g. *They could eat all the plants the other fish need to survive.* *The new fishes will overbreed.*
	2.4	Mentions new species introducing diseases (bacteria, parasites, etc.) E.g. *The new species could carry viruses which kill off the other fish.* *They release disease and bacteria.*
	14.1	Mentions effects of predation (new species killing off existing species or vice-versa). E.g. *The fish can destroy other species in the water.* *The new species gets eaten by the fish already in the lake.*
	32.3	Mentions that the new species cannot survive in the lake (extinction due to inhospitable habitat). E.g. *They could have trouble adapting and die out.* *The fish may die out immediately because it is not the right type of water.* *The new fish might catch a disease in the lake and die.*
	1.7	Mentions upsetting the food web or ecological balance. E.g. *The species could damage the food web. It could alter the environment.* *The food chain is disturbed.*
	1.7	Mentions mating with existing species. E.g. *Could give rise to new species of fish.* *The new species might mate with other fish.*
	0.9	Other correct.
		Incorrect Answer (Boys: 36.5%; Girls: 35.1%)
35.8	21.9	Response too general.
	13.9	Other incorrect answers.
		No Answer (Boys: 10.3%; Girls: 5.4%)
7.9	7.9	Blank, crossed out/erased, illegible, or impossible to interpret.

Response 2 (i.e. the second unwanted outcome stated in the answer)

Percentage		Response
Overall	Breakdown	
		Correct Answer (Boys: 43.9%; Girls: 43.5%)
43.8 (46.8)*	5.9	Mentions competition with native species (e.g. overpopulation, eating the limited food supply). E.g. *They could eat all the plants the other fish need to survive.* *The new fishes will overbreed.*
	6.2	Mentions new species introducing diseases (bacteria, parasites, etc.) E.g. *The new species could carry viruses which kill off the other fish.* *They release disease and bacteria.*
	12.8	Mentions effects of predation (new species killing off existing species or vice-versa). E.g. *The fish can destroy other species in the water.* *The new species gets eaten by the fish already in the lake.*
	14.1	Mentions that the new species cannot survive in the lake (extinction due to inhospitable habitat). E.g. *They could have trouble adapting and die out.* *The fish may die out immediately because it is not the right type of water.* *The new fish might catch a disease in the lake and die.*
	1.4	Mentions upsetting the food web or ecological balance. E.g. *The species could damage the food web. It could alter the environment.* *The food chain is disturbed.*
	1.0	Mentions mating with existing species. E.g. *Could give rise to new species of fish.* *The new species might mate with other fish.*
	2.4	Other correct.
		Incorrect Answer (Boys: 38.7%; Girls: 40.7%)
39.7	23.9	Response too general.
	15.8	Other incorrect answers.
		No Answer (Boys: 17.4%; Girls: 15.7%)
16.5	16.5	Blank, crossed out/erased, illegible, or impossible to interpret.

*International average

Overall result combining Responses 1 and 2

	Percentage		
	Boys	**Girls**	**Overall**
Full Credit (2 marks)	27.8	29.8	28.8 (40.3)*
Partially Correct Answer (1 mark)	41.7	43.6	42.6
Incorrect Answer	20.5	21.2	20.8
No Answer	10.0	5.4	7.9

* International average

The Question (2003)

The following question is one of the four sub-questions of a multi-step / structured question. Students are first introduced, by drawings, to the geographical location of the Galapagos Islands with the following description: "They are a group of volcanic islands in the Pacific Ocean about 1000 kilometres from South America. When these islands first formed they consisted of lava only. Eventually, living organisms (plants and animals) inhabited the islands. These organisms arrived long before people settled on the islands."

When settlers came to live on the Galapagos Islands, they brought with them a number of new animals such as cats and goats. Write down one effect the introduction of cats and goats could have on the animals and plants already living on the islands.

A. One effect of cats
B. One effect of goats

Result

Response for effect of cats:

Percentage		Response
Overall	**Breakdown**	
		Correct Answer (Boys: 43.9%; Girls: 44.3%)
44.3 (35.1)*	40.7	Refers to cats preying upon other organisms, or similar (resulting in a reduction in population). *E.g. They will eat the birds and other animals.* *The cats help them by eating the rats and mice.* *Their prey could become extinct.*
	3.4	Other correct answers. *E.g. They might pass on diseases to other animals.*
		Incorrect Answer (Boys: 32.8%; Girls: 33%)
32.8	8.0	Refers only to an effect on the cat with no explicit effect on other organisms.
	24.9	Other incorrect answers.
		No Answer (Boys: 23.3%; Girls: 22.6%)
22.9	22.9	Blank

*International average

Response for effect of goats:

Percentage		Response
Overall	Breakdown	
		Correct Answer (Boys: 64.6%; Girls: 69.3%)
69.1 (44.3)*	63.7	Refers only to the goats eating plants (resulting in a reduction of the amount of plant life on the island). *E.g. The goats will eat all the grass on the island.* *It could lead to erosion if the goats clear the land by eating all the plants.* *Large pieces of grass will disappear as the goats eat it.*
	5.1	Refers to an effect of the goat on other animals (e.g. competition for food/habitat, as food source for predators, etc.). *E.g. The animals that eat goats would have more food.* *They might become a source of food.* *The goats will eat up the plants and the populations that depend on plants will decrease.*
	0.5	Other correct answers.
		Incorrect Answer (Boys: 17.6%; Girls: 15.6%)
15.8	4.4	Refers only to an effect on the goat with no explicit effect on other organisms. *E.g. Goats would have more babies.* *They would die because they don't have any food.*
	11.2	Other incorrect answers. *E.g. Goats might eat the cats.*
		No Answer (Boys: 18%; Girls: 15.1%)
15.1	15.1	Blank

*International average

Noteworthy Points on Students' Performance

The three questions test students on the same concept. HK students' performance improved significantly from just over 2% correct in 1995 (Int. Av. = 13%) to over 44% correct in 2003 (Int. Av. = 35%). Possible reasons for this occurrence:

1. ***Question rubric* –** Difficulty of the questions decreased from 1995 to 2003 as a result of how the rubrics were written. The 1995 question was the most difficult – it is abstract, provided no specific context for students to relate to when answering the question, and has complex vocabulary[1] (e.g. species and consequence). In contrast, the 1999 question provided a specific context (lake) and an example of a species (fish); it also used simpler words ('outcome' instead of 'consequence'). These changes make the question much easier than the previous one. In 2003, the rubric of the question was further improved by providing more contextual information and asking specifically about the effect of introducing cats and goats on 'the *animals* and *plants* already living on the islands'. Clearly, this is much

[1] Actually, the test paper is bilingual. The English version and the Chinese version of the question appear side by side on the same page or opposite pages.

more specific (and hence easier) than asking for 'unwanted outcomes' or 'unwanted consequences'.

2. ***Curriculum / Teaching emphasis –*** The old curriculum placed little weight on the environmental issues which are now emphasized in the new S1 curriculum. However, some teachers (misguided by textbooks) might have placed undue emphasis on certain specific examples of extinct and endangered species rather than on the general principle of inter-dependence of life. Actually, it is quite difficult to teach this principle without mentioning ideas about the food chain and food web which, however, are in the S2 curriculum.
3. ***Fragmented teaching*** – When teachers teach ideas of food chain and food web in S2, they often miss the chance of recalling, consolidating and extending what students have learned in S1 about the inter-dependence of life.
4. ***Everyday life experience –*** 69% of HK students gave an answer that 'goats eat plants', whereas only 44% were able to point out that 'cats prey upon other organisms'. It seems that our students know more about goats than cats. It is likely that they would have come across the Chinese term 'grass' goat; and that they would regard cats as pets which only eat cat food.

Implications / Suggestions for Improvement

1. In setting test/exam questions, use simple words and phrases, refer to specific contexts and examples. These are particularly helpful for the weaker students.
2. Revise the curriculum to have food chain and food web taught before/together with the principle of inter-dependence of life.
3. Consolidate and extend students' learning:
   i. Tell stories or show videos which illustrate the unwanted consequences of introducing a new species or polluting an environment. In discussion of such effects on the food chain/web, draw students' attention to all the organisms instead of just focusing on humans.
   ii. Discuss the advantages and disadvantages of introducing new species into a certain area with real life scenarios. For example, Gambusia affinis (食蚊魚)[2]/ Mile-a-minute weed (薇金菊) in HK[3], Chinese Mitten Crabs (大閘蟹) in

---

[2] Teachers and students may find the following articles useful:
http://hk.news.yahoo.com/050629/12/1e120.html;
http://www.mingpaoweekly.com/htm/1891/BB01_12.htm

[3] Online article: http://ihome.cuhk.edu.hk/~b105713/article/Mikania.html?s2

River Thames in London[4]. Then teach students to apply knowledge learned from individual cases to general or novel situations. This could stimulate their minds to ideas outside the syllabus and build up their interest in the subject. Teachers should exercise their discretion in deciding when to discuss these examples with the students (e.g. in class or during extracurricular activities) and in what format (e.g. as extension materials in the form of a video show or in the form of posters in the laboratories, or even posing them as thought-provoking questions for the brighter students). This could stimulate students (probably more able students) to examine and think about food chain related issues after class.

iii. Use computer simulation programs/websites to demonstrate the principle of inter-dependence of living things in an ecosystem. For example: World Maker from HKU CITE[5]. The program could be used as a tool to support teaching and learning of how one species may affect the population of other species in the ecosystem. Users can manipulate the conditions of the ecosystem – for example, the population of various species of organisms, a set of simple rules to regulate their interactions, e.g. growth and death rate, etc. In sum, this provides students with a model to visualize how the stability of an ecosystem will be affected if there are any changes in the organisms living there.

iv. Apply the P-O-E strategy (predict, observe and explain) to teach new/unfamiliar concepts of this kind. That is, first, ask students to predict what will happen after initial presentation of the scenario. Then, show them a video or tell them the consequence, followed by the explanation.

[4] Online article: http://news.bbc.co.uk/2/hi/uk_news/england/4690988.stm

[5] Website: http://worldmaker.cite.hku.hk/worldmaker/pages/

2.2.5 Air pollution – Global issues

Content / Concept assessed: Global warming

Relevant Science (S1-3) topic: Living things and air (Unit 7)

The Question (1999)

What is predicted to be a result of global warming?

A. Rising ocean level
B. More severe earthquakes
C. Larger volcanic eruptions
D. Thinning ozone layer

Result

	HK (%)
(A)	58.7
B	2.2
C	11.8
D	26.1

% correct			
HK			International
Boys	Girls	Overall	
64.7	52.8	58.7	33.4

The Question (2003)

The burning of fossil fuels has increased the carbon dioxide content of the atmosphere. What is a possible effect that the increased amount of carbon dioxide is likely to have on our planet?

A. A warmer climate
B. A cooler climate
C. Lower relative humidity
D. More ozone in the atmosphere

Result

	HK (%)
(A)	72.4
B	3.7
C	5.6
D	18.3

% correct			
HK			International
Boys	Girls	Overall	
76.2	68.6	72.4	44.9

Content / Concept assessed: Importance of ozone layer
Relevant Science (S1-3) topic: Not found in the curriculum

The Question (1995)

Write down one reason why the ozone layer is important for all living things on Earth.

Result

Percentage		Response
Overall	Breakdown	
		Correct Answer (Boys: 60.4%; Girls: 49.5%)
55.6 (52.8)*	36.2	Refers to the protection against UV radiation from the sun.
	11.1	Refers to protection against dangerous or too strong radiation from the sun but does not mention UV. *E.g. Because it keeps the sun's ray from being too strong.*
	1.6	Mentions that the ozone layer protects humans so we do not get sunburned/skin cancer.
	10.5	Other correct answers.
		Incorrect Answer (Boys: 32.2%; Girls: 42%)
36.4	0.7	Confuses the effect of the ozone layer with the greenhouse effect. *E.g. It keeps the heat in.*
	3.3	Confuses protection against heat *E.g. Everything will melt without it.*
	2.9	Refers only vaguely to protection. *E.g. All living things will die without the ozone layer. It protects the Earth/us.*
	16.9	Refers to or confuses oxygen, O_2 with ozone, O_3. *E.g. It is needed for respiration.*
	3.5	Sees the ozone layer as a barrier for the atmosphere. *E.g. It keeps the air around the earth.*
	9.1	Other incorrect answers.
		No Answer (Boys: 7.4%; Girls: 8.4%)
7.8	7.8	Blank, crossed out/erased, illegible, or impossible to interpret.

*International average

The Question (2003)

Write down one reason why a "hole" in Earth's ozone layer may be harmful to people.

Result

Percentage		Response
Overall	Breakdown	
		Correct Answer (Boys: 62.5%; Girls: 47.7%)
54.7 (42.1)*	28.4	Both radiation from the sun or similar (ultraviolet or UV) AND a specific harmful effect (e.g. sunburn, skin cancer, skin damage, eye damage, burning). *E.g. Too much ultraviolet rays could damage people's eyes.* *Ultraviolet rays are not blocked so people get skin cancer.* *People get cancer because more of the Sun's rays come through.*
	24	Refers to radiation only with no harmful effect given. *E.g. Because it is letting ultraviolet rays in.* *The deadly rays from the sun get through.* *The rays from the sun get through.*
	1.5	Refers to humans getting sunburned/skin cancer with no explicit mention of radiation. *E.g. Increase in skin cancer.* *It gives people sunburn.*
	0.9	Other correct answers.
		Incorrect Answer (Boys: 29.7%; Girls: 42.6%)
36.5	10.3	Associate the ozone layer with greenhouse effect or protection against heat. *E.g. The hole will let in too much heat, which will melt the ice caps.* *A hole in the Earth's ozone layer could harm people because it is very hot.* *They could get burned.* *Because it might get hot and some people will die from the hole.*
	7.8	Refers to ozone layer as a barrier for the atmosphere: holding the atmosphere in OR protecting the Earth from materials in space getting though (other than UV radiation). *E.g. Rocks might hit the Earth.* *People would be sucked out of the hole.* *If there was a hole the atmosphere might escape.* *It keeps the oxygen from escaping.*
	18.4	Other incorrect answers.
		No Answer (Boys: 7.8%; Girls: 9.6%)
8.8	8.8	Blank

*International average

Noteworthy Points on Students' Performance

Though most HK students performed quite well on the two multiple choice questions that test concepts about global warming, a substantial number of them wrongly associated the phenomenon with ozone. Also, boys outperformed girls in both multiple choice questions. The possible explanations could be:

1. ***Common misconceptions / Curriculum emphasis –*** Students erroneously associate global warming with depletion of the ozone layer. Students gain knowledge about

thinning of the ozone layer from the media (because ozone depletion is not in the curriculum), however they then mistakenly associate the 'hole' in the ozone layer with the greenhouse effect and global warming.

2. ***Gender difference*** – It could be that boys take more interest in physical phenomena than girls do.

Implications / Suggestions for Improvement

1. Revise the curriculum to include the effect of ozone depletion.
2. Teaching should be focused on the "cause and effect" relationship in the issue of global warming and ozone depletion. As an activity, ask students to tabulate the causes and effects of global warming and ozone depletion.

	Causes	**Effects**
Global warming		
Ozone depletion		

3. Alternatively, ask students to cut newspaper articles on global warming and ozone depletion and to sort the articles under the correct topics in a table format.
4. Provide extra practice in answering open-ended questions and develop analytical skills by giving students open-ended questions in textbooks as homework.
5. Provide pictorial illustration of the causes and effects of global warming.
6. Relate issues about ozone depletion and global warming to daily matters. Talk about the function and purpose of UV sunscreen by showing students pictures of skin cancer patients, or let them watch animations simulating rising sea levels due to an increase in Earth's temperature caused by global warming. Pay attention to both local and global pollution/environmental examples, such as deforestation, the hole in the ozone above Australia, etc.
7. Discuss with students the Kyoto protocol - an international commitment to reduce the emission of carbon dioxide and five other greenhouse gases. Talk about the motives and consequences of switching on air conditioners only when the temperature is 25℃ or higher.
8. Ask students to do a research project on ozone depletion and global warming. Require more able students to concentrate on both local and international examples and students of lower ability to concentrate solely on local examples. Then present findings to the whole class.

2.2.6 Air pollution – Regional concerns

Content / Concept assessed: Acid rain

Relevant Science (S1-3) topic: Living things and air (Unit 7)

The Question (1995)

One of the principal causes of acid rain is

A. waste acid from chemical factories being pumped into rivers.
B. acid from chemical laboratories evaporating into the air.
C. gases from burning coal and oil dissolving in water in the atmosphere.
D. gases from air conditioners and refrigerators escaping into the atmosphere.

Result

	HK (%)
A	29.6
B	26.3
(C)	37.7
D	6.2

% correct			
HK			International
Boys	Girls	Overall	
35.1	40.8	37.7	35.1

The Question (2003)

One of the main causes of acid rain is

A. waste from nuclear power plants.
B. spills form chemical manufacturing plants.
C. gases from burning fossil fuels.
D. gases from aerosol spray cans.

Result

	HK (%)
A	12.4
B	24.6
(C)	56.5
D	5.9

% correct			
HK			International
Boys	Girls	Overall	
54.8	58.0	56.5	32.8

Noteworthy Points on Students' Performance

It is evident that students were misled by the term "acid" in distractor options A (30%) and B (26%) in 1995. The two answers mentioned "acid" from factories and laboratories but did not state how acid rain is formed. A similar situation occurred in

2003, with a quarter of the students misled by the option with chemical manufacturing plants in the answer. This could be caused by:

1. ***Inadequate teaching / learning –*** Students cannot visualize how gases from burning acid fossil fuels cause acid rain.
2. ***Everyday life experience –*** In recent years, more emphasis (both in the curriculum and in society) has been placed on acid rain and related environmental issues such as air pollution. This increased exposure resulted in a better performance in 2003. As supported by the fact there is very little difference in performance observed between band 1 and 2 students (data not shown here), it seems students learned the knowledge through daily life interactions rather than specific syllabus topic.

Implications / Suggestions for Improvement

1. Conduct one of the following to demonstrate the phenomenon of acid rain:
   i. Collect exhaust gases from a car and test it with pH test paper (changes colour in contact with sulphur dioxide). Otherwise extract the gases into a delivery tube and pass it through litmus solution to test the acidity. In addition, explain that it is the sulphur dioxide that causes the acidity.
   ii. Perform a test to show that water with nitrogen dioxide added to it will become acidic. First, measure the pH of water to check if it is neutral. Then, mix copper and nitric acid inside a small, sealed bottle. Use a syringe to take out the brown nitrogen dioxide gas produced and transfer the nitrogen dioxide gas into a conical flask with water. Measure the pH of the solution, it should have become acidic. The teacher can relate this experimental observation to the formation of acid rain.
   iii. Videotape the above demonstrations and show to students in future without having to perform the demonstration every year.
2. With the aid of pictorial representations or graphic simulations, discuss in a step-by-step manner the processes of acid rain formation and its effects on the environment, including erosion of surfaces of buildings.
3. Have students write down the causes and effects of global warming, acid rain and ozone depletion in a table. This also clearly demonstrates that all three issues are related to air pollution. For students of lower ability, prepare the causes and effects in statements and ask students to place them in appropriate boxes.

	Causes	**Effects**
Global warming		
Acid rain		
Ozone depletion		

2.3 Important Lessons Learned from Students' Performance

As evident from above, the following are important lessons learned from students' performance in TIMSS:

1. Students' common **misconceptions** need to be taken into consideration in planning and teaching science lessons (see 2.2.1, 2.2.5).
2. Biology teaching (**curriculum emphasis**) needs to look at the organism as a whole rather than emphasizing too much on structural details of individual organs (see 2.2.3, 2.2.4).
3. Science learning can be enhanced by helping students to **connect what they have learned** at different times of the course. This can give them a holistic picture of what they are learning (see 2.2.2, 2.2.4, 2.2.6).
4. Students' **everyday life experience** needs to be taken into consideration in planning and teaching science lessons (see 2.2.1, 2.2.2, 2.2.5, 2.2.6).
5. Science teaching can be enhanced by carrying out **simple experiments / student activities** (see 2.2.3, 2.2.6).
6. A variety of **teaching aids** (including models, pictorial or graphic representations, specimens, toys, photos, videos, computer simulation programmes) need to be employed in teaching science (see 2.2.1, 2.2.3, 2.2.4, 2.2.5, 2.2.6).
7. Science teaching can be enhanced by **out of school activities** e.g. visits to ocean park, science museum (see 2.2.1, 2.2.3).
8. Attention needs to be paid to cultivate **higher order thinking skills** in our students (see 2.2.2, 2.2.4, 2.2.5, 2.2.6).
9. Attention needs to be paid to improve our students' **communication skills** (see 2.2.3).
10. Attention needs to be paid to the vocabulary and phrasing used in the **rubic of the questions** (see 2.2.1, 2.2.4).
11. Science teaching should cater for **students of different abilities** (see 2.2.1, 2.2.6).
12. **Gender differences** in science learning should be considered in planning and teaching science lessons (see 2.2.2, 2.2.5).

Chapter 3

Learning from Students' Performance in Physics-related Questions

Siu Ling Wong

3.1 Performance in General

Table 2.1 (Chapter 2) shows that Hong Kong (HK) students did particularly well in the physics domain. Their achievement outperformed the International Average by 81 scale-points, one of the highest among the five science domains. However, further analysis of the released items in the physics domain indicated a significant gender difference, as shown in Table 3.1 below. Boys outperformed girls in all the three years and in questions of different assessment formats. There is, however, an encouraging sign that girls are catching up in recent years.

Table 3.1 Percentage of released items in which boys outperformed girls in TIMSS 1995, 1999 and 2003

Year	Percentage of Questions in which Boys Outperformed Girls		
	Multiple Choice	**Short Constructed Response Question**	**Long Constructed Response Question**
1995	77%	100%	100%
1999	63%	No released items	80%
2003	58%	67%	63%

A detailed item analysis of the released physics-related questions shows that boys performed significantly better than girls in questions requiring higher order thinking skills, e.g. understanding complex information, reasoning and analyzing. Girls showed slightly better performance over boys in questions requiring factual recall of knowledge. For constructed response questions, girls showed a lower 'no response' rate, indicating that they were more willing to attempt this type of question. This might be associated with their better language skills. Boys also performed much better than girls in questions that tested concepts outside the formal science curriculum. In fact, most of the released physics-related questions are not covered by the General Studies curriculum for primary schools or the Science curriculum for Secondary 1-3. In other words, the better performance of boys could be related to other factors, such as the differences in hobbies and outside class activities between boys and girls.

3.2 Performance in Individual Items

The physics-related topics covered by the HK Science (Secondary 1-3) curriculum are *energy* (Energy – Unit 4), *matter* (Matter as Particles – Unit 6), *electricity* (Making Use of Electricity – Unit 8), *force and motion* (Space Travel – Unit 9), and *optics* (Light, Colours and Beyond – Unit 15). Among the TIMSS released items, students showed very good performance in questions related to *electricity*. The released items under electricity are all in multiple-choice format. Similarly for questions related to *matter*, students' performance in the multiple-choice format outperformed their international counterparts. However, students showed very poor performance in two other released items that required them to explain a given phenomenon using the concepts of particle theory. This may indicate that our students have difficulty in applying the concepts and are weak in communicating in an organized manner. In the following, we focus on the analysis of the remaining three topics, namely *force and motion*, *optics* and *energy*. Six sets of related items from different years are discussed in detail.

3.2.1 Force and motion

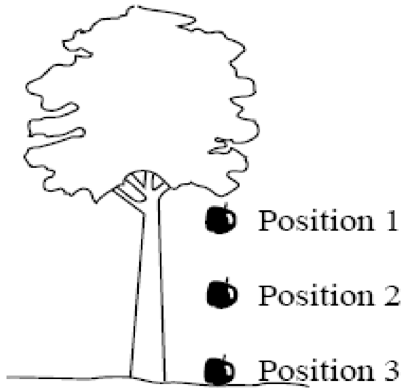
Content / Concept assessed: Gravity

Relevant Science (S1-3) topic: Space Travel (Unit 9)

The Question (1995)

The drawing shows an apple falling to the ground. In which of the three positions does gravity act on the apple?

A. 2 only
B. 1 and 2 only
C. 1 and 3 only
D. 1, 2 and 3



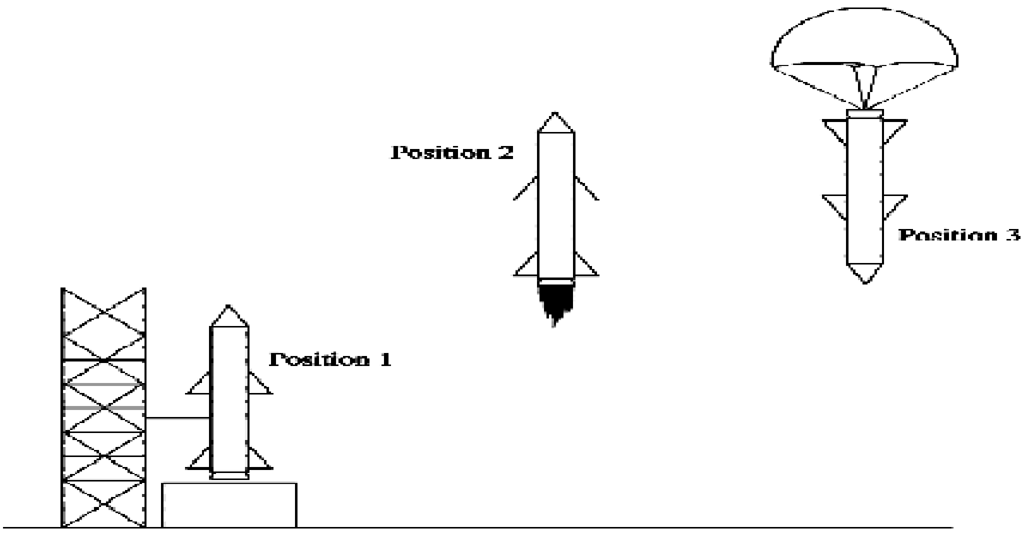
Result

	HK (%)
A	3.4
B	18.6
C	2.6
(D)	74.4

% correct			
HK			International
Boys	Girls	Overall	
76.8	72.0	74.4	54.6

The Question (1999)

The drawings show a rocket being launched from Earth and then returning.



In which of the three positions does gravity act on the rocket?

A. 3 only
B. 1 and 2 only
C. 2 and 3 only
D. 1, 2 and 3

Result

	HK (%)
A	22.1
B	24.9
C	28.6
(D)	23.9

% correct			
HK			International
Boys	Girls	Overall	
25.7	22.1	23.9	36.1

Noteworthy Points on Students' Performance

Almost 75% of HK students answered the question on force and motion correctly in 1995. However, 19% chose the distracter option B. In 1999, students chose the four options in almost equal percentage – an indication of random choice. These outcomes could be attributed to the following reasons:

1. ***Common misconceptions –*** Students who chose option B (19%) in 1995 might think that gravity no longer acts on an object when it is on the ground. They do not realize that gravity always acts on an object regardless of whether it is in mid air or on the ground.
2. ***Context of question –*** In the 1999 question, many students might have been confused by the different moving directions of the rockets and erroneously thought that the direction of the motion has an influence on the gravitational attraction on the rocket. This would partly explain the almost random response to the 1999 question. On the other hand, the scenario used in the 1995 question was more related to daily life and hence easier for students to comprehend. The question in 1999 was more difficult because the context was much more complex and abstract for the students.
3. ***Curriculum coverage –*** Gravity was not covered in the junior science curriculum before 2000. Space travel is a new topic since 2000. However, there was no similar released item in TIMSS 2003 available for comparison. Nevertheless, the knowledge of students' common misconception about gravity should help teachers in designing appropriate learning and teaching activities to address the issue. It will be interesting to see if students have better performance in future years.

Implications / Suggestions for Improvement:

1. To help students understand the concept that gravity acts on objects even when they are stationary on the ground, teachers can perform the following simple demonstration. Have students hold two heavy objects (i.e. weights) in both hands. The use of force (from students) to hold the objects in a stationary position could illustrate to students that gravity is acting on objects even when they are stationary.
2. Use magnetic force, another type of non-contact force, as an analogy for gravitational force. Place a piece of strong magnet on the floor and ask a student to hold a piece of metal directly above the magnet. When the student lets go of the metal, the magnetic force pulling the metal down is compared to the pull of gravity on an object. To demonstrate the existence of the downward gravitational pull even when the object is sitting on the ground, a similar analogy can be used. Place a piece of metal on a piece of paper sitting on a piece of magnet (i.e. imagine

the piece of paper as the ground surface). Then ask the students to try pulling the metal up. By doing so, students should be able to feel the downward attractive magnetic force acting on the piece of metal. The analogy of magnetic force could thus be used to highlight that gravitational force is acting on an object whether it is in mid-air or it is stationary on the ground.

3. Other simple demonstrations could also be performed:
   i. Get a piece of paper and soak it well. Stretch it horizontally between clamps. Place a light object on top of the paper and ask if students believe gravity is acting on it. Place a second object on it and so on until the paper splits. Ask them once more if they suppose gravity is acting on it. If students answer yes, ask them what is the difference between before and after the paper split.
   ii. Put a piece of scrub on the table. Place an object of certain mass on top, e.g. apple, cup, etc., and let students observe how the scrub is indented by the object. Then explain to students that the scrub is indented because of gravity.

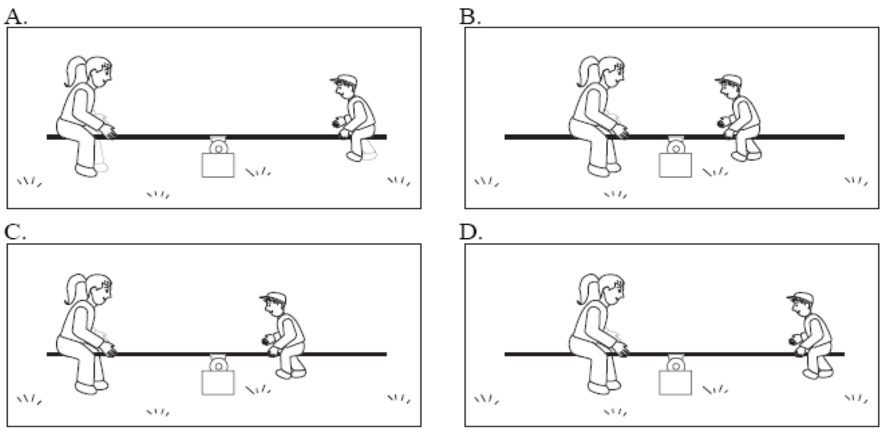
3.2.2 Force and Motion

Content / Concept assessed: Moment

Relevant Science (S1-3) topics: Not found in the curriculum

The Question (1995)

A girl wanted to play on a seesaw with her little brother. Which picture shows the best way for the girl, who weighed 50 kg (kilograms), to balance her brother, who weighed 25 kg?



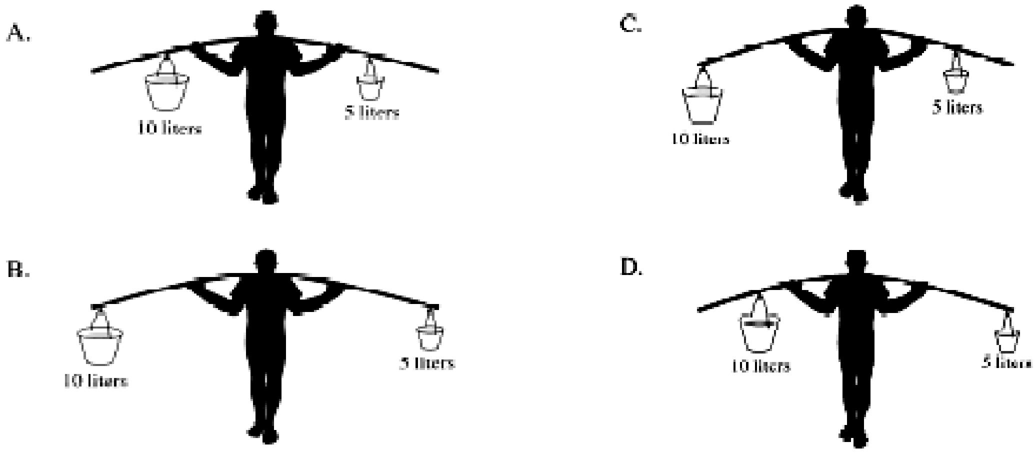
Result

	HK (%)
A	3.1
B	3.9
C	18.5
(D)	72.8

% correct			
HK			International
Boys	Girls	Overall	
77.0	67.6	72.8	71.5

The Question (1999)

Which picture shows the best way for the man to balance a ten-litre bucket water and a five-litre bucket of water?



Result

	HK (%)
A	9.2
B	7.5
C	10.9
(D)	71.8

% correct			
HK			International
Boys	Girls	Overall	
77.3	66.5	71.8	71.0

Noteworthy Points on Students' Performance

In both years, HK students performed similarly to the International Average of around 70%. For the remaining 30% of students, the sources for their mistakes might be:

1. ***Common misconceptions*** – Students could have assessed moment subjectively (intuitively) rather than objectively. They might intuitively relate bigger and heavier objects with longer distances, thus they paired up the heavier bucket with the longer end. Students at this level might be more familiar with direct proportional relationship than inverse relationship.

2. ***Everyday life experience* –** As there are no seesaws in any parks in HK anymore, students have no experience of playing with a seesaw. Also, people do not use the traditional Chinese scale for weighing anymore.
3. ***Gender difference* –** An example of questions with an unfamiliar context in which boys performed significantly better than girls. In this case, boys might have more experience about moment from building models, playing with toy cars, etc.

Implications / Suggestions for Improvement

1. Teachers should emphasize the importance of analyzing data objectively. When conducting experiments on moment, allow students to move the location of the two objects and ask them to draw their own conclusion.
2. Have students construct a leverage model using a ruler, a prism and two pieces of eraser of different size. It works a lot better if the 'weights' are hung from the ruler (using cotton thread) in small 'baskets' or 'cradles'.
3. Set questions on moment using real life scenarios, for example, pushing a door to open or opening the lid of a can by means of a spoon. In the first case, teachers could draw students' attention to the position of the handle on a door. In the latter case, teachers could ask students at which position they should hold the handle of the spoon for opening the lid with least effort.
4. Through science club activity, introduce the Chinese scale to students and ask them to practice using it. This also allows students to appreciate how ancient Chinese developed artifacts by applying knowledge about moment in a real life context.
5. Girls should be encouraged to take part in more hands-on activities both in and outside classrooms.

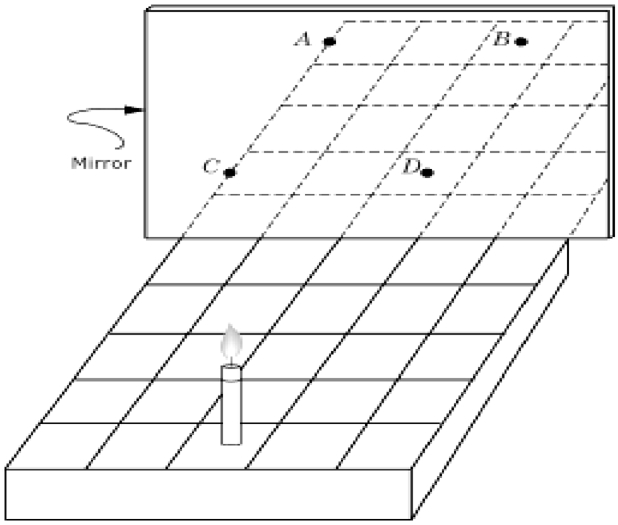
3.2.3 Optics

Content / Concept assessed: Reflection of plane mirror – Image formation

Relevant Science (S1-3) topic: Light, colour and beyond (Unit 15)

The Question (2003a)

A candle is placed on a ruled grid in front of a mirror, as shown. At what point will the reflection of the candle appear to be?



A. Point A

B. Point B

C. Point C

D. Point D

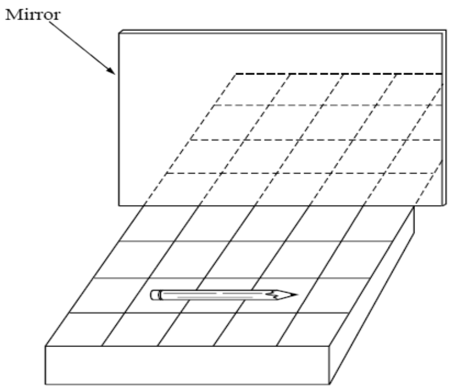
Result

	HK (%)
A	5.1
(B)	80.2
C	2.4
D	12.4

% correct			
HK			International
Boys	Girls	Overall	
82.3	77.9	80.2	64.6

The Question (1995)

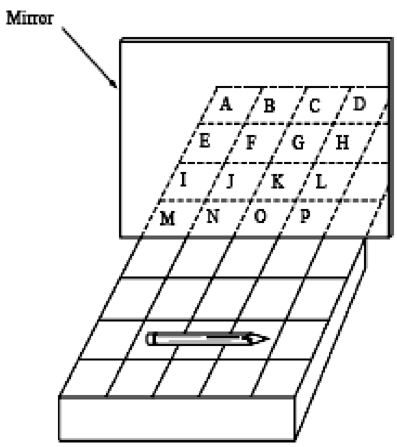
The picture shows a pencil that is lying on a shelf in front of a mirror. Draw a picture of the pencil as you would see it in the mirror. Use the patterns of lines on the shelf to help you.



Result

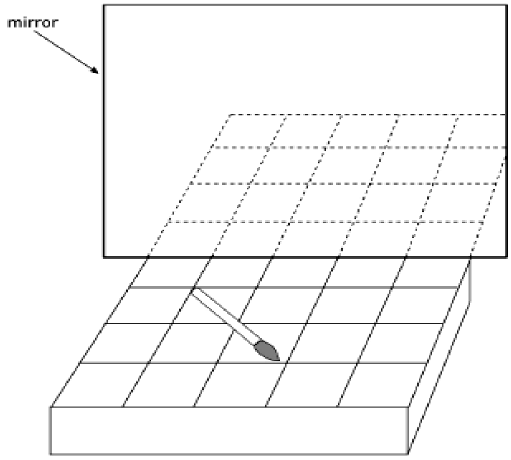
Percentage		Response
Overall	Breakdown	
		Correct Answer (Boys: 85.7%; Girls: 76.6%)
81.3 (68.9)*	78.3	FGH; pencil point to the right.
	0.3	FGH; point not shown.
	2.7	FG or GH, (point to the right either shown or not shown) OR any other in the row E, F, G, H as long as the point is not clearly turned to the left.
		Incorrect Answer (Boys: 10.5%; Girls: 21.2%)
15.6	7.2	FGH, FG or GH, pencil point clearly turned to the left OR other in the row E, F, G, H.
	0.5	Lists all or some part of the row: ABCD
	5.5	Lists all or some part of JKL; pencil point to the right may or may not be shown.
	2.4	Other incorrect answers.
		No Answer (Boys: 3.8%; Girls: 2.2%)
3.1	3.1	Blank, crossed out/erased, illegible, or impossible to interpret.

*International average



The Question (2003b)

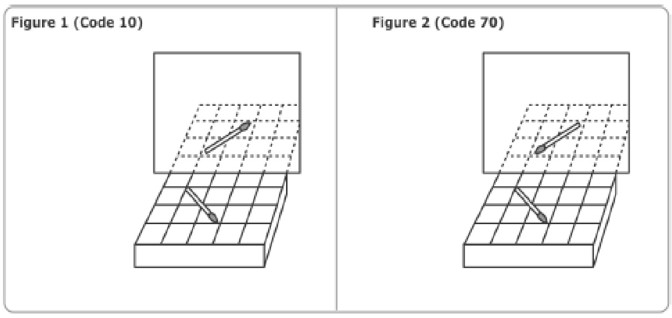
The picture shows a paint brush that is lying on a shelf in front of a mirror. Draw a picture of the paint brush as you would see it in the mirror. Use the pattern of lines on the shelf to help you.

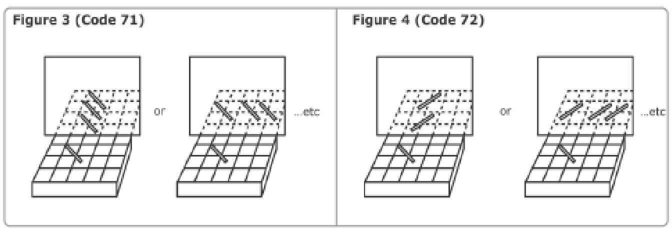


Result

Percentage: Overall	Percentage: Breakdown	Response
		Correct Answer (Boys: 60%; Girls: 60.9%)
60.5 (44.4)*	60.5	Correct placement: angle, grid position and direction (hairs to the right). (See Figure 1.)
		Incorrect Answer (Boys: 35.5%; Girls: 33.4%)
35.4	3.7	Correct angle and placement, but image flipped with hairs clearly turned to the left. (See Figure 2.)
	23.3	Image parallel to original (hairs to the right or left). (See Figure 3; other rows/columns are possible).
	3.6	Correct angle but image translated (hairs to the right or left). (See Figure 4; other rows/columns are possible).
	4.8	Other incorrect answers.
		No Answer (Boys: 3.7%; Girls: 4.6%)
4.1	4.1	Blank, crossed out/erased, illegible, or impossible to interpret.

*International average





Noteworthy Points on Students' Performance

Students performed consistently well on questions about reflection when the object was placed parallel to the mirror (about 80% in both 1995 and 2003a). Although optics is only addressed at Secondary 3, the topic of reflection is briefly covered in primary level. Thus students should be aware that the image is laterally inverted, and possibly aware that the distance of the image from the mirror is the same as that of the object from the mirror. However, when the object was placed diagonally in front of the mirror (as in 2003b), students' performance dropped considerably (60%).

1. ***Common misconception*** – When the object is placed diagonally, 23% of the students intuitively thought that the image is *parallel* to the object.
2. ***Application of knowledge*** – When the object is placed diagonally, students were distracted and failed to apply the two properties of the image: namely, lateral inversion and image distance *at every point* on the object is equal to that of the object.
3. ***Gender difference*** **–** Boys are better in spatial visualization and thus outperformed girls by almost 10% in 1995. However, when the situation became more complicated, as in the diagonally placed paint brush in 2003b, the gender difference disappeared.
4. ***Lack of practice*** - Students lack of practice in drawing reflection of objects with different orientations, let alone real objects.

Implications / Suggestions for Improvement

1. Emphasize the two properties of the image, namely lateral inversion and image distance at every point on the object is equal to that of the object. Then coach students to first locate the two end points of a mirror image and then connect the two points to create the image.
2. To address the common misconception, teachers could make use of the 2003 question with a P-O-E approach. Teachers could ask the students to first draw the predicted mirror image (Predict). Instead of just giving out the correct answer, teachers could provide students with a mirror and a pen and let the students find out for themselves (Observe). Next, ask them if there is any difference between the actual mirror image and their drawing. Then ask them to explain the difference (Explain).
3. Drawings of the mirror image of an object commonly found in textbooks are often restricted to that of an arrow or a candle. This is not authentic enough. Teachers may extend the image drawing exercise by including objects seen and used by students in their everyday life, such as a pen, a letter, a car, etc.

4. The revised Mathematics Curriculum since 2000 (both primary and secondary) also covers drawing of reflection, translation, rotation of a figure. Teachers should help students integrate knowledge learned from mathematics and science to reinforce their understanding of the properties of a mirror image.

3.2.4 Energy

Concept assessed: Energy efficiency – an analytical problem in a physics context
Relevant Science (S1-3) topic: Energy (Unit 4)

The Question (1995)

Machine A and Machine B are each used to clear a field. The table shows how large an area each cleared in 1 hour and how much gasoline each used.

	Area of field cleared in 1 hour	**Gasoline used in 1 hour**
Machine A	2 hectares	3/4 litre
Machine B	1 hectare	1/2 litre

Which machine is more efficient in converting the energy in gasoline to work? Explain your answer.

Result

Percentage		**Response**
Overall	**Breakdown**	
		Correct Answer (Boys: 29.3%; Girls: 22.8%)
26.3 (36.1)*	20.5	A. Because 3/8 < 1/2, or a similar expression.
	5.8	A. Other correct answers.
		Incorrect Answer (Boys: 49.7%; Girls: 55.8%)
52.6	15.5	B. With an explanation.
	1.8	B. No explanation.
	0.2	States that A and B are the same, with an explanation.
	4.3	A. Mentions that it clears the most area per hour.
	20.4	A. Other wrong explanations.
	4.2	A. No explanation.
	6.2	Other incorrect answers.
		No Answer (Boys: 21.1%; Girls: 21.4%)
21.2	21.2	Blank, crossed out/erased, illegible, or impossible to interpret.

*International average

The Question (1999)

Machine A and Machine B are each used to pump water from a river. The table shows what volume of water each machine removed in one hour and how much gasoline each of them used.

	Volume of Water Removed in 1 Hours (litres)	**Gasoline Used in 1 Hour (litres)**
Machine A	1000	1.25
Machine B	500	0.5

a) Which machine is more efficient in converting the energy in gasoline to work?

Answer: ________________

b) Explain your answer.

Result

Percentage		**Response**
Overall	**Breakdown**	
		Correct Answer (Boys: 33.4%; Girls: 30.2%)
31.8 (30.7)*	31.8	With correct explanation based on the concept of energy efficiency (B uses less gasoline than A for an equivalent volume of water pumping). *E.g. B. Because it used 1L of gas per 100L of water and A used 1.25L for 100L of water.* *B. Because it uses less gasoline per litre of water.* *B. Because it can pump the same amount of water using less gas.*
		Incorrect Answer (Boys: 60.7%; Girls: 65.8%)
63.3	17.7	B. It uses less gasoline (no comparison of efficiency based on volume of water pumped). *E.g. B. It used 1L of gas in an hour and A used 1.25L in an hour.* *B. It uses less gas in an hour.* *B. The engine used .75L less gas.*
	15.2	B. Other incorrect/inadequate or no explanation.
	9.4	A. It removes more water in 1 hour (no comparison of efficiency based on gas used). *E.g. A. It pumps more water.* *A. It pumps 1.25 litres and B only pumps 0.5 litres.* *A. It pumped 500L more than B.*
	18.9	A. Other incorrect/inadequate or no explanation.
	2.1	Other incorrect answers.
		No Answer (Boys: 6%; Girls: 3.9%)
5.0	5.0	Blank, crossed out/erased, illegible, or impossible to interpret.

*International average

Noteworthy Points on Students' Performance

HK students performed 10% below the International Average in 1995 and almost two times less than the top-performing country (UK – 51%). Whereas, in 1999, students' performance was comparable to the International Average, it was still 20 % below the highest percentage of correct responses (Korea – 52%). This result could be due to the following reasons:

1. ***Unfamiliar terminologies / Context of question –*** HK students are not familiar with the term "hectare" and "efficiency", therefore they left the question blank (about 20 % in 1995). Also, the context of mowing the lawn or clearing a field is unfamiliar to HK students.
2. ***Misinterpretation of data –*** One quarter of the students in 1995 chose the wrong answer (machine B), perhaps because they assumed less gasoline used equals less energy used, and greater efficiency. It is likely that they experienced confusion between rate and efficiency and thus regarded the time the machines spent to clear the field as an indicator of efficiency. Student perhaps considered both questions to be more mathematics related rather than science related. Students also found it hard to comprehend the implications of positive and negative correlation.
3. ***Weakness in calculation –*** Students are generally weak in solving problems involving proportionality and rate. The situation is worse since the change in the junior secondary mathematics curriculum in 2000. In addition, HK students are weak in fraction calculation, but somewhat more comfortable with decimal calculations.

Implications / Suggestions for Improvement

1. Teachers can try constructing questions on efficiency of machines based on scenarios/frameworks with which students are familiar. Such as, how many miles can a car travel with a tank full of fuel? What would be the difference between using unleaded and diesel oil? For the same amount of washing detergent, which type, normal or concentrated, will last for a longer period of time?
2. Many students regard it as a mathematics problem, unrelated to science. However, teachers could encourage students to use multi-perspectives logical thinking when solving problems. The concept of fair test could be applied in solving this problem by *creating* a fair condition for comparison. Thus for solving the 1999 question, students could compare the 'amount of gasoline required' for the *same* 'volume of water removed'. So by simply considering the gasoline required when 1000 L of water is to be removed by Machine B, an instant comparison can be made.
3. Recent education reforms advocate inter-disciplinary co-operation among subject panels in schools. And to achieve this objective, teachers of different disciplines should gather to discuss the needs of different subjects. Teachers can then adjust the syllabus/teaching schedule accordingly so they are working collaboratively to facilitate and enhance students' overall learning and their appreciation of the interconnection of different disciplines, as in real life situations.

3.2.5 Energy

Content / Concept assessed: Light energy

Relevant Science (S1-3) topic: Energy (Unit 4)

The Question (1995)

A flashlight close to a wall produces a small circle of light compared to the circle it makes when the flashlight is far from the wall. Does more light reach the wall when the flashlight is further away?

___Yes

(Check one)

___ No

Explain your answer.

Result

Percentage		Response
Overall	Breakdown	
		Correct Answer (Boys: 19.6%; Girls: 12.7%)
16.6 (22.7)*	3.8	No. Explains that the same amount reaches the wall.
	10.9	No. Explains that the same amount of light reaches the wall but more spread out.
	0.7	No. Less light reaches the wall because of light absorption by the air.
	1.2	Other correct.
		Incorrect Answer (Boys: 59.1%; Girls: 66%)
81.1	19.4	Yes. Because light covers a bigger area. *E.g. Because the light covers a bigger area.* *Because light can expand.* *Because light spreads out.*
	24.0	Yes. No explanation.
	5.4	No. There is less light at a greater distance.
	13.2	No. No explanation.
	19.1	Other incorrect answers.
		No Answer (Boys: 21.4%; Girls: 21.3%)
2.3	2.3	Blank, crossed out/erased, illegible, or impossible to interpret.

*International average

The Question (1999)

James turns on a flashlight in his bedroom and shines it on his wall one metre away to produce a small circle of light. He then shines the flashlight on his ceiling two metres away to produce a larger circle of light.

a) Does more light reach the ceiling than the wall? (check one)

☐ Yes

☐ No

b) Explain your answer.

Result

Percentage		Response
Overall	**Breakdown**	
		Correct Answer (Boys: 41.7%; Girls: 37.8%)
39.9 (24)*	8.1	No. Explains that the same amount of light reaches the wall and ceiling or refer to light being More spread out on the ceiling or less concentrated/focused/bright.
	30.5	No. Explains (or shows in a diagram) that light is (only) More spread out (less bright) at a greater distance. (Does not explicitly state that the light is the same.)
	1.1	No. Explains that less light reaches the ceiling because of More air absorption/scattering at a greater distance.
	0.2	No. Other correct explanation.
		Incorrect Answer (Boys: 57%; Girls: 61.5%)
59.3	5.9	No. States that less light reaches the ceiling with inadequate explanation related to distance from source. (Does not include explanation of less light due to air absorption or scattering as in Code 12).
	28.1	No. Other incorrect/inadequate or no explanation. (Includes explanations that merely paraphrase the stem).
	8.6	Yes. Explanation based on light being bigger or more spread out.
	15.1	Yes. Other incorrect/inadequate or no explanation.
	1.6	Other incorrect answers.
		No Answer (Boys: 1.3%; Girls: 0.7%)
1.0	1.0	Blank, crossed out/erased, illegible, or impossible to interpret.

*International average

Noteworthy Points on Students' Performance

In 1995, students' performance was weak on the question about light energy (16%) and was below the International Average (23%), the situation improved greatly in 1999 (40%). Several possible explanations for this occurrence:

1. ***Curriculum emphasis*** **–** Though students have learned the concept of energy, it is limited to its various forms and conversion from one form to the other, there is little discussion on light intensity or sound intensity. Since students have little

factual knowledge, they could only use subjective quantifying terms such as *more/less* or *have/do not have* to judge the question. As a result they reasoned, the larger the area, the greater the amount of energy.

2. ***Context of question –*** The context of the question in 1995 was a bit more abstract than the one in 1999, which makes reference to a concrete everyday life experience. The latter is easier for students to comprehend.
3. ***Gender difference –*** Boys' observation skills seem to be better than girls.

Implications / Suggestions for Improvement

1. The question has informed us about the possible intuition of our students about light intensity. Teachers should stress the difference between the amount of energy and energy intensity.
2. The explanation of light intensity can be assisted by the use of an analogy. For example, the analogy of water coming through a shower may give a better visualization of the abstract concept of light intensity. With the same water flow rate, the water spreads over a smaller wall surface and makes the area 'wetter' when the shower is placed close to the wall. On the other hand, the water spreads over a larger wall surface and makes the area 'less wet' when the shower is placed further from the wall. Other similar analogies include: (a) a toy gun emitting many tiny pellets, sprinkling pepper on the surface of a bowl of soup from different heights, etc.
3. Link to everyday life experience by asking students to compare the brightness of a certain area of a surface when it is illuminated with a torch held (i) perpendicular to a wall and (ii) tilted at an angle. The fact that boys performed slightly better than girls may be due to their prior observation of a similar situation when they played with a torch. Perhaps the analogy with the example of 'sprinkling pepper' may work better for girls. There will be a more extensive discussion on the context-dependent gender difference in Chapter 6.

3.2.6 Energy

Content / Concept assessed: Different forms of energy

Relevant Science (S1-3) topic: Energy (Unit 4)

The Question (1995)

Electrical energy is used to power a lamp. Is the amount of light energy produced more than, less than, or the same as the amount of electrical energy used?

The amount of light energy produced is

____ more than

____ less than (check one)

____ the same as the amount of electrical energy used.

Give one reason to support your answer.

Result

Percentage		Response
Overall	Breakdown	
		Correct Answer (Boys: 10.3%; Girls: 3.9%)
7.3 (7.7)*	5.1	Less. Mentions that (much) energy is transformed to heat.
	0.2	Less. Mentions that energy is needed to warm up the lamp.
	0.1	Less. Mentions that energy (heat) is lost to the surroundings.
	2.0	Less. Other correct explanation.
		Incorrect Answer (Boys: 86.6%; Girls: 92.3%)
89.2	34.1	The same. With erroneous explanation. *E.g. Energy is always preserved.* *When the sun is out you don't need electrical energy.*
	15.6	The same. No explanation is given.
	16.5	More. With or without explanation.
	8.4	Less. No explanation.
	3.3	Less. Energy is lost in transport.
	10.6	Less. Other erroneous explanations.
	0.7	Other incorrect answers.
		No Answer (Boys: 3.1%; Girls: 3.9%)
3.4	3.4	Blank, crossed out/erased, illegible, or impossible to interpret.

*International average

The Question (1999)

Electrical energy is used to power a lamp. How does the amount of electrical energy used compare to amount of light energy produced?

a) The amount of electrical energy used is (Check one)

□ more than the amount of light energy produced.
□ less than the amount of light energy produced.
□ the same as the amount of light energy produced.

b) Give a reason to support your answer.

Result

Percentage		Response
Overall	**Breakdown**	
		Correct Answer (Boys: 14.4%; Girls: 8.5%)
11.5 (7.5)*	6.0	More. With an explanation based on (much) energy being converted to heat. *E.g. More. When a lamp is on it heats up. So some of the electrical energy goes into heat instead of light.* *More. Some of the electrical energy is changed into heat energy to make the lamp work and only a small amount of energy is actually changed into light.*
	5.1	More. With other correct explanation of energy (power) loss or low efficiency. [Must clearly indicate that "some" energy goes elsewhere]. *E.g. More. Because some of the energy is lost in conversions.* *More. The lamp is not very efficient, so not all of the electrical energy is changed into light.*
	0.4	More. Other correct explanation.
		Incorrect Answer (Boys: 82.5%; Girls: 89.5%)
85.9	19.7	More. Incorrect/inadequate or no explanation. *E.g. More. The light gives out lots of energy.* *More. The people have got to make the electricity first.* *More. Because of all the electricity from the wall socket.*
	34.0	Same. Explanation is based on the concept of conservation of energy without considering energy losses. *E.g. Same. Energy is always preserved.* *Same. You cannot create or destroy energy, so it has to be the same.*
	2.4	Less. Explanation based on heat, energy losses or low efficiency but with an incorrect application. *E.g. Less. Some of the electrical energy is changed into heat.* *Less. The lamp is not very efficient.*
	14.7	Less. Any other incorrect or no explanation. *E.g. Less. Today's lamps do not use up much voltage.* *Less. It does not take much electricity to power a lamp.*
	15.1	Other incorrect answers.
		No Answer (Boys: 3%; Girls: 2.1%)
2.5	2.5	Blank, crossed out/erased, illegible, or impossible to interpret.

*International average

Noteworthy Points on Students' Performance

A high percentage of students in both years (50% in 1995; 34% in 1999) answered that energy input and output were the same. Possible explanations include:

1. ***Common misconceptions*** – Students were familiar with the concept of conservation of energy but failed to consider energy 'loss' in the process of conversion. As mentioned above, the concept of energy efficiency is usually not elaborated in the junior science classroom. For the processes of energy conversion, teaching mainly focuses on the change from one form of energy to other form(s) of energy, e.g. electrical energy → light energy + heat energy (strictly speaking, internal energy in senior form). There is little discussion of which form of the energy is 'useful' and which form(s) of the energy is (are) undesirable but often unavoidable.
2. ***Lack of knowledge*** **–** Some students may not be aware that there is heat lost to the surroundings in the situation described in the questions.

Implications / Suggestions for Improvement

1. Ask students to put their hands close to a light bulb to demonstrate to them that electrical energy is converted to both light energy and heat.
2. Another way of explaining energy loss through conversion could be done by discussing how energy-saving light bulbs work. The 'energy saving' principle is by reducing the amount of heat energy produced, hence requiring a lesser amount of electricity.
3. Use analogies or model representations to illustrate this concept. For example, hypothesize a cake as the total amount of energy. The manner in which you can give different people various amounts of the cake represents the ways in which energy is transformed into different forms. The debris that falls off in the process of transfer is similar to heat lost to the surroundings.
4. Use examples of household appliances and energy labels when talking about energy efficiency. Teachers may explain the unwanted energy loss of different electrical appliances, e.g. unwanted sound energy generated in vacuum cleaners and washing machines. Teachers may also explain the inevitable energy dissipation of most electrical appliances due to the heating effect of electrical wire and various electrical components. Consequently, teachers could ask students why awareness of energy efficiency is important as a responsible citizen.

3.3 Important Lessons Learned from Students' Performance

As evident from above, the following are important lessons learned from students' performance in TIMSS:

1. Students' **common misconceptions** need to be taken into consideration in planning and teaching science lessons (see 3.2.1, 3.2.2, 3.2.3, 3.2.5, 3.2.6).
2. Students' **everyday life experience** needs to be taken into consideration in planning and teaching science lessons (see 3.2.1, 3.2.2, 3.2.5, 3.2.6).
3. Students' learning can be enhanced by **use of analogies** (see 3.2.5, 3.2.6).
4. Provide students' with **more opportunities to practice** drawing reflection of objects (see 3.2.3).
5. Students' learning can be enhanced with a **P-O-E approach** (see 3.2.3).
6. Science teaching can be enhanced by carrying put **simple experiments / student activities** (see 3.2.1, 3.2.2, 3.2.6).
7. Attention needs to be paid to **curriculum emphases** (see 3.2.5) that may lead to students' inadequate understanding. Teachers should exercise their professional judgement to decide whether they should cover some extended topics for the purpose of completeness so that students can learn the topics better.
8. **Curriculum coverage** should be closely examined by taking into consideration students' performance in international comparative studies (see 3.2.1).
9. Students' understanding can vary a lot with a slight change in the **context of the questions** (see 3.2.1, 3.2.5). **Unfamiliarity of terms / context of question** can affect students' understanding of the question (see 3.2.4).
10. Efforts need to be made to enhance students' ability to **apply knowledge** learned (see 3.2.3).
11. Efforts need to be made to improve students' **data interpretation skills** when answering questions (see 3.2.4).
12. Attention needs to be paid to improving students' **communication skills** and addressing **weakness in calculation** (see 3.2.4).
13. **Gender differences** in science teaching should be considered in planning and teaching science lessons (see 3.2.2, 3.2.3, 3.2.5).

Chapter 4

Learning from Students' Performance in Chemistry-related Questions

Man Wai Cheng

4.1 Performance in General

As shown in Table 2.1 (Chapter 2), HK students' achievement in the chemistry domain outperformed the International Average by 68 scale-points. While this is encouraging, it is also noted that the gap between the HK score and the International Average in this domain was the smallest of the five content domains. This may be attributed to the mismatch between the HK Science (S1-3) curriculum and the assessed curriculum of TIMSS. Topics such as 'Subatomic Particles' and 'Element (Unit 13)/Mixture (Unit 14)/Compound (Unit13)' are either not explicitly addressed in the local curriculum and textbooks or not covered by the Secondary 2 level, when students took the assessment. It follows that many students would have difficulties in answering questions which demand an understanding or application of concepts such as 'atom', 'ion', 'molecule' and 'compound'. The following two questions and their corresponding item statistics (see Table 4.1) illustrate the performance of HK students compared with the International Averages and those of the top-performing countries in TIMSS.

1. When chlorine gas reacts with sodium metal, what type of substance is formed?
   A. A mixture
   B. A compound *
   C. An element
   D. An alloy
   E. A solution

2. The nucleus of MOST atoms consists of
   A. neutrons only
   B. protons and neutrons *
   C. protons and electrons
   D. neutrons and electrons

* the correct answer

Table 4.1 Students' achievements on Items 1 and 2 among the top-performing countries

Country (in order of ranking in TIMSS 2003)	**Percentage correct**	
	Item 1	**Item 2**
1. Singapore	58.1	63.6
2. Taiwan	59.7	63.9
3. Korea	48.8	42.6
4. Hong Kong	**31.9**	**29.9**
5. Estonia	57.2	68.8
6. Japan	56.6	36.5
International Average	**42.1**	**46.8**

It should be emphasized that although HK students' performance in these two items were the lowest among the top-performing countries and also lower than the International Average, their overall performance in the chemistry domain (542 scale-points) was still significantly higher than the International Average (474 scale-points). One might conclude that such a mismatch between the content of the curriculum and the content of the TIMSS assessment items is unfair to HK students. However, in such a large scale international comparative study, in which validity is important, it would be infeasible just to assess the curriculum that is covered by *all* countries. Inevitably, mismatches occurred in other countries, too. Interestingly, research shows that overall ranking does not change much when items with content not covered in the local curriculum are excluded from the analyses. In other words, if the test is really 'unfair' to HK students, it is equally 'unfair' to all participating countries.

4.2 Performance in Individual Items

We shall now examine students' performance in greater detail in order to ascertain the common misconceptions they have on topics like: (1) energy changes associated with chemical reaction, (2) properties of carbon dioxide and oxygen, (3) filtration, (4) water cycle, (5) composition of air, and (6) particulate structure of matter. The implications and suggestions on teaching and learning will then follow.

4.2.1 Energy change

Content / Concept assessed: Energy changes associated with chemical reactions

Relevant Science (S1-3) topic: Energy (Unit 4)

The Question (1995)

When oil is burning, the reaction will

A. only release energy

B. only absorb energy

C. neither absorb nor release energy

D. sometimes release and sometimes absorb energy depending on the oil

Result

	HK (%)
(A)	82.1
B	4.3
C	5.4
D	7.1

% correct			
HK			International
Boys	Girls	Overall	
83.3	81	82.1	50.6

The Question (1999)

If you are burning wood, the reaction will

A. release energy

B. absorb energy

C. neither absorb nor release energy

D. sometimes release and sometimes absorb energy depending on the kind of wood

Result

	HK (%)
(A)	69.9
B	4.4
C	7.5
D	17.7

% correct			
HK			International
Boys	Girls	Overall	
73.6	66	69.9	55.4

The Question (2003)

Some chemical reactions absorb energy, while others release energy. Of the chemical reactions in burning coal and exploding fireworks, which will release energy?

A. Burning coal only
B. Exploding fireworks only
C. Both burning coal and exploding fireworks
D. Neither burning coal nor exploding fireworks

Result

	HK (%)
A	11.7
B	10.5
(C)	74
D	3.6

% correct			
HK			International
Boys	Girls	Overall	
76.3	71.6	74	51.9

Noteworthy Points on Students' Performance

1. ***Everyday life experience*** – Although the 1995 and 1999 questions differ only in the context – burning oil and burning wood – the percentage of students who opted for the correct answer differed by more than 10% (82% vs. 70% respectively). This might be attributed to students' experiences in burning these two substances. While oil is flammable, we usually have to heat up a piece of wood for an extended period of time (the wood absorbs energy) before it catches fire. With this in mind, some students (17.7%) might then have wrongly chosen D as the correct answer in the 1999 item.
2. ***Use of terminology*** – Some students might have confused the term 'burning' and 'heating' and used them interchangeably. This is further complicated by the Chinese word '燒' which in English could mean either 'to burn' or 'to heat'. The word '燒' in '燒熱水' means heating up water, in which water absorbs energy. On the other hand, '燒煤' is burning of coal, which releases energy. The dual meaning of the Chinese word could have confused some students in thinking that 'burning' and 'heating' are identical in meaning[6].

[6] For more information about the use of Chinese terms in science/chemistry context, please refer to: 岑紹基 (2003) 中學會考化學科專科語體資料冊. 香港: 香港大學教育學院母語教學教師支援中心

Implications / Suggestions for Improvement

1. Everyday life experiences of students have an impact on students' understanding of abstract scientific concepts. Teachers could foster students' interest and understanding by discussing more daily examples or by linking concepts of 'energy change' to other science topics. Examples include: (a) the 'pop' sound in testing hydrogen gas could be used as an example of a sudden release of energy, (b) a barbecue could be used as an example to address the energy changes associated with chemical reactions, (c) energy changes associated with a firework display could be investigated. To cater for gender differences, teachers may use examples related to beauty care, such as hair blowing. Moreover, some students might have never used a match or made a fire. Teachers could organize a 'lighting a match' activity for students, particularly for girls, in lower forms.
2. Some teachers indicated that students are not always clear about the subject/object involved in energy change and energy transfer. For example, we feel warm when we burn charcoal. Teacher should make it explicit that it is the charcoal which releases energy while we, the human body, absorb the heat energy. Failing to identify the subject/object or holding an anthropocentric view may lead to confusion. Similar examples include the energy flow from our hand to the can when we hold a can of cold drink.
3. Using calorimetry to determine the energy value of various kinds of food (by burning a piece of food to heat up a known volume of water) is a very common student activity in the topic 'food as a source of energy'. Prior knowledge required for comprehending the activity includes at least (a) burning food releases energy and (b) chemical energy stored in food substances is finally converted to heat, which raises the temperature of the water. From students' performance in the above TIMSS items, it can be inferred that some students do not regard burning food as energy-releasing. Rather, they tend to perceive this as energy-absorbing because they often have to spend great effort in igniting the food. With this misconception, learning from calorimetry becomes shaky and questionable. It is suggested that teachers should clarify such a misconception before the activity is conducted or in a discussion held after the activity.

4.2.2 Properties of carbon dioxide and oxygen

Content / Concept assessed: The role of carbon dioxide and oxygen in combustion

Relevant Science (S1-3) topic: Living things and air (Unit 7)

The Question (1995a)

Carbon dioxide is the active material in some fire extinguishers. How does carbon dioxide extinguish a fire?

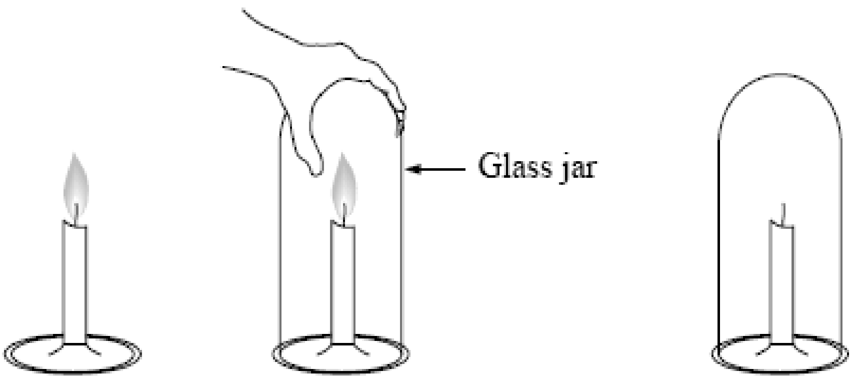
Result

<table>
<tr><th colspan="2">Percentage</th><th rowspan="2">Response</th></tr>
<tr><th>Overall</th><th>Breakdown</th></tr>
<tr><td colspan="2"></td><td>Correct Answer (Boys: 40.4%; Girls: 32%)</td></tr>
<tr><td rowspan="3">36.8
(50.2)*</td><td>27.2</td><td>Mentions that carbon dioxide keeps oxygen away; response includes explicit reference to oxygen.
E.g. Fire doesn’t get oxygen.
Carbon dioxide is heavier than oxygen and will cover the burning materials.</td></tr>
<tr><td>1.6</td><td>Mentions that carbon dioxide keeps “air” away.
E.g. It strangulates the flame.
The fire cannot “breathe”.</td></tr>
<tr><td>8.0</td><td>Other correct.</td></tr>
<tr><td colspan="2"></td><td>Incorrect Answer (Boys: 45.8%; Girls: 48.3%)</td></tr>
<tr><td rowspan="3">46.9</td><td>1.0</td><td>Mentions that carbon dioxide cool down the fire.</td></tr>
<tr><td>1.1</td><td>Refers to a material in carbon dioxide.</td></tr>
<tr><td>44.8</td><td>Other incorrect answers.</td></tr>
<tr><td colspan="2"></td><td>No Answer (Boys: 13.8%; Girls: 19.7%)</td></tr>
<tr><td>16.3</td><td>16.3</td><td>Blank, crossed out/erased, illegible, or impossible to interpret.</td></tr>
</table>

*International average

The Question (1995b)

When a glass jar is placed over a lighted candle, the flame goes out.



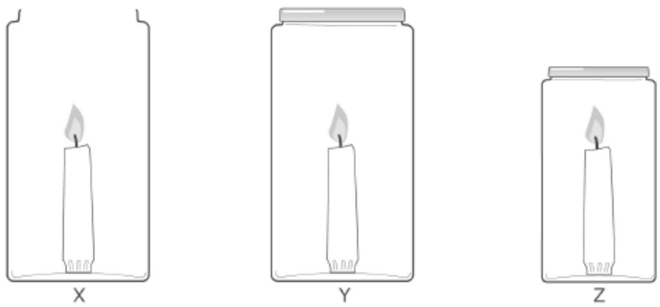
Why does this happen?

Result

Percentage		Response
Overall	**Breakdown**	
		Correct Answer (Boys: 93.2%; Girls: 88.2%)
91.1 (88.1)*	71.7	Refers to the need for oxygen. *E.g. Fire does not get enough oxygen.* *The oxygen will be used up.*
	18.5	Refers to the need for air. *E.g. Fire does not get enough air.*
	0.2	Refers to the need for air, using non-scientific language. *E.g. The fire will be “strangulated”.* *The fire cannot “breathe”.*
	0.7	Other correct.
		Incorrect Answer (Boys: 3.1%; Girls: 8.6%)
5.6	0.3	Refers to its getting too hot.
	0.3	States that the gas (smoke, vapor, carbon dioxide…) is trapped inside the jar. *E.g. The smoke cannot come out.*
	0.2	Refers to the properties of the glass. *E.g. The glass makes it cold.*
	4.8	Other incorrect. *E.g. You put it on too fast and the wind makes it go out.*
		No Answer (Boys: 3.8%; Girls: 2.8%)
3.3	3.3	Blank

*International average

The Question (2003)



Three identical candles are placed in the three jars shown above and lit at the same time. Jars Y and Z are then sealed with lids, and Jar X is left open.

Which candle flame will go out first (X, Y or Z)? ___________
Explain your answer.

Result

<table>
<tr><th colspan="2">Percentage</th><th rowspan="2">Response</th></tr>
<tr><th>Overall</th><th>Breakdown</th></tr>
<tr><td colspan="2"></td><td>Correct Answer (Boys: 61.1%; Girls: 64.8%)</td></tr>
<tr><td rowspan="2">63.1
(47.3)*</td><td>62.8</td><td>Z. Explanation refers to the need for oxygen (for burning).
E.g. Z. The flame in the smaller jar will go out first since it has the least oxygen in it.
Z. Oxygen is needed for the candle to burn.
Z. It has less oxygen.</td></tr>
<tr><td>0.2</td><td>Other fully correct.</td></tr>
<tr><td colspan="2"></td><td>Partial Answer (Boys: 22.6%; Girls: 19.5%)</td></tr>
<tr><td rowspan="3">21</td><td>18.8</td><td>Z. Explanation refers to lack of air (gas) explicitly or using non-scientific language (e.g. suffocation, smothering, choking, etc.). (No explicit mention of oxygen).
E.g. Z. The flame in the smallest jar will be suffocated first.
Z. It does not have enough air to breath.
Since Z is the smallest jar, it will have less air in order to burn.
Z. It gets smothered as the carbon dioxide increases.</td></tr>
<tr><td>0.4</td><td>Indicates both Y AND Z (Y, Z; Y or Z; Y/Z etc.). Explanation based on the need for oxygen or air.
E.g. Y and Z. The flame needs oxygen for it to burn, and both of these jars will run out of it.
Y, Z. The closed jars do not get any air.
Y or Z. They do not get any oxygen.</td></tr>
<tr><td>1.8</td><td>Other partially correct.</td></tr>
<tr><td colspan="2"></td><td>Incorrect Answer (Boys: 15.5%; Girls: 15.4%)</td></tr>
<tr><td rowspan="4">15.3</td><td>9.6</td><td>Z with no explanation or an incorrect explanation.
E.g. Z. This jar will have smallest flame since it is in the smallest jar.
Z. The smoke cannot escape, so the flame dies.
Z. The carbon dioxide level builds up too much.
Z. The candle wants to let off heat, so it bursts.</td></tr>
<tr><td>3.1</td><td>X. Explanation based on the candle being blown out (or similar).
E.g. X. A person walking past the candle might cause it to blow out
X. If the jar is not closed, it goes out from the wind.</td></tr>
<tr><td>1.9</td><td>X OR Y with no explanation or any other incorrect explanation.</td></tr>
<tr><td>0.7</td><td>Other incorrect answers.</td></tr>
<tr><td colspan="2"></td><td>No Answer (Boys: 1%; Girls: 0.2%)</td></tr>
<tr><td>0.6</td><td>0.6</td><td>Blank</td></tr>
</table>

*International average

Noteworthy Points on Students' Performance

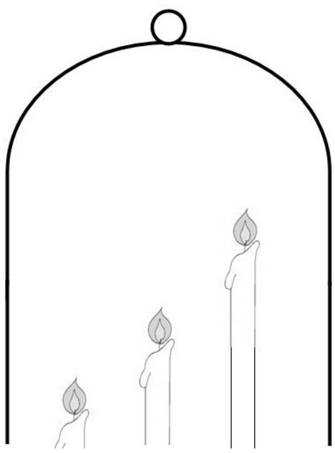
1. ***Role of different gases in combustion*** – HK students performed poorly in Question 1995a, with only 36.8% answering correctly and 16.3% failing to give any answer. Performance improved significantly when the gas identified in the question was changed, i.e. from the role of carbon dioxide to the role of oxygen in combustion (as in Question 1995b and Question 2003).
2. ***Curriculum coverage –*** The fire triangle and ways of extinguishing fire were not covered in the 1995 curriculum. Hence, students might not have sufficient knowledge to answer Question 1995a (46.9%) or may leave the answer blank

(16.3%). The 1995a question also reveals that some students might not know that carbon dioxide is denser than air. Hence they could not explain why carbon dioxide can extinguish fire. Students perhaps did not understand the term "active material".

3. ***Written communication skill*** – In 2003, 91.4% of the students chose the correct answer, i.e. Z. Only 63.0% gave a correct explanation of their option, while the others were unable to justify their choice. In other words, these students failed to explain their reasoning in written form. Indeed, 18.8% of students in 2003 opted for Z but failed to address the role of oxygen, and many students did not use scientific terms precisely ('air' for 'oxygen').

Implications / Suggestions for Improvement

1. Demonstrate how to use a carbon dioxide fire extinguisher. However, merely observing the demonstration does not ensure learning – learning does not naturally come about with only observation[7]. Teachers need to discuss with students the scientific principles involved in putting out a fire.
2. Ensure students understand the relationship between the three elements in the fire triangle (fire, heat and fuel). They should be able to explain fire-fighting by using the framework of the fire triangle.
3. To enhance students' written communication skill, teachers could help them to present their ideas and thinking in complete sentences and connected prose[8].
4. For more capable students, teachers could conduct different experiments to demonstrate the properties of air and challenge students' thinking skills and their application of knowledge:



7 For a concise analysis of the role and use of practical work in school science, see Hodson, D. (1990) A critical look at practical work in school science. *School Science Review*, 70(256), 33-40

8 Please see Wellington, J. & Osborne, J. (2001) *Language and literacy in science education*. Buckingham; Philadelphia: Open University Press for further discussion and suggestions.

i. Align three candles of different height on a bench. Light the candles and cover them with a bell jar as shown in the diagram above; and then ask students which candle would go out first and why. If the wicks of the candles are identical, the tallest candle should go out first – because carbon dioxide produced from burning has a higher temperature and would rise to the top of the jar. Then the carbon dioxide gases would cool down, fall and extinguish the tallest candle first. (To ensure the expected results, teachers should make sure that the wicks are identical in size and length, and that the height of the candles differ significantly.)

ii. To demonstrate the change in concentration of carbon dioxide gases inside the gas jar, teachers might fill three plastic bottle caps with lime water and use Blu-tack to stick them at different levels inside the jar.

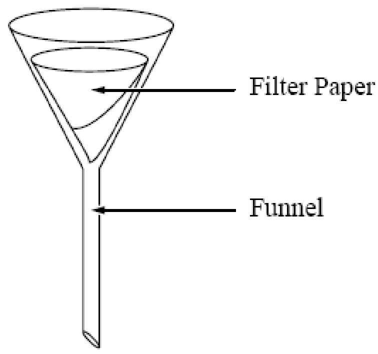
5. Draw daily life application by asking students questions such as: What are the necessary actions to be taken when a fire breaks out? Why do we need to bend towards the floor when we are caught in smoke? [It is because smoke rises up and a better supply of air (oxygen) is near the ground.]

4.2.3 Using tools, routine procedures and science processes

Content / Concept assessed: Filtration

Relevant Science (S1-3) topic: The wonderful solvent – water (Unit 5)

The Question (1995)



Filtration using the equipment shown can be used to separate which materials?

A. A solution of copper sulphate and water

B. A solution of sodium chloride and water

C. A mixture of alcohol and water

D. A mixture of mud and water

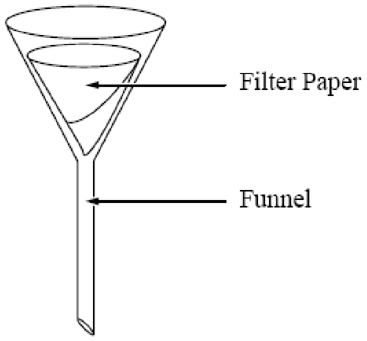
E. A mixture of sand and sawdust

Result

	HK (%)
A	4.9
B	7.1
C	4.8
(D)	62.2
E	20.9

% correct			
HK			International
Boys	Girls	Overall	
66.7	56.7	62.2	51.9

The Question (1999)



Filtration using the equipment shown can be used to separate which materials?

A. A mixture of salt and pepper
B. A mixture of pepper and water
C. A mixture of oxygen and water
D. A solution of silver nitrate in water
E. A solution of sugar in water

Result

	HK (%)
A	21.9
(B)	38.4
C	8.6
D	17.3
E	13.6

% correct			
HK			International
Boys	Girls	Overall	
34.8	42.2	38.4	38.7

Noteworthy Points on Students' Performance

1. ***Context of the question*** – The two questions are very similar regarding what a simple filtration apparatus can do. It is only the keys that differ – one is about separating mud/water (1995) and the other is about pepper/water (1999). Data from

the participant countries indicate that students did better in the 1995 question than in 1999. This decline in performance goes against the intuition of those people who suggest that students know more about and learn from their daily experience, and might be attributed to students' experience in their science lessons. Separating mud/water is a typical science laboratory activity while dealing with pepper/water is not, though it is more daily-relevant.

2. ***Gender differences*** **–** Although boys outperformed girls in many TIMSS items, the reverse is true in the 1999 question. Girls were more capable than boys in pointing out that the setup could separate a mixture of pepper and water. It might be due to the greater familiarity of girls with materials used in the kitchen. Also, due to their learning experience in Home Economics, girls may be more familiar with the properties of pepper.
3. ***Question rubric*** – Some teachers argued that Option A of the 1999 question could be a correct answer. That is, the setup, when supplemented by adding water to the mixture of salt and pepper, can be used to separate the two substances.

Implications / Suggestions for Improvement

1. To make learning more authentic and meaningful, it is suggested that, in addition to traditional mud and water, students could be asked to plan and/or carry out practical activities to separate different kinds of substances. Suitable examples are pepper/water, starch/water, chalk/water or things that can be found in every household. Similarly, students could be asked to compare fresh orange juice (with some orange pulp inside) and its filtrate.
2. In order to make learning in the laboratory accountable and strengthen students' practical skills, some schools have started to arrange practical examinations for junior secondary students. Where appropriate, teachers could provide students with a wide range of apparatus, from which they have to select the items they need to complete a given task. They could also make reference to the practice, used in some schools already, of asking senior students to assist with assessment procedures.

4.2.4 Water cycle

Content / Concept assessed: Movement of water and its energy source

Relevant Science (S1-3) topic: The wonderful solvent – water (Unit 5)

The Question (1995a)

Draw a diagram to show how the water that falls as rain in one place may come from another place that is far away.

Result

Percentage		Response
Overall	**Breakdown**	
		Correct Answer (Boys: 31.1%; Girls: 23.6%)
27.7 (31.8)*	27.7	Response includes the three following steps: i. Evaporation of water from a source. ii. Transportation of water as vapour/clouds to another place. iii. Precipitation in other places.
		Partial Answer (Boys: 17.2%; Girls: 15.9%)
16.6	8.0	As in the correct answer, but response does not mention evaporation.
	5.8	As in the correct answer, but response does not mention transportation.
	0.7	As in the correct answer, but response does not mention precipitation.
	2.1	Other partially correct answers.
		Incorrect Answer (Boys: 41.1%; Girls: 49.8%)
45.1	11.8	Response indicates precipitation only; it may use vertical or diagonal lines.
	33.3	Other incorrect answers.
		No Answer (Boys: 10.7%; Girls: 10.6%)
10.7	10.7	Blank

*International average

The Question (1995b)

The source of energy for the Earth's water cycle is the

A. wind
B. Sun's radiation
C. Earth's radiation
D. Sun's gravity

Result

	HK (%)
A	27.3
(B)	46.8
C	10.3
D	13.9

% correct			
HK			International
Boys	Girls	Overall	
51.4	41.8	46.8	40.9

Noteworthy Points on Students' Performance

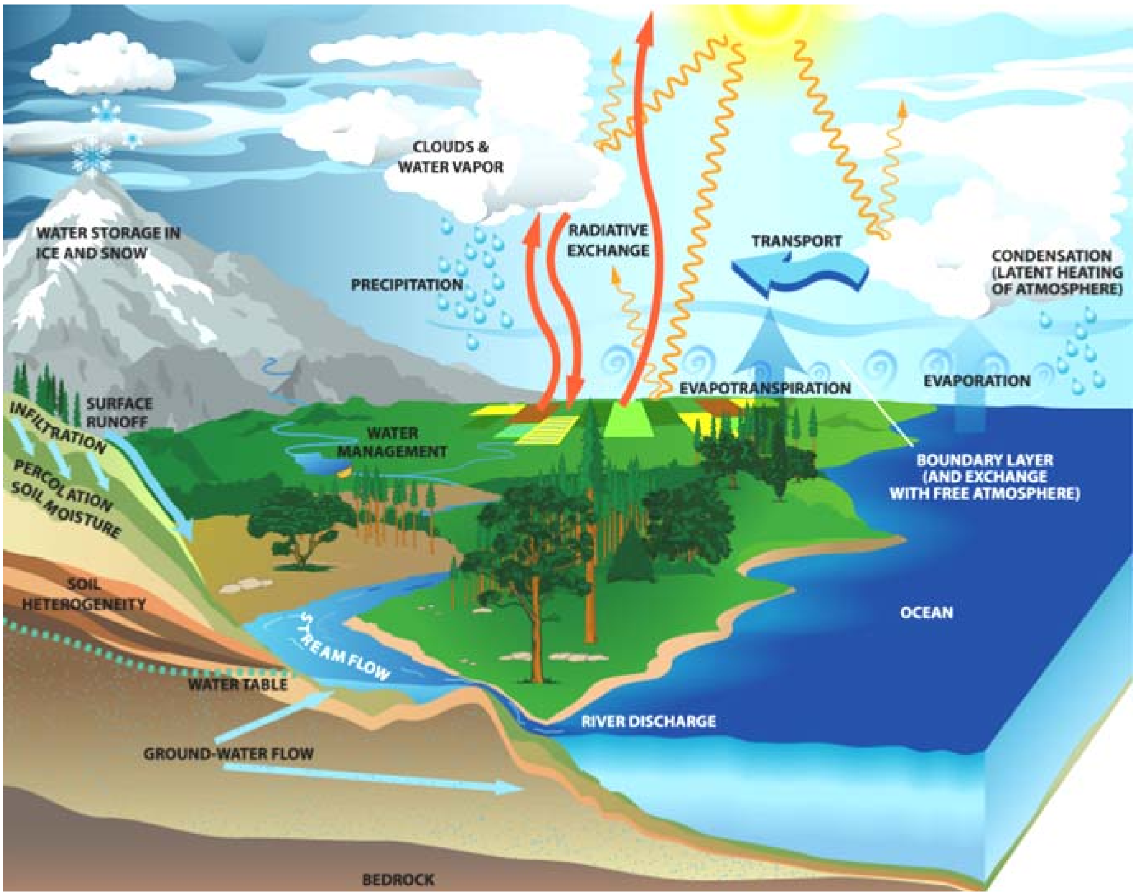
1. ***Weak in drawing diagrams –*** HK students are weak in using diagrams to present their ideas. According to the observation of some teachers, students tend to regard diagram drawing as copying what the teacher presents to them. Seldom do students use diagrams to represent their own understanding.
2. ***Common misconception –*** The water cycle is often represented by a diagram of a cycle between water, cloud and land. Perhaps students consider that cloud and rain are transported by wind from one place to another. Therefore, in question 1995b, 27.3% of students opted for the distracter 'wind' as the source of energy for the water cycle without appreciating that the Sun is the ultimate source of energy on planet Earth.
3. ***Unfamiliar terminology –*** While the concepts of solar energy, light energy and heat energy are common in the junior secondary science curriculum, the term 'radiation' is usually not introduced to S2 students.

Implications / Suggestions for Improvement

1. Where appropriate teachers might like to draw connections between different science topics. This would help to prevent compartmentalization of knowledge and learning. The topic 'water cycle' could well be linked with 'energy'.
2. Alternative approaches to give students a more holistic view of water cycle:
   i. Use of videos to help students see and appreciate the role of the Sun on planet Earth. *Energy on Earth* (Madison, Wis.: Hawkhill, 2001) and *Prisoners of the Sun* (London: BBC, 1992) are excellent resources for teachers and students. They are available for loan at HKU library.
   ii. Get a big bucket filled with water and cover the top with a plastic bag. Then, shine a big spotlight directly on top of the bucket and observe the condensation of water on the plastic bag. Once the light is removed water droplets fall.
3. It has been assumed that diagrams in textbooks or in other teaching resources assist students' learning. Recent research in science education suggests that such an assumption may not hold in many instances. It has been found that some students have difficulties in interpreting diagrams which teachers might have regarded as self-explanatory. Consider, for example, the problems associated with reading the diagram[9] on the next page. First, students may focus on some concrete elements (the wind that moves the cloud) rather than more abstract elements (the Sun as the ultimate source of energy of water cycle, and other physical and chemical

9 Source: http://www.usgcrp.gov/usgcrp/images/ocp2003/ocpfy2003-fig5-1.htm

processes of the Earth). Second, it is common for a diagram to use many different notations to represent some, or even the same, processes. In the following diagram, there are many different types of 'arrows'. Some 'arrows', though with different shapes and thickness, represent the movement of water, while other 'arrows' represent solar radiation. A further complication is caused by the use of 'whirl' which, like other 'arrows', also means the movement of water (evaporation). Sometimes, the direction of 'arrows' also represents the movement of something, whether it be water or energy, from one place to another. It could be unclear to students where exactly that 'something' comes from and goes to. Therefore, it is very important that teachers should make sure that students can understand and interpret the 'visual language' of diagrams.



4.2.5 Composition of Air

Content / Concept assessed: Nitrogen and oxygen

Relevant Science (S1-3) topic: Living things and air (Unit 7)

The Question (1995)

Air is made up of many gases. Which gas is found in the greatest amount?

A. Nitrogen
B. Oxygen
C. Carbon dioxide
D. Hydrogen

Result

	HK (%)
(A)	49.7
B	28.1
C	15.4
D	6.3

% correct			
HK			International
Boys	Girls	Overall	
59.1	37.7	49.7	27.2

The Question (1999a)

Why do mountain climbers use oxygen equipment at the top of the world's highest mountains?

A. There is less oxygen in the air at great heights.
B. There is little nitrogen in the air at great heights.
C. There is a hole in the ozone layer.
D. There is no air at the top of very high mountains.

Result

	HK (%)
(A)	94.5
B	1.0
C	0.6
D	3.6

% correct			
HK			International
Boys	Girls	Overall	
94.6	94.5	94.5	78.8

The Question (1999b)

Diana and Mario were discussing what it might be like on other planets. Their science teacher gave them data about Earth and an imaginary planet Proto. The table shows these data.

	Earth	Proto
Distance from a star like the Sun	148 640 000 km	902 546 000 km
Atmospheric pressure at surface of planet	101 325 Pa	100 Pa
Atmospheric conditions:		
gas components	21% oxygen 0.03% carbon dioxide 78% nitrogen	5% oxygen 5% carbon dioxide 90% nitrogen
ozone layer	yes	no
cloud cover	yes	no

Write down one important reason why it would be difficult for humans to live on Proto, if it existed. Explain your answer.

Result

Percentage		Response
Overall	**Breakdown**	
		Correct Answer (Boys: 69.4%; Girls: 71%)
70.3 (65.9)*	40.8	States there would be insufficient (too little, less, not enough, etc.) oxygen (to breathe). E.g. *People could not survive on Proto because there is not enough oxygen to breathe.* *There is only 5% oxygen and that is not enough to survive.*
	0.2	States that the atmospheric pressure would be too low with an explanation based on low/air oxygen levels or effects of lower boiling point, etc. E.g. *The atmosphere is too thin, so we cannot breat*he.
	10.9	States that there is no ozone layer to protect people (against star's UV radiation). E.g. *There is not enough ozone on Proto to give humans adequate protection from UV rays.* *No protection (from Sun) by an ozone layer.*
	1.5	States that no cloud cover means no (low) water. E.g. *Proto may not have any water on it since it lacks clouds*
	14.5	States that it is too cold (due to distance from the star). E.g. *Proto is too far from the star, and therefore it would be very cold.* *People would freeze to death on Proto.*
	2.4	Other correct.
		Incorrect Answer (Boys: 24.2%; Girls: 25.8%)
24.9	11.8	Merely repeats information in table and/or stem with no explanation. E.g. *Proto has no ozone layer.* *The pressure is too low to survive.* *The oxygen is 5% and the nitrogen is 90%.*
	0.3	States that it is too hot due to no ozone layer. E.g. *Without an ozone layer the planet would heat up too much.*
	0.4	States that there is little (no) gravity (confuses atmospheric pressure with gravity). E.g. *People would float on Proto because the pressure is too low.* *There is not enough pressure to keep people on the ground.*
	12.4	Other incorrect answers.
		No Answer (Boys: 6.4%; Girls: 3.1%)
4.7	4.7	Blank

*International average

The Question (2003a)

Three gases found in Earth's atmosphere are carbon dioxide, nitrogen, and oxygen. What is their order of abundance from greatest to least?

A. Nitrogen, oxygen, carbon dioxide
B. Nitrogen, carbon dioxide, oxygen
C. Oxygen, nitrogen, carbon dioxide
D. Carbon dioxide, oxygen, nitrogen

Result

	HK (%)
(A)	55.5
B	18.0
C	13.2
D	13.1

% correct			
HK			International
Boys	Girls	Overall	
57.3	53.9	55.5	27.8

The Question (2003b)

The table shows some information about the planets Venus and Mercury

	Average Surface Temperature (℃)	Atmospheric Composition	Mean Distance from the Sun (millions of km)	Time to Revolve Around the Sun (Number of Days)
Venus	470	Mostly Carbon Dioxide	108	225
Mercury	300	Trace amounts of gases	58	88

Which of the following best explains why the surface temperature of Venus is higher than that of Mercury?

A. There is less absorption of sunlight on Mercury because of the lack of atmospheric gases.
B. The high percentage of carbon dioxide in the atmosphere of Venus causes a greenhouse effect.
C. The longer time for Venus to revolve around the Sun allows it to absorb more heat from the Sun.
D. The Sun's rays are less direct on Mercury because it is closer to the Sun.

Result

	HK (%)
A	7.3
(B)	68.8
C	13.8
D	9.1

% correct			
HK			International
Boys	Girls	Overall	
69.8	67.8	68.8	36.0

Noteworthy Points on Students' Performance

1. ***Common misconception* –** Many students did not know that air contains 78% nitrogen gas, or that nitrogen makes up the largest proportion in air. Some students regarded oxygen as the main component of air (28.1% in 1995a and 13.2% in 2003a).
2. ***Curriculum / Teaching emphasis* –** Teaching of photosynthesis and respiration mainly focuses on the functions of carbon dioxide and oxygen. Students might thus infer that they make up a more significant proportion in air.
3. ***Reasoning skill* –** From the 1999b question, 11.8% of students merely repeated the information given in the question but did not come up with a conclusion. Some teachers indicated that students sometimes are not able to draw conclusions from given data.

Implications / Suggestions for Improvement

It might be a surprise to some teachers that only half of the students are able to recall the plain fact that the percentage of $N_2 > O_2 > CO_2$ in the Earth's atmosphere. Nevertheless, from the constructivist perspective of learning – (a) that students do not learn by simply being told, (b) that meaningful learning occurs when students are able to link their prior experience with new knowledge, and (c) that learning is more likely to occur when students can see the significance of their learning – students could be provided with the following learning experiences in order that they would find it easier to retain the knowledge.

1. Teachers could assess students' understanding of the composition of air by asking them to draw a diagram to present their ideas. Instruct them to use red dots for nitrogen, blue dots for oxygen, etc. This could assess students' knowledge of (a) the composition of air, such as the proportion of nitrogen to oxygen being 4:1, (b) the particle model of gases, i.e. air particles are randomly and well separated within a given space.
2. Discuss with students the significance of the composition of air on Earth. For example, what would happen if air was composed of only oxygen? Why is nitrogen

the most abundant element in air? What would happen if the proportion of carbon dioxide in air increased? Conversely, since the greenhouse effect is caused by an increase in carbon dioxide concentration in the atmosphere, would it be desirable if the concentration of the gas was significantly reduced? In short, it is important to develop students' ability to relate facts with their significance. This could be done by first listing the properties of different gases and then asking students what would happen if the percentage of different gases changed (e.g. percentage of oxygen increased).

4.2.6 Particulate structure of matter

Concepts assessed: Particulate structure of matter and conservation of matter

Relevant Science (S1-3) topic: Matter as particles (Unit 6)

The Question (1995a)

Animals are made up of many atoms. What happens to the atoms after an animal has died?

A. The atoms stop moving.
B. The atoms recycle back into the environment.
C. The atoms split into simpler parts and then combine to form other atoms.
D. The atoms no longer exist once the animal has decomposed.

Result

	HK (%)
A	11.8
(B)	34.1
C	25.2
D	28.4

% correct			
HK			**International**
Boys	**Girls**	**Overall**	
37.5	30.5	34.1	26.0

The Question (1995b)

The words *cloths, thread*, and *fibre* can be used in the following sentence: *cloth* consists of *threads* which are made of *fibre*.

Use the words *molecules, atoms,* and *cells* to complete the following sentence:

_______________consist of _______________ which are made of _______________.

Result

Percentage		Response
Overall	Breakdown	
		Correct Answer (Boys: 34.6%; Girls: 29.2%)
32 (31.4)*	32.0	Cells – Molecules – Atoms.
		Incorrect Answer (Boys: 59%; Girls: 68.8%)
51.9	26.9	Cells – Atoms – Molecules.
	5.5	Molecules – Atoms – Cells.
	7.9	Molecules – Cells – Atoms.
	11.6	Atoms – Molecules – Cells.
	4.3	Atoms – Cells – Molecules.
	7.3	Other incorrect answers.
		No Answer (Boys: 6.4%; Girls: 2.1%)
4.4	4.4	Blank

*International average

The Question (2003)

If you took all of the atoms out of a chair, what would be left?

A. The chair would still be there, but it would weigh less.

B. The chair would be exactly the same as it was before.

C. There would be nothing left of the chair.

D. Only a pool of liquid would be left on the floor.

Result

	HK (%)
A	32.7
B	8.3
(C)	46.5
D	11.8

% correct			
HK			International
Boys	Girls	Overall	
47.7	45.4	46.5	51.3

Noteworthy Points on Students' Performance

1. ***Curriculum coverage*** – as noted in Section 4.1, the concepts assessed were not covered in the local curriculum when the students took the assessment. This is regarded as the main factor attributing to the rather unsatisfactory results. Nonetheless, it does not mean that students at S2 level are not capable of mastering the topic / relevant concepts. This is supported by the following international data:
   i. While only 32% of HK students answered the 1995b question correctly, 66% of Singapore students did so.

ii. Band 1 (58%) students outperformed band 2 and 3 students (29% and 36% respectively) in the 2003 question – a cursory survey indicates that concepts like 'atom', 'molecule' and 'ion' are introduced in some higher banding schools. Nevertheless, the figure was still lower than the national average of Korea (66%), Estonia (69%), Japan (65%) and England (68%).

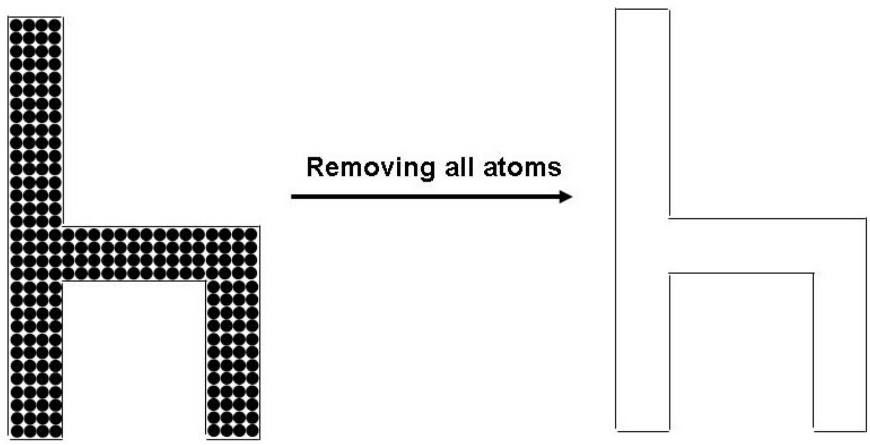
2. ***Common misconceptions*** – As reflected in students' responses to the 2003 question, it is likely that they knew that matter is made up of atoms/particles, but might consider that the atoms/particles are contained within 'something'. Research in a similar context found that some students believe 'there is liquid water between water molecules' or 'an aqueous state exists between water molecules'.

   Some students were not able to relate the concept of conservation of mass to the 1995a question and thought that once an animal has died, the atoms will no longer exist (perhaps misled by TV games in which a figure/character disappears when deceased). As a result, 28.4% of students wrongly opted for D as the correct answer.

3. ***Unfamiliar terminologies –*** Students might not understand the meaning of 'recycle' (in 1995a) and 'fibre' (in 1995b).

Implications / Suggestions for Improvement

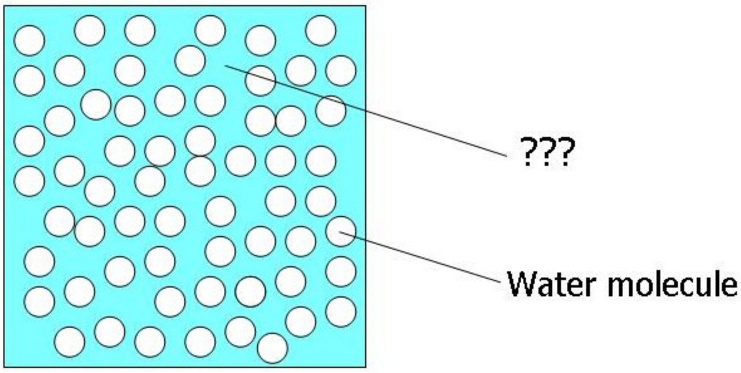
1. In illustrating the particulate nature of substances, some textbooks, while showing particles, often use a 'physical border' to outline the shape of that object. It is natural that students will believe that there is 'something' left behind after all the atoms are taken away:



   Therefore, it is very important for the teacher to discuss with students the limitations of such visual representations (what are the other limitations of the above diagram in depicting chair being made of particles?).

2. Similarly, students might be misled by the shaded, usually blue, background of the

following diagram, which tries to illustrate liquid water at the microscopic perspective:



Instead of indicating that the space between the particles is 'aqueous state' or 'liquid water', it should be brought to students' attention that there is empty space between the particles[10].

3. To demonstrate the conservation of mass and/or the particle theory of matter, teachers could try the following:
   i. Show or build an atomic model of an object
   ii. Use Lego blocks as atoms to build a chair. On completion, dismantle it by removing the Lego blocks one by one until there is nothing left. This demonstration should enable students to visualize that matter is only made up of atoms without any 'borders'.
   iii. Use PowerPoint or other visual methods to illustrate the idea to students.

10 For more discussion on how students understand physical phenomena from a microscopic perspective, please refer to: Lijnse, P.L., Licht, P., de Vos, W. & Waarlo, A.J. (1990). *Relating macroscopic phenomena to microscopic particles: a central problem in secondary science education*. Utrecht, Netherlands: CD-[Beta] Press

4.3 Important Lessons Learned from Students' Performance

As evident from above, the following are important lessons learned from students' performance in TIMSS:

1. Students' **common misconceptions** need to be taken into consideration in planning and teaching science lessons (see 4.2.4, 4.2.5, 4.2.6).
2. Student misconceptions may be induced by unbalanced **curriculum / teaching emphasis** on certain topics at the expense of the others (see 4.2.5).
3. Teachers should help students to **draw connections between different science topics** (see 4.2.4). Avoid compartmentalization of knowledge.
4. Students' **everyday life experiences** need to be taken into consideration in planning and teaching science lessons (see 4.2.1, 4.2.2, 4.2.3).
5. Science teaching can be enhanced by carrying out **simple experiments / student activities** (see 4.2.2, 4.2.3, 4.2.4, 4.2.5, 4.2.6).
6. Science learning can be enhanced by helping students acquire the skills of **reading diagrams** and **drawing diagrams** (see 4.2.4, 4.2.5, 4.2.6).
7. Effort needs to be made to cultivate in students' **reasoning skills** (see 4.2.5).
8. Attention needs to be paid to improve students' **written communication skills** (see 4.2.2).
9. **Curriculum coverage** should be closely examined and could be informed by students' performance in international comparative studies (see 4.2.6).
10. Care should be taken when introducing new **scientific terms / terminologies** to students (see 4.2.1, 4.2.4, 4.2.6).
11. Students' understanding can vary a lot with a slight change in the **context of the questions** (see 4.2.3).
12. Attention needs to be paid to the **rubric of the questions** (see 4.2.3).
13. **Gender differences** in science teaching should be considered in planning and teaching science lessons (see 4.2.3).

Chapter 5

Learning from Students' Performance in NOS-related Questions

Benny Hin Wai Yung and Siu Ling Wong

5.1 Performance in General

There have been slight variations in the categorization of items in TIMSS across the years. For example, in 1999, because of increasing interest in students' performance on questions related to the nature of science (NOS) and scientific inquiry, efforts were made to single out these items from the rest for a separate analysis. Table 5.1 shows the results for Hong Kong (HK) as compared to the International Average (Int).

Table 5.1 Average percent correct in science content areas in TIMSS 1999

Science Content Area	Multiple-choice Items				Constructed-response Items			
	No. of Items	Average % correct			No. of Items	Average % correct		
		HK	Int	HK-Int		HK	Int	HK-Int
Chemistry	15	57	50	+7	7	47	41	+6
Earth Science	16	65	56	+9	4	61	54	+7
Environmental & Resource Issues	7	62	49	+13	10	66	54	+12
Life Science	28	64	58	+6	14	57	51	+6
Physics	28	67	59	+8	11	41	35	+6
Scientific Inquiry and the Nature of Science	9	60	48	+12	3	27	28	-1
Overall	103	63	55	+8	49	52	45	+7

As shown in Table 5.1, HK students' overall performance (i.e. HK-Int) in multiple-choice items (+8%) was slightly better than in free-response items (+7%). That is, there was only a very small 'performance gap' between these two kinds of items. However, this 'performance gap' varied from 'no noticeable difference' in Life Science (+6% vs. + 6%) to 'marked differences' in Scientific Inquiry and the Nature of Science

(+12% vs. –1%). Our below-international-average performance in free-response items in this area (-1%) might have been regarded as a worrying sign, especially when our performance in multiple-choice items in the same area was well above the International Average (+12%). However, it may well be attributable to students' inadequate communication skills, rather than their poor NOS understanding.

In 1999, the questions were also subject to further analyses by re-grouping them according to the 'Science Performance Expectations', as shown in Table 5.2. The data were compared to the International Averages to obtain their differences. HK students' performance was above the International Average in each of the following areas in both multiple-choice and constructed-response items:

- Applying Scientific Principles to Solve Quantitative Problems (+12% vs. +8%)
- Gathering Data (+9% vs. +12%)
- Complex Information (+6% vs. +8%)

It seemed that HK students performed better in Gathering Data and Solving Quantitative Problems than dealing with Complex Information.

Table 5.2 Average percent correct in science performance expectations in TIMSS 1999

Science Performance Expectation	Multiple-choice Items				Constructed-response Items			
	No. of Items	Average % correct			No. of Items	Average % correct		
		HK	Int	HK-Int		HK	Int	HK-Int
Applying Scientific Principles to Develop Explanations					17	52	47	+5
Applying Scientific Principles to Solve Quantitative Problems	5	68	56	+12	3	36	28	+8
Complex Information	30	63	57	+6	21	59	51	+8
Conducting Routine Experimental Operations	1	38	39	-1	1	45	25	+20
Designing Investigations	3	57	47	+10	1	21	40	-19
Formulating conclusions from investigational data	1	49	48	+1	1	17	19	-2
Gathering Data	2	48	39	+9	4	47	35	+12
Interpreting Data	6	72	56	+16	1	49	49	0
Making Decisions	55	64	56	+8				
Overall	103	63	55	+8	49	52	45	+7

Though, by and large, HK students performed well above the International Average (as measured by constructed-response items) in Conducting Routine Experimental Operations (–1% and +20%), they were comparatively weak at Formulating Conclusions from Investigational Data (+1% and –2%). More importantly, they lagged far behind (as measured by constructed-response items) in Designing Investigations (+10% and –19%). This is because HK students were rarely asked to design their own experiment. Instead, they carried out a lot of recipe-type practical work (see also 5.2.5). The comparatively large performance gaps between the multiple-choice and constructed-response items in Interpreting Data (+16% vs. 0%) and especially in Designing Investigation (+10% vs. –19%) are noteworthy. This might be attributed to students' inadequate communicative skills. However, it should be noted that the interpretation of the data needs to be undertaken with caution whenever the number of items for analysis is limited to one or two. Nonetheless, these findings, coupled with those derived from the following analyses on students' performance in individual items, can shed light on areas to which we need to pay more attention if we are to improve science education in HK.

5.2 Performance in Individual Items

In this section, we examine more closely HK students' performance in questions that test their understanding about concepts related to scientific inquiry, namely, those associated with (1) formulating hypothesis, (2) controlling variables in designing experiments, (3) choice of instruments in carrying out scientific investigations, (4) the need for repeated measurements, (5) data interpretation and (6) drawing conclusions. It should be pointed out that these concepts ought not to be introduced and discussed only at the beginning of the junior secondary science course (i.e. Unit 1 - Introducing science). Rather, the relevant ideas and attitudes should be an integral part of the course, and opportunities taken to address them wherever possible.

5.2.1 Formulating hypothesis

Content / Concept assessed: scientific investigation

Relevant Science (S1-3) topic: Introducing science (Unit 1)

The Question (1999)

Alexander Fleming noticed that bacteria growing on a plate of agar did not grow next to a mold that was growing on the same plate. He wrote in his laboratory report: “The mold may be producing a substance that kills bacteria.” This statement is best described as

A. an observation
B. a hypothesis
C. a generalization
D. a conclusion

Result

	HK (%)
A	26.6
(B)	31.5
C	11.0
D	30.6

% correct			
HK			International
Boys	Girls	Overall	
30.2	32.8	31.5	35.4

Noteworthy Points on Students’ Performance

HK students’ performance (31%) was below the International Average (35%). Almost 27% of them chose the wrong answer (A) – observation; while approximately 31% wrongly regarded the statement as a conclusion (D). These confusions might be accounted by:

1. ***Curriculum emphasis*** – Aspects about NOS and scientific inquiry were not emphasized before 1999.
2. ***Teaching emphasis*** **–** In the past, both teaching as well as the arrangement of textbooks tended to put emphasis on the importance of making observations and drawing conclusions in the process of scientific investigation. Though more attention is now paid to other components of scientific method, the discussion is often superficial. In most cases, it entails just a description of the scientific method as comprising four steps, namely, making observation, setting a hypothesis, experimentation and drawing conclusions. Little discussion is given to what constitutes a hypothesis, or how a conclusion is different from the results. Often, the different stages of a scientific investigation are discussed in a decontextualized

manner, rather than illustrated and discussed through actual examples or historical case studies, as in this question.

Implications / Suggestions for Improvement

1. Illustrate/discuss the different stages of a scientific investigation through actual examples or case studies. For example:
   i. Have students read a historical case of a scientific investigation (e.g. Alexander Fleming's discovery of penicillin, Doctor Marshall's discovery of the bacteria causing stomach ulcers). Ask them to identify which part of the story is describing the observations, which part is talking about the hypothesis, and which parts are about experimentation, results and conclusions.
   ii. Bring up relevant daily life examples to consolidate the concepts of observation, hypothesis and conclusion. For example: (1) Talk about the different hypotheses put forward by various organizations for the possible causes of the SARS outbreak, (2) Have students imagine what hypotheses they would come up with if their MP3 players are not working properly (e.g. low battery, the machine is broken, etc), and how would they find out which of their hypotheses is correct (i.e. designing a fair test).
   iii. Provide more opportunities to help students distinguish between observation, result and conclusion. For example, pass a beam of light through a prism (where white light is splitting into a mixture of colours) and then ask students to make conclusions based on their observations. Through discussion, help students to distinguish the differences among the three terms: observation, result and conclusion.

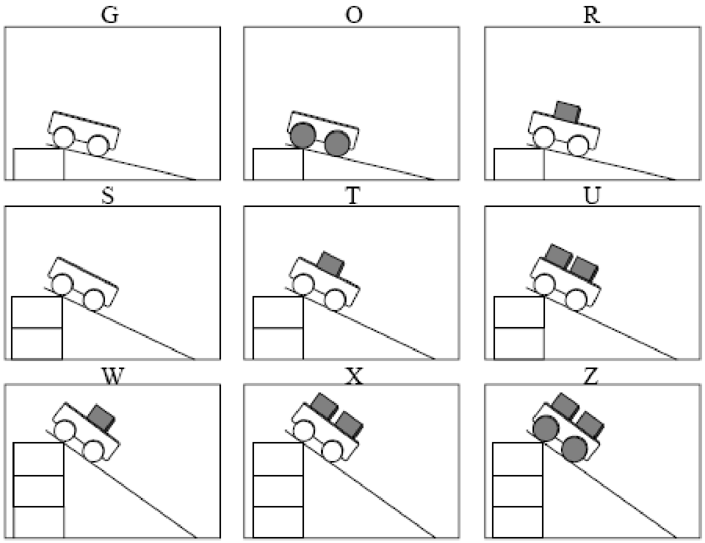
5.2.2 Experimental design - controlling variables

Content / Concept assessed: Fair test

Relevant Science (S1-3) topic: Space Travel (Unit 9)

The Question (1995)

The diagrams show different trials Abdul carried out with carts having different-sized wheels. He started them from different heights and the blocks he put in them were of equal mass.



He wants to test this idea: The heavier a cart is, the greater its speed at the bottom of a ramp. Which three trials should he compare?

A. G, T, and X
B. O, T, and Z
C. R, U, and Z
D. S, T, and U
E. S, W, and X

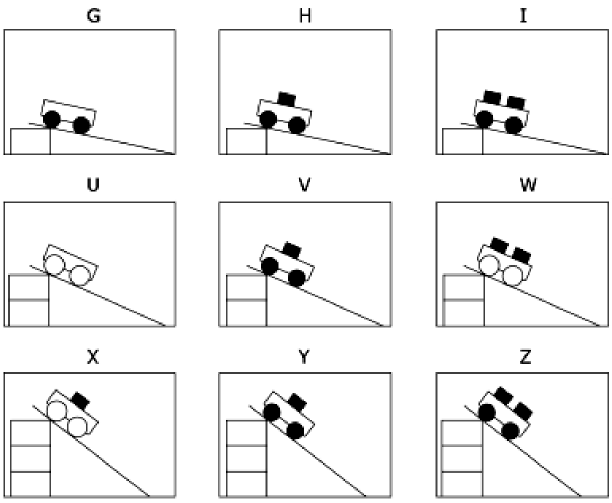
Result

	HK (%)
A	18.0
B	10.0
C	22.7
(D)	41.7
E	6.1

% correct			
HK			**International**
Boys	**Girls**	**Overall**	
48.0	34.3	41.7	36.8

The Question (2003)

The diagrams show different trials Usman carried out with wheels of two different sizes and different numbers of block of equal mass. He used the same ramp for all trails, starting the carts from different heights.



He wants to test this idea: The higher the ramp is placed, the faster the cart will travel at the bottom of the ramp. Which three trials should he compare?

A. G, H, and I
B. I, W, and Z
C. I, V, and X
D. U, W, and X
E. H, V, and Y

Result

	HK (%)
A	5.2
B	20.6
C	4.5
D	3.5
(E)	65.8

% correct			
HK			**International**
Boys	**Girls**	**Overall**	
64.1	67.6	65.8	43.3

Noteworthy Points on Students' Performance

Students' grasp of the notion of a fair test in designing experiments improved considerably in 2003 (65%) as compared to 1995 (41%). Other data not shown here indicate that students of lower ability gave up more easily. Several reasons could have contributed to these outcomes:

1. ***Curriculum emphasis –*** Before 2000, students rarely had the opportunity to design their own experiments. Experiments were usually of a cookbook recipe-type or were teacher demonstrations. The great improvement shown in 2003 could be attributed to the curriculum revision in 2000, when a great deal of attention was put on nature of science and scientific investigation, in particular, the idea of a fair test.
2. ***Comprehension skill –*** Students misunderstood the requirement of the question, thinking that it concerned merely the speed of the cart on the ramp.

Implications / Suggestions for Improvement

1. Valuable knowledge can be learned and experience gained when students are required to design a fair test themselves. By testing in a trial and error manner, students learn what needs to be changed and what not in a fair test. Through this process they gain a better and long-lasting understanding of what the necessary conditions for a fair test are.
2. Allowing more opportunities for students to design and involve themselves in science learning would cultivate a genuine interest in science. This is considerably better than simply copying and memorizing what the teacher has said.
3. The concept of experimental design can be quite challenging for students of lower ability because it is abstract and requires higher order thinking skills. Teachers can carry out a fair test design exercise after each topic, so as to familiarize and make students comfortable in dealing with the concept of fair tests.
4. Have students develop the habit of using a table to display the various variables of an experiment in a systematic manner, as shown in the example below. This enables students to see if all other variables are kept constant except the one being investigated. (In the example below, the aim is to study the effect of changing the amount of baking powder on the rise of the cake. It can be easily spotted from the third column of the table that it would be problematic if the temperature is not kept constant). Teachers should also discuss with students that tables are more than ways of presenting data after they have been collected. They can be used as ways of organizing the design, and subsequent data collection and analysis, in advance of the experiment.

	Amount of baking powder (spoonful)	Temperature for keeping the baking powder-flour mixture before baking (^{o}C)	Amount of flour (g)	Time of keeping the baking powder-flour mixture before baking (min)
1	1	10	2	30
2	2	25	2	30
3	3	40	2	30
4	4	55	2	30

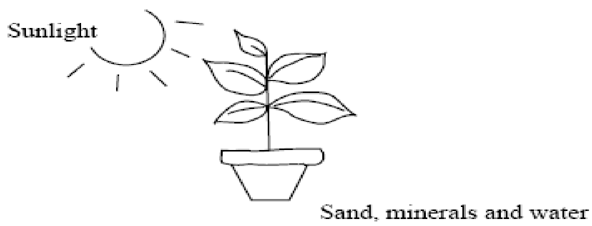
5.2.3 Experimental design - controlling variables

Content / Concept assessed: Conditions favourable for plant growth

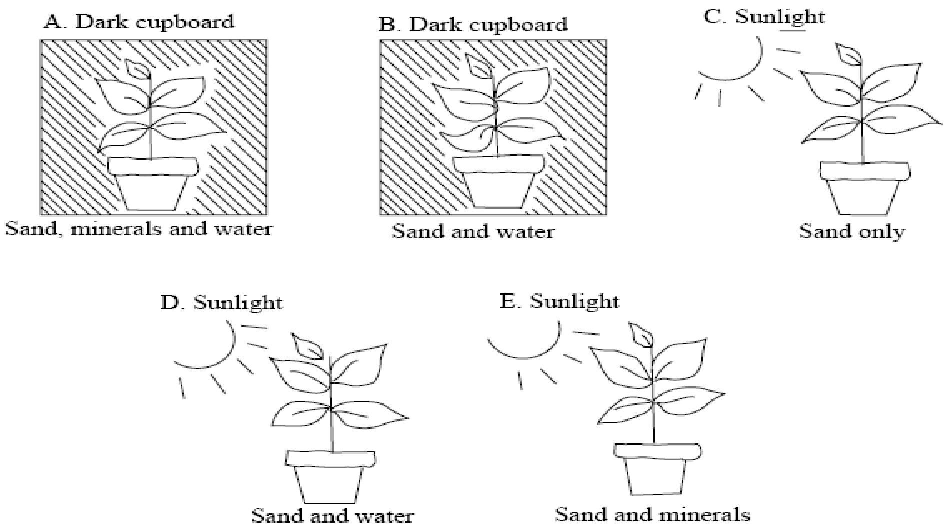
Relevant Science (S1-3) topic: Living things and air (Unit 7)

The Question (1995)

A girl had an idea that plants needed minerals from the soil for healthy growth. She placed a plant in the Sun, as shown in the diagram below.



In order to check her idea she also needed to use another plant. Which of the following should she use?



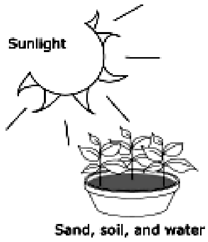
Result

	HK (%)
A	22.2
B	7.4
C	4.0
(D)	56.7
E	9.2

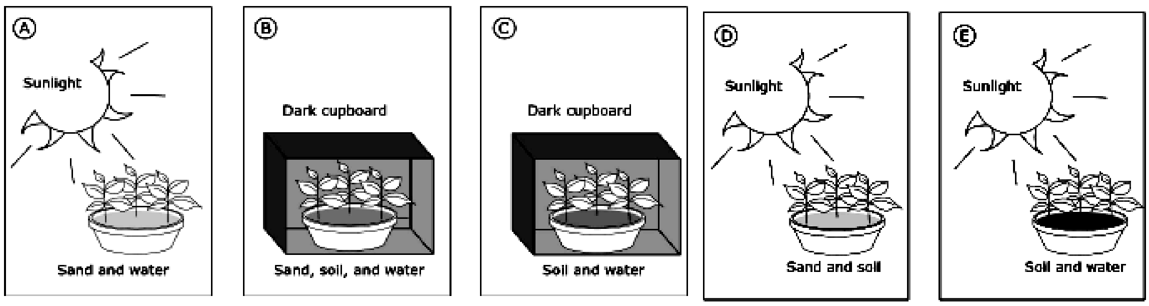
% correct			
HK			International
Boys	Girls	Overall	
59.0	54.0	56.7	44.9

The Question (2003)

A girl has an idea that green plants need sand in the soil for healthy growth. In order to test her idea she uses two pots of plants. She sets up one pot of plants as shown below.



Which ONE of the following should she use for the second pot of plants?



Result

	HK (%)
A	5.2
B	11.1
C	3.0
D	3.5
(E)	76.7

% correct			
HK			**International**
Boys	**Girls**	**Overall**	
77.1	76.5	76.7	59.0

Noteworthy Points on Students' Performance

The two questions tested the same concept i.e. applying the concept of controlling variables in experimental design. In 1995, a substantial number of students were misled by the distractor options (A-22%, B-7%), both of which had the experimental setup placed inside a dark cupboard. A similar pattern was observed in the 2003 question, though the percentage correct improved from 57% to 77%. Plausible reasons are:

1. ***Rote learning*** – Students probably did not read the question carefully. They anticipated that the question concerned the dependence of photosynthesis on sunlight. So, many were attracted to the options with the setup housed inside a dark cupboard. That is, they tended to focus on the effect of sunlight and neglected the influence of other factors. In other words, they tended to recall the kind of standard

experiments they had carried out or had been taught on this topic. This reflected that they rote-learned about photosynthesis rather than applying the concept of controlling variables in experimental design, as required by this question.

2. ***Curriculum emphasis* –** The better performance shown in 2003 can be attributed to the curriculum emphasis on NOS since 2000.

Implications / Suggestions for Improvement

1. Avoid relying too much on cookbook recipe-type experiments. Have students design their own scientific investigations whenever possible.
2. Explain clearly the general principles of experimental design – the importance of identifying the hypothesis, the dependent variable and the independent variable. Subsequently, it is essential to practice relating the general principles to specific examples. For example, the experimental design for investigating a plant's need for minerals has to be different from that for investigation of its need for light.
3. Discuss concepts related to experimental design using contexts that are familiar to students. Start with simple examples (e.g. testing the properties of different kinds of toilet paper), and then move on to more complex examples that involve controlling of various other variables (e.g. factors affecting plant growth or the rusting of iron).

5.2.4 Experimental design - choice of instrument

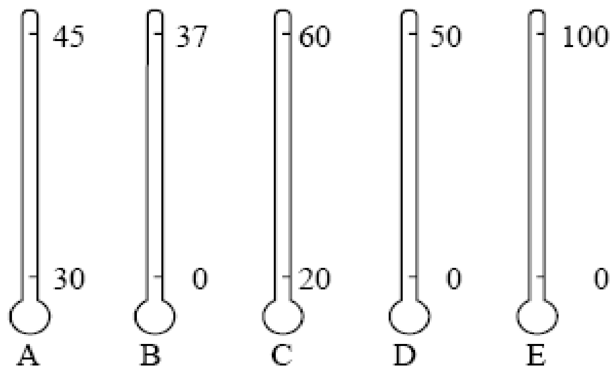
Content / Concept assessed: Percentage error

Relevant Science (S1-3) topic: Introducing science (Unit 1)

The Question (1995)

The diagram shows five different Celsius thermometers. The body temperature of sick people ranges from about 36℃ to 42℃. Which thermometer would be most suited for accurately measuring body temperature?

A. thermometer A
B. thermometer B
C. thermometer C
D. thermometer D
E. thermometer E



Result

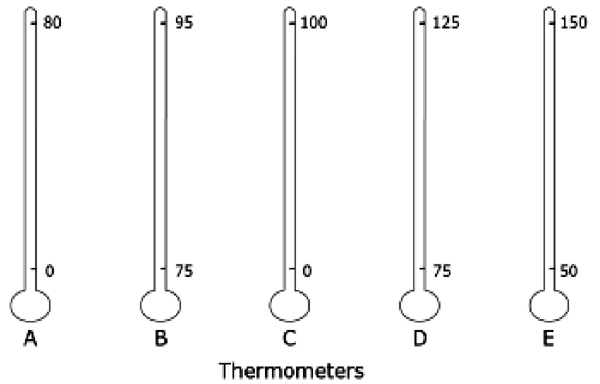
	HK (%)
(A)	48.3
B	12.9
C	10.0
D	13.2
E	14.9

% correct			
HK			International
Boys	Girls	Overall	
52.9	43.1	48.3	61.1

The Question (2003)

At different altitudes, the boiling points of water ranges from about 80℃ to 100℃. Which of the Celsius thermometers shown below would give the most accurate measurement of the boiling point of water at different altitudes?

A. Thermometer A
B. Thermometer B
C. Thermometer C
D. Thermometer D
E. Thermometer E



Result

	HK (%)
A	1.4
B	3.4
C	43.9
(D)	37.8
E	13.4

% correct			
HK			International
Boys	Girls	Overall	
39.6	36.2	37.8	20.1

Noteworthy Points on Students' Performance

Both in 1995 and 2003, the percentages of HK students who answered correctly were low (48% and 38% respectively). This could be due to several reasons:

1. ***Everyday life experience –*** Students may be more used to thermometers in daily life for measuring room temperature than body temperature. In fact, nowadays digital thermometers are more common than traditional mercury thermometers so students do not have much exposure to scaled thermometers. Some students might not realize that the thermometers shown in the options were simply hypothetical and thus tended to reject those thermometers with a lowest marking of 50°C and

75°C.

2. ***Common misconception –*** Students lack the concepts related to accuracy and precision of an instrument. The concept on percentage error is also weak. They wrongly believe that the wider the range of the thermometer, the more accurate it is.
3. ***Curriculum / Teaching emphasis –*** First, the HK curriculum rarely requires students to choose between different types of thermometers. Second, students are used to memorizing information but lack the crucial skills to apply the knowledge. Third, students are relatively weak in experimental design and the concept of instrument sensitivity.

Implications / Suggestions for Improvement

1. Widen the scope of teaching from just focusing on thermometers used in the laboratory, to other types of thermometers used in different settings e.g. clinical thermometer, cooking thermometer, etc. Ask students to note the upper and lower limits of temperature that can be measured by each kind of thermometer. To bring out the idea of accuracy and precision, discuss why there is a necessity for different kinds of thermometer for different purposes.
2. Discuss how instruments with a small range can measure more precisely. For example, ask students to measure a set amount of liquid with measuring cylinders of different capacities.
3. Carry out an experiment to demonstrate that thermometers with a smaller range can measure temperature more accurately than thermometers with a larger range. Do this by arranging students into different groups and give each group one of each type of thermometer – large, medium and small scale range – to measure the temperature of a substance.
4. Allow students to choose instruments of different scales/sizes/capacities for their experiments, e.g. ammeters, voltmeters, measuring cylinders, balances, etc. Require students to justify their choices.
5. Science Panel may collaborate with the Mathematics Panel to help students solve problems on percentage errors and make decisions on the choice of measuring instruments. (Graphing scales would be another opportunity for collaboration.)

5.2.5 Replication of measurements

Content / Concept assessed: Experimental error

Relevant Science (S1-3) topic: Introducing science (Unit 1)

The Question (1999)

The primary reason scientists repeat the measurements they take during experiments is so that they can

A. check that the equipment is working

B. list all the results in a table

C. estimate experimental error

D. change the experimental conditions

Result

	HK (%)
A	9.8
B	24.5
(C)	55.7
D	8.7

% correct			
HK			International
Boys	Girls	Overall	
55.5	56.8	55.7	40.3

Noteworthy Points on Students' Performance

One quarter of HK students erroneously thought 'listing all the results in a table' as the main reason for scientists to repeat measurements during experiments. Possible explanations of students acquiring such a misconception include:

1. ***Lack of knowledge / Curriculum / Teaching emphasis –*** Students might not know what experimental error is. Even though the notion of accuracy in measurement is mentioned in the 2000 curriculum, the concept of experimental error is rarely addressed explicitly in textbooks and/or during teaching.
2. ***Routine recipe-type practical work –*** Many teachers who attended the TIMSS workshops laughed at this set of results. But they soon stopped laughing when they realized that this set of results was reflective of how some of the teachers conducted their laboratory classes. For example, before 2000, students were rarely asked to design their own experiments; instead, they were asked to follow predetermined procedures and to enter the results into predetermined tables in worksheets. Even worse, there was no discussion of why things need to be done in a particular way. Such an approach has conditioned our students to think 'listing all the results in a table' is the main reason why scientists repeat the measurements they take during experiments.

Implications / Suggestions for Improvement

1. Introduce the concept of experimental errors using daily life examples. For example, the time records of a 100-metre race of the same athlete taken by two timers are often different.
2. Provide more opportunities for students to design their own investigations. Require them to justify their design, including the number of readings that needs to be taken for a particular variable.
3. Conduct post-lab discussions on experimental results obtained and bring out the concept of experimental error and the need for repeated measurements. Below are examples of investigations that can serve these purposes:
   i. The water rocket "水火箭" experiment – Divide the class into groups and assign a specific role to each student (e.g. starter, measurer, etc.). In terms of investigative method, control the volume of water and gas pressure inside the water rocket, and keep all other conditions constant. The aim is to measure the height reached by the water rocket in several trials and observe the differences in results. Guide students to understand that the variation in results is affected by a number of environmental factors which they cannot manipulate. Ask what they can do to improve the reliability of the result.
   ii. This experiment involves measuring the time taken for a table tennis ball to fall to the ground from a certain height. First, take a table tennis ball and hold it at a distance from the ground. Measure the time taken for the ball to reach the ground, and repeat this a few times. It is most likely that the time taken would vary with each measurement. This could prove to students the existence of random error and convey the message that random errors can be reduced by taking the average of all the measurements.
   iii. Another demonstration entails decolourization of DCPIP solution by orange/lemon juice. First, divide students into groups, then hand out identical amount of orange/lemon juice and DCPIP solution to each group. Ask each group to measure the number of DCPIP droplets that can be decolourized by an agreed amount of orange/lemon juice. Get each group to put their result on the blackboard to compare with the results of the whole class. Students can readily see that, because of random errors, the result can vary between different groups, even though they are using an identical experimental set-up.

5.2.6 Data Interpretation

Content / Concept assessed: Conditions necessary for photosynthesis

Relevant Science (S1-3) topic: Living things and air (Unit 7)

The Question (1995)

Amount of oxygen produced in a pond	
Location	**Oxygen produced**
Top Metre	4 grams/cubic metre
Second Metre	3 grams/cubic metre
Third Metre	1 gram/cubic metre
Bottom Metre	0 gram/cubic metre

Which statement is consistent with the data in the table?

A. More oxygen production occurs near the surface because there is more light there.

B. More oxygen production occurs near the bottom because there are more plants there.

C. The greater the water pressure, the more oxygen production occurs.

D. The rate of oxygen production is not related to depth.

Result

	HK (%)
(A)	48.9
B	19.6
C	16.9
D	13.7

% correct			
HK			International
Boys	Girls	Overall	
49.7	49.8	48.9	52.5

Noteworthy Points on Students' Performance

The HK Average (49%) was below the International Average (53%), and those students who answered incorrectly chose the three distractor options, B, C and D, in very similar proportions. This could be because:

1. ***Comprehension skill* –** Students might not understand the meaning of 'second metre' and 'third metre', and that 'location' equals 'depth of water'.
2. ***Data interpretation / Application of knowledge* –** Students were weak in data interpretation, especially in recognition of trends / patterns in data. In this case, students have, first, to discern the pattern of the data before moving on i.e. the shallower the water, the more oxygen is produced, or vice versa. Without identifying and establishing this pattern, it is unlikely that the students will be able

to explain why such a pattern occurs. The latter requires students to apply their knowledge of photosynthesis. In short, this question is testing students' understanding about the inter-relationship among the three variables: depth of water, light intensity and rate of photosynthesis. It requires students to work with the data step by step. That is, data analysis → pattern recognition → explanation / drawing conclusions.

Implications / Suggestions for Improvement

1. Teaching should emphasize the importance of information processing skills: data analysis → pattern recognition → explanation / drawing conclusions. It is important for teachers to model to students (by thinking aloud, drawing a graph on the board, etc.) how to discern the relationship between two variables from a given set of data. This should develop students' abilities to convert data into graphical format for easy identification of patterns and for drawing conclusions.
2. Have students practice the skills of identifying and describing the relationship among variables/patterns of data in different forms, such as a graph or a table.
3. Provide students with ample opportunities to verbalize and describe the patterns of data they observe from graphs and tables of results.

5.2.7 Drawing conclusion

Content / Concept assessed: Factors affecting evaporation

Relevant Science (S1-3) topic: The wonderful solvent - water (Unit 5)

The Question (1995)

A cupful of water and a similar cupful of gasoline were placed on a table near a window on a hot sunny day. A few hours later it was observed that both cups had less liquid in them but that there was less gasoline left than water. What does this experiment show?

A. All liquids evaporate.
B. Gasoline gets hotter than water.
C. Some liquids evaporate faster than others.
D. Liquids will only evaporate in sunshine.
E. Water gets hotter than gasoline.

Result

	HK (%)
A	7.7
B	14.8
(C)	67.6
D	5.5
E	3.9

% correct			
HK			International
Boys	Girls	Overall	
69.7	65.0	67.6	62.0

The Question (1999)

Two open bottles, one filled with vinegar and the other with olive oil, were left on a window sill in the Sun. Several days later it was observed that the bottles were no longer full. What can be concluded from this observation?

A. Vinegar evaporates faster than olive oil.
B. Olive oil evaporates faster than vinegar.
C. Both vinegar and olive oil evaporate.
D. Only liquids containing water evaporate.
E. Direct sunlight is needed for evaporation.

Result

	HK (%)
A	6.3
B	3.1
(C)	48.9
D	16.5
E	24.6

% correct			
HK			International
Boys	Girls	Overall	
50.1	47.8	48.9	47.7

Noteworthy Points on Students' Performance

Students were generally weak at drawing conclusions. In fact, the performance worsened in 1999 compared to 1995. Possible explanations are:

1. ***Learning principles versus learning facts / rote-learning –*** Students could not generalize the concept of evaporation from the examples they learned to other unknown liquids. Perhaps water is always used as an example both for teaching and in carrying out experiments. Students are not familiar with evaporation of liquids other than water; as a result, they believe that only substances containing water will evaporate (hence 17% chose the distractor option D in 1999 as the correct answer). Students were also unclear of the factors that affect evaporation. Many (25%) just rote-learned that a high temperature can speed up evaporation,

and hence got trapped by the distractor E in the 1999 question. Obviously, this group of students had overlooked the fact that both setups were left under direct sunlight and that there was no other control setup to allow them to draw the conclusion that ‘direct sunlight is needed for evaporation’. In other words, this question is testing students’ concepts related to scientific inquiry rather than those related to evaporation per se. Students who tend to rote learn are more likely to get it wrong.

2. ***Drawing conclusions versus speculation –*** Students who wrongly opted for B in the 1995 question seemed unable to come to grips with what constitutes a conclusion, which ought to be evidence-based, or to recognize that anything that goes beyond an evidence-based conclusion becomes a speculation (of which B is an example). Their speculation was based on their intuitive thinking that the hotter the liquid, the faster is its rate of evaporation.
3. ***Lack of practice –*** Students are weak in drawing conclusions because they tend to wait for and copy the answers given by their teacher.

Implications / Suggestions for Improvement

1. Consolidate students’ understanding of the concept of evaporation by repeating the evaporation experiment with different types of liquids. For example, place both water and alcohol on students’ hands. This demonstrates that alcohol, like water, can evaporate. The greater cooling effect of alcohol also demonstrates that alcohol evaporates at a faster rate.
2. Teachers teach students by modelling to them (thinking aloud with the help of blackboard drawing) how to make evidence-based conclusions. At the moment, most experiments conducted in schools only require students to verify taught theories rather to make genuine investigations.
3. The ability to accurately draw conclusions from observations made or data collected is difficult. To master this skill it is essential for students to practice more. In addition to carrying out more genuine investigations, teachers can ask students to answer more structured / data-response questions, and a wider variety of questions, in order to practice how to deduce and conclude. Train students to read the question carefully and highlight its main points.

5.3 Important Lessons Learned from Students' Performance

As evident in the preceding discussions, the important lessons learned regarding improving science education in HK in general, and the teaching of nature of science and scientific inquiry, in particular, are as follows:

1. The science **curriculum** needs to be reviewed and **updated** continually so as to keep pace with the international trend towards increasing **emphasis** on teaching of NOS and scientific inquiry (see 5.2.1 - 5.2.7).
2. It is highly undesirable to teach NOS and scientific inquiry:
   - in a superficial manner (see 5.2.1)
   - in a decontextualized manner (see 5.2.1)
   - through routine cookbook recipe-type practical work (see 5.2.2, 5.2.3, 5.2.5)
3. Instead, efforts have to be made to teach NOS and scientific inquiry by:
   - providing students with opportunities to carry out **genuine investigations** (see 5.2.1, 5.2.2, 5.2.3) where they are allowed to choose from a variety of instruments and are required to justify their choices (see 2.2.4)
   - using **daily life examples** (see 5.2.1, 5.2.3, 5.2.5)
   - using **historical cases** (see 5.2.1)
   - providing students with **scaffolding activities** to help them acquire the necessary concepts, skills or habits, for example:
     i. distinguishing between hypothesis, observation and conclusion (see 5.2.1)
     ii. drawing a table of variables when designing an investigation (see 5.2.2)
     iii. designing an investigation with reference to the hypothesis to be tested (see 5.2.3)
     iv. discussing the choice of instruments in relation to minimizing experimental errors (see 5.2.4)
     v. discussing the need for repeated measurements (see 5.2.5)
     vi. modelling to students (by thinking aloud) the proper way to analyze and interpret data (including identification of patterns and trends in data) (see 5.2.6)
     vii. modelling to students how to draw evidence-based conclusions (see 5.2.7)
4. Efforts should be made to rectify students' **rote-learning** habit (see 5.2.3, 5.2.7).

Chapter 6

Gender Differences in Students' Performance

Man Wai Cheng, Benny Hin Wai Yung and Siu Ling Wong

6.1 Introduction

Gender differences in science achievement have received serious attention in science education research for at least two decades. Indeed, this is an area that has been closely monitored in international surveys like TIMSS. The TIMSS 2003 International Science Report [11] not only includes evidence of gender differences in science achievement across different nations or regions in 2003, it also provides information on gender achievement differences across the countries as observed in TIMSS 2003, 1999 and 1995. The trend revealed by the data indicates that gender differences in science achievement remains a contentious issue to be further explored.

Gender differences in science achievement raise concerns about the issue of equity in science education. According to research studies, the better performance of boys in science has been attributed to a variety of factors, such as girls' lack of exposure to science-related activities outside the classroom, gender biases of teachers with respect to strategies for asking questions and probing answers, cultural influences from society and school, gender differences in spatial abilities, cognitive abilities, learning styles, mathematics background and test item response format. Nevertheless, more detailed analysis and further research is still needed in order to better understand the reasons and hence to improve practice.

In TIMSS 1995, Hong Kong (HK) showed the largest gender difference in science performance in favour of boys at the 8th grade among all participating countries that satisfied the sampling and participation requirements for international comparison. Though the situation has greatly improved in TIMSS 1999 and 2003, a significant performance gap between girls and boys still exists. In the present study, data from TIMSS 1995, 1999 and 2003 were used to explore what is evidently still an important

---

11 Martin, M., Mullis, I.., Gonzalez, E. and Chrostowski, S. (2004). *TIMSS 2003 International Science Report*. Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Boston College.

issue for HK science education.

The following sections begin with a brief introduction of the situation in HK regarding gender differences in science achievement over recent years. Next, we analyze the performance of girls and boys in the different content domains. This is followed by an analysis of gender differences in performance at different ability levels, the interaction between teacher gender and student gender and its effect on student achievement. Focus is then directed to individual test items in which gender differences in performance were particularly evident. Subsequently, implications and suggestions for improvement are addressed.

6.2 General Trends

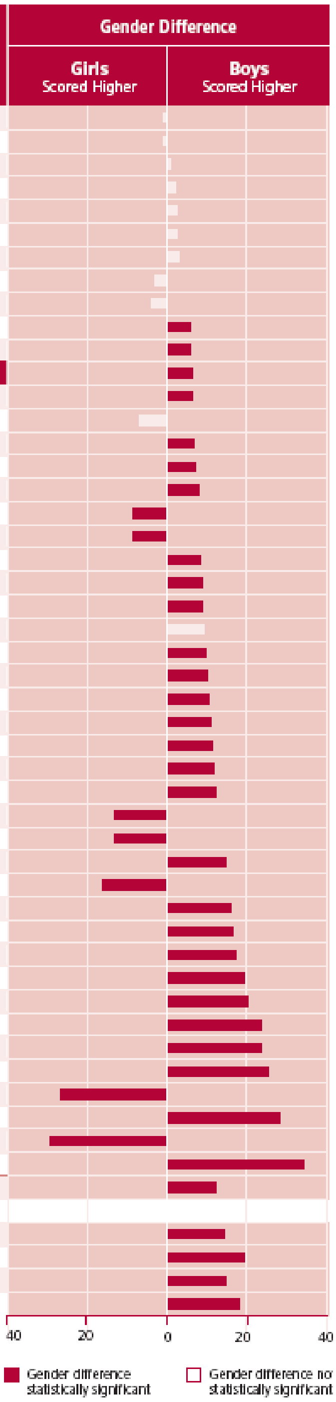
Figure 6.1 shows the gender differences in science achievement in 2003. It presents average achievement separately for girls and boys for each of the TIMSS 2003 countries, as well as the difference between the means. Countries are shown in increasing order of their gender difference. The gender difference for each country is shown by a bar, indicating the amount of the difference, whether the direction of the difference favoured girls or boys, and whether the difference is statistically significant (indicated by a darkened bar). The darkened row in the middle of the figure shows the international average. The International Average is the weighted mean of all participating countries. There was an International Average of 6 scale-score points favouring boys, a statistically significant difference. For HK, there was a 9 scale-score points favouring boys, a statistically significant difference. However, it should be noted that the results varied considerably from country to country. In many countries the gender difference was negligible. These suggest that there may be cultural factors specific to individual countries that account for the varying gender differences across the countries.

Figure 6.2 shows the trends in average science achievement by gender across the participant countries. In the figure, the achievement differences between that of TIMSS 2003 and 1999, as well as those between TIMSS 2003 and 1995 are presented separately for girls and for boys. Across all countries the girls showed, on average, a ten-point improvement from 1995 to 2003, while the boys showed no improvement. Fifteen participant countries showed significant improvement for girls, and just eight for boys.

Figure 6.1   Average science achievement by gender in TIMSS 2003

(Adopted from TIMSS 2003 International Science Report, Martin et. al, 2004)

Countries	Girls		Boys		Difference (Absolute Value)
	Percent of Students	Average Scale Score	Percent of Students	Average Scale Score	
Egypt	46 (2.7)	422 (4.8)	54 (2.7)	421 (5.5)	1 (6.8)
Iran, Islamic Rep. of	40 (4.1)	454 (3.9)	60 (4.1)	453 (3.7)	1 (6.1)
Chinese Taipei	48 (1.0)	571 (3.8)	52 (1.0)	572 (3.8)	1 (3.1)
Botswana	51 (0.7)	364 (3.2)	49 (0.7)	366 (3.4)	2 (3.3)
South Africa	51 (0.9)	242 (7.2)	49 (0.9)	244 (7.7)	2 (6.1)
Lebanon	57 (1.8)	392 (4.8)	43 (1.8)	395 (6.0)	3 (6.4)
Singapore	49 (0.8)	576 (4.0)	51 (0.8)	579 (5.0)	3 (3.1)
Estonia	50 (1.0)	554 (2.8)	50 (1.0)	551 (2.9)	3 (2.8)
Cyprus	49 (0.6)	443 (2.3)	51 (0.6)	440 (2.8)	4 (3.0)
1 Lithuania	50 (0.9)	516 (2.7)	50 (0.9)	522 (2.4)	6 (2.5)
1 Serbia	49 (0.8)	465 (2.9)	51 (0.8)	471 (2.6)	6 (2.5)
International Avg.	**50 (0.2)**	**471 (0.7)**	**50 (0.2)**	**477 (0.7)**	**6 (0.6)**
Slovenia	50 (0.9)	517 (2.4)	50 (0.9)	524 (2.3)	7 (3.0)
Philippines	58 (0.9)	380 (5.9)	42 (0.9)	374 (6.4)	7 (4.1)
Latvia	49 (0.8)	509 (2.6)	51 (0.8)	516 (3.0)	7 (2.4)
Sweden	51 (0.9)	521 (3.2)	49 (0.9)	528 (2.7)	8 (2.5)
Norway	50 (0.8)	490 (2.2)	50 (0.8)	498 (3.0)	8 (2.9)
2 Macedonia, Rep. of	49 (0.9)	454 (3.7)	51 (0.9)	445 (4.2)	8 (3.3)
Moldova, Rep. of	51 (0.8)	477 (3.5)	49 (0.8)	468 (3.7)	8 (2.6)
Romania	52 (0.9)	465 (5.5)	48 (0.9)	474 (4.9)	9 (3.5)
† Hong Kong, SAR	50 (2.4)	552 (3.4)	50 (2.4)	561 (3.8)	9 (3.9)
Japan	49 (1.2)	548 (3.0)	51 (1.2)	557 (2.7)	9 (4.5)
New Zealand	52 (1.7)	515 (4.8)	48 (1.7)	525 (6.7)	9 (5.7)
Malaysia	50 (1.8)	505 (4.3)	50 (1.8)	515 (4.0)	10 (4.0)
Italy	50 (0.9)	486 (2.7)	50 (0.9)	496 (3.8)	10 (2.5)
Russian Federation	49 (1.2)	508 (3.7)	51 (1.2)	519 (4.2)	11 (3.1)
1 ‡ Morocco	50 (1.8)	392 (3.2)	50 (1.8)	403 (3.8)	11 (4.6)
1 Indonesia	50 (0.7)	415 (3.9)	50 (0.7)	426 (4.6)	11 (2.7)
† Scotland	50 (1.3)	506 (4.0)	50 (1.3)	517 (3.5)	12 (3.6)
↔ Korea, Rep. of	48 (2.8)	552 (2.1)	52 (2.8)	564 (1.9)	12 (2.5)
Armenia	53 (0.7)	468 (4.0)	47 (0.7)	455 (3.4)	13 (2.8)
Palestinian Nat'l Auth.	55 (2.4)	441 (3.7)	45 (2.4)	428 (5.2)	13 (6.2)
† Netherlands	49 (1.2)	528 (3.3)	51 (1.2)	543 (3.8)	15 (3.5)
Saudi Arabia	43 (2.3)	407 (6.2)	57 (2.3)	391 (5.4)	16 (8.2)
Bulgaria	48 (1.3)	470 (6.3)	52 (1.3)	487 (5.2)	16 (5.2)
‡ United States	52 (0.7)	519 (3.2)	48 (0.7)	536 (3.4)	16 (2.1)
Slovak Republic	48 (1.3)	508 (3.8)	52 (1.3)	525 (3.4)	18 (3.1)
2 Israel	52 (1.6)	479 (3.2)	48 (1.6)	498 (4.1)	20 (4.2)
Australia	51 (2.2)	517 (4.6)	49 (2.2)	537 (4.6)	20 (5.6)
Belgium (Flemish)	54 (2.1)	505 (3.0)	46 (2.1)	528 (3.4)	24 (4.2)
Tunisia	53 (0.7)	392 (2.3)	47 (0.7)	416 (2.6)	24 (2.6)
Hungary	50 (1.0)	530 (3.4)	50 (1.0)	556 (3.0)	26 (3.0)
Jordan	49 (1.7)	489 (4.5)	51 (1.7)	462 (5.6)	27 (6.9)
Chile	48 (1.6)	398 (3.2)	52 (1.6)	427 (3.6)	29 (4.0)
Bahrain	50 (0.4)	453 (2.7)	50 (0.4)	423 (2.3)	29 (3.5)
Ghana	45 (0.9)	236 (6.4)	55 (0.9)	271 (6.5)	35 (4.7)
ǂ England	50 (2.4)	538 (4.7)	50 (2.4)	550 (5.1)	12 (5.3)
Benchmarking Participants					
Basque Country, Spain	49 (1.7)	481 (3.0)	51 (1.7)	496 (3.3)	15 (3.6)
Indiana State, US	49 (1.2)	521 (4.7)	51 (1.2)	540 (5.3)	19 (3.1)
Ontario Province, Can.	51 (0.9)	526 (3.1)	49 (0.9)	540 (2.8)	15 (2.6)
Quebec Province, Can.	50 (1.6)	522 (3.7)	50 (1.6)	540 (3.2)	18 (3.1)



† Met guidelines for sample participation rates only after replacement schools were included (see Exhibit A.9).

‡ Nearly satisfied guidelines for sample participation rates only after replacement schools were included (see Exhibit A.9).

ǂ Did not satisfy guidelines for sample participation rates (see Exhibit A.9).

1 National Desired Population does not cover all of International Desired Population (see Exhibit A.6).

2 National Defined Population covers less than 90% of National Desired Population (see Exhibit A.6).

↔ Korea tested the same cohort of students as other countries, but later in 2003, at the beginning of the next school year.

( ) Standard errors appear in parentheses. Because results are rounded to the nearest whole number, some totals may appear inconsistent.

Figure 6.2 Trends in average science achievement by gender

(Adopted from TIMSS 2003 International Science Report, Martin et. al, 2004)

Countries	Girls			Boys		
	2003 Average Scale Score	1999 to 2003 Difference	1995 to 2003 Difference	2003 Average Scale Score	1999 to 2003 Difference	1995 to 2003 Difference
Australia	517 (4.6)	– –	10 (6.0)	537 (4.6)	– –	18 (7.1) ▲
Belgium (Flemish)	505 (3.0)	-21 (5.4) ▼	-19 (9.2) ▼	528 (3.4)	-16 (7.9) ▼	-14 (9.7)
Bulgaria	470 (6.3)	-41 (8.6) ▼	-78 (8.8) ▼	487 (5.2)	-38 (8.3) ▼	-56 (7.6) ▼
Chile	398 (3.2)	-11 (5.6)	◇ ◇	427 (3.6)	-5 (6.2)	◇ ◇
Chinese Taipei	571 (3.8)	10 (5.5)	◇ ◇	572 (3.8)	-6 (6.6)	◇ ◇
Cyprus	443 (2.3)	-11 (4.2) ▼	-11 (3.6) ▼	440 (2.8)	-26 (4.2) ▼	-11 (3.8) ▼
Hong Kong, SAR	552 (3.4)	29 (5.7) ▲	60 (7.4) ▲	561 (3.8)	24 (6.2) ▲	36 (7.4) ▲
Hungary	530 (3.4)	-10 (5.5)	5 (4.8)	556 (3.0)	-10 (5.4)	7 (4.7)
Indonesia	415 (3.9)	-12 (7.7)	◇ ◇	426 (4.6)	-18 (6.7) ▼	◇ ◇
Iran, Islamic Rep. of	454 (3.9)	24 (6.9) ▲	6 (7.0)	453 (3.7)	-7 (5.7)	-22 (5.8) ▼
Israel	479 (3.2)	18 (6.8) ▲	– –	498 (4.1)	23 (7.0) ▲	– –
Italy	486 (2.7)	1 (4.9)	– –	496 (3.8)	-7 (7.2)	– –
Japan	548 (3.0)	5 (4.0)	3 (3.5)	557 (2.7)	0 (4.1)	-7 (3.6) ▼
Jordan	489 (4.5)	29 (6.8) ▲	◇ ◇	462 (5.6)	20 (8.3) ▲	◇ ◇
Korea, Rep. of	552 (2.1)	14 (4.4) ▲	22 (3.2) ▲	564 (1.9)	5 (4.0)	6 (3.4)
Latvia (LSS)	511 (3.2)	16 (5.9) ▲	48 (5.0) ▲	515 (3.3)	5 (6.0)	25 (5.4) ▲
Lithuania	516 (2.7)	38 (5.2) ▲	64 (5.2) ▲	522 (2.4)	23 (5.6) ▲	45 (5.1) ▲
Macedonia, Rep. of	454 (3.7)	-4 (7.1)	◇ ◇	445 (4.2)	-13 (6.6) ▼	◇ ◇
Malaysia	505 (4.3)	17 (7.1) ▲	◇ ◇	515 (4.0)	18 (7.1) ▲	◇ ◇
Moldova, Rep. of	477 (3.5)	22 (5.7) ▲	◇ ◇	468 (3.7)	3 (6.2)	◇ ◇
Netherlands	528 (3.3)	-8 (8.0)	0 (6.5)	543 (3.8)	-11 (8.2)	-11 (8.3)
New Zealand	515 (4.8)	9 (7.0)	18 (7.5) ▲	525 (6.7)	11 (9.7)	1
Norway	490 (2.2)	◇ ◇	-16 (3.4) ▼	498 (3.0)	◇ ◇	-25 (4.8) ▼
Philippines	380 (5.9)	29 (10.2) ▲	◇ ◇	374 (6.4)	35 (11.3) ▲	◇ ◇
Romania	465 (5.5)	-3 (8.0)	2 (7.7)	474 (4.9)	-1 (8.0)	-4 (7.5)
Russian Federation	508 (3.7)	-11 (8.0)	-7 (5.9)	519 (4.2)	-21 (7.3) ▼	-12 (6.4)
Scotland	506 (4.0)	◇ ◇	19 (6.6) ▲	517 (3.5)	◇ ◇	3 (7.5)
Singapore	576 (4.0)	19 (8.8) ▲	3 (7.8)	579 (5.0)	1 (10.9)	-8 (8.6)
Slovak Republic	508 (3.8)	-17 (5.0) ▼	-12 (5.7) ▼	525 (3.4)	-21 (5.6) ▼	-20 (4.7) ▼
Slovenia	517 (2.4)	– –	13 (3.8) ▲	524 (2.3)	– –	0 (4.0)
South Africa	242 (7.2)	8 (11.6)	– –	244 (7.7)	-9 (10.8)	– –
Sweden	521 (3.2)	◇ ◇	-26 (6.0) ▼	528 (2.7)	◇ ◇	-31 (5.5) ▼
Tunisia	392 (2.3)	-25 (3.9) ▼	◇ ◇	416 (2.6)	-26 (4.5) ▼	◇ ◇
United States	519 (3.2)	14 (5.8) ▲	14 (6.3) ▲	536 (3.4)	11 (6.3)	16 (6.9) ▲
‡ England	538 (4.7)	16 (7.9) ▲	15 (6.3) ▲	550 (5.1)	-4 (7.3)	7 (8.0)
International Avg.	**486 (0.7)**	**7 (1.2) ▲**	**3 (1.3) ▲**	**495 (0.8)**	**0 (1.2)**	**-5 (1.4) ▼**
Benchmarking Participants						
Indiana State, US	521 (4.7)	-3 (8.4)	◇ ◇	540 (5.3)	-5 (9.3)	◇ ◇
Ontario Province, Can.	526 (3.1)	17 (4.7) ▲	38 (4.7) ▲	540 (2.8)	13 (4.3) ▲	35 (5.5) ▲
Quebec Province, Can.	522 (3.7)	-14 (7.7)	16 (8.5)	540 (3.2)	-5 (5.6)	26 (8.1) ▲

▲ 2003 significantly higher

▼ 2003 significantly lower

‡ Did not satisfy guidelines for sample participation rates (see Exhibit A.9).

Trend notes: Because of differences in population coverage, 1999 data are not shown for Australia and Slovenia, and 1995 data are not shown for Israel, Italy, and South Africa. Korea tested later in 2003 than in 1999 and 1995, at the beginning of the next school year. Similarly, Lithuania tested later in 1999 than in 2003 and 1995. Data for Latvia in this exhibit include Latvian-speaking schools only.

( ) Standard errors appear in parentheses. Because results are rounded to the nearest whole number, some totals may appear inconsistent.

A dash (–) indicates comparable data are not available.

A diamond (◇) indicates the country did not participate in the assessment.

For HK, a dramatic increase of 60 scale-points from 1995 to 2003 was recorded for girls. Between 1999 and 2003, there was an increase of 29 scale-points. During the corresponding periods, significant improvements were also shown by the boys, though not as great as those observed in the girls. The exact figures for the boys are an increase of 36 scale-points from 1995 to 2003, and an increase of 24 scale-points from 1999 to 2003. Overall, the 60 scale-points improvement in the girls' performance, compared with 36 scale-points for boys, has significantly reduced the gender performance gap from a difference of 33 scale-points in 1995 to a difference of only 9 scale points in 2003, though still in favour of boys. This was the most remarkable improvement in girls' performance among all participant countries.

6.3 Gender Differences in Different Content Domains

In the last section, we reported that there was a 9 scale-points difference in science achievement in favour of the boys in TIMSS 2003. In this section, we attempt to explore the gender differences in science achievement further. Table 6.1 shows the mean science achievement score in each content domain by gender. HK boys outperformed girls in all content domains. The differences in mean score ranged from 19 scale-points in Earth Science to 2 scale-points in Life Science and Chemistry. Significant gender differences in performance were found in Earth Science and Physics. In these two content domains, boys outperformed girls by 19 and 12 scale-points, respectively. These results are consistent with those found in the USA and the UK. That is, the male advantage is largely in the physical sciences, rather than life sciences.

Table 6.1 Gender differences in performance in different content domains

Science Content Domain	Girls		Boys		Boys - Girls
	Mean	S.E.	Mean	S.E.	
Earth Science	539	3.4	558	3.5	19*
Physics	549	3.6	561	3.6	12*
Environmental Science	554	3.0	557	3.6	3
Life Science	550	3.2	552	3.7	2
Chemistry	541	3.2	543	3.4	2
All Content Domains	552	3.4	561	3.8	9*

Note: * indicates that there is significant gender effect at $p < 0.05$.

According to the sociological interpretation of some research studies, the advantage of boys in physical sciences is mostly due to their previous experience from hobbies and games; and greater motivation, interest, and positive attitude towards science fostered by gender stereotypes that science is still a male dominant subject. We believe that it is important for science teachers to be well informed about this relative female underachievement in certain subject areas. In particular, teacher of physics and earth science have a responsibility to be aware of the effects of their teaching styles on male and female students, to be sensitive to the effects of competitive and collaborative work on students of different gender, and to illustrate the curriculum in ways that appeal to female students.

The differential performance of boys and girls in the different subject domains has implications for the concept of fairness in science assessment. Indeed, the presence of gender differences on any question or test raises the issue of bias. Should we include more life science questions than physical science questions in order to address the issue of gender differences in science achievement? What constitutes a gender-equitable assessment? On this issue we believe that it is inappropriate to attribute such differences to bias (i.e. systematic errors in measurement) when differences are a reflection of what actually occurs. Given the differences in attitude between girls and boys and their different early childhood experience, it is unsurprising that there are gender differences in certain areas of science education. There are, inevitably, some curricular areas such as electrical circuits or earth/space topics where it might be impossible to construct items for which male students had not had more relevant experience, given current attitudes and preferences. Therefore, test content should be selected on grounds of educational importance, not by the consideration of gender differences that it might or might not produce. For example, if it is deemed to be important for students to learn about planetary orbits, then it is appropriate to test this topic, even if consistent gender differences are found.

Research findings also suggest that boys are favoured by multiple-choice tests and girls by constructed-response items. Table 6.2 shows the gender differences in performance in the different content domains by response format. The results show that boys outperformed girls in the multiple-choice items and the difference is statistically significant. Overall, girls performed slightly better than boys in the constructed-response items, though the difference is not statistically significant. Their better performance than boys in constructed-response items on life science and chemistry is noteworthy. In sum, our results are supportive of previous research findings that boys are favoured by multiple-choice items. On the other hand, it is marginal to say that girls are favoured by constructed-response items.

Table 6.2 Gender differences in performance in different content domains by response format

Science Content Domain	**Gender Difference (Boys - Girls' Achievement)**			
	Constructed-Response		**Multiple-Choice**	
	Mean	**S.E.**	**Mean**	**S.E.**
Earth Science	3.61	2.25	6.08*	1.66
Physics	3.35	1.67	3.00*	1.24
Environmental Science	0.12	1.63	2.96	1.66
Life Science	- 1.64	1.06	1.72	1.01
Chemistry	- 0.76	1.39	2.32	1.32
All Content Domains	-0.52	0.71	3.15*	0.61

Note: * indicates that there is significant gender effect at $p < 0.05$.

Another way to analyze gender differences in science achievement is to compare the performance of girls and boys in the different ability areas, namely, Factual Knowledge, Conceptual Understanding, Analysis and Reasoning. Table 6.3 shows the gender differences in performance in these areas across the different content domains. Overall, boys outperformed girls in all the three abilities tested, with statistically significant results found in Factual Knowledge and Conceptual Understanding, but not in Analysis and Reasoning. At a more detailed level, we found that actually boys were outperforming girls in all aspects except Life Science (Analysis and Reasoning), Chemistry (Conceptual Understanding) and Environmental Science (Analysis and Reasoning), though a statistically significant gender effect, in favour of the boys, was only observed in Earth Science (Factual Knowledge & Analysis and Reasoning) and Physics (Analysis and Reasoning). Earth Science (Conceptual Understanding) is marginally significant (p=.054), as is Physics (Conceptual Understanding) (p=.052).

With respect to Analysis and Reasoning, although boys outperformed girls significantly in both Earth Science and Physics, and moderately in Chemistry, their overall performance lead was reduced by their poorer performance, relative to girls in Life Science and Environmental Science. This suggests the existence of an underlying effect of subject domain. The subject domain being tested seems to be playing a crucial role here. That is, Life Science and Environmental Science are feminine-oriented subjects.

Table 6.3 Gender differences in performance in different content domains by ability tested

Science Content Domain	Gender Difference (Boys - Girls' Achievement)					
	Factual Knowledge		Conceptual Understanding		Analysis and Reasoning	
	Mean	**S.E.**	**Mean**	**S.E.**	**Mean**	**S.E.**
Earth Science	7.06*	2.87	3.69	1.73	5.30*	1.86
Physics	1.59	3.64	3.14	1.52	3.79*	1.17
Environmental Science	2.20	2.16	5.63	1.40	- 0.31	1.66
Life Science	.779	1.15	0.47	1.19	- 2.04	1.83
Chemistry	3.67	2.09	- 0.13	1.46	1.79	1.77
All Content Domains	2.71*	0.97	1.89*	0.72	1.44	0.79

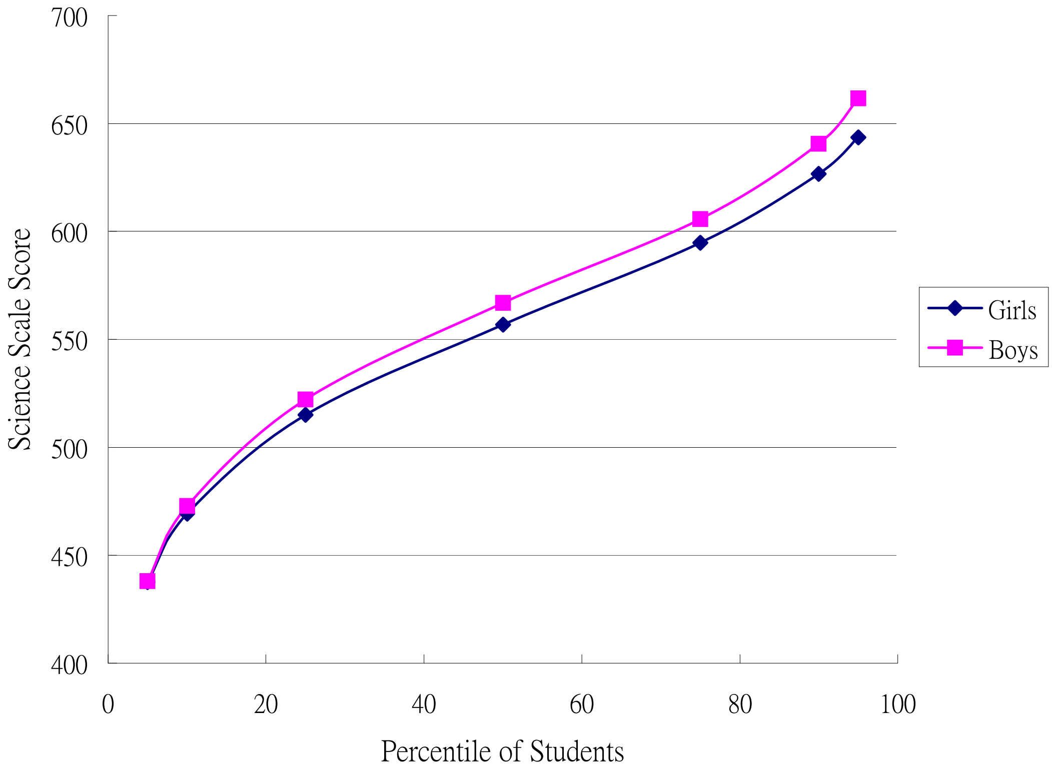
Note: * indicates that there is significant gender effect at $p < 0.05$.

6.4 Gender Differences across Different Ability Levels

Another way of examining the pattern of gender differences in science achievement is by looking at student performance in the different ability ranges. Figure 6.3 shows the distribution of achievement scores by gender at the different ability levels (in terms of percentile ranking of the students in the test). The results suggest that there is an interaction between gender and achievement level. The widening performance gap found at the upper percentiles indicates that boys not only performed better than girls on average, they were also performing significantly better than girls at the upper percentiles range (i.e. 50^{th} and above, at $p < 0.05$). This finding is consistent with a study conducted in the USA[12] which found that the gender gap on physical science questions in a test taken by 13-year-olds is more pronounced for more able students. This also has implications for teachers teaching more able students.

12 Lee, V. E. and Burkam, D. T. (1996). Gender differences in middle grade science achievement-subject domain, ability level and course emphasis. *Science Education*, 80, 613-650.

Figure 6.3 Student scores in science at different percentiles by gender



6.5 Interaction between Teacher Gender and Student Gender

Teachers are important mediators of student learning. Analyses of students' performance should be complemented with teacher variables. In this section, we examine the relations between teacher gender and student achievement.

Secondary analysis of the HK data in TIMSS 1995 and 1999 indicated that there is interaction between teacher gender and student gender on student achievement scores. Table 6.4 shows the student mean achievement scores by teacher gender and student gender from 1995 to 2003. For the girls, according to the 2003 data, teacher gender appears to have no effect on mean achievement scores. That is, girls performed equally well regardless of the teacher's gender. However, this finding is different from those derived from the 1995 and 1999 data. In these two earlier surveys, girls taught by male teachers performed much more poorly than girls taught by female teachers (by 25 scale-points less in 1995 and 13 scale-points less in 1999, respectively). In other words, the performance gap between the ‘girls taught by men’ and ‘girls taught by women’ has been narrowing from 1995 to 2003. In fact, this narrowing performance gap has been brought about by the improving performance of the ‘girls taught by men’.

A different picture is observed on the effect of teacher gender on boys’ achievement scores. In 1995, ‘boys taught by women’ performed better than ‘boys taught by men’ by 11 scale-points. However, beginning from 1999, the female teachers seemed to be losing ground to the male teachers. ‘Boys taught by women’ performed more poorly (instead of better, as in 1955) than ‘boys taught by men’ by 6 scale-points in 1999. The performance gap increased to 18 scale-points in 2003.

Table 6.4 Effect of teacher's gender on student's average science achievement

Year	**Teacher**	**Average Scale Score (S.E.)**			
		Girls		**Boys**	
		% Students	**Score**	**% Students**	**Score**
2003	Female	22	552(6)	19	550(7)
	Male	28	552(5)	31	568(5)
1999	Female	22	530 (7)	17	533 (11)
	Male	27	517 (5)	34	539 (6)
1995	Female	19	513 (11)	15	530 (8)
	Male	27	478 (9)	39	519 (9)

In order to depict the complicated interactions and trends, the results in Table 6.4 are plotted and shown in Figure 6.4. The dotted lines represent the trends in mean score obtained by students taught by female teachers. The solid lines represent the trends in mean score obtained by students taught by male teachers. Several major trends are summarized as follows.

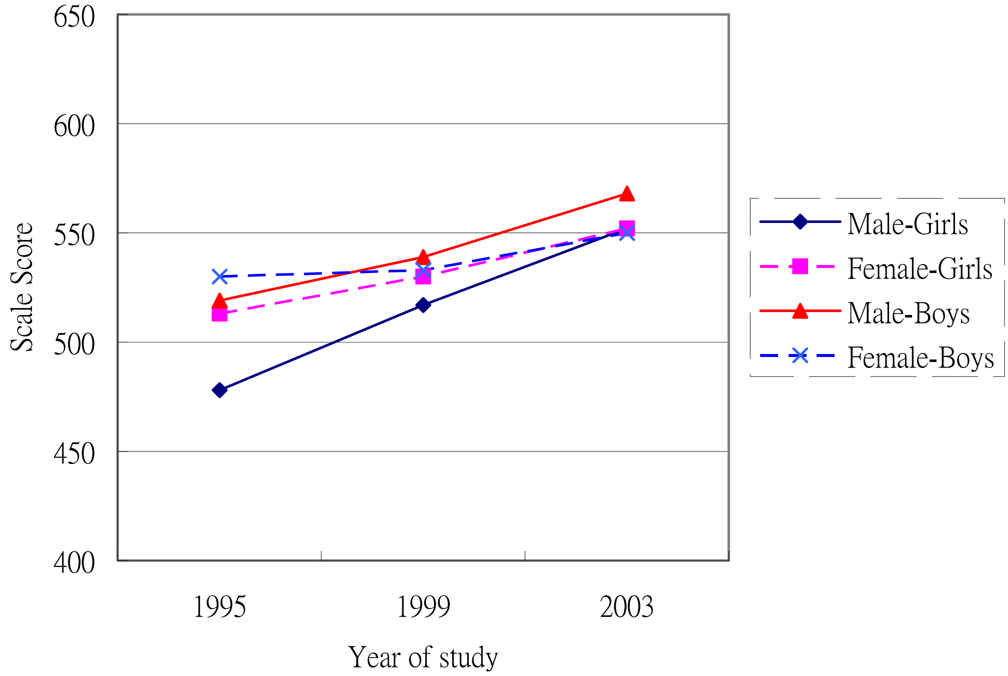
First, as shown by the upward trends of all the graphs, it can be concluded that regardless of the different combinations of teacher gender and student gender, HK students showed consistent improvements from 1995 to 2003. That is, be they male or female students, taught by male or female teachers, improvements were shown by students in all possible settings regarding teacher-student-gender combination.

Second, the steeper slopes of the two solid lines (compared with those of the dotted lines) indicate that the improvement in mean scores obtained by students taught by the male teachers was greater than those students who were taught by female teachers. In other words, the male teachers seemed to be doing a better job than their female counterparts in terms of raising students' achievement scores in science during this period of time.

Third, the solid line at the bottom of Figure 6.4 suggests that girls who were taught by male teachers had always been the group that performed least satisfactorily since 1995, though the performance gap had narrowed down in 1999 and even became on par with the other two teacher-gender-and-student-gender combinations in 2003. Indeed, the significant improvement of this group of female students contributed much to reducing the gender gap from 23 scale-points in 1995 to only 9 scale-points in 2003.

Fourth, there is a crossover between the Female-boys line and the Male-boys line at the top of the figure. This suggests that the 'beneficial effect' (of improving achievement score) of associating boys with female teachers has now shifted to that of associating with male teachers.

Figure 6.4 Interaction effect of teacher gender and student gender on student science achievement



Lastly, but more importantly, there is a large gap between the two solid lines whereas the gap between the two dotted lines is relatively much smaller. This suggests that there is teacher gender and student gender interaction. That is, boys and girls attain comparable mean scores when they are taught by the female teachers. On the other hand, boys and girls attain quite different mean scores when they are taught by male teachers, with the boys attaining a higher score and the girls a lower one. In short, the interaction between the gender of the teacher and that of their students is marked in the case of a male teacher, but not so for a female teacher. These findings concur with other research studies concluding that most of the variance in student achievement is due to the gender of teachers, not students. Such findings lead us to speculate whether the varying gender differences in student achievement scores observed among some of the TIMSS participating countries (refer to Figure 6.1) might also be explained, at least in part, by this teacher-student- gender interaction. That is, in addition to the other factors, be they cultural or not, the different teacher-student combinations, and their relative proportions within a country - might account, in part, for the observed gender differences in students' science achievement. Could this teacher-student-gender interaction explain why, for the same test instrument, there are such wide variations in gender differences across the participating countries, with some favouring boys while others favouring the girls? Could this be due to the different ratios of teachers of different sexes teaching male and female students. This is a factor that has been overlooked in studies on gender differences in learning.

6.6 Revisiting Gender Differences through Item Analysis

Through the above discussions, readers should now have a general picture of how various factors would play out and interact with one another to bring about gender differences in students' science achievement. In this section, we examine these factors in greater details by examining the item statistics of questions with profound gender differences in students' achievement. We begin with looking at items in which boys outperformed girls. Then, we examine items in which girls did better. This is followed by examples of items in which the gender difference in achievement actually varied with the changing social context. In the last part, we present an item with a huge gender difference in achievement and invite readers to explore with us the possible underlying factors and their implications for science teaching.

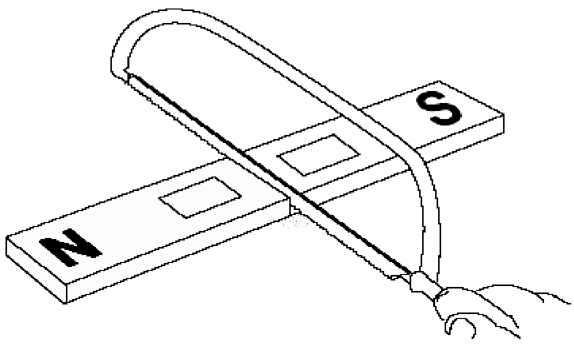
6.6.1 Context/domain areas in which boys performed better

As pointed out in section 6.3, boys outperformed girls in all content domains in TIMSS 2003, with the greatest performance gap found in earth science and physics (see also 3.2.2, 3.2.3, 3.25 for specific items). While this is generally true, readers may also recall that there were instances where the superior performance of boys over girls disappeared when the question becomes a bit more complicated. For example, in the series of questions on formation of image by reflection in a plane mirror (see 3.2.3), the better performance of boys over girls by 10% in the easier questions was attributed to the generally better ability of boys in spatial visualization. However, when the situation became more complicated, in the diagonally placed paint brush in 2003, no apparent gender difference was observed. Such a ‘vanishing gender difference upon complicated situations’ is also evident in the following example.

6.6.1.1 Polarity of broken magnet

Magnetism is not covered in the junior secondary science curriculum. The two questions shown below are examples of TIMSS items that are located in contexts beyond the formal science curriculum, and in which the boys outperformed the girls. The better performance of boys in these questions may be attributed to a difference in outside class activities between boys and girls. Boys might have played with magnets and heard about the phenomenon of “like poles repel and unlike poles attract”. However, when the situation becomes more complicated, as in the 2003 question, both boys’ and girls’ performance fell significantly. More importantly, not only did the gender difference favouring boys disappear, a reverse trend was observed instead. That is, girls did better than boys in the more complicated situation.

The Question (1995)



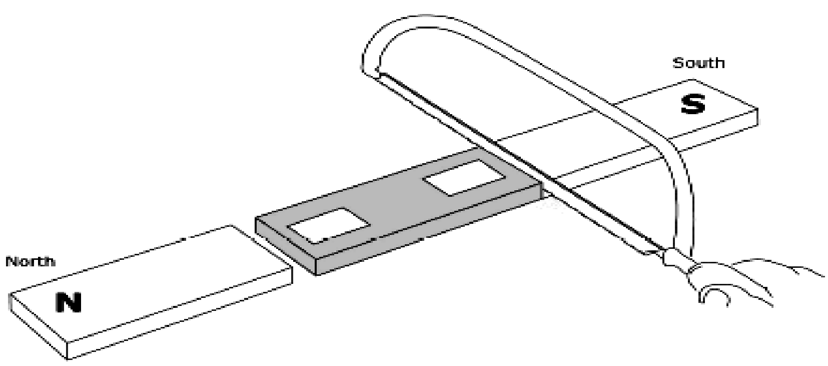
A bar magnet is cut in two with a hacksaw. Write an “N” or an “S” in each box on the diagram to show the polarity of the cut ends.

Result

% correct			
HK			
Boys	**Girls**	**B – G**	**Overall**
67.3	57.0	10.3	62.8

The Question (2003)

The diagram shows a bar magnet which is cut into three pieces with a hacksaw.



Write an "N" or an "S" in each box on the diagram to show the polarity of each end of the center piece.

Result

% correct			
HK			
Boys	**Girls**	**B – G**	**Overall**
37.4	42.5	-5.1	40.0

6.6.2 Context/domain areas in which girls performed better

Research studies indicate that, in general, girls perform better than boys in life sciences. However, this was not the case for TIMSS 2003: HK boys outperformed girls by 2 scale-points, though the difference is statistically not significant. A close examination of some of the life science items, as follows, may shed light on such a close performance between the two genders. We begin by looking at examples where girls did better than boys, followed by examples in which boys did better.

6.6.2.1 Food nutrition

HK girls outperformed boys in the following three questions. All of them tested students' knowledge on food nutrition, a topic in which girls usually show more interest than boys. Girls learn about this topic in their home economics lessons as well as in science lessons. Also, they may have more chances than boys to learn about these things at home, by assisting with cooking. (Please also refer to 4.2.3 for a discussion of which girls were found more capable in tackling questions in the context of kitchen materials.)

The Question (1995a)

What is the BEST reason for including fruits and leafy vegetables in a healthy diet?

A. They have a high water content.
B. They are the best source of protein.
C. They are rich in minerals and vitamins.
D. They are the best source of carbohydrates.

Result

	HK (%)
A	17.1
B	2.3
(C)	67.3
D	12.8

% correct			
HK			
Boys	**Girls**	**B – G**	**Overall**
63.1	72.4	-9.3	67.3

The Question (1995b)

Which of these meals would give you most of the nutrients that you need?

A. Meat, milk and a piece of chocolate.
B. Bread, vegetables and fish.
C. Vegetables, fruits and water.
D. Meat, fish and bread.

Result

	HK (%)
A	7.5
(B)	52.7
C	24.4
D	15.2

% correct			
HK			
Boys	**Girls**	**B – G**	**Overall**
48.5	57.8	-9.3	52.7

The Question (2003)

Eating leafy vegetables is important for human health. This is because leafy vegetables are a good source of which of the following?

A. Protein
B. Carbohydrates
C. Minerals
D. Fat

Result

	HK (%)
A	7.1
B	51.1
(C)	40.8
D	1.0

% correct			
HK			
Boys	**Girls**	**B – G**	**Overall**
36.1	45.8	-9.7	40.8

6.6.2.2 Health and diseases

In the question below, girls outperformed boys by about 12%. This is in line with other research findings that girls tend to be more interested in their own body than the boys. Because the question required students to answer in their own words, the better communication skills of girls might have also contributed to their better performance.

The Question (1995)

Jose caught influenza. Write down one way he could have caught it.

Result

% correct			
HK			
Boys	Girls	B – G	Overall
61.4	73.2	-11.8	28.5

6.6.2.3 Territorial behaviour

Though girls, in general, do better than boys in life science topics, as illustrated above, there are exceptions. The two questions below are cases in point. They test students on the same concept (i.e. the territorial behaviour of animals) though they are referring to different animals (bird in 1995; wolf in 1999). In both cases, boys performed better than girls by at least 10%. In addition, the gender difference was much more prominent in the case of wolves than that of birds. In fact, territory behaviour of animals was not in the curriculum. Students answered these questions based on their own knowledge. The results seemed to suggest that because boys are, in general, more aggressive than girls, they like to learn about animals' territorial behaviour. When the wolf was the animal in focus, the performance gap became even larger. A plausible explanation is that wolves are often projected as cunning animals in stories and are not popular among girls; who therefore show little interest in learning more about them, including their territorial behaviour.

The Question (1995)

When a bird sings, it is likely singing in order to

A. frighten away other types of birds

B. mark the bird's territory against the same type of birds

C. attract insects

D. wake up other animals

Result

	HK (%)
A	13.6
(B)	38.4
C	36.2
D	11.8

% correct			
HK			
Boys	Girls	B – G	Overall
43.1	32.9	10.2	38.4

The Question (1999)

When male wolves place their scent on trees, they most likely are doing this in order to

A. attract female wolves

B. attract prey

C. mark their territory against other wolves

D. mark the location of food supplies

Result

	HK (%)
A	37.5
B	6.6
(C)	54.1
D	1.4

% correct			
HK			
Boys	**Girls**	**B – G**	**Overall**
62.2	45.7	16.5	54.1

6.6.3 Gender differences and the changing social context

As pointed out earlier, boys outperformed girls in most physics related questions. However, it is interesting to note a shift in boys' and girls' performance from 1995 to 1999 in the following set of questions on sunburn/sunscreen and ultraviolet radiation. We attribute this shift in performance to a change in social context. In the late 80s to early 90s, advertisements concerning sunscreen were mostly related to water sports, to which boys may be more attracted. Hence boys outperformed girls in the 1995 question by 7%. In the late 90s, advertisements of this sort started to be related to skin care for cosmetic purposes, to which girls may be more attracted. This resulted in the better performance of girls over that of the boys by 5.4% in the 1999 question.

The Question (1995)

Which form of solar radiation causes sunburn?

A. Visible

B. Ultraviolet

C. Infra-red

D. X-rays

E. Radio waves

Result

	HK (%)
A	6.8
(B)	83.3
C	7.0
D	1.2
E	1.4

% correct			
HK			
Boys	**Girls**	**B – G**	**Overall**
86.6	79.5	7.1	83.3

The Question (1999)

Sunscreen is used to protect the skin from exposure to which type of solar radiation?

A. Visible

B. X-rays

C. Infra-red

D. Ultraviolet

E. Microwaves

Result

	HK (%)
A	6.8
(B)	1.6
C	6.1
D	84.5
E	0.9

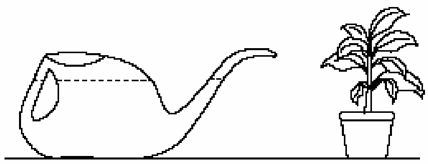
% correct			
HK			
Boys	**Girls**	**B – G**	**Overall**
81.8	87.2	-5.4	84.5

6.6.4 The greatest gender difference observed

It should be pointed out that many of the explanations we have put forward to explain gender differences are speculative. Sometimes, it is just too difficult to find a plausible explanation for the observed differences. Consequently, in this part, we present an item with the greatest gender difference in achievement and invite readers to explore with us the possible underlying factors and their implications for science teaching.

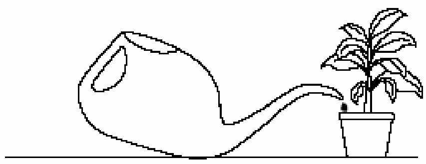
The Question (1995)

A watering can is almost filled with water as shown.



The watering can is tipped so that the water just begins to drip through the spout.

Draw a line to show where the surface of the water in the can is now.



Result

% correct			
HK			
Boys	**Girls**	**B – G**	**Overall**
82.2	56.3	25.9	70.5

We are surprised by the huge gender difference exhibited in this question, which is in fact the greatest among all the released items in the three years. The most common incorrect response is a line drawn parallel to the bottom of the container. The disappointing performance by girls may be an indication that they are less observant than boys to the details of daily life phenomena. Perhaps, under such circumstances, boys would attend to the water level inside the container (i.e. the physical phenomenon) and girls to the plant (i.e. the living organism). This seems to be a logical argument in view of our earlier discussion that boys perform better in physical science and girls do better in life sciences.

Of course, readers may not agree with these speculations. However, we are quite sure that you would agree with us that gender differences in students' science achievement is an issue that every teacher must take on board and must deal with in their day-to-day teaching.

6.7 Concluding Remark

Knowledge about the patterns of gender differences revealed in this chapter is important for the development of informed teaching strategies and decisions related to science education in junior secondary school. The results support implications from previous research about the necessity of diagnostic assessment and treatment that can compensate for disparities in boys' and girls' science related experiences outside school. Most importantly, the existing gender gap in science achievement may be ameliorated by a greater awareness of the various factors that may result in such a gender difference, including, notably, the possible effect of teacher-student-gender interaction - a factor often overlooked in studies of gender difference in student science achievement.

Chapter 7

What Have We Learned from TIMSS?

Benny Hin Wai Yung

As evident from the previous chapters, through discussions with teachers, we have learned a lot from the TIMSS data. The important lessons deriving from individual items can be found at the end of the corresponding chapters and will not be repeated here. Instead, attention is focused on a few major issues that run through the different chapters and discuss their implications for curriculum development as well as for teaching and learning of science.

7.1 Lessons on curriculum development

The first important lesson learned from TIMSS regarding curriculum development is the need for continual updating and renewal of our science curricula. In the process, we need to keep ourselves abreast of developments elsewhere. This is especially important if we are to compete with neighbouring countries in the increasingly competitive global economy. The importance of continual renewal of curriculum was fully reflected in the great improvement shown by our students in TIMSS 2003. They came 4^{th} in the performance list, which was a huge jump from their previous 14^{th} and 15^{th} positions in 1995 and 1999, respectively. Such a great improvement has been attributed to the major curriculum revision that took place in 2000, with increased emphasis on nature of science (NOS) and scientific inquiry. Indeed, our students' performance in this area has greatly improved, as indicated by evidence from item analysis of questions in this area before and after 2000 presented in Chapter 5.

Of course, it would be unwise to revise our curriculum just for the sake of improving our ranking in international comparative studies of this kind. In keeping ourselves abreast of international trends, we must not blindly follow what other countries are doing. Instead, we have to take into consideration the local context and be selective in making changes. We have to be sure that the changes made to the curriculum are for the good of our students in learning science. The inclusion of NOS and scientific inquiry discussed above is a case in point. Through scientific inquires, students come to acquire high order skills like problem solving, creativity and communication skills, all of which are important for lifelong learning and for students' future careers.

However, at times, it is difficult to decide whether certain topics should be included or excluded from the curriculum simply by making reference to the TIMSS data. For example, it is noted in 4.1 that Hong Kong (HK) students' relatively poor performance in some of the chemistry questions is attributed to the mismatch between HK Science curriculum (Secondary 1-3) and the assessed curriculum of TIMSS. Topics such as 'Atomic theory / Subatomic particles' and 'Element / Mixture / Compound' are either not explicitly addressed in the local curriculum and textbooks or had not yet been covered at Secondary 2 level when students took the assessment. That explains why so many students had difficulties in answering questions which demanded an understanding or application of concepts such as 'atom', 'ion', 'molecule' and 'compound'.

Actually, topics related to atomic theory / subatomic particles were once in the HK curriculum. The decision to remove this topic from the curriculum might have been based on the intuition that it is too abstract and difficult for junior secondary students. Or, because the English terminologies are just too difficult for students to master. The TIMSS data provides some evidence to test whether such intuitive thinking holds true (or not) and so provides a further opportunity to consider the suitability of this topic for the curriculum (see 4.2.6). Other curriculum areas that fall into this category include depletion of the ozone layer (see 2.2.5) and moment (see 3.2.2). The TIMSS data reveal that our students are capable of mastering those concepts.

The third important lesson learned regarding curriculum development is not so much about what topics to include or exclude, but rather what the curriculum emphasis should be. For example, as discussed in 2.2.3, our curriculum (and hence teaching) emphasizes structural details of the sense organs (i.e. eyes and ears) and pays insufficient attention to the advantages of having two eyes and two ears in relation to the survival of an organism as a whole. This has resulted in very distorted learning; students can remember the names of the detailed structures of the eyeball and their functions but are unable to explain the advantage of having two eyes. In other words, our curriculum has put too much emphasis on the learning of details and has overlooked the learning of principles. We might not have become aware of these shortcomings in our curriculum if we have not participated in the TIMSS studies and, more importantly, had not undertaken the present secondary analysis by drawing on teachers' expertise.

7.2 Lessons on classroom teaching and learning

There are at least two important lessons learned from TIMSS regarding teaching and learning of science. First, as pointed out in Chapter 5, knowledge about the patterns of gender differences in science achievements is important for the development of informed teaching strategies. That should be an issue that every teacher takes on board and deals with in their day-to-day teaching.

Second, there is ample evidence pointing to the necessity for HK science teachers to find ways to consolidate student learning. This includes not only avoiding fragmented teaching, but also helping students to cross-link topics that are learned in different parts of the curriculum. In other words, helping them to form a holistic picture of what they are learning. For example, after learning about algae in S1 in the topic Classification, which focuses on external features rather than mode of nutrition, they can be reintroduced when students learn about photosynthesis in S2. Actually, the more able students may be challenged to find out why different kinds of algae (namely, green, brown and red) are found at different depths of the sea, and how this is related to photosynthesis (see 2.2.2). However, teachers often miss this opportunity of bringing algae into the picture again to consolidate students' learning. Similarly, when teachers teach ideas of food chain and food web in S2, they often miss the chance of recalling, consolidating and extending what students have learned in S1 about the inter-dependence of life (see 2.2.4).

There is also ample evidence pointing to the need for more efforts to rectify students' habit of rote learning, as evident in the following two cases. First, as pointed out in 5.2.3, many students wrongly answered the question by simply recalling the standard kind of experiments they had carried out or been shown instead of applying the principle of controlling variables in experimental design, as required by the question. Second, as discussed in 5.2.7, many students just rote-learned that high temperature can speed up evaporation, and hence got trapped by a question that required them to draw conclusions based on the results presented. In short, much of this rote learning is related to how students are taught the process skills involved in scientific investigations. For example, students can recite the major steps of a scientific inquiry but without knowing exactly what each step entails (see 5.2.1 - distinguishing between observation, hypothesis and conclusion), or why certain procedures need to be taken (see 5.2.5 - reasons for repeated measurement). Little wonder that one quarter of HK students erroneously thought that 'listing all the results in a table' is the main reason why scientists repeat measurements during experiments.

While we can easily put all the blame onto the students, we should, as reflective practitioners, consider whether these results are related to our own teaching practices. In this connection, the following data from TIMSS 1999 may provide further food for thought. In that survey, teachers were asked how much weight do they put on different types of assessment in their science teaching. Table 7.1 presents the types of assessment to which teachers in several participating countries give ‘quite a lot’ or ‘a great deal’ of weight.

Table 7.1 Types of assessment to which teachers in TIMSS 1999 give quite a lot or a great deal of weight

Country	% Students by Type of Assessment (S.E.)						
	External Standardized Tests	Teacher-Made Tests Requiring Explanations	Teacher-Made Objective Tests	Homework Assignments	Projects or Practical exercises	Observations of Students	Students’ Responses in Class
Chinese Taipei	36 (4.1)	43 (4.5)	69 (4.1)	67 (3.6)	55 (4.1)	67 (3.8)	76 (3.4)
Singapore	28 (3.9)	70 (4.2)	67 (3.5)	39 (4.5)	61 (4.2)	40 (4.2)	36 (4.5)
Hungary	52 (2.7)	80 (1.9)	31 (2.0)	29 (2.0)	47 (2.2)	72 (2.3)	92 (1.3)
Japan	15 (2.6)	64 (4.3)	55 (4.3)	48 (4.3)	81 (3.6)	74 (3.9)	66 (3.5)
Korea, Rep. of	51 (4.1)	84 (2.8)	76 (3.6)	89 (2.5)	99 (0.6)	92 (2.2)	81 (3.1)
Australia	16 (2.6)	71 (3.6)	67 (3.4)	37 (3.8)	66 (2.9)	38 (3.8)	30 (3.3)
HK, SAR	17 (3.1)	58 (4.2)	76 (3.5)	33 (3.8)	23 (3.8)	23 (3.6)	30 (4.1)
United States	18 (2.5)	70 (2.8)	60 (3.2)	66 (2.8)	82 (2.7)	49 (3.6)	49 (2.6)
International Average	33 (0.5)	76 (0.5)	60 (0.6)	58 (0.6)	65 (0.6)	68 (0.5)	75 (0.5)

It is worrying to note that HK teachers did not put as much weight on those formative kinds of assessment as their overseas counterparts. For example, compared to the International Average of over 60% of students being assessed by their teachers using those formative kinds of assessment methods, HK teachers placed little weight on Projects or Practical Exercises (23%), Observation of Students (23%) and Students’ Response in Class (30%) as assessment tools. This is certainly a worrying sign in view of our earlier recommendation concerning the need for teachers to take account of students' misconceptions in their teaching. Failing to attend to student responses in class and not putting weight on observation of students are surely impediments to realizing that recommendation, and hence betterment of science education in HK. More professional development work in this area needs to be done. This is especially important in the context of the present education reform, which calls for 'assessment for learning'.

Nonetheless, we want to conclude this study on a positive note. The encouraging sign is that we had more than 250 enthusiastic teachers coming to our workshops, working hard for the whole day, analyzing the data, searching for reasons, reflecting on their own practices, sharing their expertise and teaching experiences. Most importantly, they all worked with a heart, contributed to the collective wisdom in the making of this book for the betterment of science education in HK. For this, we express our greatest appreciation to them once again. With so many hardworking, enthusiastic and dedicated teachers, we have confidence that our science education will continue to improve in the years to come.

Appendix

TIMSS Released Items

About one-third of the items used in each of the TIMSS studies are made available to the public after the test. The released items can be downloaded from the following websites:

http://timss.bc.edu/   (in English)