
Title	Primary science assessment and its contribution to the nurturing of creativity
Author(s)	Boo Hong Kwen
Source	<i>Australian Association for Research in Education Conference, Sydney, Australia, 1-5 December 2002</i>

This document may be used for private study or research purpose only. This document or any part of it may not be duplicated and/or distributed without permission of the copyright owner.

The Singapore Copyright Act applies to the use of this document.

Boo, H.K. (2002). Primary science assessment and its contribution to the nurturing of creativity. Paper presented at the *International Education Research Conference organised by the Australian Association for Research in Education* (1-5 December 2002), Brisbane, Australia ISSN 1324-9320.

Paper code: BOO02255

Primary science assessment and its contribution to the nurturing of creativity

Boo Hong Kwen

National Institute of Education, Nanyang Technological University, Singapore

Abstract

In 1997, a nation-wide initiative to promote creativity, lateral thinking and problem solving skills was launched in Singapore. Since then schools have been paying special attention to the development of creativity, lateral thinking and problem solving skills in our school going population. Teachers have been trained to teach such thinking skills explicitly as well as indirectly, through infusion into their science lessons. Teachers have risen to the challenge of nurturing creativity, lateral thinking and problem solving ability among their students. However, an important question needs to be raised: Does formal school science assessment contribute positively towards the drive towards the stimulating of creativity, lateral thinking and problem solving skills among pupils?

Introduction and background

In 1997, a nation-wide initiative to promote creativity, lateral thinking and problem solving skills was launched in Singapore. Since then schools have been paying special attention to the development of creativity, lateral thinking and problem solving skills in our school going population. Teachers have been trained to teach such thinking skills explicitly as well as indirectly, through infusion into their science lessons. Teachers have risen to the challenge of nurturing creativity, lateral thinking and problem solving ability among their students.

The word creativity has a wide scope. However, it generally means the production of something new (de Bono, 1986). There is some overlap between “creativity” and “lateral

thinking”. Lateral thinking, which is specifically concerned with the generation of new perceptions and new ideas, overlaps with creativity since both are concerned with producing something new, such as new ways of viewing a phenomenon or alternative ways of asking questions. If a person produces something new which is not good, then s/he is not considered creative. This means that the use of the term “creativity” is usually a value judgment of a result while lateral thinking is a process. It can be said that the ability to think laterally is involved in creativity.

Assessment and examinations are integral elements of the teaching and learning process at all levels of education across all environments and contexts. Whilst a wide range of assessment and examination methodologies have been developed, the most common remains the paper-and-pencil test which seeks to assess student understanding of scientific concepts and processes across topics through a series of tightly bounded questions covering the curriculum area.

Whilst the curriculum in any subject domain may evolve only slowly, there is a constant need to keep generating new assessment papers in keeping with the challenges of nurturing pupils who can think critically and innovatively while at the same time overcoming such problems as the limited number of concepts/topics in the primary science syllabuses. In addition, there is the need to present each cohort with a fresh examination paper so that model answers are not simply handed down from one cohort to the next, and rote-memorized and regurgitated.

All of this leads to a heavy demand on teachers to regularly generate new examination papers for use in the classroom. At the same time, in the Singapore context in particular, this has created a whole new market segment for the publishing industry. The heavy demand on teachers and the tendency to assume that material that is in print is accurate sometimes leads to the incorporation of inappropriate or weak test items from commercially available assessment books. In many cases, the same weak test items are found in several schools examination papers.

This paper is based upon an examination of about one hundred sets of school science examination papers from different primary schools. Each set of examination papers comprises P3-P6 papers, each comprising typically 30 multiple choice type questions and 16 open-ended supply type or free response questions.

Problematic test items which surfaced in this study

A scrutiny of these school examination papers reveal that while many of the assessment items were well-crafted and tested higher order thinking, there are, on virtually every paper, some test items and/or their mark schemes which would cause pupils, and in particular, creative pupils, to have difficulties or problems. These items are categorised and discussed in the following sections.

Group 1: Items which involve perceptual mismatch between question setter and creative pupils

One of the root problems is that creative pupils see issues, alternatives and ambiguity in examination questions which the setter did not intend or was unaware of. In other words there is a perceptual mismatch between the question setter and the creative pupils. In fact, in many of the other examination questions, there is some perceptual mismatch even between the setter and the non-creative pupils. What may appear to be a well-bounded and precise question on the part of the question setter can often be interpreted quite differently by the pupils.

For the creative pupils, there is an additional obstacle to performing well on such conventional paper-pencil test items since they are likely to have even more alternative perceptions or interpretations of a question. This is particularly severe in the case of multiple-choice questions (MCQs) where there is supposed to be one and only one correct answer out of four given options for each question and the pupils' answers are marked by computers. In some MCQs (such as the questions 1.1 and 1.2), creative pupils find situations in which all or none of the alternatives are correct while non-creative pupils who know the material in the conventional way simply select the most appropriate (often most obvious) answer and get marked correct.

Group 2: Test items which are weak in scientific content or processes

A pre-requisite to creative science (or technology) is sound knowledge of science content and processes such as scientific problem solving or experimenting. Another problem surfaced in the scrutiny of examination papers is that fact that some questions reflect weaknesses or misconceptions about either scientific content or problem solving procedures or processes. If pupils are subject to examination questions (as well as in homework and class tasks/assignments) which do not exemplify sound scientific concepts or problem solving procedures or processes, then they could be consciously or subconsciously be influenced and their grasp of scientific concepts and processes could be adversely affected.

Group 3: Test items which test trivial or esoteric facts

Test items which focus on trivial or esoteric facts instead of important science concepts or processes can impede the fostering of creativity in pupils. An examination which contains too many such test items are a concern because they could give the wrong signal to pupils, especially the conscientious ones, that learning science and doing well in science is about memorising a huge number of such facts. This would impede the development of creativity in them as their minds would not be free to roam and create, but instead would be distracted and cluttered with facts which should be more appropriately stored in places such as internet servers, books and CD ROMs.

Group 4: Test items with rigid and/or narrowly based mark schemes

In the case of supply type items it is common to find mark schemes which contain a narrow range of acceptable answers, and in some cases, even rigid, in terms of requiring a specific word or term to be given by pupils in order to earn marks or credit. It is also common to find mark schemes which explicitly state penalty for incorrect spelling or grammar.

In some supply type questions, creative pupils see alternatives which are often not seen by the setter, and if the mark schemes are followed rigidly, creative pupils end up giving

correct answers in terms of their own perceptions of the question but is/are the wrong answer(s) as per the question setter's mark schemes, which have too narrow a range of possible answers, and which fail to anticipate valid but innovative/unusual answers given by pupils who are more creative or critical in thinking. In such a case creative pupils can become confused or even de-motivated in science learning.

Specific examples for illustration and discussion

The following are specific examples put into the four groups which illustrate and discuss specific problems that could impede the drive to foster creativity in pupils.

Group 1 Examples: Items which involve perceptual mismatch between question setter and creative pupils

Question 1.1

Which one of the following animals should not be in the same group as the others?

- 1 goat
- 2 hen
- 3 rabbit
- 4 tiger

In this question, the setter's answer key is option 2. To the setter, who has taught pupils the classification of vertebrates into groups such as mammals, birds, fish, reptiles, amphibians and so forth, the given item is a very easy question, almost a give-away, since option 2 the hen is clearly the odd one out, the only bird among the mammals. However, because the basis of grouping is not made explicit in the question, a pupil who is thinking laterally could opt for any of the other alternatives as the answer:

- Option 1 could be the answer key since goat is the only animal with horns or in the question.
- Option 3 could be the answer key since rabbit is the only animal that burrows

- Option 4 could be the answer key since tiger is the only carnivore or only one not domesticated, and so forth.

Question 1.2

Which one of the following animals is unlike the others in terms of how it moves?

1. frog
2. kangaroo
3. rabbit
4. lion

In this question, the setter's answer key is option 4. To the setter, who has taught pupils the topic of "animals and their movement", the given item is a very easy one, almost a give-away, since the lion clearly does not hop while the others do.

However, a creative pupil could see alternatives not seen by the setter and the non-creative pupils, and could choose any of the other options as the answer.

- Option 1 could be the answer key since the frog is the only animal that could swim underwater.
- Option 2 could be the answer key since the kangaroo is the only animal that uses only two legs which moving at optimum speed.
- Option 3 could be the answer key since the rabbit is the only animal that can burrow or move underground.

Group 2 Examples: Items which are weak in scientific content or processes

Question 2.1: *A question that does not exemplify good experimental design/procedure*

Janet carried out an experiment in her living room with 4 stalks of flowers cut from the same plant. She recorded the results of her experiment in a table as shown below.

VASE	A	B	C	D
Number of stalks of flowers	1	1	1	1
Type of chemicals	P	Q	R	S
Amount of water (ml)	500	500	500	500
Number of days flowers stayed fresh	7	5	6	3

- (a) What was Janet trying to find out from her experiment? [1]
- (b) Besides using similar vases in size and quality, state 2 other variables which she must keep the same for the experiment. [2]

It should be noted that in experiments dealing with organisms (whether plant or animal) a sample of one (1) is not acceptable. Unlike physical science experiments where non-living things are involved, and where relationships between variables can often be described unambiguously (“proportional to”, “inversely proportional” or defined in terms of an equation) living things such as flower stalks incorporate many uncontrollable variables and their behavior cannot be reliably predicted. The exploration of possible relationships between causes and supposed effects is not as straightforward as physical science experiments. Correlational reasoning is the process used to assess the strength of relationships between variables, and a sample size for each condition investigated should be as large as is feasible. A sample size of one (1) for each condition of the experiment is unacceptable. The improved version of this test item suggested below has the sample size increased from one (1) to ten (10) and a very minor modification made to the last row of

data where instead of “Number of days flowers stayed fresh” was changed to “*Number of flowers stayed fresh after 10 days*”.

The original item is also lacking in the use of an experimental control. A control group receives the same treatment as the experimental group except the factor being tested. This again is corrected for in the improved version by removing chemical P in vase A and leaving the rest of the chemicals Q, R and S intact in their respective vases.

A suggested improved version

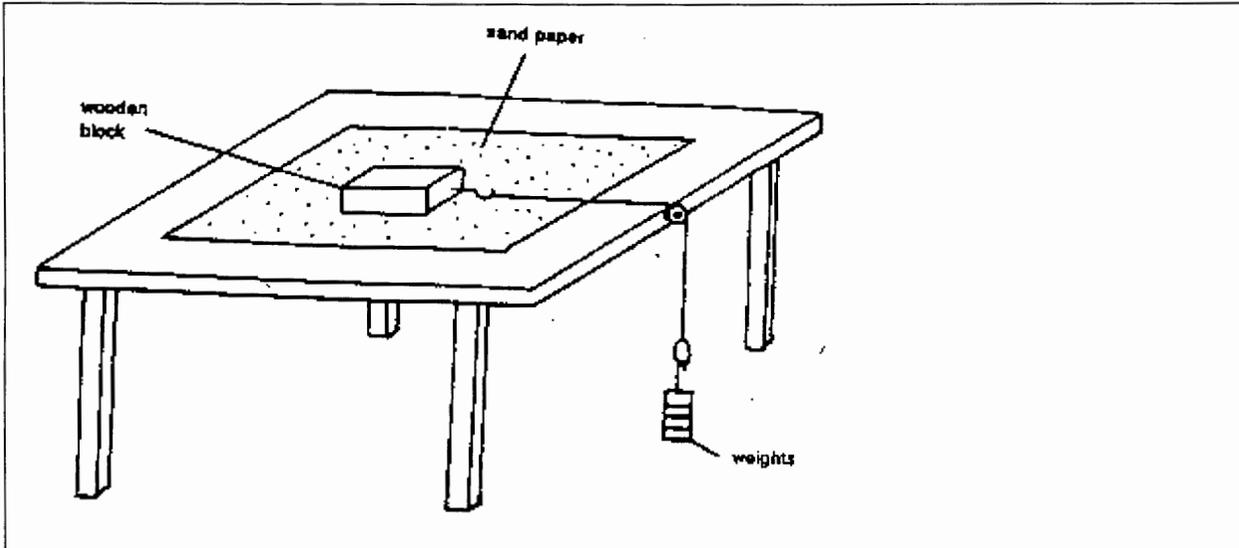
Janet carried out an experiment in her living room with 40 stalks of the same kind of cut flowers. She recorded the results of her experiment in a table as shown below.

VASE	A	B	C	D
Number of stalks of flowers	10	10	10	10
Type of chemicals	nil	Q	R	S
Amount of water (ml)	500	500	500	500
Number of flowers stayed fresh	7	5	6	3

After 10 days

- (a) What was Janet trying to find out from her experiment? [1]
- (b) Besides using similar vases in size and quality, state 2 other variables which she must keep the same for the experiment. [2]

Question 2.2: A question that does not exemplify good experimental procedure



In the experiment shown above, weights were added until the wooden block began to move over the sandpaper.

The experiment was repeated by replacing the sandpaper with the following surfaces: carpet, wood and glass.

The results are shown in the table below:

Type of surface	Sandpaper	Carpet	wood	glass
Force required	190g	180g	160g	135g

What was the purpose of this experiment? [2marks]

It was to find out _____

The most obvious flaw in this test item is that an incorrect unit for force has been used. Instead of using the unit “g” which represents “gramme” which is a unit for mass, a scalar quantity, the unit for the force required ought to have been stated as “gf” which

stands for “gramme force” a unit of force or “gwt” which stands for “gramme weight” yet another unit for force, which is a vector quantity. Repetition of the unit for force required could have been avoided by stating the unit in the first column, row 2 of the data table (see improved version below).

Unlike the previous example, this test item involves non-living things. While non-living things are more predictable, in any physical science experiment, there are various uncertainties or errors associated with each trial of the experiment. These uncertainties can be minimized by carrying out repeated trials for each type of surface. Thus the test items as it stands does not exemplify good scientific problem solving procedure. A more acceptable format of the table of data is suggested below.

Improved version of table of data

Type of surface		Sandpaper	Carpet	wood	glass
Force required (gf)	Trial I	189	184	161	136
	Trial II	185	183	157	137
	Trial III	187	185	163	133
Average force required (gf)		187	184	160	135

Group 3 Examples: Items which test trivial or esoteric facts

Question 3.1

Which of the following is the fastest swimmer in water?

- 1 guppy
- 2 sailfish
- 3 clownfish
- 4 shark

Question 3.2

Which of the following has leaves that fold up at night?

- 1 angsana
- 2 flame of the forest
- 3 orchid
- 4 rain tree

Questions 3.1 and 3.2 are examples of items which test trivial or esoteric facts, and not important learning outcomes in science such as concepts or processes. Items such as these may give the wrong signal to pupils, especially the more conscientious ones, that learning science and doing well in science involves memorising a large number of such facts. This is counterproductive to the drive to stimulate creativity in pupils.

Group 4 Examples: Items with narrow or rigid mark schemes

Question 4.1

Given a diagram showing a solar powered car and a hot-air balloon and the following table:

Mode of transport	Energy source
Hot-air balloon	Burns fuel
Solar-powered car	Sun's energy

- a) Which mode of transport is environmentally friendly? (1m)
- b) Explain your answer in (a). (1m)

Answer given on setter's mark scheme:

- a) solar-powered car
- b) Sunlight/solar battery/cell/solar energy does not cause pollution/conserves natural resources

Such a narrow mark scheme is disadvantageous to the creative pupil who can see have alternative perspectives. For example, a creative child could argue that the hot-air balloon is more environmental friendly because it does not require as much energy to manufacture as the solar-powered car and that it does not require the clearing of forests to make roads. Another additional point that could be argued in favour of the hot-air balloon is that it could use hydrogen fuel which would produce non-polluting substances (i.e. water).

Question 4.2

Given diagrams of a plant labeled X with large leaves and flowers and another plant of the same species labeled Y with smaller leaves and flowers.

- a) Which of the above plants was likely to be found in the deserted field? (1 mark)
- b) Explain your answer in a) (1 mark)

Answers given on setter's mark scheme

- a) Plant X
- b) Bigger leaves/flowers, must have received more sunlight

Such a narrow mark scheme is disadvantageous to the creative pupil who can see have alternative perspectives. For example, a creative child could argue that plant Y, and not plant X, is more likely to be found in the deserted field because "deserted" means "barren or having poor soil conditions" .

How assessment can contribute positively towards nurturing of creativity

The following are suggestions to address the problems highlighted in this paper, which could serve to align assessment more positively in terms of its contribution to the drive towards fostering creativity in pupils. For addressing the problem of perceptual mismatch between the question setter and the creative pupils, it would be worth considering the inclusion of an open-ended section following every multiple-choice question. This would ensure that the perspectives of creative children are taken into consideration during marking. It would mean that a longer time would be spent in the marking of the test items. However, this additional “cost” in time taken can be justified in terms of the gain that resulted from its positive contribution in taking into account the creativity of pupils. For addressing the rest of the problems (i.e. those associated with items weak in scientific content or processes, items testing trivial or esoteric facts and mark schemes which are too narrow or rigid) careful vetting and stringent quality control in the process of crafting and drafting of mark schemes should be exercised. Moreover during the marking process itself, markers should be open and be prepared to modify mark schemes as and when valid answers which are not anticipated in the mark schemes are given by pupils. Also, there should be no penalty for incorrect spelling if the words are recognizable and no penalty for grammatical errors if the answers make sense. The bottom line in the marking of supply type items should be on the correctness of the science concepts and the quality of reasoning or thinking that are revealed in the answers.

References

De Bono, E., (1986). Cort 4 Thinking: Creativity (Teachers notes)