Pupils' Test-Wiseness And Its Implications In Writing Multiple-Choice Items In Mathematics

by

Tan Hui Leng
Maktab Perguruan Ilmu Khas
Kuala Lumpur

ABSTRACT

Multiple-choice items (MCQs) are a common form of test items having been in use in the Malaysian examination system since the 1970s. This wide usage has led pupils to be very alert to particular features in such test items that seem to operate as cues to the correct answers. This paper examines this alertness or test-wiseness by subjecting 50 Form Four pupils to a sample of 40 MCQs. Analyses of pupils' responses to these MCQs indicate that they are particularly alert to features such as "seen number", heterogeneous options in form and length and order of options. Other features that operate as cues include distribution of answers and repeated (equivalent) options. These findings are used to critique some mathematics test items in an effort to lend direction and provide guidelines to teachers on what they should avoid doing in their effort to write good MCQs in mathematics.

1. Introduction

Since 1970, multiple-choice items or MCQs, often loosely referred to as objective items in the Malaysian context, have served in national examinations such as the *Ujian Penilaian Sekolah Rendah (UPSR), Penilaian Menengah Rendah (PMR),* earlier known as the *Sijil Rendah Pelajaran (SRP)* and *Sijil Tinggi Pelajaran Malaysia (STPM)).* With the introduction of MCQs to the *Sijil Pelajaran Malaysia (SPM)* mathematics examination this year, there is indication that this form of test items is here to stay.

In view of this, it should be recognized that teachers should be able to write good MCQs. This task may be particularly challenging to teachers as pupils have long been accustomed to this form of test and are likely to be alert to some features in MCQs that inadvertently operate as cues and enable pupils to guess at the correct answers. This alertness among pupils is referred to as test-wiseness in this paper.

Guessing has often been described as the bane to MCQs (Dunstan, 1971, Trigwell, 1992). Irrespective of how suitable an MCQ is in ascertaining or measuring a particular learning outcome, guesswork will definitely yield data or measurements that will mislead about pupils' success in learning. Thus it is imperative that in writing good MCQs, teachers are mindful of possible cues in their items so that an all-out effort is directed towards

eliminating guessing. It is to be noted that the quality of an objective test is determined by the skill of the constructors of the test (Trigwell, 1992).

2. Test-Wiseness: What and Why

In an effort to examine pupils' test-wiseness, 50 Form 4 pupils were subjected to a sample of 40 MCQs. Form 4 pupils are recognized as the most wise pupils in answering MCQs, having sat most recently for the *PMR* and the *UPSR* three years before that. The 40 MCQs are five sets of eight items with particular features that are likely to act as cues. These items were written in-gibberish to simulate a content-free situation which pupils definitely have no knowledge of. It is believed that under such a situation, pupils can only look for cues and guess at the correct answers. These guesses are analyzed to ascertain what features pupils perceive to be cues to the correct answers and their reasons are also sought as to why these features function as cues.

Table 1 summarizes pupils' responses to the eight types of test items with their particular features that are possible cues.

Table 1
Pupils' Responses to MCQs with Particular Features

Rank	Item Feature	Number of Correct Responses					Total Correct (%)	Possible Total
		Set 1 Items 1-8	Set 2 Items 9-16	Set 3 Items 17-24	Set 4 Items 25-32	Set 5 Items 33-40		
1.	Seen number in text	50	48	45	46	40	**229 (91.6)**	250
2.	Seen number in diagram/ illustration	50	47	42	43	40	**222 (88.8)**	250
3.	Heterogeneous in numerical form	46	45	44	42	40	**217 (86.8)**	250
4.	Heterogeneous in word form	44	44	43	42	41	**214 (85.6)**	250
5.	Heterogeneous in length/ complexity	42	40	41	40	42	**205 (82.0)**	250
6.	Order of placement of options	46	44	34	40	40	**204 (81.6)**	250
7.	Equivalent options	35	35	36	40	48	**194 (77.6)**	250
8.	Order of answers	35	30	30	40	40	**175 (70.0)**	250
Total Correct (%)		**348 (87.0)**	**333 (83.3)**	**315 (78.8)**	**333 (83.3)**	**331 (82.8)**	**1660 (83.0)**	**2000**
Possible Total		400	400	400	400	400		

From Table 1, it is evident that pupils are able to guess correctly the answers to MCQs that contain the particular features described in the table such as "seen numbers" and heterogeneous options. It is indeed rather surprising (and worrisome) that pupils obtained

a high mean score of 83.0% which translates to getting about 33 items correctly answered. It may be argued that the way the particular features are presented in the given 40 MCQs cues or prompts pupils more readily than when these features are buried in mathematical content. However, it is indeed the aim of this study to identify the worst scenario in guessing. The high mean score of 83.5 should be taken to indicate that pupils pick up cues very readily, particularly when they have no knowledge of the mathematics tested. This means that teachers have to be very mindful of these cues or revealing features in an MCQ. Otherwise, the MCQ will appear as a give-away to weak pupils in particular, and no longer function as a test item.

2.1 Seen numbers

From the rankings in Table 1, it appears that pupils (about 90% and above) are very alert to seen numbers in either the text (stem) or diagram/ illustration. They tend to pick these seen numbers as answers. Indeed in the first set of items where there is only one seen number, all the pupils chose that one given number as the answer. In the second and third sets of items, most of the pupils (80% and above) still chose a seen number as the answer although two or three numbers were seen. However in the fourth and fifth sets, slightly fewer pupils guessed correctly when the derivatives (product in Set 4 and quotient in Set 5) of the two seen numbers, rather than the seen number itself, are the answers. The responses also seem to suggest that a product of two seen numbers is slightly more readily recognizable than its quotient.

It is evident that pupils are very likely to pick seen numbers as answers. Even simple obvious derivatives of the seen numbers are easily identifiable as answers. Pupils reason that the seen number, particularly in the absence of any knowledge on the content tested, seems like the only sensible link to the question and hence is the answer. This kind of reasoning suggests that pupils who do not have any idea how to do a sum is likely to pick on any seen number as the possible answer if it appears among the options.

2.2 Heterogeneous options

In the face of heterogeneous options, pupils seem to be more alert to heterogeneity of form (more than 85%) rather than length (82%). Among the heterogeneous forms, pupils are just a little more likely to pick up the numerical (86.8%) rather than word (85.2%) heterogeneities. Pupils give the reason that the odd-looking number is a more likely answer as the other more homogeneous forms are too similar to have only one of them admissible as an answer. By such elimination, the odd-looking number is selected. Several pupils add that they tend to pick the “only whole number” or “the most complicated form of number” as the answer.

2.3 Order of placement of options

Pupils also seem to be alert to the order or placement of options. In the first and second sets, about 90% of the pupils picked B and C as answers to items where the words second and third were seen. In the fourth and fifth sets, 80% of the pupils guessed by selecting A

(smallest value among the options) and D (largest value among the options) for items that asked for minimum and maximum values respectively. Pupils seem to be least sensitive to the order word “middle” (Set 3) probably because B and C are both possibilities. Pupils reason that “it is logical” and “it is safe” to select the smallest or largest number listed among the options when the test item requires the minimum or maximum values. In the absence of a comprehensible context, they readily read the item as the second/ third/ minimum/ maximum value among the options instead. This may be taken to mean that pupils who do not know how to calculate the second/ third/ minimum/ maximum values in a given mathematical situation are likely to pick answers to such items based on the order of placement of the options.

2.4 Equivalent options

It will appear as if this form of cue should have been the most obvious to pupils as they should be able to eliminate overlapping and equivalent forms of options. Pupils seem unable to recognize equivalent forms too well, particularly word equivalents such as pyramid and cone. The success rate of 77.6% would have been lower if the high scores in Sets 4 and 5, where obvious numerical equivalents were given, did not boost the rate. It appears that pupils recognize order of numbers more readily than equivalent forms.

Its low rank among other cues suggests that pupils do not think too deeply or eliminate options when they guess. Pupils tend to look at the items superficially and focus on the numbers given, especially the obviously unusual looking ones, without mathematically processing them. This notion is supported by the slightly smaller number of correct guesses in Sets 4 and 5 of seen numbers where the answers, the product and quotient of the seen numbers, require some processing.

2.5 Order of answers

However, pupils are fairly able to recognize patterns in the distribution of answers, particularly the obvious ones such as A, B, C, D, A, B, C, D in set 4 and A, A, B, B, C, C, D, D in Set 5. The other patterns such A, B, C, D, D, C, B, A in Set 1 and A, D, B, C, A, D, B, C in Set 2 and A, C, A, C, B, D, B, D in Set 3 are slightly less obvious so they are less used to prompt correct guesses to the eighth item in each set. This is again suggestion that pupils process superficially when they guess.

Pupils relate that they do not consciously seek out patterns in answer distribution but “when patterns are so obvious as in the latter two sets, they will conform to the pattern when they have no other guess options”. Other pupils who did not guess correctly relate that they could not discern the patterns as they had made wrong guesses in the earlier items. With the latter explanation, it is understandable that this form of a cue is least discernible as it is dependent on pupils answering previous items correctly.

3. Implications

In the section, the implications of the findings will be explained by appraising some flawed MCQs in terms of its inherent cues. Suggestions are put forth as to how these cues may be removed.

Implication #1: Seen number

The following sum is flawed in that a seen number is the answer.

Find the average of the numbers 30, 48, 55 and 59.

A. 45.5 B. **48 (answer)** C. 51.5 D. 96

Instead the given numbers should be changed, a possibility being 30, 46, 57 and 59 for the answer to remain as 48 but is now unseen. The options remain functional.

Likewise, the next sum also contains a possible cue as the answer 5 is a seen number.

January	☎☎☎☎
February	☎☎
March	

☎ represents 5 telephone calls

The table above shows the number of telephone calls Seri makes in each month. If she has made 55 calls in the three months, how many ☎ must be drawn for the month of March?

A. 5 (answer) B. 6 C. 11 D. 49

Instead one less of ☎ should be drawn for the month of January so that the answer will be 6, an unseen number. Options A and C are likely to function still. Option D will function better replaced with 50.

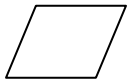
Implication #2: Heterogeneous options

The following sums are not good items because their answers are dissimilar in form to the other options.

2 1/3 + 1/6 ÷ 1/4 =

A. 2 9/24 B. 2 ¼ C. 2 ¾ **D. 3 (answer)**

The option D, being the only whole number, is likely to be chosen. To avoid this cue, replace A with 10, another whole number. The options should be rearranged to read as:
A. 2 ¼ B. 2 ¾ C. 3 D. 10



The figure shown above is a

A. rhombus (answer) B. pentagon C. hexagon D. octagon

To avoid the odd word as the answer, replace options with heterogeneous forms eg.
A. kite B. rhombus C. square D. rectangle

Implication #3: Order of placement of options

The following items may enable pupils to guess the answers because the order of placement of the answer coincides with the order required by the mathematical task.

Find the third term for the series; 5, 11, ….., 23, 29,…

A. 15 B. 16 **C. 17 (answer)** D. 18

Avoid C, the third option, as the answer. Instead, a possible set of options is:
A. 17 B. 18 C. 19 D. 20

Both x and y are integers such that $2 < x < 5$ and $20 < y < 30$. Find the maximum value of y/x.

A. 4 B. 6 C. 10 **D. 15 (answer)**

The answer D is indeed the largest value among the options. To avoid this, a possible set of options is:
A. 10 B. 15 C. 100 D. 150

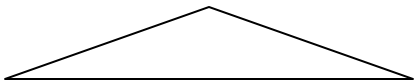
Implication # 4: Repeated or equivalent options

The two items contain equivalent forms as options. To a discerning pupil, these equivalent forms cannot be the answer.

1/5 + 3/10 =

A. 1/2 (answer) B. 4/5 C. 0.6 D. 0.8

Options B and D are equivalent, and pupils will eliminate them as possible answers. To avoid this, a possibility is to replace C with 0.4 and D becomes 0.6



What is the shape shown in the diagram?

A. Cone B. Pyramid **C. Triangle (answer)** D. Tetrahedron

Options A, B and D are equivalent since A and D are subsets of B. To avoid this, ask for the kind of triangle shown and the options may read as:

A. Scalene B. Isosceles (answer) C. Equilateral D. Right-Angled

4. Conclusion

It is evident that an MCQ has to be written carefully as features such as seen numbers, heterogeneous options and order of options operate as strong cues. It appears such cues favour pupils who have no knowledge or skills in mathematics. Thus it should be of particular concern that pupils who guess "without mathematical processing" can get at the correct answers to such items with a good measure of success. It is imperative that test writers are mindful to *avoid* these features. Otherwise, multiple-choice items, particularly the more challenging ones, will yield negative discriminating indices- a situation that should not arise in testing in our classrooms.

References:

Dunstan, M. (1971) *A Guide to the Planning Writing and Review of Multiple Choice Tests*, Tertiary Education Research Centre, University of New South Wales

Trigwell, K. (1992) *Information for UTS staff on Assessment*. Paper compiled for the Working Party on Assessment, Autumn Semester, January 1992, University of Technology Sydney.