Contemplation on marking scheme for Type X multiple choice questions, and an illustration of a practically applicable scheme

Nazeem Ishrat Siddiqui, Vinayak H. Bhavsar¹, Arnav V. Bhavsar², Sukhwant Bose³

ABSTRACT

Ever since its inception 100 years back, multiple choice items have been widely used as a method of assessment. It has certain inherent limitations such as inability to test higher cognitive skills, element of guesswork while answering, and issues related with marking schemes. Various marking schemes have been proposed in the past but they are not balanced, skewed, and complex, which are based on mathematical calculations which are typically not within the grasp of medical personnel. Type X questions has many advantages being easy to construct, can test multiple concepts/application/facets of a topic, cognitive skill of various level of hierarchy can be tested, and unlike Type K items, they are free from complicated coding. In spite of these advantages, they are not in common use due to complicated marking schemes. This is the reason we explored the aspects of methods of evaluation of multiple correct options multiple choice questions and came up with the simple, practically applicable, nonstringent but logical scoring system for the same. The rationale of the illustrated marking scheme is that it takes into consideration the distracter recognition ability of the examinee rather than relying on the ability only to select the correct response. Thus, examinee’s true knowledge is tested, and he is rewarded accordingly for selecting a correct answer and omitting a distracter. The scheme also penalizes for not recognizing a distracter thus controlling guessing behavior. It is emphasized that if the illustrated scoring scheme is adopted, then Type X questions would come in common practice.

KEY WORDS: Distracter recognition, guessing behavior, marking schemes, Type X multiple choice questions

Introduction

The tests using prototypes of multiple choice questions (MCQs) were first developed by Edward Thorndike, early in 20th century.[1] Frederick J. Kelly was the first to use MCQs for a large-scale assessment at Kansas State Normal School (now Emporia State University) in 1914.[2] The tests were used to assess the intelligence of World War I military recruits. MCQs serve a good objective method for evaluation. Nowadays, MCQs are commonly used to evaluate the performance of the students in various kinds of competitive examinations claiming to measure knowledge, thinking, and reading comprehension.[3] They are very commonly used as an assessment tool in formative and summative assessment in medical education.[3]

Evaluation system is the part of any educative process; it is commonly used to determine the final grade or marks.
and promotional decision. Besides this, there are other purposes served by the process of evaluation. They include communication about the important aspects of the course content, motivational aspects for the students, identification of the areas of deficiencies in curriculum, and remedial measures to be planned for the same.14–16

There are various evaluation methods such as written examination including long answer question, short answer question, essay type questions, modified essay type questions, and viva voce. Element of subjective bias is always an issue in these methods. To overcome this issue, objective methods in evaluation are emphasized. MCQ is one of the objective methods and has gained popularity over the period of time.3,7,8

There are various types of MCQs, labeled alphabetically (Type A, B, C, D, E, H, I, K, R, and X)3 with their chronological introduction over the period of time. The most commonly used MCQs are of Type A. Here, only one option serves as the key, viz., the most appropriate of all the options. The quality of the Type A MCQ depends on the proper framing of the item, and most importantly on the appropriate distracter options. In the absence of properly framed distracter options, these may look meaningless. Importantly, it is also felt that usually only the recall of the data is evaluated in this type, and a limited span of a particular topic is covered in this type of MCQ.3,5 This limitation can be overcome by the multiple (correct) responses MCQ (Type X).

As compared to Type A, instead of a single best option, there may be more correct options in an item. Importantly, this multiple correct option item does not follow the code of Type K.13 Type K MCQs, though reliable and able to test higher cognitive abilities as compared to other formats, their use is limited because of the complexities of the codes used for answering.14,16–18 However, in case of Type X the examinee is to select the correct options (which, in a 4-option item, may be one or many from 1, 2, 3, or 4). Thus, operationally, the examinee has to follow a simple instruction that there may be one or more correct options for the item.

At this juncture, it would be worth to consider the approach of examinee while solving the multiple choice item. While attempting the Type A MCQ, the examinee may be looking for the correct option by overviewing the options. He may be excluding “wrong” options first and then boiling down to less number of options to decide for the correct one. In case, he finds the correct answer at the first glance, he may not look properly for the remaining options.

When there are possible multiple correct options in an item, the approach of examinee bound to be different. Here, he may not look for “exclusion” of the option but he will look for the possibility of correctness of each option. Correctness of one option will not deter him from paying proper attention to the next option.

Thus, X-Type of item is neither cumbersome in framing nor is associated with any complex code. This is to further note that certain lacunae of single correct answer MCQ can be covered by X-Type of question. In X-Type question, multiple concepts/application/facets of a topic can be tested, and cognitive skill of various level of hierarchy can be tested. Even the students’ strategies are different while solving this type of questions.10–13

Despite all these good features, X-Type questions are not in common use. One wonders why so? We believe that the scoring system is an important reason for its uncommon use. Lack of appropriate practicable method for evaluation is the responsible factor for nonuse of X-Type of questions. This is the reason we explored the aspects of methods of evaluation of multiple correct options MCQs and came up with the logical and practicable method for the same. In subsequent section, the scoring schemes in this regards are reviewed.

Note on Existing Marking Schemes for Multiple Correct Options Multiple Choice Questions (Type X)

For any MCQ, the marking is based on which of the options is selected (e.g., by a “tick mark”). Marking in single correct option MCQ (where only one option is correct, e.g., Type A), is rather straightforward. A full credit (1 mark) is awarded for tick marking the correct option.14–16

A tick mark on a distracter leads to loss of mark. In addition, tick marks on more than one option also leads to loss of marks, irrespective of whether one of them is correct or not.17 A tick mark on a distracter indicates guesswork for the right answer.15 In some cases, to reduce the temptation of using guesswork, provision of deducting a fraction of mark is made. This is labeled as “negative marking (NM)” better called “correction for guesswork.”16,18 Typically, when NM is used, a tick mark on a distracter results into −1/4 mark (when 1 mark is allotted to the item, and there are four options in the item).15,16,20,21

The scenario is not as straightforward when one considers multiple correct options MCQs (those where more than one option are correct). This is due to the fact that there can be multiple ways that one can consider crediting the selection or rejection of the options in the question. One straightforward way is that if all correct options are tick marked then full score is to be awarded for the item. However, even one of the correct options is not tick marked, no mark is awarded. This kind of the scoring system appears very stringent, and perhaps, it is an important reason which discourages the use of multiple correct options MCQs (X-Type items). However, we note that the above scheme is not the only scheme for marking multiple correct options MCQs (Type X items), and literature is available on the various marking system for X-Type items.15,22–24 However, the marking systems in literature22,23 are described either briefly and/or in mathematical terms, which are difficult to interpret by the medical personnel who do not have good mathematical/statistical background.

In this section, we discuss in an elaborate manner, for the benefit of the medical readers, the marking approaches which are reviewed in the above-referred studies.22–24

An important point, which is assumed, but not explicitly mentioned in,23 is that in the first five methods below, if the examinee does not select a particular option, it is considered as rejecting that option. Thus, if an option is not explicitly “tick marked,” it is considered that that option is rejected. The significance of such a consideration is apparent in the last approach, where the examinee also has a choice of omitting an option (neither selecting nor rejecting). The implication of this assumption can also be further appreciated in our discussion in the next section.

Keeping the above aspects in mind, hereby we will take a brief overview of the existing marking schemes for multiple

Indian Journal of Pharmacology | April 2016 | Vol 48 | Issue 2 | 115
correct options (multiple responses [MR]) MCQs. In the following
discussion, “k” denotes the total number of options, “i” denotes
the number of correctly chosen options, and “f” denotes
the final score. The titles of the methods are same as those in.\textsuperscript{[12,13]}
Moreover, in the discussion below, we follow the convention
that maximum marks for a question is one (01).

**Multiple-response Method or Number Right Method**

Scoring of MCQs have been done using these conventional
methods in which correct responses will be awarded with a
positive value whereas incorrect and omitted responses were
awarded zero (00) value.\textsuperscript{[14,15]} These schemes are straight
forward extension to scoring in single response MCQs.
Mathematically, it is stated as follows:

\[
f = \begin{cases} 
1 & \text{if } i = k \\
0 & \text{otherwise}
\end{cases}
\]

This simply means that complete marks are awarded only
if all the correct options are marked correctly, otherwise no
marks are awarded.

**Count-3 Method (for } k = 4\)**

This scoring approach is used when the question has four
options. It is slightly relaxed than the MR method. It is stated
mathematically as follows:

\[
f = \begin{cases} 
0 & \text{if } j = 0 \text{ or } 1 \\
\frac{i-2}{k-2} & \text{otherwise}
\end{cases}
\]

The above equation implies that if the examinee selects
<3 correct options, no marks are awarded. However, if the
examinee selects 3 out of 4 correct options, then a partial
credit is given (which is 1/2 marks). If all the selections are
correct, then a full credit of 1 mark is awarded. Note that the
correct selection means (a) tick marking a key/s and (b) not
tick marking a distractor/s. Furthermore, note that while the
above equation of this scoring scheme is valid for the case of 4-
option questions, it can be easily modified to accommodate
questions with a different number of options.

**Count-2 Method (for } k = 4\)**

This scoring approach is similar to the count-3 method
and is also used when the question has four options. It is
somewhat more relaxed than the above approach. It is stated
mathematically as follows:

\[
f = \begin{cases} 
0 & \text{if } j = 0 \\
\frac{i-1}{k-1} & \text{otherwise}
\end{cases}
\]

Similar to the above, in this case also, the examinee gets
partial credit, but the relaxation allows the partial credit to be
awarded starting from two correct answers. More specifically,
if the examinee selects <2 correct options, no marks are
awarded. However, for 2 correct options out of 4, a partial credit
of 1/3 marks is given. For 3 correct options, a partial credit of
2/3 is given and for all correct selections, a full credit of 1 mark
is awarded. As in the above case, the equation of this scoring
scheme is valid for the case of 4-option questions, but it can
be easily modified to accommodate questions with a different
number of options.

**Multiple True-false Method**

This method is a further relaxation of the above methods,
in which each correct selection is given credit. In mathematical
terms, the scoring is stated as follows:

\[
f = \frac{i}{k} \text{ for } (k \geq i \geq 0)
\]

This equation implies that every correct selection is
awarded partial marks. For the case of 4-option questions, this
means that 1, 2, 3, and 4 correct selections will be awarded,
respectively, 1/4, 2/4, 3/4, and 1 mark. Note that the equation
is general enough for questions with any number of options.
The interval of marks between the consecutive partial credits
will differ accordingly, i.e., the range from 0 to 1 mark will be
divided equally for deciding the partial credit depending on the
number of correct responses.

In these marking algorithms, student gain marks by their
guessing behavior, and thus, test reliability and validity go down,
which is a matter of concern.\textsuperscript{[14,15,24-27]}

**Correction for Guessing or Negative Marking**

To control guessing behavior of examinee, various formulas
have been given for correction of guessing.\textsuperscript{[15]} In such marking
schemes, a penalty is imposed and examinee loses marks for
marking an incorrect response (NM). Thus, test scores are the
reflection of true knowledge of the examinee since the penalty
for incorrect selections reduces guessing behavior.\textsuperscript{[15,17]} The
mathematical statement for this approach is as follows:

\[
f = \frac{i - (k - i)}{k}
\]

Elaborating the equation, we find that as in the above case of
multiple true-false (MTF), a marking scale is divided at regular
intervals for assigning credit. However, unlike that in the MTF
scheme (where the scale was 0–1), in the correction for guessing
(CFG) scheme the scale is from −1 to 1. Thus, for a special
case of 4-option questions, the credit will be awarded as follows:

For 0 correct selection (which means four incorrect selections),
−1 mark will be given. For 1 correct selection (hence, three
incorrect selection), −1/2 mark will be given. For 2 correct (and
2 incorrect) selections, 0 will be given for 3 correct (1 incorrect)
selection, 1/2 will be awarded, and for 4 correct (and 0 incorrect)
selection, a complete 1 mark will be awarded. This system does
not consider the possibility of examinee to answer on the basis
of partial information or misinformation.\textsuperscript{[16]}

**Let Omit Method**

Finally, the let omit (LO) method also considers negative
marks, but differs from CFG approach, in that it also allows
the candidate to omit options. Denoting “i” as the number of
options, the candidate has omitted, this scoring scheme can
be stated as:

\[
f = \frac{i - j}{k} \text{ for } (k \geq i, j \geq 0) \text{ and } j \leq (k - i)
\]

One can note that there can be many combinations with
various values of i and j, for listing the various scores in this
scheme. For brevity, we discuss the distinction of this scheme
with respect to the CFG scheme, considering two cases for i.

Considering that the candidate has selected 2 correct
options, this yields 0 marks in the CFG scheme (2 correct and
2 incorrect). However, in the LO case, the possibilities can be
(a) 2 correct, 2 incorrect, (b) 2 correct, 1 incorrect, 1 omit, and
(c) 2 correct, 2 omit. These possibilities will, respectively, be
credited as 0 (same as CFG), 1/4 and 1/2.

Similarly, considering the case of 3 correct options, yields
possibilities in the LO method as a) 3 correct, 1 incorrect
Siddiqui, et al.: Practically applicable scheme for marking X-Type MCQs

b) 3 correct, 1 omit. These will be, respectively, credited as 1/2 (same as CFG), and 3/4.

An interesting aspect to note here is that both cases of “2 correct, 2 omit,” and “3 correct, 1 incorrect,” get the same credit of 1/2. Thus, even if the number of correct selections is less, a positive credit is given for not selecting an incorrect choice. Hence, rather penalizing for an incorrect response, the student is awarded for not guessing, and he is given credit for not selecting an incorrect response.[28] This is a psychological advantage to the examinee since it promotes a desired behavior rather penalizing an undesired behavior; i.e., guesswork. Examinee does not feel threatened for getting a reward for omission than a penalty for wrong selection.[29]

Similarly, there can be other combinations where some incorrect cases may get equal credit as some omissions. This scenario is one example which indicates that there can be multiple ways to consider how to award credit. This partial credit system is preferable to dichotomous system, especially in Type-X items.[13,28]

Similar to the last approach mentioned above, the work in[28] reviews and reports some methods which allow for the omission of certain options in the items. This is different from the first five cases discussed above where there is no notion of omission. These methods are described below:

Morgan Method
- This was the first marking algorithm which is a relatively simple partial crediting approach which considers penalizing for wrongly marked options.[29] More specifically, the examinee achieves
  - 1/n marks for an option marked which is actually correct
  - −1/m marks for an option marked which is actually incorrect
  - 0 marks if the option is not marked.
- Here, n is the number of actually correct options in the item, and m is the number of distracters. This approach, while simple, seems somewhat skewed as the penalty and the positive markings are not balanced.

Ripkey Method
- This approach does not consider any negative penalty but awards 0 marks in case of over guessing. Denoting n as the number of actually correct options in the item, and c denotes the number of chosen/marked options; the examinee achieves,[13]
  - 1/n marks for an option marked which is actually correct
  - 0 marks for an option which is not marked or incorrectly marked
  - 0 marks for the complete item, if c > n.
- Thus, if the total number of options chosen by the examinee is less than or equal to the total number of correct options in the item, he is awarded partial credit depending on how many correct options he has chosen. However, if he chooses more than the number of actually correct options in the item (regardless of whether they are correct or incorrect), he is awarded no marks. Clearly, although there is no negative penalty, this approach strongly discourages over-guessing. However, the tradeoff is that if the number of actually correct options in the item is very less (e.g., 1 or 2 in a 4-option item), then this approach becomes quite stringent and similar to the MR method, which gives either complete marks or 0 marks for the item.

Balanced Scoring Method
- This scheme is a modified version of the Ripkey method, where instead of the 0 marks for the case of c > n, the method follows partial negative penalty for over guessing.[27] More specifically, denoting k as the number of actually incorrect options in the item, the examinee is awarded:
  - 1/n marks for an option marked which is actually correct
  - 0 marks for an option which is not marked or incorrectly marked
  - −(c− n)/k marks as a penalty for the item in case of c > n.

Looking closely, one can notice that the method is same as the Ripkey method for the positive scoring part.[13,27] However, while the Ripkey’s method awards 0 marks for over-guessing; this approach awards some partial penalty, and thus is less stringent. The amount of penalty is dependent on the number of extra options chosen over the number of actually correct samples. Note that in the case of c ≤ n, the penalty is not applied. However, this approach is somewhat complex to implement as it requires keeping track of how many options the examinee has over guessed. Moreover, similar the Morgan and the Ripkey’s method, the positive and NMs are not equal.

The aim of formula scoring methods is to determine examinee’s degree of knowledge with respect to each option presented in the item; however, none of them was found to be superior to the conventional marking methods in terms of reliability and validity of the test.[21,31] Studies have shown that NM fails in discouraging the guessing behavior of students.[17,22] They further added variance in the test scores, which is related to the risk-taking behavior of the student.[24,32] High-risk-takers score more marks in this system as compared to students of equal ability but less risk taking behavior. Thus, NM schemes do not reflect true knowledge; rather they reflect risk-taking ability and answering strategies of the examinee.[6,15,32] However, they can discriminate between examinee’s full knowledge, the absence of knowledge, and misinformation.[14,21,32,34]

Based on the above discussion, one can observe and argue that the existence of multiple marking schemes (of which the above nine are indicative but clearly not exhaustive examples), highlight that there can be various different perspectives on the awarding of credit in MR MCQs. Choosing a scheme may depend on various factors such as the number of options in an item, the number of items in the examination, difficulty level an examiner wants to set, and even, perhaps, consideration of the previous performance of the examinees. The important underlying aspect to note here is that the different marking schemes had the low and contradictory power of discrimination of examinee’s knowledge.[14,21,31,33,35,36]

A Practically Applicable Scheme for Scoring of X-Type Items: Illustration and Discussion

We provide below an elaborate illustrative examples and discussion on the scheme which we believe is more practicable.

As seen above, the various schemes for marking are available, but some element of distortion either toward recognizing true knowledge of the student or toward penalizing for guessing is present in them. The traditional binary marking system had its own inherent weaknesses of increasing test error variance due to guessing work. Alternative methods
try to overcome such shortcomings by awarding credit to the examinee’s partial knowledge but they failed on practical grounds, making the test time-consuming, and difficult for the examinee, and they are shown to have low discriminative power thus decreasing the test validity.\[37\] Further, these methods have not been considered as practical and balanced since none of them took into consideration the significance of the examinee’s ability to recognize distracter.\[35\]

Keeping all these factors into consideration, we have tried to formulate a practical and nonstringent but rational scoring scheme. This marking scheme tries to overcome shortcomings of existing marking schemes by laying more emphasis on distracter recognition ability of the examinee rather based only on their ability to select the correct response. An elaborate discussion of this marking scheme with illustrative examples is given below. The scheme is based on MTF method; however, the approach takes into account the shortcomings of MTF method. The illustration is specifically meant for the medical teachers who are not very good at mathematical expressions.

Let us take an example of X-Type items. Each item has four options; out of which 1 or more than 1 may be correct. For some items, number of correct option will vary from 1 to 4. The examinees are given the instruction that there will be 1 or more than 1 correct option/s for each item. They are supposed to tick mark the correct option/s.

The marking system is illustrated as under:

Suppose for an item options a, c and d are keys and b is distracter. An examinee (Student-A) has tick marked options a, b, and c as correct options. In this case, as the examinee has tick marked options a and c which are keys, he should get the credit for that. He did not mark option d which is also a key, so he misses credit for that. Option b being distracter, a tick mark on this option will result into discredit (missing the credit). Thus, the examinee gains credit for tick marking two keys and misses credit for wrongly tick marking a distracter.

As there are four options and item carries 1 mark, each credit will be equivalent to \(\frac{1}{4}\) marks. Thus, in the given example [Table 1], the examinee scores \(\frac{1}{2}\) marks.

A simple logic is followed in the scoring system.

• If the key is tick marked, a credit is given, and if key is not tick marked, credit is missed (as instead of one key, multiple keys are there in an item, the same logic is applicable)
• If distracter is tick marked, credit is missed
• Importantly, if the distracter is not tick marked a credit is to be counted. This is an important aspect of marking system, enabling examinee to get reward for discerning some options as distracters.\[33\] Thus, a balance between reward and penalty is maintained, and partial knowledge of the examinee is rewarded.

In another instance, we consider that for the same item (key: a, c, and d; distracter: b) another examinee (Student-B) has tick marked following options: c and d.

This examinee should receive two credits for tick marking two keys; misses a credit for not marking a key (option a). For not tick-marking the distracter (option b), he gains credit. Thus, score would be 3/4 marks [Table 2].

Another item has option c as key, and a, b, and d are distracters. An examinee (Student-C) has a tick-marked key (option c) only. He gets credit for that. For not marking the distracters, he gets three credits. Thus gets total credits of 4, meaning by that he gets full one mark for the item [Table 3].

Here are more examples, where responses of the various students are included, and credit for each option and total mark for the item is shown [Tables 4-6].

It is speculated that while solving this kind of MR item, the examinee is reading each option carefully and deciding if the option is correct/true or otherwise. The examinee’s ability to select the correct answer and recognize a detractor indicates his knowledge. If this fact is not taken into account, then it can jeopardize the discriminating power of the test and subsequently would affect reliability and validity.\[32,33\]

### Table 1:

| Options of item | Examinee’s response | Credit |
|-----------------|---------------------|--------|
| a. Key          | ☑                   | 1/4    |
| b. Distracter   | ☐                   | Missed |
| c. Key          | ☑                   | 1/4    |
| d. Key          | ☐                   | Missed |
| Total marks obtained |                  | 1/2    |

### Table 2:

| Options | Examinee’s response | Credit |
|---------|---------------------|--------|
| a. Key  | Missed              |        |
| b. Distracter | ☐               | 1/4    |
| c. Key  | ☑                   | 1/4    |
| d. Key  | ☑                   | 1/4    |
| Total   |                     | 3/4    |

### Table 3:

| Options | Examinee’s response | Credit |
|---------|---------------------|--------|
| a. Distracter |                 | 1/4    |
| b. Distracter |                 | 1/4    |
| c. Key       | ☑                   | 1/4    |
| d. Distracter | ☑                   | 1/4    |
| Total        |                     | 1      |

### Table 4:

| Options | Examinee’s response | Credit |
|---------|---------------------|--------|
| a. Distracter |                 | 1/4    |
| b. Distracter |                 | 1/4    |
| c. Key       | ☑                   | 1/4    |
| d. Distracter | ☑                   | Missed |
| Total        |                     | 3/4    |
Table 5:
Example: Responses and scores in a type X item (Student-E)

| Options | Examinee’s response | Credit |
|---------|---------------------|--------|
| a. Distractor | 1/4 | Missed |
| b. Distractor | 1/4 | Missed |
| c. Key | | Missed |
| d. Distractor | X | Missed |
| Total | | 1/2 |

Table 6:
Example: Responses and scores in a type X item (Student-F)

| Options | Examinee’s response | Credit |
|---------|---------------------|--------|
| a. Distractor | X | Missed |
| b. Distractor | X | Missed |
| c. Key | | Missed |
| d. Distractor | X | Missed |
| Total | | 0 |

Thus, of course, a separate decision is needed for each option. The decision related to the earlier option does not influence the decision for the next option. Thus, decision-making procedure regarding correctness of the each option appears to be different than what it is in the case of single correct option item.

What about making the guess about correctness about the option? That aspect may appear as unavoidable, but guessing again is not influenced by correctness or otherwise of other option. Can correction about guessing is possible? It seems so. This issue is involved when a distractor is tick marked. On the logical basis, the tick mark on the distractor causes loss of a possible credit. If additional emphasis is desired to curb the guessing by tick marking a distractor; an extra discredit is to be considered. Thus, a tick mark on the distractor should result into loss of an extra discredit (instead of 1, one should count 2 discredits) for final calculation of the score for an item.

There can be tricky situations with the present scheme. The examinee may tick mark all the options (as guess work) and will get the benefit. He will get the score for all the keys in the item though he will lose the score for marking the distractors. He will be beneficiary if there are 3 or 4 keys in the item. Thus, the examinee will get unnecessary marks without any skill on his part. Similarly, when he does not tick mark any option, he will get the score for not tick marking the distractors; a fallacy of getting the marks without any knowledge.

To avoid such problems, the item should be considered as “not attempted” if (a) no option is tick marked or (b) all the options are tick marked. No marks are given in this situation. Thus, the disadvantage of MTF described by Tarasowa and Aver[23] is taken care of. However, the examinee should be given relevant instructions in advance.

Thus, the prerequisite for this type of MCQ is that in each item, there must be at least one key or one distracter. Number of options may be four (4) or more.

Below, we summarize the steps for calculation of the marks in the proposed scheme. Such a step-wise formulation makes the evaluation easier to understand.

**Calculation of overall marks**

Operationally, one has to follow the following steps to get final scores:

- If all or no options are ticked, then no marks are given. If this is not the case, then the following steps are to be followed for evaluation
- Count down the options as key/s and distracter/s; each may be one or more than as per the construct of the item
- Count the key/s which are tick mark and note the credits for the same
- Count the distracter/s not tick marked and count the credits for the same; observe carefully if distracter/s is/are tick marked and accordingly discredit is to be counted
- Count the total credit for the item; divide the same by 4 to get the score for the item.

**Comparison Among Various Marking Schemes**

When compared to other scoring systems,[30] the illustrated system is not rigorous (scoring is higher than the binary/dichotomous system), there is no unnecessary penalty for over guessing or mistakes in tick marking, the system is rational and can be clearly understood by the examinee. Herewith, in case of each of the examples that are given above, marks are calculated based on the various methods described earlier. A comparison is presented in Table 7.

Comparison of marks by different methods shows that marks obtained by MTF method are well matching with the suggested scheme. Other methods appear stringent and cumbersome to calculate the marks. Even with the MTF method, calculation is bit cumbersome using the mathematical formula. Moreover, we have already highlighted the situation of getting disproportionate marks in the MTF method when the examinee put correct marks at all options or leave every option unmarked. That aspect is taken care of in the suggested method as pointed out earlier. Moreover, the steps mentioned above for calculation of the overall marks, provide a simpler understanding of the proposed scheme and can help in its easy implementation.

**Additional Practical Considerations**

There are relevant practical questions about the use of X-Type questions. Will the X-Type questions take care of the credit for partial knowledge? Will it be easy for the teacher to construct the X-Type of questions and evaluate them using the suggested method? Will this type of questions will be understood and accepted by the students?

An example is presented herewith, to explain the construction of the X-Type question that will give ease to the teachers for formulating the questions.

Suppose one decides to form an X-Type question for students of pharmacology (second M.B.B.S.) on the following aspects of malaria: Relapse, recrudescence, species of Plasmodium having the possibility of relapse, antimalarial drugs likely to result into recrudescence, prevention of the same by a combination of drugs, drug therapy that will prevent relapse.

Four statements are constructed where the above aspects are included. For example:
Table 7:

Comparison of marks by various methods

| Method for scoring                        | Number 1 | Number 2 | Number 3 | Number 4 | Number 5 | Number 6 |
|-----------------------------------------|----------|----------|----------|----------|----------|----------|
| Suggested method                        | 1/2      | 3/4      | 1        | 3/4      | 1/2      | 0        |
| Multiple-response method or Number right method | 0        | 0        | 1        | 0        | 0        | 0        |
| Count-3 method (for k=4)                 | 0        | 1/2      | 1        | 1/2      | 0        | 0        |
| Count-2 method (for k=4)                 | 1/3      | 2/3      | 1        | 2/3      | 1/3      | 0        |
| Multiple true-false method               | 1/2      | 3/4      | 1        | 3/4      | 1/2      | 0        |
| Correction for guessing or negative marking | 0        | 1/2      | 1        | 1/2      | 0        | −1       |
| Morgan method                           | −1/3     | 2/3      | 1        | 2/3      | −1/3     | −1       |
| Ripkey method                           | 2/3      | 2/3      | 1        | 0        | 0        | 0        |
| Balanced scoring method                  | 2/3      | 2/3      | 1        | 3/4      | 0        | −1/2     |

• Use of primaquine will prevent relapse of *Plasmodium falciparum* (False)
• Artesunate is an efficacious drug for *P. falciparum*, but recrudescence is the problem (True)
• Chloroquine administered for 3 days will treat acute clinical manifestation as well will prevent the relapse of *Plasmodium vivax* malaria (False)
• Co-administration of lumefantrine and artesunate is recommended to avoid recrudescence in *P. falciparum* infection (True).

Some statement/s is/are true (it/they will serve as key).
Some statement/s is/are false (it/they will serve as distracter).
It is obvious that multiple concept/application/facets of malaria and antimalarial drugs are included in one question in the above example.

Thus, if all true and false statements are identified; a full credit will be due. In case, all true or false statements are not identified, some marks are given for identified true or false statements; thus, student gets the credit for partial knowledge.

The various examples for the scoring based on the suggested method, if studied properly, would lead one to get a full grasp on the scoring system.

By practicing on this line, the teacher would develop the insight of constructing a good variety of the X-Type questions and scoring them with ease. However, training of the teacher in respect to X-Type questions construction and evaluation can be included in the one of the sessions on MCQ in Basic Course Workshops conducted by Medical Education Unit of the Medical Colleges and Regional Training Centers, if the need is felt.

It is our experience that there is no problem of accepting the X-Type questions by the students; of course, the students are given instructions about multiple correct options and method of evaluation prior to the test.

We have experience of using the different sets of multiple correct option items adopting the present marking scheme with various batches of the examinees; we noted that the large number of examinees obtained the score of 40–70%. Number of examinees who scored more than 80% is limited and so also is the case regarding low range of marks (below 30%).

**Conclusion**

The system described above for scoring for the X-Type item is rational, simple, and behaves better than the existing marking algorithms. If it is adopted, use of this type of MCQs will be common.

It is felt that the problems of creating proper distracters, which is usually seen in the case of single response items, can be easily dealt with in MR items.[39]

In the X-Type items, wide areas and wider aspect of the topic under question can be covered. Possibilities of the flaws while writing the MCQs are depicted in the literature,[39,40] and due care can be taken to avoid the same. Skillful use of options would increase the utility of this type of item to evaluate the higher level of cognitive skill.

It is pointed out in the literature that MTF type of questions are not suitable as options may be lengthy and may not be suitable in a time frame.[35,37] To avoid this difficulty, the allotted time can be increased and that may not be a great deal if the quality of the items improves.

It is felt that this kind of items will be more suitable for formative evaluation where coverage of various facets of the topics can be well appraised; the students will have a chance to look at the level of their learning process.

Further to state that the acceptance of these items by the examinees will have no problem, and they will develop the habit of looking to each option carefully.

**Financial Support and Sponsorship**

Nil.

**Conflicts of Interest**

There are no conflicts of interest.

**References**

1. Goodenough FL. Edward Lee Thorndike, 1874-1949. Am J Psychol 1950;63:291-301.
2. Lemann N. The Big Test: The Secret History of the American Meritocracy. New York: Farrar, Straus and Giroux; 1999.
3. Albanese MA, Sabers DL. Multiple true-false items: A study of inter item correlations, scoring alternatives, and reliability estimation. J Educ Meas 1988;25:111-23.
4. Newble DI, Jaeger K. The effect of assessments and examinations on the learning of medical students. Med Educ 1983;17:165-71.
5. Roife J, McPherson J. Formative assessment: How am I doing? Lancet 1995;345:837-9.
6. Moss E. Multiple choice questions: Their value as an assessment tool. Curr Opin Anaesthesiol 2001;14:661-6.
7. Case SM, Swanson DB. Constructing Written Test Questions for the Basic and Clinical Sciences. 3rd ed. Philadelphia: National Board of Medical Examiners; 2001.
8. DiBattista D, Kurzawa L. Examination of the quality of multiple-choice items on classroom tests. Can J Scholarsh Teach Learn 2011;2:1-23.
9. Fenderson BA, Damjanov I, Robeson MR, Veloski JJ, Rubin E. The virtues of extended matching and uncued tests as alternatives to multiple choice questions. Hum Pathol 1997;28:526-32.
10. Cronbach LJ. An experimental comparison of the multiple true-false and multiple choice multiple-choice tests. J Educ Psychol 1941;32:533-43.
11. Hohensinn C, Kubinger KD. Applying item response theory methods to examine the impact of different response formats. Educ Psychol Meas 2011;71:732-46.
12. Pomplun M, Omar MH. Multiple-mark items: An alternative objective item format? Educ Psychol Meas 1997;57:949-62.
13. Ripkey DR, Case SM, Swanson DB. A "new" item format for assessing aspects of clinical competence. Acad Med 1996;71 10 Suppl: S34-6.
14. Bereby-Meyer Y, Meyer Y, Flascher OM. Prospect theory analysis of guessing in multiple choice tests. J Behav Decis Mak 2002;15:313-27.
15. Kurz TB. A Review of Scoring Algorithms for Multiple-choice Tests. Paper Presented at the Annual Meeting of the Southwest Education Research Association. San Antonio, TX; January, 1999. Available from: http://files.eric.ed.gov/fulltext/ED428076.pdf. [Last accessed on 2016 Feb 26].
16. Jennings S, Bush M. A comparison of conventional and liberal (Free-Choice) multiple choice tests. Pract Assess Res Eval 2006;11:(8). Available from: http://papersonline.net/getvn.asp?v = 11 and n = 8. [Last accessed on 2015 Aug 15].
17. Betts LR, Elder TJ, Hartley J, Trueman M. Does correction for guessing reduce students' performance on multiple-choice examinations? Yes? No? Sometimes? Assess Eval High Educ 2009;34:1-15.
18. Rowley GL, Traub RE. Formula scoring, number-right scoring, and test-taking strategy. J Educ Meas 1997;14:15-22.
19. Lesage E, Valcke M, Sabbe E. Scoring methods for multiple choice assessment in higher education – is it still a matter of number right or negative marking? Stud Educ Eval 2013;39:188-93.
20. Karandikar RL. On multiple choice tests and negative marking. Curr Sci 2010;99:1042-5.
21. Frary RB. Formula scoring of multiple choice tests (Correction for guessing). Educ Meas: Issues Pract 1988;7:33-8.
22. Tsai FJ, Suen HK. A brief report on a comparison of six scoring methods for multiple true-false items. Educ Psychol Meas 1993;53:399-404.
23. Tarasowa D, Aver S. Balanced Scoring Method for Multiple-mark-questions Published in 5th International Conference on Computer Supported Education (CSEDU 2013); 2013. Available from: http://scholar.google.com/scholar?cluster =1667572386692368560&hl=en&oi=scholar. [Last accessed on 2016 Feb 26].
24. Gross LJ. Considerations in scoring multiple true-false tests. Health Prof Educ Bull 1978;7:26-30.