Abstract -- An assessment methodology based on “Reverse Multiple-Choice Method,” is described. The novel method combines the ease, potential for automation, uniformity, objectivity and scalability of familiar multiple-choice format with the depth of feedback generally associated with essay or open format testing. The method is flexible and general, and may be adapted for educational testing in most subject areas at achievement levels from kindergarten through college.

Keywords -- testing, multiple-choice, training, in-depth assessment

By way of a warning, Reverse Multiple-choice Method presented here, is not an easy alternative to traditional multiple-choice. It is not suitable to conduct an entire test such as the SAT. But it can be effectively used in several situations, including for instance, an instructor looking for reliable feedback from the students, or for an assessment imperative when mastery of the material is crucial, e.g., for the teacher trainees. For the SAT’s and ACT’s a few RMCM questions could be added to provide “control” of validity, something the “essay” portion was intended to do.

1. Introduction

Frederick J. Kelley invented multiple-choice tests exactly 100 years ago in 1914. In these one hundred years, multiple-choice testing method has been exulted and maligned, but steadily grown to become the world’s most common method of educational testing, and the norm for examinations taken by large numbers of students. This run-away use of multiple-choice is beyond the inventor’s original intent. The format was devised to address the then need to deal with an eight-fold increase in the number of college bound high school students within two decades. The numbers, no doubt are behind its dominance to this day, along with the convenience, low cost, high speed, and uniformity with which these tests can be administered and graded.

But, as we show here, the format has other non-obvious virtues not commonly appreciated. In fact, despite the oxymoronic juxtaposition of “multiple-choice” and “in-depth” in the paper’s title, multiple-choice paradigm can be used for in-depth learning and assessment, by a change of perspective.

2. Multiple-Choice Format and Alternatives

2.1 Multiple-choice

A typical multiple-choice question comprises three sections: a set of presumed facts (e.g., a narrative, an expression, an equation, a geometric figure), an interrogative sentence (a.k.a., the "call of the question"), and answer choices, typically between three and five, one which would be graded as “correct” or “best” answer.

2.2 Open Format, Long/Short Essay

The classical methods of testing require a student to demonstrate in his own words mastery of the material learnt, by writing an essay on a given topic, by detailed answers to pointed questions or by giving solutions to problems given in a test etc.

In order to keep the students from rambling and keep answers within reasonable lengths, limits are often placed on the length of an answer.

2.3 Other formats

Several other formats have been devised and employed, for instance “filling in the blanks,”
“verifying statements,” often as true or false, “matching answers to similar or related questions” etc. In some cases, a multiple-choice question may admit more than one correct answer, in other cases “all of the above” or “none of the above” are provided as possible answers, thereby discouraging random selection of answers. These methods are generally either variations of multiple-choice or attempts to replicate the ease and scalability of multiple-choice tests. They may also aim for automatic grading since computers are widely available for the purpose. We do not discuss these alternatives further but our analysis applies with modifications.

3. Advantages and Drawbacks of Multiple-Choice

3.1 Advantages for Examiners, and for Examinees

Multiple-choice tests are scalable and “objective,” in that the grader may not contaminate grading and mark an answer correct or incorrect based on subjective bias; other advantages, for instance, uniformity, automatic grading, or low cost flow as byproducts. In fact, “democratization” by “objective” tests was recognized early as a virtue of the format.

The format has advantages for the test maker as well as the test-taker. For an examinee it means no tired fingers writing long answers and no time wasted in making the answer presentable stylistically or in penmanship.

For the examiner the advantages are immense: the drudgery of reading a stream of similar answers is reduced, even eliminated by machine grading. Due to the brevity of answers, a larger swath of the subject matter can be tested, leading to wider-scaled feedback. Pressure on test-takers is reduced as well, since spotting a correct answer is often easier than generating one.

3.2 Disadvantages of Multiple-choice and Its Alternatives

Advantages of multiple-choice format mentioned above come at a cost and lead to its drawbacks. The format provides little room for demonstrating the soundness of one’s underlying analysis or the accuracy of computation. Total loss of credit may result from incorrect reading of a single word, phrase, value or fact in the question on which the answers turn. The very ease and simplicity of selecting and recording an answer to a test question may blur the difference between knowledge and ignorance.

Additionally, the format can encourage students to learn the material superficially. And worse, the examiner cannot easily spot cheating when it occurs.

On the other hand, while providing opportunity to test-takers to demonstrate their command of the subject matter, open format testing is time consuming to take and to grade, as well as expensive, subjective and non-uniform in measurement. Several other alternatives listed above suffer from similar disadvantages.

3.3 Reverse Multiple-choice Method: A Promising Alternative

The Reverse Multiple-choice Method (RMCM) offers a promising assessment alternative that combines the uniformity, efficiency and grading ease of "objective" or "standardized" multiple-choice tests with reliable "measure of knowledge and understanding" generally associated with open format tests.

RMCM assessment methodology lends itself to automatic grading which makes it attractive for distance learning and MOOCs, massively large online courses.

4. The Reverse Multiple-choice Method (RMCM)

4.1 How RMCM Works

The Reverse Multiple-choice Method utilizes multiple-choice questions, with a twist: It requires a student/examinee to consider each answer choice, and inquire how the given facts would be modified to make that answer choice the "correct" or "best" answer.

The distinguishing steps typically undertaken by an RMCM system are as follows:

- prompt the examinee to select an answer choice as the correct answer;
- record the examinee's selection and assign credit for the selection;
- prompt the examinee to select at least one of answer choices not selected as correct, then prompt the examinee to provide a follow-up query to which this selected answer choice is a correct answer;
• match the follow up query against stored queries for which the answer is correct and provide a score for the question depending on the match.
• The system can flag the answer for human evaluation when the follow up query does not match a stored query – so as to address a situation where an examinee’s follow up query may be unanticipated but appropriate.

4.2 An Example

Let us consider as an example one of Frederick Kelly’s original questions:

Q. Which of following animals is a farm animal?

(A) a cow. (B) a tiger. (C) a rat. (D) a wolf.

The correct answer is (A), which the standard multiple-choice question seeks. Here, RMCM question would further ask the student to determine how the query would have to be changed so that the correct answer would be (B), (C) or (D).

All answers here are animals, so the answer turns on the qualifier “farm.” Therefore, changing the word “farm” is necessary, which requires one to think of the unique properties of cow, tiger, rat and wolf, not shared by other animals named. A rat is small, for example, unlike the other three animals. Tiger and wolf are both wild animals, but they can be distinguished possibly as being feline and canine respectively.

Therefore, the new, RMCM follow up queries might be:
For (B): Which of following animals is a feline animal?
For (C): Which of following animals is a small animal?
For (D): Which of following animals is a canine animal?

This approach indeed ratchets up complexity of the question. But, with reason: We can be confident that a student who can answer this RMCM question knows his animals and did not pick the answer to the original question at random.

This simple example illustrates when, where and how to use Reverse Multiple-choice. The original question of this example may have been aimed at elementary school students, and they may not have the vocabulary or maturity to come up with the reverse queries. But the same question and modified queries may be appropriate for older students, particularly for English Language Learners. RMCM items tend to be harder. And, not every multiple-choice question is useful to generate RMCM items.

5. Implementation Notes for RMCM

5.1 Question Creation: Expectations from an Examiner

While RMCM can reduce the burden on the examiner by mechanizing grading, much of the onus for effective use of the method is on the examiner.

A Reverse Multiple-choice question may be written ab initio, or created from an existing, imported multiple-choice question. In either case, an examiner has to put in forethought to examine suitability of the narrative stem as well as the answer choices of the question. Completely irrelevant answer choices may fit in traditional multiple-choice questions as “easy” but be totally inappropriate for query reversal. Examiners must also decide the purpose of each answer choice in terms of what is tested.

5.2 Context, Fact Objects and Fact Values

Implicit in any type of question and answer is the background or context of the narrative for underlying matter.

A rarely mentioned but notable merit of multiple-choice format is that it captures interpretative context better than other testing formats, for the simple reason that it is possible to view each putative answer of a multiple-choice question as adding contextual information for the interpretation of the narrative.

Furthermore, the selection of one answer choice as “correct” over the other choices generally turns on a few syntactic elements, such as, words, phrases, operations, numbers and symbols etc. In RMCM terminology, the syntactic elements that make an answer choice correct are called Fact Objects (FO). In our example of animals, the qualifier for the word “animal” is a fact object.
The value (perhaps a character string) of a fact object in an answer choice is Fact Value (FV) of the fact object, a concept akin to assignment of a constant value to a variable in algebra. In our example, farm, feline, small and canine may be regarded as the Fact Values of the qualifier fact object for the four answers.

When creating a question, the examiner specifies fact objects and fact values for all the answer choices. The system provides the platform and editorial support for question creation, later uses examiner’s specifications to automatically evaluate student answers or flag unexpected student answers for human evaluation.

In a test taken by very large number of students, the proportion of answers flagged for human evaluation would be relatively small if the question is well constructed and appropriate to the students’ academic level.

5.3 Question Creation Administration: Mode and Type of RMCM Question

Mode and type information for a question is specified by the examiner, assisted by the system in administrative functions and in generating reverse multiple-choice tasks.

Under RMCM regime, a question may be posed in three possible modes: A mode wherein the student is asked to select an incorrect answer, then asked to provide a modified or follow up query for which that answer would be correct; another mode wherein the student is asked additionally to identify the correct answer to the original query; and the third mode of traditional multiple-choice questions where only the correct answer is identified. This last mode is intended for administrative purposes of integration with standard multiple-choice tests.

The “type” of a question refers to the form in which the task of modification of query is specified for the examinee. There are many different types of RMCM question.

For each answer choice that the test-taker regards as incorrect answer to original query, such tasks include one or more of the following: identify the fact objects that need to be changed; identify the fact objects to change from a given list; write in the fact objects that need to be changed; write in the fact objects to change from a given list; identify the fact values of a fact object that need to be changed; identify the fact values to change from a given list; write in the fact values of a fact object which need to be changed; write in the fact values to change from a given list.

The write-in type questions are similar in spirit to “fill-in-the-blank,” and identification type to “matching” questions. But the creation of such question in RMCM protocol is systematic and offers potential advantages of automation at various levels.

We note that graders of classical open format questions typically look for segments in the nature of “fact object” and “fact value” in a student’s answer; Reverse Multiple-choice paradigm makes it possible to deploy computers.

5.4 Answering RMCM Question: Expectation from Examinees

In purely formal terms, for a test-taker answering RMCM questions may be a bit unfamiliar at first but not much harder than answering multiple-choice questions, and probably easier than answering long or short essay type questions.

The RMCM task may be specified in simple, familiar terms. For example:

“Find the words/phrases/symbols or other segments of the query which, if they are changed, will make your selected incorrect answer the correct answer for the changed question.”

Or,

“Your selected answer is incorrect because at least one query segment has the wrong value; identify which value from the given list should be assigned to the query segment(s) so that your selected incorrect answer becomes the correct answer for the changed question.”

Similar language may be used for write-in answers for fact objects or fact values.

Answering Reverse Multiple-Choice questions, however, puts different kind of pressure on students used to traditional multiple-choice questions – learning the subject matter in-depth. Students must acquire the skill to deconstruct and reassemble a question, and learn to focus on closely reading the fact pattern, critically evaluating the answer choices and recognizing the critical pieces of information in the fact pattern on which the answers turn.

In other words, the students must be prepared for in-depth assessment. We may also note that RMCM format is less prone to cheating by test-takers than standard multiple-choice format.
5.5 Grading

The examiner’s up-front work in question creation pays off in grading, by shifting a lot of the drudgery and cost to machines or to assistants who cost less.

5.6 A Detailed Example

Q. Divide and simplify the expression: \( \sqrt{10} \div \sqrt{2} \) =
(A) \( \sqrt{8} \)  (B) \( 2\sqrt{2} \)  (C) \( \sqrt{5} \)
(D) \( 2\sqrt{5} \)  (E) \( 2\sqrt{3} \)

The Correct answer is (C).

Here RMCM would task the student to modify the question suitably to make one or more of incorrect answers correct for the changed question. Thus, the incorrect answer:

(A) would be correct if we: change the division “\( \div \)” to a subtraction “\(-\)” and drop the second square root operation “\( \sqrt{\} \)” as well as extend the first “\( \sqrt{\} \)” to cover both the numbers;
(B) would be correct if we: change the “\( \div \)” to a “\( -\)” drop the second square root operation “\( \sqrt{\} \)” and extend the first “\( \sqrt{\} \)” to cover both the numbers;
(D) would be correct if we: change the division “\( \div \)” to “\( \ast \)” a multiplication; and
(E) would be correct if we: change the division “\( \div \)” to an addition “\( +\)” and drop the second square root operation “\( \sqrt{\} \)” as well as extend the first “\( \sqrt{\} \)” to cover both the numbers.

This question is designed to test as well the observation that the answer choices A and B are equivalent; hence the collated information for answer choices A and B above is identical.

The list of segments/syntactic elements here may be listed as: the first \( \sqrt{\} \), the first paren (, the first number, the second paren ), the arithmetic operation, the second \( \sqrt{\} \), the second number.

A significant point that we note is that Fact Objects are semantic entities encapsulated in syntactic elements. Context for interpretation of the narrative of the question imbues it with meaning, but in written language syntax is the bridge over which meaning is communicated.

Answer choices of a multiple-choice question append the contextual lexicon along with the query narrative and call of the question in a natural, efficient manner. By emphasizing the role of Fact Objects and Fact Values, RMCM can utilize this strength of the format for in-depth testing.

In practice, this may be achieved by calls to databases and table look ups, etc.

Therefore, the system may maintain several related tables, for example, the illustrative Table 1 below for compilation of Fact Objects and Fact Values; as well as, Table 2 for grading purposes that shows fact objects/fact values for answer choices with the score breakdown. In a question where examinees are asked to identify or modify the fact objects for a given answer choice, fact object selection score column of Table 2 is relevant. When the question asks for the fact values for one or more fact objects, fact value selection scores are relevant.

Tables similar to Table 1 and Table 2, below are expected to be provided by examiner when making the questions.
### TABLE 1. Fact Object/Fact Value Table

<table>
<thead>
<tr>
<th>Fact Object</th>
<th>Ans. choice</th>
<th>Fact Value</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>[the first square root], √</td>
<td></td>
<td></td>
<td>No change for any answer</td>
</tr>
<tr>
<td>[the first half paren], (</td>
<td></td>
<td></td>
<td>No change for any answer</td>
</tr>
<tr>
<td>[the first number]</td>
<td></td>
<td></td>
<td>No change for any answer</td>
</tr>
<tr>
<td>[the second half paren], )</td>
<td>(A)</td>
<td>(M, move symbol to location after second number)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(B)</td>
<td>(M, move symbol to location after second number)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(E)</td>
<td>(M, move symbol to location after second number)</td>
<td></td>
</tr>
<tr>
<td>[arithmetic operation]</td>
<td>(A)</td>
<td>- [subtraction]</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(B)</td>
<td>* [multiplication]</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(D)</td>
<td>+ [addition]</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(E)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>[the second square root ], √</td>
<td>(A)</td>
<td>30%</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(B)</td>
<td>40%</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(E)</td>
<td>30%</td>
<td></td>
</tr>
<tr>
<td>[the second number]</td>
<td></td>
<td>No change for any answer</td>
<td></td>
</tr>
</tbody>
</table>

### TABLE 2. Answer choice/Fact Object/Fact Value/Scoring

<table>
<thead>
<tr>
<th>Ans. Choice</th>
<th>Fact Object</th>
<th>FO selection score</th>
<th>Fact Value</th>
<th>FV selection score</th>
<th>comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>(A)</td>
<td>[the second half paren], )</td>
<td>25%</td>
<td>(M, move location after second number)</td>
<td>30%</td>
<td>40%</td>
</tr>
<tr>
<td></td>
<td>[arithmetic operation]</td>
<td>50%</td>
<td>- [subtraction]</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>[the second square root ], √</td>
<td>25%</td>
<td>(D, delete symbol)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(B)</td>
<td>[the second half paren], )</td>
<td>25%</td>
<td>(M, move location after second number)</td>
<td>30%</td>
<td>40%</td>
</tr>
<tr>
<td></td>
<td>[arithmetic operation]</td>
<td>50%</td>
<td>- [subtraction]</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>[the second square root ], √</td>
<td>25%</td>
<td>(D, delete symbol)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(D)</td>
<td>[arithmetic operation]</td>
<td>100%</td>
<td>* [multiplication]</td>
<td></td>
<td>100%</td>
</tr>
<tr>
<td>(E)</td>
<td>[the second half paren], )</td>
<td>25%</td>
<td>(M, move location after second number)</td>
<td>30%</td>
<td>40%</td>
</tr>
<tr>
<td></td>
<td>[arithmetic operation]</td>
<td>50%</td>
<td>+ [addition]</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>[the second square root ], √</td>
<td>25%</td>
<td>(D, delete symbol)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
6. Classroom Experience

The following is a simple example from a quiz in an introductory Statistics course.

Q. (i) The common intelligence quotient (IQ) scale is Normally distributed with mean 100 and standard deviation 15.

What proportion of population has IQ scores between 115 and 130?

A. 68%  B. 95%  (C) 13.5%  (D) 34%?

(ii) Answer choice _____ is incorrect because ____________________________.

(iii) Select one incorrect answer choice given above. Then change the question so that your selected answer choice is the correct answer to the changed question.

Herebelow is one student’s answer to this question:

[Answer choice] B is incorrect

• B is incorrect because to cover 95% of the graph the IQ score has to be between 70 and 130.

[New question]

• The common intelligence quotient (IQ) scale is normally distributed with mean 100 and standard deviation 15. What proportion of population has IQ scores between 70 and 130?

Such questions throughout the course proved helpful to the examiner for grading since the job of matching answers against model answers was carried out by an assistant. Also, in cases where the students did not select the correct answers, the one-line reason and the changed question gave a window into the students’ thinking and helped give partial credit where due.

7. Conclusion, Direction for Further Work and Research

Reverse Multiple-choice provides a viable assessment methodology, and an economical alternative to standard multiple-choice and open format testing techniques. It is particularly useful in situations where it is necessary to evaluate the extent to which the test-takers have mastered the art of recognizing critical pieces of information in the subject matter tested.

RMCM questions tend to be harder for the students, but their complexity deters both, mindless filling of the answer bubbles and cheating. These questions are harder for an examiner to create as well, but they can be reused over a longer shelf life.

Construction of the create and answer question platform is on its way to completion. We have planned an item library of RMCM questions for several subject area at various levels that could ease examiners’ burden.

8. References

