

# Medical Education

## Multiple choice questions: A literature review on the optimal number of options

RASHMI VYAS, AVINASH SUPE

### ABSTRACT

**Background.** Single, best response, multiple choice questions (MCQs) with 4 options (3 distractors and 1 correct answer) or 5 options (4 distractors) have been widely used as an assessment tool in medical education in India and globally. Writing plausible distractors is time consuming and the most difficult part of preparing MCQs. If the number of options can be reduced to 3, it will make preparing MCQs less difficult and time consuming, thus reducing the likelihood of flaws in writing MCQs. We reviewed the literature to find out if the number of options in MCQ test items could be reduced to 3 without affecting the quality of the test.

**Methods.** A systematic database search was done using the following question as a framework: How many options are optimal for multiple choice questions? Theoretical, analytical and empirical studies, which addressed this issue, were included in the review.

**Results.** There was no significant change in the psychometric properties of the 3 options test when compared with 4 and 5 options. MCQs with 3 options improved the efficiency of the test as well as its administration compared with 4- or 5-option MCQs. MCQs with 3 options had a higher efficiency because fewer distractors needed to be prepared, took up less space and required less reading time, decreased the time required to develop the items and the time to administer, and more items could be administered in a given time thus increasing the content sampled.

**Conclusion.** Our review of the literature suggests that MCQs with 3 options provide a similar quality of test as that with 4- or 5-option MCQs. We suggest that MCQs with 3 options should be preferred.

Natl Med J India 2008;21:130-3

### INTRODUCTION

Multiple choice questions (MCQs) are used globally as test items for assessment in various fields of education.<sup>1-4</sup> In India, MCQs have been commonly used for undergraduate and postgraduate medical entrance and university examinations.<sup>5-8</sup> These are also used as an assessment tool for paramedical courses and formative

assessment in the medical curriculum.<sup>9-11</sup>

The main challenge in preparing MCQs is to construct good test items.<sup>12-15</sup> This requires a good knowledge of the content and understanding of the objectives of assessment as well as good skills in writing the items.<sup>15-17</sup> Though there are many guidelines for writing good test items,<sup>18-21</sup> these are usually not adhered to; leading to the preparation and administration of faulty MCQs.<sup>8,22,23</sup>

The single, best-response type MCQ (the format that asks the student to choose the best answer from a list of possible answers) with 4 or 5 options is the most commonly used format.<sup>14,22,24-28</sup> For a distractor (wrong option) to be useful, it should represent a common misconception among students about the correct answer.<sup>26</sup> Writing plausible distractors is time consuming and the most difficult part of preparing MCQs.<sup>17,26,29,30</sup> Some of the main flaws in writing distractors include implausible distractors, more than one or no correct answer, the longest option being correct, and the use of 'all of the above' and 'none of the above' options.<sup>22,23</sup> Lack of training of the faculty in preparing MCQs and lack of time have been identified as important reasons that contribute to flaws in writing distractors/items.<sup>1,8</sup> Training faculty in the skills required for writing items and reducing the number of distractors in an MCQ could be some possible remedial measures.

We hypothesized that if the number of options in an MCQ can be reduced to 3 from the commonly used 4 or 5, it would make MCQ writing less difficult and less time consuming, thus reducing the likelihood of flaws in writing MCQs. We, therefore, reviewed the existing literature to determine if the number of options in an MCQ can be reduced to 3 without affecting the quality of the test.

### METHODS

A systematic database search was done using Science Direct, Pubmed, Ovid and ERIC search engines for the period 1960-2007. Specific articles were also looked for in the Sage publications, JStor and Blackwell Synergy electronic resources. The articles were searched using the following question as a framework: How many options are optimal for multiple choice questions? The search terms included the keywords: MCQ, multiple choice questions, optimal distractors, optimal options, number of options, number of distractors, one best response, item analysis and item writing guidelines.

All studies that assessed the 3-option format were included in the review. Studies not available in electronic format were excluded from the review. The initial results were screened for pertinence, which yielded 15 articles for reading. A review of the reference lists identified 8 additional articles of interest.

### RESULTS

Studies on the optimal number of options were grouped into (i) theoretical and analytical studies, and (ii) empirical studies,

Christian Medical College, Vellore 632002, Tamil Nadu, India

RASHMI VYAS Department of Physiology

Seth G.S. Medical College and K.E.M. Hospital, Parel, Mumbai 400012, Maharashtra, India

AVINASH SUPE Department of Surgical Gastroenterology

Correspondence to RASHMI VYAS; [rashmifyas@yahoo.co.in](mailto:rashmifyas@yahoo.co.in); [rashmifyas@gmail.com](mailto:rashmifyas@gmail.com)

© The National Medical Journal of India 2008

based on the approach of the study. These studies are described based on the parameters used by them for recommending 3 options over 4 and 5 options in a single, best-response MCQ.

#### Reliability

*Theoretical and analytical study.* Grier in 1975 provided mathematical proof of increased reliability with 3 options, provided the total number of test items was increased to balance the decrease in the number of alternatives.<sup>31</sup>

*Empirical studies.* In their review of 10 empirical studies, Haladyna and Downing showed that reliability increased with increasing number of options.<sup>24</sup> However, this increase was small when more than 3 options were used.<sup>24</sup> Other studies have shown little difference in reliability estimates of different test formats using a different number of options,<sup>32-35</sup> and an increase in score reliability on removal of non-functioning distractors.<sup>36</sup> Rodriguez in 2005 did a meta-analysis of 80 years of research on the number of options<sup>16</sup> and concluded that reducing the number of options decreased test score reliability (5 to 4 options, 5 to 2 options, 4 to 2 options), except when the reduction was from 4 to 3 options, which increased the test score reliability.

Haladyna and Downing,<sup>24</sup> in their review of research on writing MCQs, concluded that 'empirical research supports theoretical study in indicating that three options achieve the optimal balance between reliability and efficiency'.

#### Validity

*Empirical studies.* Green *et al.* in 1982 showed that the 5-option MCQ was the least defensible.<sup>37</sup> Owen and Froman in 1987<sup>38</sup> and Trevisan *et al.* in 1994<sup>34</sup> concluded that 3-option MCQs improved validity. Also, more test items could be added in the 3-option MCQs. This would provide the benefit of increased sampling from the domain to be tested thus increasing validity.<sup>34,36,38</sup> Rodriguez<sup>16</sup> in his meta-analysis concluded that the validity remains the same when the number of options are reduced from 5 to 3 and from 5 to 4 to 3.

#### Item discrimination

*Theoretical and analytical study.* Tversky, in 1964, in his landmark paper provided mathematical proof that 'whenever the amount of time spent in the test is proportional to its total number of alternatives, the use of three choice alternatives at each point will maximize discriminability, power and the amount of information obtained per unit time'.<sup>39</sup>

*Empirical studies.* To summarize the observations, Costin in 1970 gave empirical evidence for Tversky's mathematical proof on the number of options.<sup>40</sup> He showed higher mean discrimination indices for the 3 choice items. In 1985, Haladyna and Downing in their review of research on MCQs showed mixed results for item discrimination.<sup>24</sup> In their review, while one study showed no difference in item discrimination between 3 and 4 options, another study showed 3-option items to have better item discrimination than 4 options.<sup>24</sup> However, later studies showed an increase in item discrimination with 3-option MCQs.<sup>16,32,33</sup>

#### Item difficulty

*Empirical studies.* Haladyna and Downing reviewed studies on the number of options in terms of item difficulty and concluded that 3-4 options are optimal.<sup>24,41</sup> Studies have shown no difference in mean item difficulty, between 3 and 5 options<sup>38</sup> and between 3 and 4 options.<sup>42</sup> Yet other studies have shown that student

performance improved while at the same time the difficulty of the test item increased with the 3-option format compared with the 4-option format.<sup>32,33</sup> A small increase in difficulty index was observed when reducing the items from 4 to 3 options while there was a large increase in the difficulty index (items were easier) if the options were reduced to 2.<sup>16</sup> Hence, it can be concluded from empirical studies that there is not much difference in item difficulty between 3 and 4 options.

#### Test efficiency

*Theoretical and analytical studies.* Lord in 1977 used the item characteristic curve (ICC) theoretical model to show that the test efficiency for different groups of students may differ.<sup>25</sup> Budescu and Nevo proposed the use of the law of proportionality for deciding on the number of options.<sup>43</sup> The law of proportionality is based on Grier's assumptions of generalized proportionality. In his work, Grier showed that 'the approximate lower bound of the test reliability is maximized by using three options in tests with at least  $n=18$  items, and two options in shorter tests'.<sup>31</sup> Later, Bruno and Dirkwager<sup>27</sup> used information-based theoretical perspectives to conclude that 3 choices to a multiple choice item are optimal. This is based on 'information content or the channel capacity of the multiple-choice test items for determining the optional number of alternatives for multiple-choice items'.<sup>27</sup>

*Empirical studies.* Levine and Drasgow in 1983 showed that information is maximized by using more options per item for lower ability groups and fewer numbers of options for higher ability groups.<sup>44</sup> This approach uses the concept of 'ability' and does not show much advantage when comparing 3 and 4 options. Haladyna and Downing reviewed studies on the number of options in terms of efficiency and concluded that 3-4 options are optimal.<sup>24</sup> Compared with 4-option formats, 3-option formats have been shown to improve student performance, are considered to be a better measure of student knowledge, are preferred by students, and perceived by them to be less confusing and tricky.<sup>33,38</sup> The students' preference for the 3-option format was evaluated using ballots on which they could vote for their preferred choice.<sup>38</sup>

Removal of non-functioning distractors and reducing the number of options to 3 provided benefits such as reduced testing time,<sup>36,38</sup> more efficient use of testing time and greater score precision per unit testing time.<sup>3,4</sup> A study compared the psychometric properties of 3-option and 5-option formats.<sup>28</sup> Data were obtained from two batteries of tests administered earlier. Trained item writers had written the items using common principles of writing. The items were analysed using item-total correlations, distractor analysis, item difficulty, chi-square and latent trait techniques. 'Items were then edited or rewritten as necessary and reviewed by a 3-person expert panel. The remaining items were pilot tested to insure the adequacy of their psychometric properties.'<sup>28</sup> Removing the options with the lowest response rate from the 5-option test created the 3-option test. The results of the 3-option and 5-option tests were analysed. The means of the two tests were similar. The internal consistency reliability for the two tests was measured using coefficient alpha. Further analyses were done using the z-test to compare p values of the 5-option and 3-option items. The authors concluded that 'while there were significant differences between some items, the differences tended to be very small and, as the small means suggest, of limited practical significance.'<sup>28</sup> They also found that both formats had similar psychometric properties but the 3-option format was more cost-efficient in terms of development and administration.<sup>28</sup>

### Guessing

*Theoretical and analytical studies.* In his observation, Lord took into account the issue of guessing which is more common for low performers. He concluded that for most examinees 3 options appeared to be optimal.<sup>25</sup>

*Empirical study.* A comparison of 3- and 4-option items showed a decrease in 'test-wiseness' or guessing with 3-option items.<sup>35</sup> 'Test-wiseness' was defined by Millman *et al.*<sup>45</sup> as 'a subject's capacity to utilize the characteristics and formats of the test and/or test taking situation to receive a high score'.

### Useful options

Haladyna and Downing concluded that the 3-option format is optimal as the number of functional distractors per item was optimal.<sup>30</sup> Other studies confirmed that the 3-option format had fewer dysfunctional distractors,<sup>32,33</sup> the mean number of functioning distractors was much lower than 2 and reducing the least popular option had only a minimal effect on the performance of the remaining options,<sup>1</sup> and test items seldom contained more than 3 useful options.<sup>26</sup>

### DISCUSSION

The 3-option MCQ formats have a number of advantages without compromising the effectiveness of the test. One advantage is that it is easier to construct two plausible distractors thus reducing the likelihood of flaws in the distractors. Flaws in writing items introduce threats to validity,<sup>22,46</sup> which is a major concern for any assessment tool.<sup>46-48</sup> The strongest rationale for recommending the 3-option format is that there is no significant change in the psychometric properties of this format when compared with the 4- and 5-option formats. A recent meta-analysis showed that the 3-option format reduces item difficulty, and increases item discrimination and reliability.<sup>16</sup> Some studies have shown that the 3-option format improves validity<sup>16, 36,42</sup> though more studies are required to confirm this.

Three-option items have also been recommended in the literature because of their ease of preparation as these require fewer distractors, take up less space, require less reading time, and decrease time for item development and administration. More items can be administered in a given time, thus increasing the content that is sampled resulting in improved validity. Moreover, studies have addressed the concern that MCQs may be susceptible to 'test-wiseness' or guessing. It has been shown that the 3-option MCQ in a high stakes school-leaving mathematics examination decreases test-wiseness.<sup>35</sup>

It has also been documented that test items seldom contain more than 3 useful options. In India and globally, 4- or 5-option formats are commonly used. In many countries including India, MCQs in medical education are usually constructed by in-house faculty members.<sup>8</sup> Most of them have other commitments, which include teaching, clinical work, research and administrative responsibilities. The use of the 3-option format will save considerable time and energy without compromising on the quality of tests. At the same time, training faculty in the skills for writing items and regular pre- and post-validation of MCQ items is necessary to ensure that the guidelines for writing MCQs are followed.

### CONCLUSION

Based on our review of the published literature, we recommend that the 3-option format of MCQs should be used. We would also second the recommendations of Rodriguez<sup>16</sup> that item writing guidelines should be made more direct.

### REFERENCES

- Shizuka T, Takeuchi O, Yashima T, Yoshizawa K. A comparison of three- and four-option English tests for university entrance selection purposes in Japan. *Language Testing* 2006;**23**:35-57.
- Collins J. Education techniques for lifelong learning: Writing multiple-choice questions for continuing medical education activities and self-assessment modules. *Radiographics* 2006;**26**:543-51.
- Swanson DB, Holtzman KZ, Clauser BE, Sawhill AJ. Psychometric characteristics and response times for one-best-answer questions in relation to number and sources of options. *Acad Med* 2005;**80**:S93-S96.
- Swanson DB, Holtzman KZ, Albee K, Clauser BE. Psychometric characteristics and response times for content-parallel extended-matching and one-best-answer items in relation to number of options. *Acad Med* 2006;**81**:S52-S55.
- Kartha CC. Anguish of a medico. *Curr Sci* 2005;**89**:725-6.
- Sood R, Adkoli BV. Medical education in India: Problems and prospects. *J Indian Acad Clin Med* 2000;**1**:210-12.
- Ramakrishnan K, Singhal K. Training needs of international medical graduates seeking residency training: Evaluation of medical training in India and the United States. *Internet J Fam Pract* 2004;**3**:1.
- Ananthakrishnan N, Ananthakrishnan S, Oumachigui A. MBBS examinations: Are we asking the right questions? *J Postgrad Med* 1993;**39**:31-2.
- Singh T, Natu MV. Examination reform at the grassroots: Teacher as the change agent. *Indian Pediatr* 1997;**34**:1015-19.
- Regulations and syllabus for 2007 MDS Post Graduate Dental Course. The Tamil Nadu Dr MGR Medical University, Chennai, India. Available at [www.tmmu.ac.in/pdf/MDSREGULATIONS.pdf](http://www.tmmu.ac.in/pdf/MDSREGULATIONS.pdf) (accessed on 23 May 2008).
- Haranath PSRK. Reflections on the evolution of pharmacology in India during twentieth century. *Indian J Pharmacol* 1999;**31**:1-13.
- Walsh K. Advice on writing multiple choice questions (MCQs). *BMJ Careers* 2005. Available at <http://careers.bmj.com/careers/advice/view-article.html?id=616> (accessed on 23 May 2008).
- Walsh K. Answering multiple choice questions. *BMJ Careers* 2005. Available at <http://careers.bmj.com/careers/advice/view-article.html?id=891> (accessed on 23 May 2008).
- Epstein RM. Assessment in medical education. *N Engl J Med* 2007;**356**:387-96.
- Huang Yi-M, Trevisan M, Storfer A. The impact of the 'all-of-the-above' option and student ability on multiple choice tests. *Int J Scholarship Teach Learn* 2007;**1**:1-13.
- Rodriguez MC. Three options are optimal for multiple-choice items: A meta-analysis of 80 years of research. *Educ Meas: Issues Pract* 2005;**24**:3-13.
- Haladyna TM. MC formats. In: Haladyna TM (ed). *Developing and validating multiple-choice test item*. 3rd edn. Mahwah, New Jersey:Lawrence Erlbaum Associates; 2004:67-96.
- Haladyna TM, Downing SM, Rodriguez MC. A review of multiple-choice item-writing guidelines for classroom assessment. *Applied Meas Educ* 2002;**15**:309-34.
- Question writing guidelines. Available at [www.acvim.org/uploadedFiles/Candidates/exam/ABIMQuestionWritingGuidelines.pdf](http://www.acvim.org/uploadedFiles/Candidates/exam/ABIMQuestionWritingGuidelines.pdf) (accessed on 24 May 2008).
- Cheung D, Bucat R. How can we construct good multiple-choice items? Paper presented at the Science and Technology Education conference, Hong Kong, 20-21 June 2002. Available at <http://www3.fed.cuhk.edu.hk/chemistry/files/constructMC.pdf> (accessed on 24 May 2008)
- Haladyna TM. Guidelines for developing MC items. In: Haladyna TM (ed). *Developing and validating multiple-choice test items* 3rd edn. Mahwah, New Jersey:Lawrence Erlbaum Associates; 2004:97-126.
- Downing SM. The effects of violating standard item writing principles on tests and students: The consequences of using flawed test items on achievement examinations in medical education. *Adv Health Sci Educ* 2005;**10**:133-43.
- Tarrant M, Knierim A, Hayes SK, Ware J. The frequency of item writing flaws in multiple-choice questions used in high stakes nursing assessments. *Nurse Educ Today* 2006;**26**:662-71.
- Haladyna TM, Downing SM. A quantitative review of research on multiple-choice item writing. Paper presented at the annual meeting of the American Educational Research Association, Chicago, Illinois, April 1985. Available at [http://eric.ed.gov/ERICDocs/data/ericdocs2sql/content\\_storage\\_01/0000019b/80/30/32/ee.pdf](http://eric.ed.gov/ERICDocs/data/ericdocs2sql/content_storage_01/0000019b/80/30/32/ee.pdf) (accessed on 24 May 2008).
- Lord FM. Optimal number of choices per item: A comparison of four approaches. *J Educ Meas* 1977;**14**:33-8.
- Haladyna TM, Downing SM. How many options is enough for a multiple-choice test item? *Educ Psychol Meas* 1993;**53**:999-1010.
- Bruno JE, Dirkwzager. Determining the optimal number of alternatives to a multiple-choice test item: An information theoretic perspective. *Educ Psychol Meas* 1995;**55**:959-66.
- Sidick JT, Barrett GV, Doverspike D. Three alternative multiple choice tests: An attractive option. *Pers Psychol* 1994;**47**:829-35.
- Lowe D. Set a multiple choice question (MCQ) examination. *BMJ* 1991;**302**:780-2.
- Haladyna TM, Downing SM. Functional distractors: Implications for test-item writing and test design. Paper presented at the annual meeting of the American Educational Research Association, New Orleans, Louisiana, 5-9 April 1988. Available at [http://eric.ed.gov/ERICDocs/data/ericdocs2sql/content\\_storage\\_01/0000019b/80/1d/83/1d.pdf](http://eric.ed.gov/ERICDocs/data/ericdocs2sql/content_storage_01/0000019b/80/1d/83/1d.pdf) (accessed on 24 May 2008).

- 31 Grier JB. The number of alternatives for optimum test reliability. *J Educ Meas* 1975;**12**:109–13.
- 32 Trevisan MS, Sax G, Michael WB. The effects of the number of options per item and student ability on test validity and reliability. *Educ Psychol Meas* 1991;**51**:829–37.
- 33 Landrum RE, Cashin JR, Theis KS. More evidence in favor of three-option multiple-choice tests. *Educ Psychol Meas* 1993;**53**:771–8.
- 34 Trevisan MS, Sax G, Michael WB. Estimating the optimum number of options per item using an incremental option paradigm. *Educ Psychol Meas* 1994;**54**:86–91.
- 35 Rogers WT, Harley D. An empirical comparison of three- and four-choice items and tests: Susceptibility to test-wiseness and internal consistency reliability. *Educ Psychol Meas* 1999;**59**:234–47.
- 36 Cizek GJ, Robinson KL, O'Day DM. Nonfunctioning options: A closer look. *Educ Psychol Meas* 1998;**58**:605–11.
- 37 Green K, Sax G, Michael WB. Validity and reliability of tests having different numbers of options for students of differing levels of ability. *Educ Psychol Meas* 1982;**42**:239–45.
- 38 Owen SV, Froman RD. What's wrong with three-option multiple-choice items? *Educ Psychol Meas* 1987;**47**:513–22.
- 39 Tversky A. On the optimal number of alternatives at a choice point. *J Math Psychol* 1964;**1**:386–91.
- 40 Costin F. The optimal number of alternatives in multiple-choice achievement tests: Some empirical evidence for a mathematical proof. *Educ Psychol Meas* 1970;**30**:353–8.
- 41 Haladyna TM, Downing SM. Validity of a taxonomy of multiple-choice item-writing rules. *Appl Meas Educ* 1989;**2**:51–78.
- 42 Crehan KD, Haladyna TM, Brewer BW. Use of an inclusive option and the optimal number of options for multiple-choice items. *Educ Psychol Meas* 1993;**53**:241–7.
- 43 Budescu DV, Nevo B. Optimal number of options: An investigation of the assumption of proportionality. *J Educ Meas* 1985;**22**:183–96.
- 44 Levine MV, Drasgow F. The relation between incorrect option choice and estimated ability. *Educ Psychol Meas* 1983;**43**:675–85.
- 45 Millman J, Bishop CH, Ebel R. An analysis of test-wiseness. *Educ Psych Meas* 1965;**25**:707–26.
- 46 Downing SM, Haladyna TM. Validity threats: Overcoming interference with proposed interpretations of assessment data. *Med Educ* 2004;**38**:327–33.
- 47 Downing SM. Validity: On the meaningful interpretation of assessment data. *Med Educ* 2003;**37**:830–7.
- 48 Downing SM. Reliability: On the reproducibility of assessment data. *Med Educ* 2004;**38**:1006–12.

### *New 5-year subscription rates*

We have introduced a new 5-year subscription rate for *The National Medical Journal of India*. By subscribing for a duration of 5 years you save almost 14% on the annual rate and also insulate yourself from any upward revision of future subscription rates. The new 5-year subscription rate is:

<b>INDIAN SUBSCRIBERS:</b>	Rs 2600 for institutions	<b>OVERSEAS SUBSCRIBERS:</b>	US\$ 365 for institutions
	Rs 1300 for individuals		US\$ 182 for individuals

Send your subscription orders by cheque/demand draft payable to *The National Medical Journal of India*. Please add **Rs 75** for outstation cheques. If you wish to receive the *Journal* by registered post, please add **Rs 90 per annum** to the total payment and make the request at the time of subscribing.

Please send your payments to:

The Subscription Department  
*The National Medical Journal of India*  
 All India Institute of Medical Sciences  
 Ansari Nagar  
 New Delhi 110029