Very Short Answer Questions: A Viable Alternative to Multiple Choice Questions

THOMAS PUTHIAPARAMPIL (✉ pthomas@unimas.my)

Universiti Malaysia Sarawak <https://orcid.org/0000-0002-7939-5185>

Md Mizanur Rahman

Universiti Malaysia Sarawak

Research article

Keywords: Very Short Answer Questions, Multiple Choice Questions, Best answer questions

Posted Date: March 10th, 2020

DOI: <https://doi.org/10.21203/rs.3.rs-16617/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License. [Read Full License](#)

Version of Record: A version of this preprint was published at BMC Medical Education on May 6th, 2020.
See the published version at <https://doi.org/10.1186/s12909-020-02057-w>.

Abstract

Background Multiple choice questions, used in medical school assessments for decades, have many drawbacks, such as: hard to construct, allow guessing, encourage test-wiseness, promote rote learning, provide no opportunity for examinees to express ideas, and do not provide information about strengths and weakness of candidates. Directly asked and answered questions like Very Short Answer Questions (VSAQ) is considered a better alternative with several advantages.

Objectives This study aims to substantiate the superiority of VSAQ by actual tests and obtaining feedback from the stakeholders.

Methods Conduct multiple true-false, one best answer and VSAQ tests in two batches of medical students, compare their scores and psychometric indexes of the tests and seek opinions from students and academics regarding these assessment methods.

Results Multiple true-false and best answer test scores showed skewed results and low psychometric performance compared to better psychometrics and more balanced student performance in VSAQ tests. The stakeholders' opinions were significantly in favour of VSAQ.

Conclusion and recommendation This study concludes that VSAQ is a viable alternative to multiple choice question tests, and it is widely accepted by medical students and academics in the medical faculty.

Introduction

Multiple True/False (MTF) and One Best Answer Questions (BAQ) are widely employed in medical faculties by virtue of their advantages of instant machine scoring, freedom from examiner bias and dependable reliability [1–4]. In this article, 'MCQ' is used to refer to both these instruments of assessment. Reliability of a test is higher when the subject coverage is wider [5, 6]. Reliability refers to test reproducibility with similar results when used for different batches, but it does not ensure validity [2, 3]. MCQ tests seem to sacrifice validity for reliability [3]. A primary purpose of any assessment is to let students know what is important to learn [6]. Assessments are supposed also to enable feedback to students and direct the teaching strategies [1, 7–10], and this is applicable especially to formative assessments. MCQ tests do not provide information, which would enable feedback, as they do not require students to construct the answers [1]. It is an established fact that assessments drive the learning style [1, 6, 10, 11]. MCQ is blamed for promoting rote learning, guessing, test-wiseness, and turning students into data banks [2, 5, 7, 12–14]. MCQ tend to test trivia [2] and they are not able to test complex issues [3, 7]. MCQ does not provide any opportunity for students to express their understanding [15]. What is required in a medical school is learning that leads to the formation of competent doctors, which MCQ tests are incapable of assessing fully [3]. Constructing good MCQs with appropriate difficulty and discriminating ability needs expertise and experience, which are often lacking [6, 15, 16]. Discrimination index (DISi) in MCQ may be deceptive as good students may find them more difficult to answer, as the questions may be the authors' opinions and not well-known facts [2, 15]. True-false MCQ may even have a negative effect on students, attributable to the false statements, which they might take home as true [10]. It appears that the demerits of MCQ are numerous and

overwhelming, and there is strong backing for alternative instruments like directly asked directly answered questions.

Objectives of the study

This study aims to test directly asked directly answered questions like very short answer questions (VSAQ) in medical students, compare their performance with MCQ tests and seek opinions from participating students and academics of the Faculty of Medicine and Health Sciences, Universiti Malaysia Sarawak, regarding MCQ and VSAQ.

Materials And Methods

Thirty-seven year-3 and 39 year-5 medical students took part in the study. VSAQ papers were prepared by converting MTF and BAQ items used for end of posting examinations of the same batches of students but of different rotations of the posting. This meant that the questions were testing the same knowledge using different instruments in different students of the same standard. The VSAQ answer scripts were manually scored. The student scores of MTF, BAQ and VSAQ tests were compared using Pearson correlation coefficient. Student feedbacks were obtained using a written questionnaire about their views on MTF, BAQ and VSAQ. The item analysis of VSAQ, manually performed, was compared with those of MTF and BAQ. The academics' views were obtained using a Google Form questionnaire with 17 items.

Results

Table 1
Year-3 and year-5 Students' scores in the 6 Tests.

Test	Y3	Q	Mean (SD)	Min, Max	Y5	Q	Mean (SD)	Min, Max	Sp-value
MTF	37	20	45.5 (9.4)	25, 54	39	20	55.0 (11.1)	29, 76	p < 0.001***
BAQ	37	15	54.9(12.8)	31.7, 82.3	39	15	76.9 (11.6)	53.3, 100	p < 0.001***
VSAQ	37	53	54.1(12.4)	33.7, 77.9	39	32	62.9 (11.2)	38, 91	p < 0.01**
VSAQ vs. MTF (r = 0.714***)					VSAQ vs. MTF (r = 0.632***)				
VSAQ vs. BAQ (r = 0.435**)					VSAQ vs. BAQ (r = 0.321*)				
*p < 0.05, **p < 0.01, ***p < 0.001. Sp-value reached from independent sample t test Y3 = year-3, Y5 = year-5 number of students, Q = number of questions, Min = minimum, Max = maximum, SD = standard deviation, MTF = multiple true-false test, BAQ = one best answer test, VSAQ = very short answer questions test.									

Table 1 shows significantly higher performance of year-5 students in MTF, BAQ and VSAQ compared to year 3 (p < 0.01). Test-wise analysis reveals a strong correlation between MTF and VSAQ both in year-3 and year-5 tests (p < 0.001). However, a weak positive relationship is found between BAQ and VSAQ (p < 0.01).

The scores of 2 batches of students (year-3, 37 and year-5, 39) in each of 3 tests were combined for this chart. The scores (%) achieved are shown in 4 categories (< 50 , $50-59$, $60-74$ and ≥ 75). For VSAQ preparation the previous rotation of same batch students' question papers was used. The overall performance in MTF shows a low trend in contrast to BAQ, which shows an upward trend. However, the performance in VSAQ shows a desirable, balanced distribution (Fig. 1).

The item analysis values of 3 tests of year-3 and 3 tests of year-5 batches were combined for this figure. The bars show percentages of items falling in 3 categories: ≤ 0.3 (too difficult), $0.31-0.79$ (moderate difficulty) and ≥ 0.8 (too easy) as per Ebel and Frisbe (1972). Most MTF items fall in the moderate difficulty category, BAQ shows a trend towards more easy items, while VSAQ shows a more balanced distribution.

The discrimination index of year-3 and year-5 test items are shown in Fig. 3. MTF (20 items), BAQ (15 items, 4-option) and VSAQ (54 items Y-3 and 32 items Y-5) are combined in this figure. Different groups of students of the same batch participated in the study. The VSAQ papers were prepared by converting the MTF and BAQ questions. The bars show percentages of items falling in the three categories: ≤ 0.0 = no or negative discrimination; $0.1-0.19$ = fair and ≥ 0.2 good and excellent distribution combined. MTF shows a fair distribution, BAQ higher than MTF and VSAQ about 70% of items having the highest discrimination power. Item analysis of VSAQ was manually performed on Excel Worksheet using the UNIMAS formulae [17].

Table 2
Students' feedback on VSAQ

Statements	Year-3 (n = 37)			Year-5 (n = 39)			p-value
	In favour	Neutral	Against	In favour	Neutral	Against	
VSAQ is an efficient method of assessment of knowledge	86.5	13.5	0.0	84.6	10.3	0.0	p > 0.05
VSAQ can test more items compared to MTF and BAQ in the same amount of time	70.3	10.8	18.9	59.0	33.3	17.9	p < 0.05
VSAQ tests knowledge better as the students need to write the answers	67.6	24.3	13.5	87.2	7.7	12.8	p > 0.05
VSAQ provides an opportunity for students to express their ideas	86.5	10.8	2.7	89.7	7.7	2.6	p > 0.05
Happy to see a new assessment method	78.4	21.6	0.0	84.6	12.8	0.0	p > 0.05
I would like to see more VSAQ tests.	59.5	35.1	5.4	79.5	15.4	5.1	p > 0.05
*p < 0.05, **p < 0.01, ***p < 0.001. p-value reached from chi-square test							

Most of the students expressed opinions in favour of VSAQ. More than four-fifths (86.5%) of them opined that VSAQ is an efficient method of assessment of knowledge, and equal proportions of them opined that VSAQ provides an opportunity for students to express their ideas. About 80% in both groups were happy to see this new method of assessment, and 59.5% and 79.5% of them expressed interest to see more VSAQ tests (Table 2). Pearson's chi-square test of independence revealed no statistically significant difference between year-3 and year-5 students' opinions (p > 0.05). The results showed high acceptance of VSAQ compared to MCQ.

Table 3
Students' feedback on MTF and BAQ

Statements	Year-3 (n = 37)			Year-5 (n = 39)			p-value
	Agree	Not sure	Disagree	Agree	Not sure	Disagree	
MTF is more efficient than VSAQ for testing knowledge	16.2	64.9	18.9	15.4	46.2	38.5	p > 0.05
BAQ is more efficient than VSAQ in testing knowledge	29.7	40.5	29.7	15.4	41.0	43.6	p > 0.05
MTFs, which I encountered had been clear and unambiguous	29.7	29.7	40.5	28.2	15.4	56.4	p > 0.05
BAQs are easy, as most distractors are easy to exclude to reach the answer	62.2	10.8	27.0	53.8	20.5	25.6	p > 0.05
	Very often	Sometimes	Seldom	Very often	Sometimes	Seldom	
Did you encounter confusing items in MTF?	0.0	10.8	89.2	59.0	35.9	5.1	p < 0.001
Did you encounter confusing items in BAQ?	0.0	43.2	56.8	28.2	51.3	20.5	p < 0.001
	Very confident	OK	Not confident	Very confident	OK	Not confident	

*p < 0.05, **p < 0.01, ***p < 0.001. p-value reached from chi-square test

Statements	Year-3 (n = 37)			Year-5 (n = 39)			p-value
	Agree	Not sure	Disagree	Agree	Not sure	Disagree	
How confident have you been while answering MTF?	0.0	73.0	27.0	53.8	20.5	25.6	p < 0.001
*p < 0.05, **p < 0.01, ***p < 0.001. p-value reached from chi-square test							

There is a significant similarity in the opinions expressed by year-3 and year-5 students about MCQ tests. They did not consider MCQ tests any better in assessing knowledge compared to VSAQ. Over 62% (year-3) and 53% (year-5) students considered BAQ easy to answer, as they contained many distractors easy to exclude as the answer. A good number of students, more of them in year-5 group, had encountered confusing MTF and BAQ items in their examinations.

Lecturers' feedback on MTF, BAQ and VSAQ

1. More than 90% were involved in constructing MCQ and BAQ, but less than 10% of them mentioned it was a pleasant experience.
2. More than 90% felt writing BAQ was more difficult than writing MTF. Over 67% considered writing 3-4 plausible distractors a difficult task.
3. In MTF, 87% considered writing of false options more difficult than writing true options. Twenty per cent of them said they converted true options to false, and some said they just wrote false options without care, as they were false anyway.
4. About using item analysis to improve the questions, 60% mentioned they did not do it.
5. Cent per cent respondents said they got ideas about the students' strengths and weaknesses while marking directly asked-directly answered questions like MEQ and SAQ, and 40% of them said they do not get any such information from MTF and BAQ results. However, 97% believed such information was important to guide teaching.
6. About their openness to VSAQ, 60% answered 'surely' and 40% answered 'maybe'.

Discussion

Students scored poorly in MTF with high failure rates and very few high scorers. On the contrary, BAQ was skewed towards high scores with very few poor performers. The VSAQ scores showed a fair and balanced distribution as would be expected in a reliable test. There was a strong positive correlation between VSAQ and MTF both in year-3 and year-5 tests ($p < 0.001$), while the correlation between VSAQ and BAQ in both tests were weakly positive ($p < 0.01$). This discrepancy was somewhat reflected in the item analysis of these tests. While most MTF items fell in the moderate difficulty category, BAQ items were more easy and

VSAQ more balanced in distribution. Items with 0.2 or higher DISi were more in VSAQ compared to MTF and BAQ, while items with lower DISi were more in both MTF and BAQ tests. Our overall results showed fairer and balanced student performance and superior psychometric properties of VSAQ compared to MTF and BAQ.

Students' opinions about MCQ were generally unfavourable. Mention of many confusing items in the questions was notable. Negative opinions about MTF and BAQ were more evident in year-5 students' feedback, which reflected their seniority and the number of tests they would have taken compared to year-3 students. Generally, students' opinions were highly in favour of VSQ. Hift [18] proposed elimination of open-ended questions like MEQ in summative assessments, as context-rich well written best answer questions had higher reliability and validity than MEQ. But this author also supported open-ended questions for formative assessments or assessments for learning. Open-ended questions were favoured also by college students [1].

Validity and reliability of tests depend on how widely the subject is covered in the assessment [5, 6], but unduly prolonged tests would be counterproductive. Our VSAQ test with 53 questions and 14-item feedback was completed in 30–45 minutes. MTF of 20 questions took 50 minutes and BAQ of 15 questions took 45 minutes to complete. Our study showed that wider coverage of topics was easier to achieve in VSAQ tests compared to MCQ, as claimed in literature[3, 19]. Test-wiseness and guessing, which are unavoidable in MCQ tests are of no concern in VSAQ. It has also been observed that students found false options harder than true ones to answer, and omission rates due to uncertainty are high in MTF [20]. The adverse effect of using false options was also highlighted by Wood [10]. BAQ claims higher-order assessment than MTF but can be too easy because of the many non-functioning distractors most of them contain [21].

Assessments are of immense value in letting students know what is important to learn [6] and for providing feedback to students and directing the teaching [1, 7, 8, 15, 22]. VSAQ was seen to achieve these ends better than MCQ. Reading the VSAQ answers while marking provides the examiner's insight into students' strengths and deficiencies, which would help in providing feedback and modifying teaching [7, 19]. MCQ tests do not reveal with certainty what the students know, how much they know, and whether they are capable of using the knowledge in real-life situations, as they do not require the students to construct the answers [1]. Directly answered questions like VSAQ is an alternative, which will reveal more of students' competence. Factual knowledge, important for a doctor [15] was effectively tested by our VSAQ tests. Higher taxonomy testing is also said to be better achieved in VSAQ [1]. VSAQs are easier to construct [12]. The struggle of writing MCQ was reflected in our academics' feedback. Most of them considered writing good MTF and 3 to 4 plausible distractors in BAQ a difficult task, just as described in the literature [21]. Our students' feedback highlighted the issue of confusing items in MCQ, while no such opinions were expressed about VSAQ. The requirement to construct the answers would drive better learning, as preparation strategies would change [1, 3, 10]. The motivation for learning will be better when students are required to express their knowledge in the examination [1]. Direct questions let students express their ideas [1, 19], which is a desired purpose of assessments. MCQ can be considered an assessment of convenience far from real-life medical practice, while direct questions are considered more natural [2].

In the past, when essay questions were used, the onus of constructing the answers was on students. With the advent of MCQ, the roles reversed. Now, the teachers need to cover the syllabus in questions, while the students do not need to write a single word, but to choose the correct answers either knowingly, by guessing or using test-wiseness. In effect, a switch from MCQ to VSAQ will transfer the burden of writing all the answers (true, false and distractors) from the teachers back to the students. Constructing questions will still be an expert job, albeit less arduous than MCQ, as it won't require to create false options and plausible distractors. The only disadvantage envisaged in adopting VSAQ is the time spent on manual scoring. The time spent on constructing MCQ can now be utilised for marking scripts with more rewarding benefits. VSAQ can be scored by anyone with the help of model answers [7, 12], and the possibility of computer technology [3] making the task easier in future is real. Our academics and students responded positively to a shift from MCQ to VSAQ. The demerits of MCQ are many, and there is strong backing for VSAQ, but the transition, as usual, is slow to come.

Limitation of the study

This study was limited to one institution and the student scores and comparisons were limited to one batch of students each of year-3 and year-5.

Conclusion

VSAQ employ directly asked questions, which require students to answer briefly and directly with no scope of guessing or reaching the answer by elimination. A larger quantity of knowledge recall, understanding and application can be tested in a shorter period compared to MCQ. The information obtained by the teachers while scoring will help to modify the teaching and to give feedback to the students. Previously used MTF and BAQ items can be converted to VSAQ effectively. Our study showed fairer student performance and better psychometric indexes in VSAQ compared to MTF and BAQ. The students by and large preferred VSAQ to MTF and BAQ. Our academics also expressed the deficiencies of MCQ and showed openness to VSAQ. Last but not the least VSAQ using computers will cut the costs and the carbon footprint of the faculty drastically. In conclusion, VSAQ will be a viable alternative to MCQ with many plus points.

Abbreviations

BAQ	Best answer questions
DIFi	Difficulty index
DISi	Discrimination index
IRB	Institutional Review Board
MCQ	Multiple Choice Questions
MTF	Multiple true-false questions

VASQ Very Short Answer Questions

UNIMAS Universiti Malaysia Sarawak

Declarations

Acknowledgements

Our thanks to the dean and deputy deans of the Faculty of Medicine and Health Sciences, UNIMAS for the approval of the project and permission to use the examination questions, results and item analysis reports for this project.

Ethics approval

The Technical Review Committee approved the study proposal of the Faculty of Medicine and Health Sciences, Universiti Malaysia of Sarawak (UNIMAS). Ethical clearance was also obtained from the Institutional Review Board (IRB) of the Faculty. Strict anonymity and confidentiality of the data were maintained. No student particulars were extracted and disclosed publicly. Students' participation in this study was voluntary. The participants signed a formal consent form as per the faculty's ethics committee requirements. The scores of these tests were not used for any formal assessment.

Consent to publish

Both authors read the manuscript and consented for publication

The data and materials

The data will not be available publicly. The principal author keeps all the materials and data used in this study. However, the printed questionnaire for this could be available with permission from the first author.

Conflicts of interest

There is no conflict of interest involved in this project.

Funding

Self-funded

Authors' contributions

TP conceptualised the study, captured data and wrote the manuscript, MMR did the statistics, proofreading and editing of the manuscript.

References

1. Cirm JT. True/False versus Short Answer Questions. *Coll Teach*. 1986;34:34–7.
<https://www.jstor.org/stable/27558154>. Accessed 10 Jan 2020.
2. Pickering SG. Against Multiple Choice Questions. *Med Teach*. 1979;1:84–6.
doi:10.3109/01421597909019397.
3. Sam AH, Hameed S, Harris J, Meeran K. Validity of very short answer versus single best answer questions for undergraduate assessment. *BMC Med Educ*. 2016;16:266.
4. van der Vleuten CPM, Schuwirth LWT. Assessing professional competence: from methods to programmes. *Med Educ*. 2005;39:309–17.
5. Burton RF. Multiple-choice and true/false tests: myths and misapprehensions. *Assess Eval High Educ*. 2005;30:65–72. doi:10.1080/0260293042003243904.
6. Case SM, Swanson DB. Constructing Written Test Questions For the Basic and Clinical Sciences. *Natl Board Med Exam*. 2001;:181. http://www.medbev.umontreal.ca/docimo/DocSource/NBME_MCQ.pdf.
7. Hauer KE, Boscardin C, Brenner JM, van Schaik SM, Papp KK. Twelve tips for assessing medical knowledge with open-ended questions: Designing constructed response examinations in medical education. *Med Teach*. 2019;:1–6.
8. Martinez ME, Lipson JI. Assessment for Learning. *Educ Leadersh*. 1989;46:73–5.
9. Van Der Vleuten CPM, Schuwirth LWT, Driessen EW, Govaerts MJB, Heeneman S. Twelve Tips for programmatic assessment. *Med Teach*. 2015;37:641–6.
10. Wood T. Assessment not only drives learning, it may also help learning. *Med Educ*. 2009;43:5–6.
11. Wass V, Van der Vleuten C, Shatzer J, Jones R. Assessment of clinical competence. *Lancet Lond Engl*. 2001;357:945–9.
12. Chan CKY. Assessment: Short answer question, assessment resources,. 2009.
https://ar.cetl.hku.hk/am_saq.htm. Accessed 12 Jan 2020.
13. McCoubrie P. Improving the fairness of multiple-choice questions: a literature review. *Med Teach*. 2004;26:709–12.
14. Sim S-M, Rasiyah RI. Relationship between item difficulty and discrimination indices in true/false-type multiple choice questions of a para-clinical multidisciplinary paper. *Ann Acad Med Singapore*. 2006;35:67–71.
15. Anderson J. For Multiple Choice Questions. *Med Teach*. 1979;1:37–42.
doi:10.3109/01421597909010580.
16. Dutke S, Barenberg J. Easy and Informative: Using Confidence-Weighted True–False Items for Knowledge Tests in Psychology Courses. *Psychol Learn Teach*. 2015;14:250–9.
doi:10.1177/1475725715605627.
17. Puthiaparampil T. Assessment Analysis: How It Is Done. *MedEdPublish*. 2017;6.
doi:10.15694/mep.2017.000142.
18. Hift RJ. Should essays and other “open-ended”-type questions retain a place in written summative assessment in clinical medicine? *BMC Med Educ*. 2014;14:249. doi:10.1186/s12909-014-0249-2.
19. Stone HL. Return to “basics” in medical education: a commentary. *Med Teach*. 1982;4:102–3.

20. Puthiaparampil T, Gudum HR, Rahman MM, Saimon R, Lim IF. True-false analysis reveals inherent flaws in multiple true-false tests. *Int J Community Med Public Health*. 2019;6:4204–8. doi:10.18203/2394-6040.ijcmph20194476.

21. Tarrant M, Ware J, Mohammed AM. An assessment of functioning and non-functioning distractors in multiple-choice questions: a descriptive analysis. *BMC Med Educ*. 2009;9:40. doi:10.1186/1472-6920-9-40.

22. Case SM, Swanson DB. *Constructing Written Test Questions For the Basic and Clinical Sciences*. 1998;:181.

Figures

Figure 1

Overall Student performance in 6 tests compared. p-value reached from Person’s Chi-square test of independence MTF vs. VSAQ (p<0.05) and BAQ vs. VSAQ (p<0.05)

Figure 2

Percentage distribution of difficulty index of 6 tests .

Figure 3

Percentage distribution of discrimination index of 6 tests