# Comparison between MCQ, Extended matching questions (EMQ) and Script concordance test (SCT) for assessment among first-year medical students – A pilot study

Nalini YC, Shivashakthy Manivasakan[1], Dinker R. Pai[2]

PhD Scholar, Centre for Health Professions Education (CHPE), Department of Physiology, Mahatma Gandhi Medical College Research Institute, Sri Balaji Vidyapeeth (Deemed to be University), Pondicherry – Cuddalore Road, Pillayarkuppam, Puducherry, India, [1]Department of Prosthodontics, Deputy Director of Institute of Health Professions Education (CHPE), Indira Gandhi Institute of Dental Sciences, Sri Balaji Vidyapeeth (Deemed to be University), Pondicherry – Cuddalore Road, Pillayarkuppam, Puducherry, India, [2]Director in Medical Simulation Centre & Professor in Department of Surgery, Sri Balaji Vidyapeeth (Deemed to be University), Pondicherry – Cuddalore Road, Pillayarkuppam, Puducherry, India

**Address for correspondence:**
Dr. Nalini YC,
PhD Scholar, Institute of Health Professions Education (CHPE), Department of Physiology, Mahatma Gandhi Medical College Research Institute, Sri Balaji Vidyapeeth (Deemed to be University), Pondicherry – Cuddalore Road Pillayarkuppam, Puducherry, India.
E-mail: drnaliniyc@gmail.com

**Abstract:**

**BACKGROUND:** Single-best response types of MCQs are commonly used tools in medical assessment. However, these are not particularly apt for the assessment of higher-order thinking skills (HOTS) among students. Assessment of HOTS and clinical reasoning skills requires unexplored tools like extended matching questions (EMQs) and SCTs. The aim is to assess HOTS and clinical reasoning skills during formative assessment among medical students post a simulation-based education (SBE) using EMQ and SCT on the topic of shock and collect student perceptions regarding new assessment tools.

**METHODS AND MATERIAL:** The research is an observational descriptive study. Fifty-two first-year medical students were asked to take a formative assessment consisting of 20 MCQs, 6 EMQs, and 2 SCT post a SBE during July 2022 on the topic of shock. MCQs were categorized into themes of aetiology, pathophysiology, and management of shock. These categorized MCQs were compared and analyzed with EMQs and SCTs prepared on the same themes. The data analysis by a one-way analysis of variance (ANOVA) was conducted to compare the difference in per cent mean scores of MCQ with EMQ and MCQ with SCT.

**RESULTS:** The overall mean scores and also under the theme of aetiology and management, the student's score in MCQ were higher than EMQ and SCT and the difference was statistically significant with the $P$ value ($\leq 0.001$).

**CONCLUSIONS:** Students scored better in familiar assessment tools like MCQ, but majority of the students were of the opinion that EMQ tested the delivered content better and SCT tested the clinical application better.

**Keywords:**

Clinical competence, education, medical students, shock, simulation training

## Introduction

The use of assessment tools in the medical curriculum is based on the domains of the assessment. Various written formats assess the cognitive domain, including multiple choice questions (MCQs) and viva voce, while OSPE/OSCE or directly observed procedural skills (DOPS) are utilized for psychomotor domain assessment.[1,2] Single-best response type of MCQs are commonly used tools in medical education, but the assessment of higher-order thinking skills (HOTS) among students with these routinely employed assessment tools is difficult.[3] Importance to assessment of clinical reasoning skills is given more during internship and post-graduation than when compared to

**How to cite this article:** Nalini YC, Manivasakan S, Pai DR. Comparison between MCQ, Extended matching questions (EMQ) and Script concordance test (SCT) for assessment among first-year medical students – A pilot study. J Edu Health Promot 2024;13:52.

under graduation.[4] Few studies have reported that clinical reasoning skills among final-year students are not thoroughly assessed. The word "clinical reasoning" here is understood as "the cognitive process whereby the information contained in a clinician's knowledge and experience and used to diagnose and manage the patient's problem."[5] It is also said that the introduction of clinical reasoning skills early in their curriculum results in better development of these skills in these students. Extended matching items or questions are a type of MCQs, they have four components, a *theme* (for ex . aetiology of shock), *lead in* (provides instructions to students for ex . for each clinical scenario listed below choose the most likely diagnosis), *questions* describing the clinical vignettes, and lastly 9–26 *potential answers*.[6] Script concordance tests (SCTs) are used to assess clinical reasoning in ambiguous or uncertain situations. It probes the multiple judgments that are made in the clinical reasoning process. Scoring reflects the degree of concordance of these judgments to those of a panel of reference experts.[7,8] In the current Indian undergraduate medical curriculum, there is an introduction of clinical content as part of early clinical exposure to bridge the understanding but our routine assessment tools are not designed for the same. As part of early clinical exposure after a teaching-learning intervention on the physiological mechanism of different types of shock, a formative assessment consisting of MCQs, extended matching questions (EMQs), and SCT was conducted to assess higher-order thinking and clinical reasoning skills among first-year physiology students. The aim of the study was to ascertain the feasibility of using the assessment tools EMQ/SCT and compare the test score of the students' MCQ with that of EMQ and SCT on the theme of aetiology, management, and pathophysiology of shock and collect student perception regarding new assessment tools. Currently, there is limited literature on utilization of these assessment tools in the pre-clinical phase among undergraduates.

## Materials and Methods

### Study design and setting

This is an observational descriptive pilot study conducted at a medical college in South Eastern part of India. The institute is a university college engaged in teaching undergraduate and postgraduate medical as well as paramedical students for more than 20 years. This pilot study was a part of PhD project on simulation-based teaching of clinically relevant topics in the cardiovascular system during their early clinical exposure hours for first-year medical students. A teaching–learning intervention with provision for student feedback based on physiological mechanisms of different types of shock was designed and conducted in the medical simulation center of our university.

### Study participants and sampling

The study included a sampling frame of a batch of 60 students who were given an open invitation by a formal announcement in the class on the previous day. Fifty-two students participated in the study by giving informed consent. The enrolled students were categorized as high (the top 26 students) and low scorers (the bottom 26 students) based on the previous cumulative scores in the assessments conducted by the department.

### Data collection tool and technique
#### Teaching learning intervention
The students were exposed to teaching learning intervention of shock consisting of a mannequin and screen simulation of the deranged physiological mechanisms, along with clinical scenarios discussion of hypovolemic, obstructive, cardiogenic, and neurogenic shock.

#### Process of preparation of assessment tools
Twenty MCQs, two sets of EMQs (each consisting of 3 questions), and 2 SCTs were prepared on the topic of shock by subject experts. SCT was prepared according to the Fournier and colleagues' guidelines on how to write SCT tests.[9] Validation of the assessment tools was done by one external subject expert and two medical education experts.

Sensitization to new assessments using simple real-life examples and analogy tools was done to the students 15 days prior to the planned intervention.

Following this teaching intervention, a formative test using google forms was conducted among the students consisting of 20 MCQs, 6 EMQs, and 2 SCTs (with 3 subsections each). Twenty-five minutes were provided to the students to complete the assessment. Time allotted for each MCQ was 40 seconds, 60 seconds for each EMQ, and 5 minutes for one SCT.

#### Categorization of themes: (modified Delphi method)
Two set of examiners independently categorised the MCQ under the themes of aetiology, pathophysiology and management of shock. The results were compared and differences in the categorization of questions were finalized after consensus among the subject experts.[10] MCQs were categorized into themes as shown in Table 1. The scores of these categorized MCQs were compared and analyzed with those of EMQs and SCTs prepared on the same themes.

#### Student feedback
Student feedback was obtained regarding the assessment tools used during this TLI through a self-administered structured questionnaire via an anonymous online survey. The questionnaire was prepared and validated through a pilot study, on a sample of students (10 in

number) who were not part of the TLI. They were asked to fill the questionnaire, and based on the inputs obtained, modifications were done to questions 8 and 9. The opinion of a medical education expert was obtained before administering the questionnaire.[11]

### Data collection and analysis

All the data were downloaded in Excel from the Google form repository. The data was analyzed using SPSS software for Windows version 20 (SPSS IBM Corp.Ltd. Armonk, NY, USA). The categorical data were represented as frequencies and percentages. The continuous data were summarized as mean with standard deviation. The comparison of the mean scores between groups was performed using one-way ANOVA. All the statistics were carried out at a 95% confidence interval and a *P* value of less than 0.05 was considered significant.

### Ethical consideration

According to the Declaration of Helsinki, institutional ethics approval was obtained for this teaching-learning intervention (TLI) vide infra letter no. MGMCRI/RAC/2021/02/IHEC/10. A signed informed consent was obtained from all the participants as per the recommendations of the Ethics Committee. The students were given the option to withdraw from the study at any time during the intervention and course of the study.

## Results

The overall mean scores obtained by the students in MCQ were higher than EMQ and SCT, and the difference was statistically significant with the *P* value of 0.0018 and 0.00325 as shown in Table 2.

On comparison of scores under the theme of aetiology, mean scores obtained in MCQ were higher than EMQ and SCT, and the difference was statistically significant as shown in Table 3. Under the theme of management, the mean scores in MCQ were higher than EMQ the difference was statistically significant with a *P* value of 0.001 [Table 4].

On comparing the mean scores of various assessment tools between the high and low scorers, it is observed that mean scores are high among high scorers when compared to low scorers thereby validating the tools. The difference obtained is statistically significant in EMQ alone with a *P* value of 0.01 [Table 5].

We looked at the Pearson Coefficient between various assessment tools used in the study and found there was a moderate correlation between MCQ and EMQ only (r = 0.46) [Table 6].

**Table 1: Categorization of 20 MCQs under various themes**

| Themes | Question number |
|---|---|
| Aetiology | 1,12,15,16 |
| Pathophysiology | 2,3,4,5,8,9,10,11,13,18,19,20 |
| Management of shock | 6,7,14 |

These categorized MCQs were compared and analyzed with EMQs and SCTs prepared on the same themes. Categorization of themes done by modified Delphi method

**Table 2: Mean score comparison of MCQ with EMQ and SCT (*n*=52)**

| Assessment tool | Mean±SD | One way ANOVA |
|---|---|---|
| MCQ | 55.6±14.6 | EMQ: MCQ |
| EMQ | 41.9±26 | Q=4.94 (*P*=0.0018) |
| SCT | 42.6±17.3 | SCT; MCQ Q=4.69 (*P*=0.00325) |

The overall mean scores in MCQ were higher than EMQ and SCT, and the difference was statistically significant with the *P* of 0.0018 and 0.00325, respectively

**Table 3: Comparison of mean scores between tools (Theme:Aetiology)**

| Parameter | Mean±SD | Std. error of mean | ANOVA | *P* df=2 |
|---|---|---|---|---|
| MCQ (aetiology) | 61.9±19.5 | 2.7 | *F*=7.4 | *P*<0.001 |
| EMQ (aetiology) | 53.9±35.7 | 5.0 | | |
| SCT (aetiology) Std. deviation | 42.7±17.2 | 2.4 | | |

The mean scores in MCQs were higher than EMQs and SCTs, and the difference was statistically significant. *df - degree of freedom

**Table 4: Comparison of mean scores between tools (Theme:Management)**

| Parameter | Mean±Std. deviation | Std. error of mean | ANOVA |
|---|---|---|---|
| MCQ (management) | 56.4±29.3 | 4.1 | *F*=20.6 (*P*<0.001) |
| EMQ (management) | 30.1±29.7 | 4.1 | |

The mean scores in MCQ were higher than EMQ; the difference was statistically significant (*P*<0.001)

**Table 5: Comparison of mean scores of different assessment tools between high and low scorers**

| Assessment tools | High scorers | Low scorers | *P* |
|---|---|---|---|
| MCQ | 58.9 | 52.2 | 0.10 |
| EMQ | 50 | 33.3 | ≤0.01 |
| SCT | 45 | 40.1 | 0.31 |

Mean scores are high among high scores when compared to low scorers. The statistically significant difference was seen in EMQ alone (*P*≤0.01)

**Table 6: Corelation coefficient for the various tools for similar assessment**

| Pearson Coefficient | *r* |
|---|---|
| MCQ and EMQ | 0.46 |
| MCQ and SCT | 0.15 |
| EMQ and SCT | 0.34 |

There was moderate correlation between MCQ and EMQ only (*r*=0.46)

71.4% of students said that MCQs were more familiar in comparison with other assessment tools (EMQs and SCTs), and 57.1% reported that they are easy to understand and answer. 57.1% of students thought

EMQs tested the delivered content better, while 71.4% were of the opinion that SCTs tested the clinical application better. [Table 7]

## Discussion

We wanted to use relatively unexplored tools like EMQ and SCT during formative assessment to assess the HOTS post a simulation-based intervention on the physiological mechanism of different types of shock and compare the scores between familiar, conventionally used tools and unfamiliar new assessment tools. We also wanted to collect student perceptions of these new assessment tools through a questionnaire.

In our study, students performed better in MCQs when compared to EMQs, as this form of assessment has been familiar. Majority of students reported that MCQs were familiar in comparison to other tools, and also, they were easy to understand and answer. A similar finding is observed in a study conducted on dental undergraduates' students in Pakistan where mean scores obtained by the students were more in MCQ when compared to EMQ.[12]

Among high- and low-scoring groups, means scores were more in all the assessment tools (MCQ, EMQ, and SCT) in the high scoring group but a statistically significant difference is observed in EMQ which is in contrast with the findings of the study by medical faculty at Dutch university which states that MCQ is superior to EMQ in differentiating a high with a low scorer.[6]

The face validity of the examination appeared to be good as students felt that the EMQ tested the delivered content better. A similar finding has been observed in a study done among veterinary students when EMQs were used to assess clinical reasoning skills.[13]

Time allotted for each single response of MCQ was 40 sec while for each EMQ was 100 sec as EMQs require more reading and processing time in concordance with various guidelines and medical literature,[14] but in our study majority of the students felt that the time allotted for EMQ is insufficient.

Since the teaching–learning intervention was a case-based discussion on different types of shock for preclinical medical students, SCT was incorporated as one of the tools for the assessment of clinical reasoning skills. In our study, majority of the students reported that SCT assessed the clinical scenarios better but also expressed that compared to MCQs, EMQ and SCT are challenging to understand and answer. Concordant findings were reported from a study conducted on final-year undergraduates of the School of Veterinary Medicine and Science (SVMS) at the University of Nottingham.[15]

Majority of the students did not want EMQ and SCT to be introduced in their university examinations; one reason for this observation is that these assessment tools are new and unfamiliar when compared to MCQs which they are introduced to from school and pre-university days. Similar observation has been noted by Beullens *et al*.[16] who found that their medical students expressed reservations about the the use of EMQ format in the final examination.

Even though newer assessment tools tested the delivered content and clinical reasoning skills among the undergraduates better than the conventional MCQs, students are reluctant to embrace the new tools. These new tools should be exposed to students during the formative assessment more frequently right from the preclinical phase till internship for better acceptance of these assessment tools among the student community during summative or university examinations.

### Limitations and recommendation
Firstly, categorizing MCQ as retention, understanding, and application questions was not possible as designed

## Table 7: Comparison of feedback responses of the participants for the various assessment tools (*n*=52)

| Questions | MCQ *n* (%) | EMQ *n* (%) | SCT *n* (%) |
|---|---|---|---|
| 1.The following assessment modality was easy to understand | 31 (59.6) | 7 (14.3) | 14 (26.9) |
| 2. The following assessment modality was easy to answer | 30 (57.1) | 15 (28.6) | 7 (14.3) |
| 3. The following assessment modality was challenging to answer | 8 (15.3) | 22 (42.9) | 22 (42.9) |
| 4. The following assessment modality was challenging to understand | 0 (0) | 34 (65.3) | 18 (34.6) |
| 5. The following assessment modality was familiar | 38 (71.4) | 7 (14.3) | 7 (14.3) |
| 6. The following assessment modality tested the delivered content better | 15 (28.6) | 30 (57.1) | 7 (14.3) |
| 7. The following assessment modality tested the clinical application better | 7 (14.3) | 7 (14.3) | 38 (71.4) |
| 8. The following assessment modality can be used in formative assessment (where feedback is given, strength and weakness of the student identified) | 30 (57.1) | 7 (14.3) | 15 (28.6) |
| 9. The following assessment modality can be used in summative assessment (including university examination) | 40 (76.9) | 7 (14.3) | 5 (9.6) |
| 10. The following assessment has sufficient allotment of time | | | |
| Yes (%) | 37 (71.4) | 15 (28.6) | 15 (28.6) |
| No (%) | 15 (28.6) | 37 (71.4) | 37 (71.4) |

MCQ were fitting only the retention and understanding category. Secondly, we could not design a SCT based on management appropriate to first-year students.

## Conclusions

Students scored better in familiar assessment tools like MCQ when compared to unfamiliar tools like EMQ and SCT as they found MCQ was easy to understand and answer. Majority of the students were of the opinion that EMQ tested the delivered content better and SCT tested the clinical application better.

### Suggestions/further directions
New assessment tools tailored to assess the relatively new and unconventional components of teaching like clinical reasoning, communication, group dynamics, and leadership in pre-clinical phase should be explored by the assessors in formative assessments. if not in summative or university examinations.

### Financial support and sponsorship
Nil.

### Conflicts of interest
There are no conflicts of interest.

## References

1. Badyal DK, Sharma M. Internal assessment in New MBBS curriculum: Methods and logistics. Int J Appl Basic Med Res 2020;10:68-75.
2. Schuwirth LW, van der Vleuten CP. ABC of learning and teaching in medicine: Written assessment. BMJ 2003;326:643-5.
3. Asani M. Assessment methods in undergraduate medical schools. Niger J Basic Clin Sci 2012;9:53-60.
4. Nazim SM, Talati JJ, Pinjani S, Biyabani SR, Ather MH, Norcini JJ. Assessing clinical reasoning skills using Script Concordance Test (SCT) and extended matching questions (EMQs): A pilot for urology trainees. J Adv Med Educ Prof 2019;7:7-13.
5. Newble D, Norman G, Van der Vleuten C. Assessing clinical reasoning. In: Jones JHM, editor. Clinical Reasoning in the Health Professions. 2nd ed. Oxford: Butterworth Heinemann; 2000. p. 156-68.
6. Eijsvogels TMH, van den Brand TL, Hopman MTE. Multiple choice questions are superior to extended matching questions to identify medicine and biomedical sciences students who perform poorly. Perspect Med Educ 2013;2:252-63.
7. Charlin B. Standardized Assessment of Ill-Defined Clinical Problems: The Script Concordance Test. Maastricht: Universitaire Pers Maastricht; 2002. p. 13.
8. Cobb KA, Brown G, Hammond R, Mossop LH. Students' perceptions of the script concordance test and its impact on their learning behavior: A mixed methods study. J Vet Med Educ 2015;42:45-52.
9. Fournier JP, Demeester A, Charlin B. Script concordance tests: Guidelines for construction. BMC Med Inform Decis Mak 2008;8:18.
10. Eubank BH, Mohtadi NG, Lafave MR, Wiley JP, Bois AJ, Boorman RS, *et al*. Using the modified Delphi method to establish clinical consensus for the diagnosis and treatment of patients with rotator cuff pathology. BMC Med Res Methodol 2016;16:56.
11. Aithal A, Aithal PS. Development and validation of survey questionnaire and experimental data–a systematical review-based statistical approach. Int J Manag Technol Soc Sci (IJMTS) 2020;5:233-51.
12. Shehzad S, Kabir SK, Waheed Z, Durrani SH, Maqsood F, Khan K. Comparison of extended matching questions (EMQ) with multiple choice questions (MCQ) in undergraduate dental studies. Pak Oral Dent J 2021;41:84-6.
13. Tomlin JL, Pead MJ, May SA. Veterinary students' attitudes toward the assessment of clinical reasoning using extended matching questions. J Vet Med Educ 2008;35:612-21.
14. Frey A, Leutritz T, Backhaus J, Hörnlein A, König S. Item format statistics and readability of extended matching questions as an effective tool to assess medical students. Sci Rep 2022;12:20982.
15. Duggan P, Charlin B. Summative assessment of 5th year medical students› clinical reasoning by script concordance test: Requirements and challenges. BMC Med Educ 2012;12:29.
16. Beullens J, Van Damme B, Jaspaert H, Janssen PJ. Are extended-matching multiple-choice items appropriate for a final test in medical education? Med Teach 2002;24:390-5.