

# Development of a New Scoring System To Accurately Estimate Learning Outcome Achievements via Single, Best-Answer, Multiple-Choice Questions for Preclinical Students in a Medical Microbiology Course †

Yodying Dangprapai<sup>1</sup>, Popchai Ngamskulrungraj<sup>2\*</sup>, Sansnee Senawong<sup>3</sup>,  
Patompong Ungprasert<sup>4</sup>, and Azian Harun<sup>5</sup>

<sup>1</sup>Department of Physiology, Faculty of Medicine, Siriraj Hospital, Mahidol University, Bangkok 10700, Thailand;

<sup>2</sup>Department of Microbiology, Faculty of Medicine, Siriraj Hospital, Mahidol University, Bangkok 10700, Thailand;

<sup>3</sup>Department of Immunology, Faculty of Medicine, Siriraj Hospital, Mahidol University, Bangkok 10700, Thailand;

<sup>4</sup>Clinical Epidemiology Unit, Department of Research and Development, Faculty of Medicine,  
Siriraj Hospital, Mahidol University, Bangkok 10700, Thailand;

<sup>5</sup>Department of Medical Microbiology and Parasitology, School of Medical Sciences,  
Universiti Sains Malaysia, 16150 Kubang Kerian, Kelantan, Malaysia

During the preclinical years, single-best-answer multiple-choice questions (SBA-MCQs) are often used to test the higher-order cognitive processes of medical students (such as application and analysis) while simultaneously assessing lower-order processes (like knowledge and comprehension). Consequently, it can be difficult to pinpoint which learning outcome has been achieved or needs improvement. We developed a new scoring system for SBA-MCQs using a step-by-step methodology to evaluate each learning outcome independently. Enrolled in this study were third-year medical students ( $n = 316$ ) who had registered in the basic microbiology course at the Faculty of Medicine, Siriraj Hospital, Mahidol University during the academic year 2017. A step-by-step SBA-MCQ with a new scoring system was created and used as a tool to evaluate the validity of the traditional SBA-MCQs that assess two separate outcomes simultaneously. The scores for the two methods, in percentages, were compared using two different questions (SBA-MCQ1 and SBA-MCQ2). SBA-MCQ1 tested the students' knowledge of the causative agent of a specific infectious disease and the basic characteristics of the microorganism, while SBA-MCQ2 tested their knowledge of the causative agent of a specific infectious disease and the pathogenic mechanism of the microorganism. The mean score obtained with the traditional SBA-MCQs was significantly lower than that obtained with the step-by-step SBA-MCQs (85.9% for the traditional approach versus 90.9% for step-by-step SBA-MCQ1;  $p < 0.001$ ; and 81.5% for the traditional system versus 87.4% for step-by-step SBA-MCQ2;  $p < 0.001$ ). Moreover, 65.8% and 87.8% of the students scored lower with the traditional SBA-MCQ1 and the traditional SBA-MCQ2, respectively, than with the corresponding sets of step-by-step SBA-MCQ questions. These results suggest that traditional SBA-MCQ scores need to be interpreted with caution because they have the potential to underestimate the learning achievement of students. Therefore, the step-by-step SBA-MCQ is preferable to the traditional SBA-MCQs and is recommended for use in examinations during the preclinical years.

## INTRODUCTION

Many types of written assessments have been used to evaluate whether medical students have obtained adequate knowledge (1) or achieved expected learning outcomes,

according to Bloom's taxonomy from level of knowledge to evaluation (2). Knowledge and comprehension have been classified as lower-order cognitive skills, while analysis, synthesis, and evaluation have been classified as higher-order cognitive skills, with application as a transition level (3). In general, constructed-response or essay-style test questions are required for assessments of higher-order cognition (4, 5). However, empirical evidence suggests that measuring complex cognitive processes by means of multiple-choice questions (MCQs) is possible with carefully-constructed questions (6). In fact, well-structured MCQs based on the cognitive levels of Bloom's taxonomy could be used to assess higher-order cognitive skills, except at the level of synthesis (3, 7, 8).

\*Corresponding author. Mailing address: Department of Microbiology, Faculty of Medicine, Siriraj Hospital, Mahidol University, 2 Wang Lang Rd., Siriraj, Bangkok Noi, Bangkok 10700, Thailand. Phone: +66 2 419 7053. Fax: +66 2 418 4148.

E-mail: [popchai.nga@mahidol.ac.th](mailto:popchai.nga@mahidol.ac.th).

Received: 18 February 2019, Accepted: 20 November 2019,

Published: 28 February 2020

†Supplemental materials available at <http://asmscience.org/jmbe>

MCQs are one of the most popular tools used by most medical schools around the world. Although MCQs permit sufficient coverage of learning outcomes, automated computer marking, and feedback for students, they can contain flaws that may provide clues to the correct answer for examination-wise candidates, or they may possess irrelevant difficulties (9). Three types of MCQ have been used, namely, true/false (TF), single best-answer (SBA) questions, and extended-matching questions (EMQs). Each type has different purposes, advantages, and drawbacks (9, 10).

Typically, the structure of each SBA-MCQ item consists of a stem, which is usually a clinical scenario, and a lead-in question (9, 11). To test taxonomically higher-order cognitive processes, a complex situation has to be used as the clinical scenario. Students need to be able to combine several aspects of their knowledge, including basic medical science and clinical practice, to answer the question correctly (12). On the other hand, TF-MCQs can only be used to assess the ability of students to recall factual knowledge (10). In the case of EMQ-MCQs, students are required to choose one or more correct answers from a “set” of options associated with a particular theme (10). EMQ-MCQs can also be used to assess higher-order cognitive processes. Regardless of its type, each MCQ item should be written to assess a specific course learning outcome in order to provide effective feedback for student learning (13–15). Furthermore, tests for upper-level courses targeting higher-order cognitive skills should focus on the cognitive levels of analysis and above, while tests for entry-level courses targeting foundational knowledge may contain higher numbers of questions targeting the cognitive levels of requisite knowledge and comprehension (16).

Although guidelines for effective MCQ writing are available (11, 17), item-writing flaws are common among MCQs from various disciplines (18, 19). Flaws related to irrelevant difficulty and test wiseness are rather common among MCQs in medical schools, and they mostly lie within the MCQ stems and options (11). While the former flaw (stems) challenges students for unrelated learning objectives, the latter flaw (options) gives cues to correct answers based on test-taking skills alone without necessarily achieving objectives. For example, two MCQ studies in Pakistan revealed that impossible distractors, extra details in correct options, and unfocused stems were the most common flaws in their MCQs (20, 21). However, as most studies determined their MCQ quality based on the guidelines of the US National Board of Medical Examiners, which focus mainly on stems and options (11), studies of the flaws specifically within lead-in questions are rare. In fact, only one study has directly assessed lead-in question flaws. Interestingly, more flaws were found in the lead-in questions than in the stems and options (22). Lead-in question flaws might therefore be more common than previously believed.

In Thailand, MCQs are used by all medical schools to assess the achievement of learning outcomes by students,

and they are used for the Medical Licensing Examination of Thailand (23). During the clinical years of study, SBA-MCQs are the most frequently employed method of assessment since they enable the assessment of higher-order cognitive processes (including the interpretation and application of knowledge) and problem solving (9, 10). However, during the preclinical years, when the majority of course learning outcomes target the cognitive levels from knowledge to simple analysis, the stems of the SBA-MCQs have to be more straightforward, as students have not yet gained sufficient knowledge and experience to formulate correct diagnoses from clinical scenarios. Stems frequently need to promptly provide the diagnosis. The complexity of the traditional SBA-MCQ therefore lies within the lead-in questions (7, 24). Although it is clearly stated in many standard guidelines that only one concept should be tested at a time, this requirement is mentioned specifically only for stems and options (11, 25). As a consequence, in Thailand, two or more outcomes are tested simultaneously by the lead-in questions of MCQs to assess higher-order cognitive processes.

Two examples of the traditional SBA-MCQ for the preclinical years in Thailand are presented in Figure 1. In the example provided of a microbiology-specific SBA-MCQ (Figs. 1A and 1C), knowledge of the microbiologic etiology of impetigo is needed (Outcome 1, Figs. 1B and 1D). Then, knowledge of the morphology of the etiological agent for SBA-MCQ1 (Outcome 2, Fig. 1B) or the virulence factors of this organism for SBA-MCQ2 (Outcome 3, Fig. 1D) are evaluated. In order to achieve a score for both SBA-MCQ 1 and 2, a student needs to achieve not only Outcome 2 or Outcome 3, respectively, but also Outcome 1. In other words, the evaluations of Outcomes 2 and 3 are “dependent” on a student’s ability to achieve Outcome 1 (Possibility A, Table 1). To illustrate, despite having knowledge of Outcome 2, a student will receive a score of zero for the SBA-MCQ1 if the student has no knowledge of Outcome 1 (Possibility E, Table 1). Similarly, a student will be given a score of zero for the SBA-MCQ1 if the same student has no knowledge of Outcome 1 (Possibility C, Table 1). Other possibilities of interpretation, with details, are demonstrated in Table 1. In this example scenario, it would be more accurate to give a score of 0.5 to the student for each achieved learning outcome, but this is not allowed under the scoring system used by the traditional SBA-MCQ. The validity of each test item is thus jeopardized, as more than one learning outcome is being assessed simultaneously.

Due to such potential flaws in the lead-in questions in Thailand, we proposed a new scoring system using a set of step-by-step SBA-MCQs (Fig. 2) to accurately assess each learning outcome of the basic medical microbiology course. We also demonstrated that the traditional SBA-MCQs underestimated the learning outcome achievements of preclinical students, and the new scoring system developed for step-by-step MCQs could improve the accuracy of the estimation of those achievements.

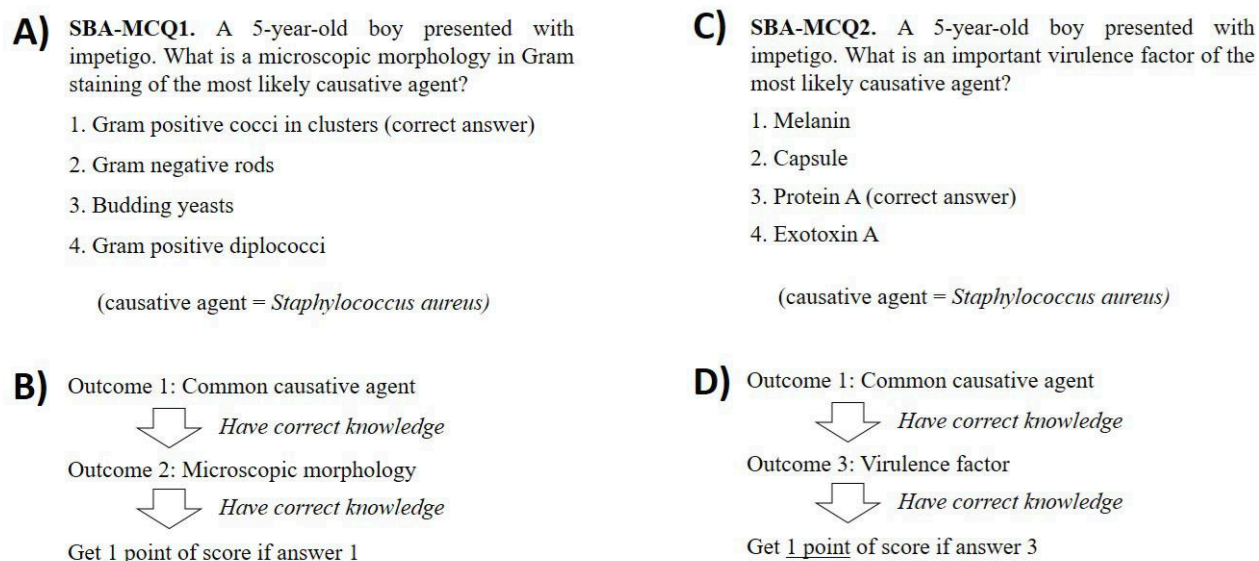


FIGURE 1. An example of the traditional SBA-MCQs used for preclinical-year students. A) SBA-MCQ1 asking two outcomes simultaneously (Outcome 1 and Outcome 2); C) SBA-MCQ2 asking Outcome 1 and Outcome 3 simultaneously; B) and D) To score 1 point from the SBA-MCQ1 or the SBA-MCQ2, a student must display correct knowledge for both outcomes. Correct knowledge for only one of the two outcomes will result in zero points. SBA-MCQ = single-best-answer multiple-choice question.

TABLE 1.  
Possible interpretations of the traditional SBA-MCQ1 scores (Figs. 1A and 1B).

Possibility	Scores (Answer)	Have Knowledge of Outcome 1?	Have Knowledge of Outcome 2?	Interpretation
A	1 point (1)	Yes (knows the correct causative agent of impetigo)	Yes (knows the correct morphology of <i>S. aureus</i> )	Has knowledge of both outcomes
B	0 point (other)	Yes (knows the correct causative agent of impetigo)	No (does not know the correct morphology of <i>S. aureus</i> )	Has knowledge only of Outcome 1
C	1 point (1)	No (does not know the correct causative agent of impetigo)	No (gets 1 point by random guessing)	No knowledge of any outcomes
D	0 point (other)	No (does not know the correct causative agent of impetigo)	No (does not know the correct morphology of any organism)	No knowledge of any outcomes
E	0 point (other)	No (does not know the correct causative agent of impetigo)	Yes (knows the correct morphology of the chosen organism*)	Has knowledge only of Outcome 2

\* For example (see also Fig. 2), a student may misunderstand that the causative agent of impetigo is *Cryptococcus neoformans* (fails to achieve Outcome 1), but knows the correct morphology of *C. neoformans* (selects choice 3 = achieves Outcome 2). SBA-MCQ = single-best-answer multiple-choice question.

## METHODS

### Structure of Thai medical curriculum

Our medical curriculum structure has been described in detail previously (26). Briefly, Thai medical students enter medical schools directly after finishing high school to study in a six-year, general medicine program. The second and third years (together termed the “preclinical years”) focus, respectively, on the foundational sciences of the normality and abnormality of human bodies. Each preclinical year is

divided into two parts: “general concepts” and “organ systems.” The first of these focuses on general principles, while the second centers on the application of concepts in order to understand human organ systems. The basic medical microbiology course is delivered in the general concept part of the third year. No real patients are involved in preclinical courses, except for the medical humanities course. The expected learning outcomes are based on the Medical Competency Assessment Criteria for National License 2012, established by the Medical Council of Thailand (27). The present study was approved by the Ethics Committee

	Step-by-step MCQ score	Traditional MCQ score
<b>Q1.</b> A 5-year-old boy presented with impetigo.		
What is the most likely causative agent (Outcome 1)?		
1. <i>Pseudomonas aeruginosa</i>	0	N/A*
2. <i>Cryptococcus neoformans</i>	0	N/A*
3. <i>Streptococcus pneumoniae</i>	0	N/A*
4. <i>Staphylococcus aureus</i>	1	N/A*
<b>Q2.</b> From ANS of Q1.		
What is its microscopic morphology in Gram staining (Outcome 2)?		
1. Gram positive cocci in clusters (correct answer of Q1-4)	1 if answer 4 for Q1**	1
2. Gram negative rods (correct answer of Q1-1)	1 if answer 1 for Q1**	0
3. Budding yeasts (correct answer of Q1-2)	1 if answer 2 for Q1**	0
4. Gram positive diplococci (correct answer of Q1-3)	1 if answer 3 for Q1**	0
<b>Q3.</b> From ANS of Q1. What is its virulence factor (Outcome 3)?		
1. Melanin (correct answer of Q1-2)	1 if answer 2 for Q1**	0
2. Capsule (correct answer of Q1-3)	1 if answer 3 for Q1**	0
3. Protein A (correct answer of Q1-4)	1 if answer 4 for Q1**	1
4. Exotoxin A (correct answer of Q1-1)	1 if answer 1 for Q1**	0

FIGURE 2. An example of a step-by-step SBA-MCQ and its scoring system. Representative scores when using the scoring system of the traditional SBA-MCQ are also presented for comparison. \*Not applicable as the traditional SBA-MCQ score only depends on the answer for Outcome 2 or Outcome 3; \*\*Score = 0 if the answer is something other than the options listed in the figure. SBA-MCQ = single-best-answer multiple-choice question.

of the Siriraj Institutional Review Board under certificate no. 289/2560 (Exempt). The study had no safety concerns as no real microorganisms were involved.

### Participants

This study was performed during the formative evaluation of third year medical students ( $n = 316$ ) in the basic medical microbiology course at the Faculty of Medicine, Siriraj Hospital, Mahidol University, Thailand, in the 2017 academic year.

### The new scoring system

We developed a new scoring system which can simultaneously estimate two learning outcomes and score each tested outcome separately. To minimize the confounding bias of knowledge variations among the students, all participants were asked to complete the same test, which had been written as multiple sets of SBA-MCQs. Each set was designed so that scores for both the traditional and step-by-step SBA-MCQ formats could be calculated and compared from the same test. Firstly, a set of three new SBA-MCQs, in the format of a four-optioned MCQ with one correct answer and three distractors, was created and named “step-by-step SBA-MCQs” (Fig. 2). While each traditional SBA-MCQ simultaneously assessed multiple learning outcomes, each new SBA-MCQ evaluated only one outcome. Each set of step-by-step MCQs was designed to evaluate the same learning outcomes as each traditional MCQ in Figure 1, namely, the causative agent of a specific infectious disease (Outcome 1), the basic characteristics

of the microorganism (Outcome 2), and the pathogenic mechanisms of the microorganism (Outcome 3). For each set of step-by-step SBA-MCQs (Fig. 2), the first question (Q1) tested Outcome 1. The second question (Q2) tested Outcome 2 specifically for the microorganism answered in Q1. Lastly, the third question (Q3) tested Outcome 3, related to the same microorganism from Q1. Nine sets (27 items; Appendix 1), four sets (12 items; items numbered 1–12 in Appendix 2), and five sets (15 items; items numbered 13–27 in Appendix 2) of the step-by-step SBA-MCQs were used to evaluate students' knowledge of medically-important bacteria, fungi, and viruses, respectively.

Moreover, the options for Q2 and Q3 were designed to ensure that the achievements of the second and third outcomes (Q2/Outcome 2 and Q3/Outcome 3) could be determined even when students chose distractors as their answers to the first question (Q1/Outcome 1). To illustrate, each of Q1's microorganism options had its corresponding compatible microscopic morphology (Outcome 2) and virulence factor (Outcome 3) listed as one of the options for Q2 and Q3, respectively. Consequently, the scores of Q2 and Q3 were able to be adjusted depending on each student's answer for Q1 in order to reflect their true learning achievement. For example, with reference to Figure 2, although choosing option 4 for Q2 was incorrect, a student was still awarded 1 point if option 3 of Q1 was chosen as the answer because gram-positive diplococci are the microscopic morphology of *Streptococcus pneumoniae*. However, as *Streptococcus pneumoniae* was not the correct answer for Q1, the student did not receive a point for Q1. In other words, it was inferred that the student only had a satisfactory knowledge for Outcome 2 (Table 1, possibility E).

TABLE 2.  
Comparison of the mean scores as percentages of the step-by-step SBA-MCQs versus the mean scores of the traditional SBA-MCQ for all questions.

Tested outcomes <sup>a</sup>	Mean ( $\pm$ SD)		P value
	Step-by-step	Traditional	
Outcome 1 + Outcome 2	90.9% $\pm$ 9.9%	85.9% $\pm$ 14.0%	< 0.001
Outcome 1 + Outcome 3	87.4% $\pm$ 9.8%	81.5% $\pm$ 13.2%	< 0.001

<sup>a</sup>Outcome 1 (causative agent), Outcome 2 (basic characteristics), and Outcome 3 (pathogenic mechanisms), respectively. SBA-MCQ = single-best-answer multiple-choice question.

### Comparing scores of the traditional and revised systems

To investigate whether the traditional SBA-MCQs correctly estimated the learning outcome achievements of the students, a comparison was made of the scores for the step-by-step and traditional SBA-MCQs. The score for Q2 without adjustment to reflect a student's chosen answer for Q1 (Fig. 1A, and Fig. 2, Q2) was considered as the score for the traditional SBA-MCQ1, which tested both Outcomes 1 and 2 simultaneously. The score for Q3 without adjustment for the student's chosen answer for Q1 (Fig. 1B, and Fig. 2, Q3) was considered as the score for the traditional SBA-MCQ2, which simultaneously tested Outcomes 1 and 3. Therefore, the full score of each step-by-step SBA-MCQ was 2 whereas the score of each traditional SBA-MCQ was 1. Finally, to normalize the scores for each SBA-MCQ type, the students' percentage scores for the traditional SBA-MCQ1 were averaged and compared with those from Q1 and Q2 of the step-by-step SBA-MCQ. In addition, the percentage scores of all students obtained from the traditional SBA-MCQ2 were averaged and compared with those from Q1 and Q3 of the step-by-step SBA-MCQ.

### Statistical analysis

Since percentage is considered to be scale data and our data were not normally distributed, descriptive statistics and the Mann–Whitney U test were used and performed using PASW Statistics for Windows, version 18. A *p* value of < 0.05 was considered to indicate statistical significance.

## RESULTS

### The traditional system generally underestimated learning outcome achievements

For the 316 medical students participating in this study, the mean of the step-by-step SBA-MCQ1 scores (expressed as the percentage of students choosing the correct answers) was 90.9% (standard deviation [SD] 9.9%). In comparison, the corresponding mean for the traditional SBA-MCQ1 was

significantly lower at 85.9% (SD 14.0%; *p* < 0.001; Table 2). Similarly, the mean of the step-by-step SBA-MCQ2 scores (again, in terms of the proportion of students choosing the correct answers) was 87.4% (SD 9.8%), which is significantly higher than the corresponding mean of 81.5% (SD 13.2%) for the traditional SBA-MCQ2 (*p* < 0.001). Moreover, an analysis of the performance of individual students found that the traditional SBA-MCQ1 scores were lower than the step-by-step SBA-MCQ1 scores for 65.8% of students; likewise, the traditional SBA-MCQ2 scores were lower than the step-by-step SBA-MCQ2 scores for 87.8% of students (Table 3). The individual test results for each student are shown in Appendix 3.

## DISCUSSION

Medical schools in Thailand have long used traditional SBA-MCQs to assess the learning outcome achievements of preclinical students. As previously mentioned, however, this MCQ format may not accurately reflect the achievement of learning outcomes by students because each MCQ item assesses more than one learning outcome simultaneously. The current study demonstrated that the traditional SBA-MCQs did indeed underestimate the achievement of students by giving an approximately 5% lower score, in terms of the percentage of students choosing correct answers, than that awarded by each comparable set of step-by-step SBA-MCQs. Moreover, our analysis of individual students showed that the majority received lower scores with the traditional SBA-MCQs. These observations suggest that the scores derived from traditional SBA-MCQs should be interpreted with caution as they potentially underestimate students' cognitive achievement. Using step-by-step SBA-MCQs coupled with our new scoring system more accurately evaluates outcome achievements. This subsequently allows more precise and individualized feedback to be given to each student about their mistakes in order to facilitate further personalized learning. Table 4 has an example of such student feedback.

The underestimation of students' learning outcome achievements by traditional SBA-MCQs has validated pre-

TABLE 3.

Percentage of students who had lower, equal, and higher scores for the traditional SBA-MCQs than for the step-by-step SBA-MCQs.

Score comparing traditional with step-by-step SBA-MCQ	Measured outcomes <sup>a</sup>		Interpretation <sup>b</sup>
	O1 + O2	O1 + O3	
Lower score	65.8%	87.0%	Have only knowledge of O1 or O2 or O3
The same score	26.3%	8.2%	Have knowledge of both outcomes
Higher score	7.9%	4.7%	No knowledge of any outcome

<sup>a</sup>O1, O2, and O3: Outcome 1 (causative agent), Outcome 2 (basic characteristics), and Outcome 3 (pathogenic mechanisms), respectively.

<sup>b</sup>See also Table 1 and Appendix 4 for more detailed interpretations.

vious suggestions that every item should reflect a single specific content and a single specific mental behavior (17, 25). However, this best practice typically focuses on stems and options of the MCQs, the lead-in question not being specifically mentioned by any standard guidelines (7, 9, 11, 13, 17). In fact, the most recent standard MCQ guidelines of the US National Board of Medical Examiners did not include lead-in questions in the list of technical item flaws (11). Therefore, it has become a common practice for Thai educators to make their SBA-MCQs more challenging for medical students by increasing the complexity of lead-in questions. Nevertheless, our study has demonstrated that, without a careful design similar to the step-by-step SBA-MCQs, implementing such

a practice potentially risks underestimating the learning outcome achievements of preclinical students.

While this study provides novel observations, it has some limitations. To begin, it was conducted among preclinical medical students, who have limited clinical knowledge. It is possible that the efficacy of the traditional SBA-MCQs may differ when assessing students in their clinical years. This could be due to the fact that the stems of clinical-year questions are usually more complicated to permit the assessment of students' skills and knowledge in formulating accurate diagnoses. This is unlike the stems of preclinical-year questions, which are usually straightforward and, frequently, already provide the diagnosis. In addition, this

TABLE 4.

An example of feedback to students<sup>a</sup>.

Organism	Feedback for 8 sample students							
	1	2	3	4	5	6	7	8
<i>S. aureus</i>	C	A	A	A	A	D	D	A
<i>S. pyogenes</i>	A	A	A	A	A	E	E	A
<i>N. gonorrhoeae</i>	A	A	A	A	A	A	B	A
<i>B. anthracis</i>	A	A	A	A	A	A	A	A
<i>C. tetani</i>	A	A	A	A	A	B	A	A
<i>E. coli</i>	A	A	A	A	A	A	A	A
<i>S. enterica</i>	A	A	A	A	A	B	A	A
<i>H. pylori</i>	A	A	A	A	A	B	B	A
<i>P. aeruginosa</i>	A	B	A	B	E	A	D	A
<i>C. albicans</i>	A	B	A	A	A	A	A	A
<i>C. neoformans</i>	A	A	A	A	A	A	A	A
<i>T. rubrum</i>	E	A	D	E	A	B	A	A
<i>R. oryzae</i>	B	A	D	A	A	A	A	A
Papilloma virus	B	A	E	A	A	B	B	A
Herpes simplex virus	B	A	A	A	A	A	B	A
Norovirus	B	B	A	A	A	B	B	A
Dengue virus	B	A	A	A	A	B	B	A
Influenza virus	B	A	A	A	A	B	A	A

<sup>a</sup>A = able to tell diseases caused by the agent and morphological characteristics; B = able to tell diseases caused by the agent but cannot tell morphological characteristics; C and D = not able to tell diseases caused by the agent or morphological characteristics; E = not able to tell diseases caused by the agent but able to tell morphological characteristics of the distractors.

study was conducted on just one course and with only one group of students. More studies involving other preclinical-year courses and different groups of students are required.

## CONCLUSION

In summary, this study demonstrated that the traditional SBA-MCQs underestimated preclinical students' learning outcome achievements for the basic medical microbiology course. Moreover, we found that using the corresponding sets of step-by-step SBA-MCQs was a better method to establish students' learning progress despite that approach being relatively more time-consuming for preclinical students in an examination setting. The step-by-step SBA-MCQs are also a potentially powerful tool for providing personalized feedback to correct students' misconceptions.

## SUPPLEMENTAL MATERIALS

- Appendix 1: Step-by-step SBA-MCQs for bacteria
- Appendix 2: Step-by-step SBA-MCQs for fungi (items numbered 1–12) and viruses (items numbered 13–27)
- Appendix 3: Scores for each student
- Appendix 4: Outcome achievements of students when tested by the step-by-step SBA-MCQ

## ACKNOWLEDGMENTS

The authors thank Dr. Kamol Suwannakarn and Dr. Jaturong Sewatanon, Department of Microbiology, Siriraj Hospital, for their critical reading of this manuscript. The authors also thank Dr. Suthan Srisangkaew and Dr. Thanamate Rujipornchaiyakij for their comments on the manuscript wording. We thank Dr. Kate Fox, DPhil, from Edanz Group ([www.edanzediting.com/ac](http://www.edanzediting.com/ac)) for editing a draft of this manuscript. The authors do not have any conflicts of interest to declare.

## REFERENCES

1. Newble D. 2016. Revisiting "The effect of assessments and examinations on the learning of medical students." *Med Educ* 50:498–501.
2. Bloom BS, Engelhart MD, Furst EJ, Hill WH, Krathwohl DR. 1956. *Taxonomy of educational objectives: the classification of educational goals*. Edward Bros., Ann Arbor, MI.
3. Crowe A, Dirks C, Wenderoth MP. 2008. Biology in bloom: implementing Bloom's taxonomy to enhance student learning in biology. *CBE Life Sci Educ* 7:368–381.
4. Yeong FM. 2015. Use of constructed-response questions to support learning of cell biology during lectures. *J Microbiol Biol Educ* 16:87–89.
5. van Hoeij MJ, Haarhuis JC, Wierstra RF, van Beukelen P. 2004. Developing a classification tool based on Bloom's taxonomy to assess the cognitive level of short essay questions. *J Vet Med Educ* 31:261–267.
6. Javaeed A. 2018. Assessment of higher ordered thinking in medical education: multiple choice questions and modified essay questions. *MedEdPublish* 7(2):60. doi.org/10.15694/mep.2018.0000128.1.
7. Scully D. 2017. Constructing multiple-choice items to measure higher-order thinking. *Pract Assess Res Eval* 22:1–13.
8. Zaidi NLB, Grob KL, Monrad SM, Kurtz JB, Tai A, Ahmed AZ, Gruppen LD, Santen SA. 2018. Pushing critical thinking skills with multiple-choice questions: does Bloom's taxonomy work? *Acad Med* 93:856–859.
9. Coughlin PA, Featherstone CR. 2017. How to write a high quality multiple choice question (MCQ): a guide for clinicians. *Eur J Vasc Endovasc Surg* 54:654–658.
10. Hayes K, McCrorie P. 2010. The principles and best practice of question writing for postgraduate examinations. *Best Pract Res Clin Obstet Gynaecol* 24:783–794.
11. Paniagua MA, Swygert KA. 2016. *Constructing written test questions for the basic and clinical sciences*, 4th ed. National Board of Medical Examiners, Philadelphia, PA, USA.
12. Al-Rukban MO. 2006. Guidelines for the construction of multiple choice questions tests. *J Family Community Med* 13:125–133.
13. Dell KA, Wantuch GA. 2017. How-to guide for writing multiple choice questions for the pharmacy instructor. *Curr Pharm Teach Learn* 9:137–144.
14. Collins J. 2006. Education techniques for lifelong learning: writing multiple-choice questions for continuing medical education activities and self-assessment modules. *Radiographics* 26:543–551.
15. Brady AM. 2005. Assessment of learning with multiple-choice questions. *Nurse Educ Pract* 5:238–242.
16. Cook S. 2008. Writing outstanding MCQs that match your objectives: why keep assessing your student's performance a secret? *SGH Proc* 17:154–159.
17. Haladyna TM, Downing SM, Rodriguez MC. 2002. A review of multiple-choice item-writing guidelines for classroom assessment. *Appl Meas Educ* 15:309–333.
18. Breakall J, Randles C, Tasker R. 2019. Development and use of a multiple-choice item writing flaws evaluation instrument in the context of general chemistry. *Chem Educ Res Pract* 20:369–382.
19. Downing SM. 2005. 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 Theory Pract* 10:133–143.
20. Baig M, Ali SK, Ali S, Huda N. 2014. Evaluation of multiple choice and short essay question items in basic medical sciences. *Pak J Med Sci* 30:3–6.
21. Tariq S, Tariq S, Maqsood S, Jawed S, Baig M. 2017. Evaluation of cognitive levels and item writing flaws in medical

- pharmacology internal assessment examinations. *Pak J Med Sci* 33:866–870.
22. Kowash M, Hussein I, Al Halabi M. 2019. Evaluating the quality of multiple choice question in paediatric dentistry postgraduate examinations. *Sultan Qaboos Univ Med J* 19:e135–e141.
  23. Kittrakulrat J, Jongjatuporn W, Jurjai R, Jarupanich N, Pongpirul K. 2014. The ASEAN economic community and medical qualification. *Glob Health Action* 7:24535.
  24. Burns ER. 2010. “Anatomizing” reversed: use of examination questions that foster use of higher-order learning skills by students. *Anat Sci Educ* 3:330–334.
  25. Rodriguez-Diez MC, Alegre M, Diez N, Arbea L, Ferrer M. 2016. Technical flaws in multiple-choice questions in the access exam to medical specialties (“examen MIR”) in Spain (2009–2013). *BMC Med Educ* 16:47.
  26. Imwattana K, Kiratisin P, Techasintana P, Ngamskulrungrroj P. 2018. An impact on medical student knowledge outcomes after replacing peer lectures with small group discussions. *MedEdPublish* 7(4):3. doi.org/10.15694/mep.2018.0000224.1.
  27. Anonymous. 2012. Medical Competency Assessment Criteria for National License 2012, on The Medical Council of Thailand. <http://www.tmc.or.th/download/medical2555.pdf>. Accessed 30 October 2017.