INTRODUCTION

The National Board of Medical Examiners (NBME) style of multiple-choice questions (MCQs) is the standardized question style for the summative assessment of the United States Medical Licensing Examination (USMLE) board examinations and so has become the standardized style for the weekly quizzes, midterm exams, and final exams of several basic medical disciplines in our school of medicine. Weekly quizzes in our educational setting are considered formative assessments, whereas midterm and final exams are summative. However, based on my experience of teaching physiology, I have learned that the NBME-style MCQs are very appropriate for summative assessment, but not always appropriate for formative assessment. Therefore, for each lecture I deliver, I attach a set of PowerPoint slides with exercise questions tailored to my lecture content and in a variety of non-NBME styles. I use these exercises year by year as a formative approach to facilitating students’ learning in a self-directed, non-competitive context. Each class I have taught has had more than 125 students, and my formative exercise questions have been greatly appreciated by the majority of them.

However, one or two students in each class usually express concern that “this teacher’s questions do not follow the NBME style.” To explain why I use multiple question styles to create my formative exercise questions, I wrote out my rationales and shared them with each of the last three classes. Once they understand my approach, I have heard no more concerns about it. The multiple question styles I use include fill in the blank, matching items, true/false, drag and drop, labeling important structures or functions, and MCQs that allow answers such as “all of the above” and “none of the above” that the current NBME style excludes.

The goal of this article is to share my rationale in terms of four important aspects that should be taken into consideration when developing formative assessment questions to enhance the quality of teaching and learning of physiology. Prashanti and Ramnarayan (2) have highlighted “Ten maxims of formative assessment.” My four aspects add more dimensions to think about and apply in the formative assessment.

1) There is a characteristic wealth-dearth dichotomy in learning in the twenty-first century. In 1971, Herbert A. Simon (4), a Nobel Prize winner in economics and one of the pioneers of computer science, cognitive psychology, and artificial intelligence, foresaw that, in the information age, “the wealth of information means a dearth of something else … a poverty of [human] attention and a need to allocate that attention efficiently among the overabundance of information sources that might consume it.” Indeed, medical knowledge is increasing explosively, but students only have 24 h in a day. They cannot spend all 24 h learning, and the learning time they do have is also taken up by e-mails, internet searches, WhatsApp messages, text messages, social media posts, new features of the various electronic products they use, and so on, not to mention the activities of daily living! Living in the fast-paced 21st century, a person’s attention is now a scarce resource that neither students nor teachers should abuse. Each should cautiously differentiate between what information is truly necessary and what will waste precious attention and time.

2) Formative and summative assessments serve different goals. Dr. Cynthia Ledford, our Associate Dean for Medical Education, often says, “Formative assessments allow us to be a student’s coach. With summative assessments, we are serving as judge.” Robert Stake, Professor Emeritus of Education at the University of Illinois said, “When the cook tastes the soup, that’s formative. When the guests taste the soup, that’s summative” (3). Hence, formative assessment facilitates the learning process, whereas summative assessment assesses the learning outcomes, i.e., how well learners have learned. Learning is a time-dependent, dynamic process and can be divided at least into early and late phases. In the early phase, students need time to digest and absorb what they were taught. In the later phase, they consolidate what they’ve learned as manifested in the formation of long-term memory or some skills. Whereas summative assessments have prefixed timing or dates (such as midterm and final exams and the USMLE exams), formative assessments can be conducted with great flexibility based on the context (the content, learning phase or timing, learners’ level of understanding, etc.). Hence, I) questions that are used for summative assessments may not always be appropriate for formative assessments; and 2) the same question may play different roles at different learning phases. I will provide two sample questions to support these two arguments.

For a normal adult, the total body water (TBW) is ~60% (50–70%) of body weight; the intracellular fluid (ICF), extracellular fluid (ECF), interstitial fluid (ISF), and plasma make up are ~40, 20, 15, and 5% of body weight, respectively. This information can also be presented using fractions: ICF is 2/3 TBW, ECF is 1/3 TBW, ISF is 3/4 ECF, and plasma is 1/4 ECF. Although these are simple facts, when the new concepts of five percentages, four fractions, and five body fluid compartments (TBW, ICF, ECF, ISF, and plasma) are presented together...
in a session, it is very easy for students to make low-level mistakes in answering test questions about them in the early phase. In this circumstance, NBME-style MCQs are not always appropriate, because they allow only one correct option, and a teacher needs to effortfully create several wrong options:

Question 1 (NBME-style MCQ): Regarding the volume of the interstitial fluid, which of the following estimations is most likely to be true?

A. 5% of ECF (wrong)
B. 15% of body weight (correct)
C. 20% of ECF (wrong)
D. 20% of TBW (wrong)
E. 40% of ECF (wrong)

Comment: In the early phase, before students form a consolidated memory of the many percentages, challenging them with many wrong percentages does not facilitate their learning, but adds an extra burden to interfere with their normal learning process. If a weekly quiz contains 10 NBME-style MCQs that test students’ memory of numerical knowledge, there will be 40 wrong options that unnecessarily confuse students, making learning much more difficult and less effective in the early learning phase. Hence, a non-NBME style MCQ that allows “all of above are correct” or only one wrong answer is appropriate in this context.

Whereas the NBME-style MCQ has a disadvantage in facilitating students’ memory formation of the simple facts about the distribution of body fluids in the early phase of learning, I do use the NBME-style for questions that require logical reasoning in the early phase:

Question 2 (semi-NBME-style MCQ): The plasma [K⁺] of a patient has a sudden increase to 7.5 meq/L. Which of the following statements is true regarding his situation?

A. The size of the K⁺ chemical gradient increases. (wrong)
B. The magnitude of the equilibrium potential of K⁺ increases. (wrong)
C. The resting membrane potentials of his cells would be more negative (i.e., farther to the threshold level). (wrong)
D. The excitability of the patient’s excitable cells reduces. (wrong)

E. The hyperkalemia needs to be treated immediately to prevent potential cardiac arrhythmia. (correct)

Comment: From an increase in the plasma [K⁺] to the possible life-threatening consequence of hyperkalemia, the logical reasoning involves all steps from answer A to E. Medical students must be able to reason through all five steps. In this particular context, the question was not designed to spare their attention, but to intentionally draw and cost their attention! Additionally, NBME style requires an alphabetical order of the answers, but this question was ordered by logic. That’s why it is a semi-NBME style question. If ordered alphabetically, this same question is appropriate to be used as a summative assessment in the final exam, which tests how well students master the important knowledge about hyperkalemia.

In brief, the above sample questions show that question styles for the formative assessment should be content appropriate, timing sensitive, or learning-phase appropriate. The following aspect will further support this summary.

3) Think about how a teacher may enhance the efficiency of the brain’s processing of lecture materials. A friend of mine once told me how she admires her teacher, a renowned computer programmer. As a new student programmer, my friend would write long codes that used a lot of memory and processing power and required a relatively long time to complete the job. After her teacher modified the code, however, only a few lines remained, but those few lines were better organized, and the computer finished the task much more efficiently and in much less time. Students’ brains are “computers,” and lecture materials are the “source code” these “computers” will process. When the brain does its job, the student gains knowledge and/or skills. To minimize the brain’s overhead, I strive to produce lecture materials that are like “source code” produced by a very sophisticated programmer. I will give three examples below:

Example 1: In the following scenarios, filling in the blanks can be the best “source code”: each lecture is like a different “forest.” To avoid students seeing the individual “trees” without seeing the “forest,” or if a specific content is long, at the end of my PowerPoint or at the end of the long content, immediately asking them to simply fill in the headings or subheadings of the lecture or the long content always effectively draws students’ attention to the forest and is better than any other question styles.

Example 2: At the end of teaching the concept of renal tubular transport, to help students to memorize where and how H⁺ is secreted along the renal tubule and prepare them to learn about the acid-base balance, a “drag and drop” question was purposely designed (question 1 in Fig. 1) to provide them with “hands-on” experience in correcting the errors. This approach turned out to be quite effective in separating the very content out from the other membrane transport mechanisms along the renal tubule that have not been digested and absorbed.

In the context above, using NBME-style MCQs with distractive options is less likely to achieve the efficiency of the “hands-on” approach for this particular content at this particular timing. The several fill-in-the-blank questions in Fig. 1 are also better “source code” than NBME-style MCQs, because, again, it is inappropriate to have students exposed to too many distractive options than correct answers in the early phase of learning.

It should be noted that, although I emphasize the effectiveness or the quality of the “source code,” this does not mean that less is always better. Sometimes deliberate redundancy also results in effective “source code,” as shown in example 3.

Example 3: After students learn about body fluids, membrane transport, and the electrophysiology of the cell membrane, to emphasize the important roles of Na⁺ and K⁺ concentration gradients, in an exercise question, each gradient is redundantly presented using numerals, the density of corresponding dots, and colored triangles (Fig. 2).

The numerical presentations of the two gradients are abstract. My experience has taught me that not all students transform the

1 The term “content appropriate” will be elaborated to be “content property appropriate” in the section on the fourth aspect.
numic gradients into concrete situations immediately, and some need some facilitated imagination to picture the situation. The differences in dot density across the cell membrane (both red and blue dots) make the two gradients concrete, and the red and blue triangles are eye-catching, which further reinforces the two gradients in their minds.

4) Different types of learning content (i.e., different information) have different physical natures and different content properties in terms of the meaning the information carries. First, information of various physical natures (visual, auditory, somatosensory, olfactory, and gustatory) reaches the brain via different nerve pathways. Second, information that carries different meanings for human life may generate different patterns of activation of brain regions. Just like different people have different personalities, characteristics, or temperaments, information that carries different meanings has different content properties or characteristics, such as factual, numerical, procedural, logical, analytical, simple, complex, comprehensive, emotion charged, or non-emotion charged, and so on. Hence, the human brain is likely to handle information with different content properties differently. The formation of long-term memory is one type of learning outcome.

Modern neuroscience research indicates that memory can be classified along two dimensions: the nature of the information stored and the duration of storage. The former includes explicit and implicit memory (or conscious and unconscious memory); the latter includes long-term and short-term memory. Long-term memory can be explicit or implicit. Explicit memory has episodic and semantic forms. Neuroscientists have also found that certain regions of the brain are much more important for some types of storage than for others (1). Hence, in terms of particular learning content or a unique combination of different contents with different properties, a thoughtful, sophisticated teacher knows which style(s) of questions is(are) the best or better “source code” for the students’ brains to process experientially. For example, learning and understanding the concept of renal clearance requires abstract and mathematical thinking, whereas learning where and how H\textsuperscript{+} is secreted into the lumen of the renal tubule requires memory, because it is factual. If fMRI or more advanced brain imaging techniques will be used to examine learners’ brains, it is likely to show the activation of different brain regions when they process information with different physical natures and content properties.

5. Which transporter is influenced by aldosterone?

![Diagram of Na\textsuperscript{+} and K\textsuperscript{+} gradients](image)

Fig. 1. A group of formative assessment questions to reinforce students’ learning. Question 1 is the “drag and drop” question that provides students with “hands-on” experience.

![Diagram of Na\textsuperscript{+} and K\textsuperscript{+} gradients](image)

Fig. 2. The important physiological roles of the Na\textsuperscript{+} and K\textsuperscript{+} concentration ([Na\textsuperscript{+}] and [K\textsuperscript{+}], respectively) gradients. Na\textsuperscript{+} gradient is in red (warm) color because its role in membrane potential is excitatory. An increase in the plasma [Na\textsuperscript{+}] results in an increase in the plasma osmolarity, which tends to holds more water in the plasma, and an increase in the plasma volume can be one of the causes of high blood pressure. K\textsuperscript{+} gradient is in blue (cold) color because it plays an inhibitory role in membrane potential. ECF; extracellular fluid; ICF, intracellular fluid.

---

**Fig. 1.** A group of formative assessment questions to reinforce students’ learning. Question 1 is the “drag and drop” question that provides students with “hands-on” experience.

**Fig. 2.** The important physiological roles of the Na\textsuperscript{+} and K\textsuperscript{+} concentration ([Na\textsuperscript{+}] and [K\textsuperscript{+}], respectively) gradients. Na\textsuperscript{+} gradient is in red (warm) color because its role in membrane potential is excitatory. An increase in the plasma [Na\textsuperscript{+}] results in an increase in the plasma osmolarity, which tends to holds more water in the plasma, and an increase in the plasma volume can be one of the causes of high blood pressure. K\textsuperscript{+} gradient is in blue (cold) color because it plays an inhibitory role in membrane potential. ECF; extracellular fluid; ICF, intracellular fluid.
In summary, my experience as a physiology teacher has taught me that question styles do matter for formative assessments. Questions for the formative assessments should be carefully designed to be attention effective, learning phase appropriate, efficient “source code” like, and content property appropriate.

ACKNOWLEDGMENTS

The author thanks Dr. Yongchao Shi, the Director of Student Assessment of our School of Medicine, for comments about formative and summative assessment and the NBME-style MCQs. The author also thanks Anna Karcher for artistic illustrations of the two figures.

DISCLOSURES

No conflicts of interest, financial or otherwise, are declared by the author.

AUTHOR CONTRIBUTIONS

S.Y.K. prepared figures; drafted manuscript; edited and revised manuscript; approved final version of manuscript.

REFERENCES

1. Kandel ER, Schwartz JH, Jessell TM, Siegelbaum ST, Hudspeth AJ. Principles of Neural Science (5th ed.). New York: McGraw Hill Medical, 2013.
2. Prashanti E, Ramnarayan K. Ten maxims of formative assessment. Adv Physiol Educ 43: 99–102, 2019. doi:10.1152/advan.00173.2018.
3. ResourcEd. Types of Summative Assessment and Formative Assessment (Online). https://resourced.prometheanworld.com/types-of-summative-formative-assessment/ [24 June 2019].
4. Simon HA. Designing organizations for an information-rich world. In Computers, Communication, and the Public Interest, edited by Greenberger M. Baltimore, MD: Johns Hopkins Press, 1971.