

ORIGINAL ARTICLE

The influence of question design on the response to self-assessment in www.elearnSCI.org: a submodule pilot study

N Liu¹, X-W Li¹, M-W Zhou¹ and F Biering-Sørensen²

Study design: This is an interventional training session.

Objective: The objective of this study was to investigate the difference in response to self-assessment questions in the original and an adjusted version for a submodule of www.elearnSCI.org for student nurses.

Setting: The study was conducted in a teaching hospital affiliated to Peking University, China.

Methods: In all, 28 student nurses divided into two groups (groups A and B; 14 in each) received a print-out of a Chinese translation of the slides from the 'Maintaining skin integrity following spinal cord injury' submodule in www.elearnSCI.org for self-study. Both groups were then tested using the 10 self-assessment multiple-choice questions (MCQs) related to the same submodule. Group A used the original questions, whereas group B received an adjusted questionnaire.

Results: The responses to four conventional single-answer MCQs were nearly all correct in both groups. However, in three questions, group A, with the option 'All of the above', had a higher number of correct answers than group B, with multiple-answer MCQs. In addition, in another three questions, group A, using the original multiple-answer MCQs, had fewer correct answers than group B, where it was only necessary to tick a single incorrect answer.

Conclusion: Variations in design influence the response to questions. The use of conventional single-answer MCQs should be reconsidered, as they only examine the recall of isolated knowledge facts. The 'All of the above' option should be avoided because it would increase the number of correct answers arrived at by guessing. When using multiple-answer MCQs, it is recommended that the questions asked should be in accordance with the content within the www.elearnSCI.org.

Spinal Cord (2015) **53**, 604–607; doi:10.1038/sc.2014.226; published online 17 February 2015

INTRODUCTION

On 4 September 2012, the web-based e-learning platform www.elearnSCI.org was launched during the 51st Annual Scientific Meeting of the International Spinal Cord Society in London.¹ The development and release of this online learning course is consistent with the goals of the International Spinal Cord Society, which include advising, encouraging, guiding and supporting the efforts of those responsible for the education and training of medical professionals and professionals allied to medicine dealing with individuals with spinal cord injury (SCI).²

The resource comprises seven modules; six modules address the educational needs of all disciplines involved in comprehensive SCI management, and the seventh module addresses the prevention of SCI. Each module consists of various submodules that cover specific topics and include an overview, activities, self-assessment questions and references. The initial purpose of the self-assessment in each submodule of the www.elearnSCI.org is to test the learning effect and enhance learning motivation.³ The commonly used designs for the self-assessment questions include single-answer and multiple-answer multiple-choice questions (MCQs). Furthermore, some single-answer questions that include the option 'All of the above' could be considered as an additional type of question that is different from conventional single-answer MCQs. To date, no study has been

conducted to investigate the self-assessment design and its influence on the number of correct answers in this online resource.

In this study, the influence of question design on the frequency of correct responses to self-assessment questions from a submodule within www.elearnSCI.org by student nurses was examined using the self-assessment questions from the same submodule. Answers to the original self-assessment questions and to adjusted questions were compared in two groups of student nurses.

PATIENTS AND METHODS

In May 2013, 28 student nurses in their third year of education from the Department of Nursing, Health Science Center, Peking University, China, were recruited into this study. The study was carried out in a 2-h class through a series of lectures on physical medicine and rehabilitation in accordance with their semester syllabus.

Training process

Twenty-eight student nurses (divided into groups A and B; 14 in each) received a print-out of the slides in the 'Maintaining skin integrity following spinal cord injury' submodule within the 'Nurses' module from www.elearnSCI.org in a Chinese translation for self-study during the first hour of the class. The translated print-out was in the same format as the original material from the website (http://www.elearnsci.org/downloads/78_6.pdf) with six screens printed on one page. Translation into Chinese was approved by the chair of the

¹Department of Rehabilitation Medicine, Peking University Third Hospital, Beijing, China and ²Department for Spinal Cord Injuries, Glostrup Hospital, University of Copenhagen, Copenhagen, Denmark

Correspondence: Dr X-W Li, Department of Rehabilitation Medicine, Peking University Third Hospital, No. 49 North Garden Road, Beijing 100191, China.

E-mail: xwli@163.com

Received 28 March 2014; revised 5 November 2014; accepted 11 November 2014; published online 17 February 2015

Table 1 Types of MCQs within the questionnaire

Type	Group A (original)	Group B (revised)	Corresponding question number
I	Conventional single-answer MCQ	Conventional single-answer MCQ	Q1, Q3, Q6, Q8
II	Single-answer MCQ with the option 'All of the above'	Multiple-answer MCQ	Q2, Q7, Q9
III	Multiple-answer MCQ	Single-answer MCQ with an 'except' added at the end of the stem	Q4, Q5, Q10

Abbreviation: MCQ: multiple-choice question

Educational Committee of the International Spinal Cord Society. Next, the first author did the initial translation, which was later scrutinized by the second author. The third author reviewed the translation and eventually the final translation was agreed upon by all three translators/reviewers.

After a 1-h self-study, the 10 self-assessment questions from the same submodule were given to the students for a 20-min test. Students in group A received all of the original questions from the website, whereas students in group B received an adjusted questionnaire, in which four conventional single-answer MCQs were the same as the original and the other six questions were revised. During the test, the students were permitted to use the print-out of the slides. Next, the answer sheets from these two groups were collected, and the correct answers were revealed to the students during the remaining 30 min.

Self-assessment questionnaires

Original self-assessment questionnaire for group A (see Appendix): Ten questions within the self-assessment of this submodule are categorized into three groups as follows: type I, conventional single-answer MCQs in which only one option is correct; type II, single-answer MCQs with the option 'All of the above', which is always the correct answer; and type III, multiple-answer MCQs in which more than one option is correct. In this submodule, group A had four type I, three type II and three type III questions. The features of the three types of questions are summarized in Table 1.

Adjusted self-assessment questionnaire for group B (see Appendix): Among the self-assessment questions for group B, the four type I MCQs were the same as in the original questionnaire. However, the other six questions were revised. The three type II MCQs with the option 'All of the above' were altered by deleting the option 'All of the above' and by making the remaining options available for multiple answers. To the three type III MCQs, an 'except' was added at the end of the stem, and the questions were reversed into the single-answer format. The features of the three types of questions are summarized in Table 1.

The number of correct answers to each question in both groups was counted (Table 2). Statistical comparisons between the two groups were not made, as the purpose was to evaluate the question design including the options in the various types of MCQs.

RESULTS

The frequency of correct responses to each question in both groups is shown in Table 2. Among the four type I questions, three (Q1, Q3 and Q6) were correctly answered by the 14 students in both group A and group B. However, only one student from group A gave a wrong answer to Q8. For type II questions, group A had a higher frequency of correct answers in answering the original version with the option 'All of the above' than group B. (For details see Table 2). Interestingly, the situation for type III questions was the opposite to that observed for type II questions. Group A had a lower number of correct answers compared with group B, which had an 'except' added at the end of the stem and the students could choose only one correct answer (Table 2).

Table 2 Number of correct responses obtained in two groups of student nurses to 10 self-assessment questions with different design for 6 of 10 questions

Question (cf. Appendix 1)	Group A (original)	Group B (revised)
<i>I. Both stem and options are the original</i>		
Q1.	14/14	14/14
Q3.	14/14	14/14
Q6.	14/14	14/14
Q8.	13/14	14/14
<i>II. Group B does not have 'All of the above' option and has multiple-answer format</i>		
Q2.	14/14	10/14
Q7.	11/14	1/14
Q9.	13/14	2/14
<i>III. Group B has an 'except' added at the end of the stem and single-answer format</i>		
Q4.	8/14	13/14
Q5.	10/14	14/14
Q10.	4/14	10/14

The original questions are from the 'Maintaining skin integrity following spinal cord injury' submodule (<http://www.elearnsci.org/sa.aspx?id=78>). There were 14 student nurses in each group.

DISCUSSION

MCQs have been widely used as an assessment tool in medical education.⁴ They can measure the most important educational outcomes, including knowledge, understanding, judgment and problem solving.⁵ However, so far no study has reported the influence of these question types on the frequency of correct responses to the assessment of elearnSCI.

According to Miller's knowledge pyramid,⁶ single-answer questions mainly test descriptive knowledge of facts (that is, the first level: 'knows'), whereas multiple-answer questions can test the application of knowledge (that is, the second level: 'knows how') among medical students. Among the 10 self-assessment questions of this submodule, 4 questions are conventional single-answer questions in which only one knowledge point is tested. Although another three type II questions were presented in single-answer format, those questions tested multiple knowledge points. The remaining three questions are typical multiple-answer questions, in which multiple knowledge points are tested.

On the basis of the knowledge evaluation by the type I questions in our study, it can be seen that the frequency of correct responses is extremely good for descriptive type of questions within this submodule. On the one hand, the study indicated that the knowledge point for these four questions was clearly mentioned in the learning materials. On the other hand, all these knowledge points are fundamental concepts of skin issue following SCI for nursing management. Although the responses to these four questions are almost perfect, it seems that using the type I question to test the basic knowledge of this submodule is somewhat an uncomplicated method for student nurses. Moreover, medical examinations are now increasingly focused on assessing higher levels of knowledge rather than the mere recall of isolated facts;⁷ hence, the efficiency and appropriateness of using these types of questions remain to be deliberated.

For three type II questions, the high frequency of correct responses in group A is a 'false-positive' result. The lower frequency of correct responses in group B without the option 'All of the above' is probably a more realistic reflection of the real situation of knowledge

transmitted from the learning material. For Q2, all four correct answers related to skin protection are the same as mentioned in the learning material, which leads to the relatively 'less' number of correct answers in group B. However, for Q7 and Q9, the correct answers are not directly revealed in the learning material. Even though students of both the groups had the learning material, group B students could not determine the correct answers. In addition, in group B, for Q7 and Q9 all four options were correct. There may have been a psychological effect making the students think that 'they can't possibly all be correct', and thus they were more prone not to select all of the answers. Thus, to choose the correct answer, students were required to understand and assimilate the knowledge in full. Therefore, the scores are markedly lower for these two questions in group B. From these results, it is recommended that single-choice 'All of the above' option be avoided when developing assessment questions. When 'All of the above' option exists, students who can identify more than one option as correct can select the correct answer even if they are unsure about other options. In this case, sophisticated students can use partial knowledge to arrive at the correct answer through guessing, which would lead to a less valid testing outcome. In addition, it is also suggested that the questions in the self-assessment closely reflect the content of the www.elearnSCL.org; otherwise, it could contribute to the inadequate selection of choices, especially when multiple answers should be selected, and decrease the frequency of correct responses. Therefore, the use of these types of questions is not appropriate, and it should be amended in the future.

For the last three type III questions, we changed the format by adding an 'except' at the end of the stem, and this adjustment resulted in a much higher number of correct answers compared with the original questions. However, this adjustment could only test the student's ability to detect incorrect answers, and not whether they knew the correct answers. Similarly, these more accurate responses are 'false positive'. Furthermore, the scores for Q10 were very low in group A, which used the original format. This MCQ differed from the previous nine questions, as it presented a clinical scenario. To tick the correct answers, a decision-making process was necessary instead of directly finding the knowledge points within the learning material. Accordingly, the adjusted format of adding an 'except' at the end of the stem is not recommended, and the original multiple-answer MCQ format should be retained.

MCQs can be used to objectively measure factual knowledge, ability and high-level learning outcomes.⁸ Unfortunately, there is a lack of evidence-based guidelines relating to the design and use of MCQs.⁹ Thus, developing a set of valid and reliable tests with MCQs targeting higher cognitive abilities and conforming to item construction guidelines presents a big challenge to the question developer.¹⁰ Neither extremely high nor extremely low frequency of correct responses should be accepted as a good question design. An extremely high number of correct answers may indicate that the questions are too simple, whereas the extremely low scores may suggest that the questions are too difficult, or could represent questions in which the answers were not directly revealed in the learning material. It is suggested that the question design require students to fully understand and assimilate the knowledge in the learning material.

From the results of our study, it is obvious that type I MCQs are too simple to answer incorrectly. In contrast, the multiple-answer MCQs could bring challenges to the students, and if not in accordance with the learning material it will create confusion as well. For type II MCQs, when 'All of the above' option exists, the frequency of correct responses is high due to the intrinsic default for this option. When this option is omitted, the scores for Q2 are still acceptable. In contrast, the

very low number of correct answers to Q7 and Q9 must be improved by matching the learning material and the question. Thus, it is suggested that during the question design, both 'All of the above' option for the 'false-positive' effect of scores and less coherence between learning material and questions for the 'false-negative' effect be avoided. The same situation is presented for type III questions. If an 'except' is added at the end of the stem, the scores will be high because the partial understanding of the knowledge will make sense. When using the original questions, the frequency of correct responses to Q4 and Q5 is still acceptable; however, the lower scores for Q10 are due to the very manner in which the decision-making process should be carried out. Thus, the original format is encouraged to be retained, although Q10 needs some refinement.

Nowadays, the most popular trend in designing medical examination questionnaires is the use of single-best-answer MCQ format, which has replaced the traditional MCQ.^{7,11} Compared with multiple true/false questions, the major difference is that the stem typically is based on a clinical scenario or vignette.¹¹ In addition, questions in which all options contain elements of correctness, which requires judgment to choose which one is the best answer, tend to be more difficult and discriminating than questions merely asking for a fact.¹² Although this format needs to increase the number of options in the MCQ to include a representative sample of the curriculum, it may not increase the examination time.¹³ This type of MCQ should be supplemented in future revisions of self-assessment.

Because we performed this study only in a submodule of www.elearnSCL.org, the results cannot be generalized to the whole learning resource. Thus, the influence of question design on the frequency of correct responses in other submodules within the seven modules of this online resource still needs to be investigated. In addition, several studies^{4,14,15} have discussed the optimal number of options in MCQs and their influence on the test; this topic was not addressed in our study. Nevertheless, we believe that our study points to specific issues that, if adjusted in future versions of www.elearnSCL.org, may enhance the benefit of self-assessment questions.

CONCLUSION

Variations in the design of MCQs definitely influence the response to questions. Well-designed questions could reflect the learning outcome more accurately. The use of type I MCQs should be reconsidered, as they only examine the recall of isolated knowledge facts. The single-choice 'All of the above' option should be avoided in the option design, because it will increase the scores of response by guessing. When using multiple answers, the MCQ tests multiple knowledge points. It is recommended that the content within the www.elearnSCL.org should be in accordance with self-assessment questions. Adjusted questions and option design should be considered in future revisions of this online learning resource. In addition, it has to be emphasized that the tests currently being used may be too simple and/or may not reflect the learning intended. Therefore, content developers should consider such factors when designing self-assessment questions.

DATA ARCHIVING

There were no data to deposit.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

- 1 Chhabra HS, Harvey LA, Muldoon S, Chaudhary S, Arora M, Brown DJ *et al*. www.elearnSCI.org: a global educational initiative of ISCoS. *Spinal Cord* 2013; **51**: 176–182.
- 2 Wyndaele JJ. Elearning: the next step in ISCoS's worldwide education on comprehensive spinal cord management. *Spinal Cord* 2013; **51**: 173.
- 3 Docherty C, Hoy D, Topp H, Trinder K. Using Elearning techniques to support problem based learning within a clinical simulation laboratory. *Stud Health Technol Inform* 2004; **107**: 865–868.
- 4 Zoanetti N, Beaves M, Griffin P, Wallace EM. Fixed or mixed: a comparison of three, four and mixed-option multiple-choice tests in a Fetal Surveillance Education Program. *BMC Med Educ* 2013; **13**: 35.
- 5 Al-Rukban MO. Guidelines for the construction of multiple choice questions tests. *J Family Community Med* 2006; **13**: 125–133.
- 6 Miller GE. The assessment of clinical skills/competence/performance. *Acad Med* 1990; **65**: 63–67.
- 7 Tan LT, McAleer JJ. Final FRCR Examination Board. The introduction of single best answer questions as a test of knowledge in the final examination for the fellowship of the Royal College of Radiologists in Clinical Oncology. *Clin Oncol (R Coll Radiol)* 2008; **20**: 571–576.
- 8 Bond AE, Bodger O, Skibinski DO, Jones DH, Restall CJ, Dudley E *et al*. Negatively-marked MCQ assessments that reward partial knowledge do not introduce gender bias yet increase student performance and satisfaction and reduce anxiety. *PLoS ONE* 2013; **8**: e55956.
- 9 Considine J, Botti M, Thomas S. Design, format, validity and reliability of multiple choice questions for use in nursing research and education. *Collegian* 2005; **12**: 19–24.
- 10 Sadaf S, Khan S, Ali SK. Tips for developing a valid and reliable bank of multiple choice questions (MCQs). *Educ Health (Abingdon)* 2012; **25**: 195–197.
- 11 McCoubrie P, McKnight L. Single best answer MCQs: a new format for the FRCR part 2a exam. *Clin Radiol* 2008; **63**: 506–510.
- 12 Fray RB. More multiple-choice item writing do's and don'ts. 4(11). *Practical Assessment, Research & Evaluation* 1995. Retrieved 18 January 2014 from <http://PAREonline.net/getvn.asp?v=4&n=11>.
- 13 Chandratilake M, Davis M, Ponnampuruma G. Assessment of medical knowledge: the pros and cons of using true/false multiple choice questions. *Natl Med J India* 2011; **24**: 225–228.
- 14 Vyas R, Supe A. Multiple choice questions: a literature review on the optimal number of options. *Natl Med J India* 2008; **21**: 130–133.
- 15 Rogausch A, Hofer R, Krebs R. Rarely selected distractors in high stakes medical multiple-choice examinations and their recognition by item authors: a simulation and survey. *BMC Med Educ* 2010; **10**: 85.

APPENDIX

Ten self-assessment questions in the 'Maintaining skin integrity following spinal cord injury' submodule (<http://www.elearnsci.org/sa.aspx?id=78>).

Note: The sentence in the first line of each question is the original version used in group A. The correct answer(s) to the original self-assessment questions is/are underlined. The sentence in parentheses in the second line explains the adjustments used in group B if any.

1. Which of the following is the largest organ of the body:
A. Liver
B. Spleen
C. Skin
D. Kidney
2. The skin provides protection from:
(Group B did not have option E, and all answers A.–D. had to be ticked)
A. Impact and pressure
B. Temperature
C. Micro-organisms
D. Radiation and chemicals
E. All of the above
3. The regulation of which vitamin is a function of skin:
A. Vitamin A
B. Vitamin B
C. Vitamin C
D. Vitamin D
E. Vitamin E
4. Which of the following are the risk factors for developing pressure sores:
(In group B changed to 'Which of the following are the risk factors for developing pressure sores except:' and in single-choice format, that is, option D to be ticked)
A. Incontinence
B. Altered skin moisture
C. Depression
D. Hyperproteinemia
E. Anemia
5. Which of the following are high-risk areas for developing pressure ulcer:
(In group B changed to 'Which of the following are high-risk areas for developing pressure ulcer except:' and in single-choice format, that is, option D to be ticked)
A. Sacrum
B. Ischial tuberosity
C. Trochanter
D. Abdominal wall
E. Occiput
6. If a reddened area is noted it should be reassessed in:
A. 15 min
B. 30 min
C. 45 min
D. 1 h
7. When should skin checks be carried out?
(Group B did not have option F and all answers A–E. had to be ticked)
A. Following each turn
B. After the patient gets into bed
C. If the patient reports a problem
D. If the patient has a pyrexia
E. During or after a shower/bath
F. All of the above
8. When should turning frequency be evaluated?
A. Daily
B. After each turn
C. Weekly
9. Where does education about skin care occur for patients?
(Group B did not have option F, and all answers A.–E. had to be ticked)
A. In the ward, delivered by the nurses
B. At the patient education lecture on skin management
C. In the gym
D. On weekend leave
E. On ward round with the doctors
F. All of the above
10. Poonam, a 42-year-old overweight paraplegic, presents with a grade 2 sacral sore. Which of the following advices would you give Poonam:
(In group B changed to 'Which of the following advices you would give to Poonam except:' and in single-choice format, that is, option D to be ticked)
A. Improve hydration
B. Reduce weight
C. Keep the area dry
D. Avoid non-vegetarian diet
E. Avoid tight clothing