# RESEARCH

# Pharmacy Student Performance on Constructed-Response Versus Selected-Response Calculations Questions

Elizabeth A. Sheaffer, PhD, MBA,<sup>a</sup> and Richard T. Addo, PhD<sup>a,b\*</sup>

<sup>a</sup>Bernard J. Dunn School of Pharmacy, Shenandoah University, Ashburn, VA <sup>b</sup>School of Pharmacy, Union University, Jackson, TN

Submitted July 5, 2012; accepted September 18, 2012; published February 12, 2013.

**Objective.** To introduce PharmD students to changes in calculations question types (constructed-response versus selected-response questions); measure and compare student performance on constructed-response and selected-response questions in a pharmaceutics course; and collect student feedback on the use of differing question types.

**Methods** A pharmaceutics/pharmaceutical calculations examination was administered that included 15 pairs of questions; each pair consisted of a constructed-response question and a similar selected-response question. An online questionnaire was conducted to collect student feedback.

**Results.** Of the 15 topics, the class scored higher on the constructed-response question for 4 topics and higher on the selected-response question for 10 topics. Eighty percent of the class preferred selected-response questions, although 47.8% felt constructed-response questions better prepared them for a career in healthcare.

**Conclusions.** Students correctly answered more selected-response questions than constructed-response questions and felt more confident in doing so. Additional constructed-response teaching and testing methods should be incorporated into pharmacy education.

Keywords: pharmaceutical calculations, NAPLEX, constructed response, selected response, question type

## **INTRODUCTION**

A milestone for doctor of pharmacy (PharmD) graduates is passing the National Pharmacy Licensure Examination (NAPLEX).<sup>1</sup> To improve performance on the examination, pharmacy programs should ensure their curriculum and instruction is reflective of what graduates will encounter on the examination. The 2010 modifications to the NAPLEX blueprint included a change to the calculations questions: instead of selected-response questions, most calculations questions were converted to constructedresponse questions.<sup>2</sup> Selected-response questions include possible answers, presented as multiple-choice, matching, true-false, etc, while constructed-response questions require respondents to construct (write or draw) an answer, such as a word or phrase, simple diagram, series of numbers, or short essay.<sup>3</sup> The authors have met many pharmaceutics faculty members who were unaware of the

**Corresponding Author:** Elizabeth A. Sheaffer, MBA, PhD. Address: Shenandoah University, Bernard J. Dunn School of Pharmacy, 45085 University Drive, Ashburn, VA 20147. Tel: 703-726-3528. Fax: 703-726-3558. E-mail: esheaffe@su.edu \*Dr. Addo was a faculty member at the Shenandoah University School of Pharmacy during the spring 2010 pilot study. He began a new position in fall 2010 at the Union University School of Pharmacy, where he completed the study. NAPLEX change. For programs that do not teach and test using the constructed-response method, its students may have difficulty when solving constructed-response calculations questions for the first time on the NAPLEX, and then subsequently in practice.

Although differences between constructed-response and selected-response question types have been studied in various educational settings,<sup>4,5</sup> selected-response questions have a more established history of successful application in low- and high-stakes test settings. This is because most educational communities have spent decades developing expertise in the use of selected-response questions, and this is apparent in the availability of consistent itemwriting guidelines, approaches to measuring higher-level thinking, and statistical item-analysis techniques.<sup>6,7</sup> Other question types, such as constructed response, have more recently been developed and implemented in computerbased tests.<sup>8</sup> According to Parshall, there is a great deal of excitement about these "innovative item types" for online educational assessment. However, expertise in the design and development of these newer types of online test items is still evolving. The use of such question types requires faculty development and student awareness and preparation to avoid a decrease in the quality of the test.<sup>7</sup> Therefore, when considering the use of alternate question types, it is important to carefully analyze the pros and cons of each type.

Selected-response questions are widely used in several licensure examinations, including the United States Medical and Licensing Examination,<sup>9</sup> NAPLEX,<sup>1</sup> the National Council Licensure Examination for Registered Nurses,<sup>10</sup> and the National Council Licensure Examination for Practical Nurses.<sup>10</sup> The use of selected-response questions seems likely to expand because their characteristics are ideal for electronic assessment. However, despite that selected-response questions can be graded quickly and impartially (constructed-response questions may involve interpretation and subjectivity), these items also present an error in the final score: the error resulting from respondents correctly answering items by guessing or using their intuition. While statistical methods exist for estimating this error, such calculations are beyond the training of many educators, and often require large data sets to determine. Simply stated, selected-response (eg, multiple-choice) questions may pose some disadvantages.

The objectives of this study were to introduce PharmD students to changes in question types that are used in NAPLEX: constructed-response versus selected-response; measure and compare student performance on constructedresponse and selected-response questions in a pharmaceutics course; and collect student feedback on the use of differing question types.

## **METHODS**

A 10-question optional online examination (review) using Questionmark Perception (Questionmark Corporation, Norwalk, CT, www.questionmark.com) testing software was pilot tested with 51 students in a Pharmaceutics II course at the Shenandoah University Bernard J. Dunn School of Pharmacy in spring 2010. A questionnaire was administered via SurveyMonkey (SurveyMonkey, Palo Alto, CA, www.surveymonkey.com) immediately following the examination. Based on a psychometric review of the pilot study results, the authors determined 2 design modifications to improve validity and reliability: (1) the examination should be required (versus optional), and (2) the number of examination questions should be increased by at least twofold. The design changes were implemented in the subsequent study.

A modified study was conducted in fall 2011 with 46 second-year PharmD students enrolled in a required Pharmaceutics II course at the Union University School of Pharmacy (the co-author changed schools between the pilot study and modified study). The course is offered in the second year and is a requirement of the traditional PharmD program. The course includes lectures and independent learning, as well as class recitation sessions where students complete assigned problems during class to reinforce learning. The skills learned in the prerequisite "Pharmacy Calculations" (Pharmaceutics I) course remain important in Pharmaceutics II, and as such, many of the quizzes and examinations in Pharmaceutics II require students to solve calculations.

The 2011 examination was conducted using Respondus testing software (Respondus, Inc, Redmond, WA, www. respondus.com) in Blackboard; the examination consisted of 30 questions (available from the authors upon request), alternating between constructed-response and selected-response formats. Fifteen pairs of similar questions (a constructed-response question paired with a selectedresponse question) were included on the examination, with 1 question displayed per Web page. Students could not return to a Web page once an answer was submitted; this feature mimics the structure of the NAPLEX.<sup>12</sup> Presenting the constructed-response question first in the topical pair prevented potential memory recall and/or focus provided by selected-response answer choices. In order not to bias responses on the proceeding questionnaire, no test results were displayed upon exiting the test.

Immediately following the examination, students completed an online questionnaire administered via Survey-Monkey. The questionnaire asked for students' quantitative and qualitative feedback on the 2 types of questions used in the examination, including the potential positive and negative aspects of each question type. The 2010 and 2011 research studies were approved by the Institutional Review Board of each institution prior to beginning the applicable research.

Examination results were imported into Microsoft Excel and IBM SPSS Statistics 19 (IBM, Somers, NY, www.ibm.com) for quantitative analyses. Correlation of the 2 question types was measured, and the Cronbach alpha was calculated for the constructed-response questions, selected-response questions, and all questions as a test of reliability. To determine and compare performance, item mean and discrimination were calculated for each question, for all constructed-response questions, for all selected-response questions, and for the entire examination (Table 1).

Pearson chi-square test is a common statistical analysis for examining the difference between nominal data. However, when determining significance (p), the small sample size required the use of the more accurate Fisher exact test.<sup>13-15</sup> The Fisher exact test was conducted on each pair of questions to determine if a difference existed between student performance on each question type (Table 2).

Questionnaire data were compiled by the survey software and exported to Excel for quantitative and qualitative

#### American Journal of Pharmaceutical Education 2013; 77 (1) Article 6.

	% Answering Correctly:	% Answering Correctly:	% Answering Correctly:		
Q# <sup>a</sup>	Whole Group	Upper 25%	Lower 25%	Discrimination <sup>b</sup>	Mean (SD)
1	84.8	100.0	50.0	0.5	84.8 (36.3)
2	95.7	100.0	83.3	0.2	95.7 (20.6)
3	89.1	100.0	83.3	0.2	89.1 (31.5)
4	84.8	91.7	58.3	0.3	84.8 (36.3)
5	93.5	100.0	83.3	0.2	93.5 (25.0)
6	76.1	83.3	58.3	0.3	76.1 (43.1)
7	82.6	100.0	66.7	0.3	82.6 (38.3)
8	80.4	100.0	58.3	0.4	80.4 (40.1)
9	84.8	91.7	75.0	0.2	84.8 (36.3)
10	93.5	100.0	83.3	0.2	93.5 (25.0)
11	91.3	100.0	83.3	0.2	91.3 (28.5)
12	95.7	100.0	91.7	0.1	95.7 (20.6)
13	73.9	91.7	58.3	0.3	73.9 (44.4)
14	89.1	100.0	91.7	0.1	89.1 (31.5)
15	56.5	100.0	16.7	0.8	56.5 (50.1)
16	84.8	100.0	66.7	0.3	84.8 (36.3)
17	78.3	91.7	66.7	0.3	78.3 (41.7)
18	95.7	100.0	83.3	0.2	95.7 (20.6)
19	19.6	41.7	0.0	0.4	19.6 (40.1)
20	34.8	58.3	8.3	0.5	34.8 (48.2)
21	65.2	66.7	50.0	0.2	65.2 (48.2)
22	87.0	100.0	75.0	0.3	87.0 (34.1)
23	84.8	100.0	50.0	0.5	84.8 (36.3)
24	84.8	100.0	66.7	0.3	84.8 (36.3)
25	43.5	58.3	33.3	0.3	43.5 (35.6)
26	45.7	75.0	25.0	0.5	45.7 (50.4)
27	65.2	83.3	16.7	0.7	65.2 (48.2)
28	89.1	100.0	83.3	0.2	89.1 (31.5)
29	43.5	66.7	33.3	0.3	43.5 (50.1)
30	23.9	41.7	8.3	0.3	23.9 (43.1)
М	73.9	88.1	56.9	0.3	73.9 (36.9)
CR	70.4	86.1	51.1	0.4	70.4 (39.4)
SR	77.4	90.0	62.8	0.3	77.4 (34.5)

Table 1. Examination results showing items Distriminations and means for ringh and Low remainers $(n=40)$	Table	1. Examin	ation H	Results	Showing	Items'	Di	scriminat	ions and	d Means	for	High	and	Low	Performers	(n = a)	46)
---	-------	-----------	---------	---------	---------	--------	----	-----------	----------	---------	-----	------	-----	-----	------------	---------	-----

Abbreviations: M = mean; CR = constructed-response; SR = selected-response.

<sup>a</sup> Paired rows indicate same topic. Odd question numbers were constructed-response; even question numbers were selected-response.

<sup>b</sup> If a question has a discrimination value of at least 0.4, it did a good job discriminating between the high and low-scoring students, ie, most students who scored in the top 25% of the class (mean test score) entered a correct answer for this question, while most students who scored in the bottom 25% of the class entered an incorrect answer for this question.<sup>18</sup>

analyses. For each quantitative question, the percent for each response option was calculated. Open-ended question responses were coded for qualitative analysis.

## RESULTS

## Examination

One hundred percent of the students (21 males and 25 females) in the 2011 class completed the examination. The Pearson correlation between the performance on constructed-response and selected-response questions was high (r = 0.82, p = 0.00). Cronbach alpha analyses showed good reliability of the test questions: constructedresponse questions ( $\alpha = 0.64$ , n = 15); selected-response questions ( $\alpha = 0.68$ , n = 15); and all questions ( $\alpha = 0.82$ , n = 30).

The mean score on the examination was 22.2 out of 30 points, or 73.9%. Based on mean comparisons, the class scored higher on the constructed-response question in 4 topics and higher on the selected-response question in 10 topics (Table 1).

CR/SR Question Pair	Topic	P <sup>a</sup>	
1	Conversions	0.16	
2	Spoonful	1.00	
3	Measurement	0.04 <sup>b</sup>	
4	Percentage	1.00	
5	Concentration	0.32	
6	Concentration 2	0.68	
7	Mixture	0.11	
8	Specific Gravity	0.01 <sup>b</sup>	
9	Ratio and Proportion	0.03 <sup>b</sup>	
10	Milliequivalents	0.16	
11	Single Dilution	0.03 <sup>b</sup>	
12	Double Dilution	1.00	
13	Endotoxins	1.00	
14	USP	0.01 <sup>b</sup>	
15	IV Infusion	0.08	

Table 2. Significance of Performance Differences BetweenQuestion Types on Each Topic

Abbreviations: CR = constructed-response; SR = selected-response; USP = United States Pharmacopeia; IV = intravenous.

<sup>a</sup> Results of Fisher exact test.

<sup>b</sup> Significant at the 0.05 level.

The Fisher exact test revealed that for 5 topics (question pairs), the results (number of correct vs. incorrect responses) differed significantly by question type (Table 2). The constructed-response results were significantly better for topic 3 ("measurement," questions 5 and 6), while the selected-response results were significantly better for the other 4 topics. No significant difference in the number of correct responses was found between question types on the other 10 topics.

#### Questionnaire

All 46 students also responded to the postexamination survey (Tables 3-5). The students reported much more confidence in correctly answering selectedresponse calculations questions than they reported confidence in correctly answering constructed-response calculations questions (76.0% vs. 28.3%) and more confidence in correctly answering constructed-response non-calculations questions than confidence in correctly answering constructed-response calculations questions (39.1% vs. 28.3%) (Table 4). The authors included the distinction between calculations and non-calculations questions to determine if confidence is affected more by question type (selected response vs. constructed response) or answer type (constructed words vs. constructed numbers).

A summary of the pros and cons that students identified regarding each question type are given in Table 5. The predominant pros of constructed-response questions were that they provide a better measure of knowledge Table 3. Post-Examination Questionnaire Student Feedback (n=46)

Questions	%
Which question type do you prefer?	
Selected response	80.4
Constructed response	0.0
It depends on the content/subject matter of the question	19.6
Which question type takes longer for you to answer/solve?	
Selected response	2.2
Constructed response	71.7
Not much difference	26.1
On constructed-response questions, your answers are typically:	
Usually correct	19.6
Usually incorrect	15.2
Balanced between correct and incorrect	65.2
If you know the quiz/test format in advance, do you study differently for a constructed-response format than for a selected-response format?	
Yes	32.6
No	67.4
Which question type do you think prepares you better for a career in health care?	
Selected response	52.2
Constructed response	47.8

(30.4%), fewer distractions/less second-guessing (21.7%), and more leeway in answers (21.7%). The predominant cons of constructed-response questions were restrictions in the software's ability to recognize their answer (17.5%), no means of cross-checking their answer (15.2%), and no focus (narrowing clue) provided for question/answer (15.2%). The predominant pros of selected-response questions were the ability to double-check answers (26.1%), the ability to select/guess the answer (19.6%), and the focus provided by the response choices (17.4%). The predominant cons of selected-response questions were hesitation/second-guessing (37.0%), and the inclination to study less and guess more (17.4%).

No students universally preferred constructedresponse questions for all subject matter. With regard to preparing them for a career in healthcare, however, 47.8% of students indicated that constructed-response questions were probably the better option. These students explained that they would not be presented with a list of potential answers on the job, while students who believed selected-response questions provided better preparation for their pharmacy career commented that they would be able to verify answers on the job using available resources.

## American Journal of Pharmaceutical Education 2013; 77 (1) Article 6.

	Responses, %						
How confident are you in correctly answering the following types of questions?	No Confidence	Little/Low Confidence	Neutral	Confident	Very Confident		
Selected-response calculations	0.0	2.2	21.7	54.3	21.7		
Constructed-response calculations	2.2	28.3	41.3	26.1	2.2		
Selected-response non-calculations	0.0	2.2	26.1	60.9	10.9		
Constructed-response non-calculations	6.5	17.4	37.0	34.8	4.3		

Table 4. Students' Confidence Levels With Each Question Type

#### DISCUSSION

There was a significant difference in pharmacy students' performance on 5 of the 15 question pairs on the examination. Ideally, no significant differences would have been found (ie, students would have performed equally on both question types). Significant differences may indicate that students need additional instruction in certain topics, additional practice with certain question types, and/or faculty need additional training to improve test development.<sup>5</sup> Because questions topics (subject matter) may impact constructed-response vs selected-response results,<sup>11</sup> further research at the program level is needed on calculations question topics to determine if students may struggle with constructed-response questions on a particular topic. If so, additional constructed response emphasis may be needed for those topics so students

Table 5. Frequencies of Coded Open-Ended Responses to Pros and Cons of Each Question Type

Question	Response Coding	Frequency No. (%) <sup>a</sup>		
What do you think are the pros of	Better measure of knowledge (vs. recognition)	14 (30.4)		
constructed-response questions?	tructed-response questions? Fewer distractions / less second-guessing			
	More leeway in answers	10 (21.7)		
	Can receive partial credit	4 (8.7)		
	Better long-term retention	3 (6.5)		
	More critical thinking required	3 (6.5)		
	No pros to constructed-response questions	3 (6.5)		
What do you think are the cons of	Restrictions in software answer recognition	8 (17.4)		
constructed-response questions?	Student has no means of cross-checking answer	7 (15.2)		
	No focus provided for question/answer	7 (15.2)		
	Student not able to guess	6 (13.0)		
	Increased subjectivity in grading	5 (10.9)		
	Student needs more time to answer questions	5 (10.9)		
	Student feels more stress/anxiety	2 (4.3)		
	Other	8 (17.4)		
What do you think are the pros	Student can double-check the answer	12 (26.1)		
of selected-response questions?	Student can guess/select the answer	9 (19.6)		
	Options can provide focus	8 (17.4)		
	Student can use process of elimination	4 (8.7)		
	Student can answer questions faster	3 (6.5)		
	Software can provide faster grading and feedback	3 (6.5)		
	Student can work backwards and find answer	2 (4.3)		
	Other	5 (10.9)		
What do you think are the cons of selected-response questions?	Must read carefully / similar answers cause hesitation / second-guessing	17 (37.0)		
	No cons to selected-response questions	9 (19.6)		
	Student may study less / guess more	8 (17.4)		
	No gray area or explanations allowed	3 (6.5)		
	Potential for unclear response choices	2 (4.3)		
	No partial credit	2 (4.3)		
	Other	5 (10.9)		

<sup>a</sup> The constructed-response percentages total > 100.0 because some responses were coded into multiple categories.

can perform well on the related NAPLEX questions/ topics.

Greater awareness that the NAPLEX calculations questions are now constructed-response is needed among pharmacy faculty and students so that they can instruct/ prepare for this question type. Students' lack of confidence in answering constructed-response calculations questions was reiterated by the questionnaire feedback on the constructed-response question type's cons, such as the inability to cross-check answers and the likelihood of second-guessing answers. Students also remarked that they must study content more thoroughly to prepare for a quiz or examination that includes constructed-response questions. These comments support the need to introduce students to constructed-response questions and allow them to practice answering this type of test item prior to taking the NAPLEX. The authors believe that introducing the constructed-response question type early in a PharmD curriculum will instill better study techniques in students, lower any stress associated with completing this type of test item, and better prepare students for taking the revised NAPLEX.

Other noted cons to the constructed-response question type, such as lack of focus (no narrowing clue like that which is available from selected-response answer choices) and restrictions in the answer-recognition capability of testing software, indicate the importance of faculty members teaching course topics clearly, writing the corresponding questions clearly, and fully utilizing available software features (eg, many online testing programs allow some flexibility in users' responses to constructed-response questions). If an instructor chooses to allow only a specific response to an item, students should be informed that this restriction was intentional by the instructor and not a shortcoming of the software.

Because this study was conducted with a single course at a private institution, the results may not be transferable to other courses or colleges and schools of pharmacy. The small number of students enrolled in the pharmaceutics course created limitations for evaluating the reliability and validity of the question items and question set.<sup>16,17</sup> The authors attempted to obtain the best possible response rate and improve students' motivation to perform well by making the test mandatory and graded. The lack of anonymity of the associated questionnaire (also mandatory) may have affected those results. For example, although students were not graded on their feedback (merely on whether it was completed), some students may have chosen responses or entered comments that they felt were expected by the professor.

## CONCLUSION

Students correctly answered more selected-response questions than constructed-response questions (77.4% vs. 70.4%) and felt more confident in doing so. However, changes to the NAPLEX calculations questions require faculty members and students to adapt to differing question types. In order to prepare students for this revision, colleges and schools of pharmacy should introduce and reinforce constructed-response calculations through teaching and formative and summative course assessments. Increasing pharmacy students' familiarity and comfort level with this calculations question type should help them successfully complete such questions on the NAPLEX.

#### ACKNOWLEDGEMENTS

The authors thank Dr. David Newton, a professor at Shenandoah University, for informing the authors of the NAPLEX change.

The ideas expressed in this manuscript are those of the authors and in no way are intended to represent the position of the American Association of Colleges of Pharmacy, the National Association of Boards of Pharmacy, or any other pharmacy organization.

#### REFERENCES

1. National Association of Boards of Pharmacy (NABP). NAPLEX. http://www.nabp.net/programs/examination/naplex/. Accessed September 16, 2012.

2. National Association of Boards of Pharmacy (NABP). NAPLEX fact sheet. http://www.nabp.net/assets/NAPLEX.pdf. Accessed September 16, 2012.

3. SRI International. Online evaluation resource library. http://oerl. sri.com/definitions.html. Accessed September 16, 2012.

4. Ventouras E, Triantis D, Tsiakas P, et al. Comparison of examination methods based on multiple-choice questions and constructed-response questions using personal computers. *Comput Educ.* 2010;54(2):455-461.

5. Lissitz R, Hou X. Multiple choice items and constructed response items: does it matter? http://www.education.umd.edu/EDMS/ MARCES/multiple choice items and constructed response items.pdf. Accessed September 16, 2012.

6. Council on Licensure, Enforcement, and Regulation (CLEAR). CLEAR Exam Review (CER). 2008;19(2). http://www.clearhq.org/ resources/CLEAR\_summer08\_4.pdf. Accessed September 16, 2012.

 Haladyna TM. Writing Test Items to Evaluate Higher Order Thinking. Allyn and Bacon; Needham Heights, MA; 1997.
Parshall CG, Spray JA, Kalohn JC, Davey T. Practical

Considerations in Computer-Based Testing. New York: Springer-Verlag; 2002.

9. United States Medical Licensing Examination (USMLE). 2012 Bulletin of Information. USMLE Secretariat, Philadelphia, PA. http:// www.usmle.org/pdfs/bulletin/2012bulletin.pdf. Accessed September 16, 2012.

## American Journal of Pharmaceutical Education 2013; 77 (1) Article 6.

10. National Council Licensure Examination for Registered Nurses (NCLEX-RN) https://www.ncsbn.org/2012\_NCLEX\_Candidate\_Bulletin.pdf. Accessed September 16, 2012.

11. Ventouras E, Triantis D, Tsiakas P, et al. Comparison of oral examination and electronic examination using paired multiple-choice questions. *Comput Educ.* 2011;56(3):616-624.

12. National Association of Boards of Pharmacy (NABP). NAPLEX and MPJE registration bulletin. http://www.nabp.net/programs/assets/NAPLEX-MPJE.pdf. Accessed September 16, 2012.

13. UCLA: Academic Technology Services, Statistical Consulting Group. Statistical analyses using SPSS: What statistical analysis should I use? http://www.ats.ucla.edu/stat/sas/whatstat/default.htm. Accessed September 16, 2012. 14. McDonald JH. Fisher's exact test of independence. In: *Handbook of Biological Statistics*, 2nd ed. Sparky House Publishing, Baltimore, MD; 2009:70-75.

15. McDonald JH. Small numbers in chi-square and G-tests. In: *Handbook of Biological Statistics*, 2nd ed. Sparky House Publishing, Baltimore, MD; 2009:80-83.

16. Traub RW, Rowley GL. Understanding reliability. *Educ Meas: Issues Pract.* 1991;10(1):37-45.

17. Crocker L, Algina J. Introduction to validity. In: *Introduction to Classical and Modern Test Theory*. Fort Worth: Holt, Rinehart, and Winston; 1986.

18. Crocker L, Algina J. Item analysis. In: *Introduction to Classical and Modern Test Theory*. Fort Worth: Holt, Rinehart, and Winston; 1986.