

RESEARCH

Pharmacy Students' Ability to Think About Thinking

Eric F. Schneider, PharmD,^a Ashley N. Castleberry, PharmD, MAEd,^b Jasna Vuk, PhD, MD,^c Cindy D. Stowe, PharmD^d

^aUniversity of Waterloo School of Pharmacy, Waterloo, Ontario, Canada

^bUniversity of Arkansas for Medical Sciences College of Pharmacy, Little Rock, Arkansas

^cUniversity of Arkansas for Medical Sciences Office of Educational Development, Little Rock, Arkansas

^dSullivan University College of Pharmacy, Louisville, Kentucky

Submitted January 4, 2014; accepted March 4, 2014; published October 15, 2014.

Objective. To investigate students' metacognitive skills to distinguish what they know from what they do not know, to assess students' prediction of performance on a summative examination, and to compare student-identified incorrect questions with actual examination performance in order to improve exam quality.

Methods. Students completed a test-taking questionnaire identifying items perceived to be incorrect and rating their test-taking ability.

Results. Higher performing students evidenced better metacognitive skills by more accurately identifying incorrect items on the exam. Most students (86%) underpredicted their performance on the summative examination (actual=73.6 ± 7.1 versus predicted=63.7 ± 10.5, $p < 0.05$). Student responses helped refine items and resulted in examination changes.

Conclusion. Metacognition is important to the development of life-long learning in pharmacy students. Students able to monitor what they know and what they do not know can improve their performance.

Keywords: metacognition, life-long learning, test-taking ability

INTRODUCTION

Metacognition is a key component of life-long learning. One of the first definitions of metacognition was offered by Flavell as the knowledge of one's own cognition.¹ Simplistically, metacognition has been referred to as "thinking about thinking."² In general, metacognition can be separated into two broad areas: knowledge of and regulation of cognition.³ Knowledge of cognition relates to an individual's self-awareness of skills and previously learned information, while regulation of cognition relates to an individual's use of strategies to think and learn new information. Examples of skills related to knowledge of cognition include declarative (awareness of what one knows), procedural (awareness of how to process information to solve problems), and conditional (awareness of when and how to apply what one knows).⁴ Conversely, regulation of cognition involves activities that control one's own learning, such as planning strategies for learning and resource allocation, evaluation and appraisal of learned information, and monitoring the awareness and

degree of comprehension.⁴ Collectively, metacognitive skills allow learners to be aware of what they know and focus on learning what they do not know. Metacognitive monitoring, which is the focus of this study, was conceptualized by Dutke, Barenberg, and Leopold as a relationship between learners' subjectively perceived behavior and their actual behavior in learning and performance.⁵

The Accreditation Council for Pharmacy Education (ACPE) emphasizes the need for life-long learning skills for a pharmacy graduate.⁶ According to the 2007 ACPE Standards, the curriculum should focus on enabling students to be active, self-directed, life-long learners in order to foster development of an effective professional.⁶ A few studies have researched pharmacy students' metacognitive skills to monitor their own learning and performance. Garrett et al assessed metacognitive skills students need to monitor their learning.⁷ Austin and Gregory explored pharmacy students' ability to self-assess their clinical knowledge and communication.⁸ To date, no studies have examined pharmacy students' ability to monitor and predict their performance on examinations. However, it is important that students can accurately do this because students with high prediction accuracy of their performance have been shown to manage their time and efforts efficiently and maximize their preparedness for

Corresponding author: Ashley Castleberry, PharmD, MAEd, University of Arkansas for Medical Sciences College of Pharmacy, 4301 West Markham Street Slot #522-4, Little Rock, AR 72205. Tel: 501-686-8383. Fax: 501-296-1168. E-mail: ancastleberry@uams.edu

examinations.^{9,10} Everson and Tobias emphasized the importance of students' ability to estimate their own knowledge, especially in "complex domains such as science and engineering, or making diagnoses in medicine and other fields. Students who can accurately distinguish between what they know and do not know should be at an advantage while working in such domains, since they are more likely to review and try to relearn imperfectly mastered materials needed for particular tasks, compared to those who are less accurate in estimating their own knowledge."¹¹ Turan, Demirel, and Sayek investigated medical students' metacognitive awareness and self-regulated learning in relation to a curricular model that medical schools employed and found that students who experienced problem-based learning demonstrated higher metacognitive awareness and self-regulated learning skills.¹²

Student monitoring of performance has also been observed in undergraduate educational psychology classrooms.^{9,13} Hacker et al compared students' predicted percentage of correct answers on an examination with their actual percentage of correct answers. They found that students who performed well demonstrated high accuracy in estimating their performance, which improved over multiple examinations, whereas students, who did not perform well, overestimated their performance with no improvement over time.⁹ Sinkavich compared students' confidence of answer correctness on multiple-choice questions with actual correctness of their answers. The results indicated a significant difference between high and low performing students in their prediction of what they knew from what they did not know on the examination.¹³ Moreover, metacognition played an important role in students' judging the correctness of their answers because answering an item depended on students' recall of information and judgment of accuracy of that information.¹³

In addition to the role of verifying a student has adequately mastered curricular expectations, the summative examination has a subtler metacognitive goal—helping the student become a life-long learner. By developing metacognitive skills, a learner can ultimately monitor their own knowledge, make a plan to fill in gaps, and evaluate new learning.¹⁴ Despite the need for such skills, there is a lack of research about pharmacy students' metacognitive monitoring on high-stakes summative examinations. The purpose of this study is: (1) to investigate students' metacognitive skills to distinguish what they know from what they do not know on a summative examination; (2) to assess students' prediction of their performance on a summative examination; and (3) to compare student-identified incorrect questions with actual examination item performance to improve examination quality.

METHODS

The second professional year progression examination, summative examination 1, is given to qualified second year pharmacy students. This examination, constructed under the guidance of the college's assessment committee, tests global knowledge and important concepts in order to distinguish competent students from those still struggling. Each student's academic record is reviewed by the Scholastic Standing Committee to determine eligibility for the summative examination. Eligible students are those deemed ready to progress to the third professional year according to the college's academic rules of progression.

The faculty-approved blueprint for the summative examination was based on the coursework from the first and second professional years of the curriculum and the college's outcome competency statements. The summative examination was developed using an examination item bank created from faculty-authored items representing this material. All test items were multiple-choice and were rigorously reviewed by groups of faculty members (both content and non-content experts) for clarity, substance, and grammar.

All items were pretested (not scored) during the first administration in order to obtain statistics on performance for inclusion as a live (scored) item on subsequent administrations. One hundred live items were identified from the examination item bank that met the blueprint criteria. All live scored items had been pretested with a previous class of students and were selected based on item analysis and modified Angoff standard-setting data. The examination was created in Respondus 4.0 (Respondus, Inc., Redmond, WA) and loaded into Blackboard 8.0 (Blackboard, Inc., Washington, DC) for delivery to the students. The examination consisted of 125 items (100 live and 25 pretest items). In an effort to pretest more items, 50 items were randomly distributed among the students.

Students had up to 3 hours to take the computerized, proctored examination. Items were presented one at a time and were randomized for each student. Note taking on scratch paper was allowed since students were able to skip and return to unanswered items.

In April 2010, the summative examination was administered to 107 qualified students. During administration, students were asked to complete a test-taking questionnaire. The questionnaire was divided into 2 sections. In the first section, an assessment of metacognitive skills, students were asked to identify up to 10 items they felt certain they had answered incorrectly and briefly explain why. These items were categorized by each student as: (1) "question was confusing or difficult to understand;" (2) "question was too detail oriented;" or (3) "other," with space provided for a written response. The

second section assessed the students' rating of their test-taking ability. Students were asked to express in a percentage how well they did on the examination and assess their ability to take multiple-choice examinations with a 5-item Likert scale (1=poor and 5=excellent).

Item performance data were collected from electronic administration of the examination and imported into an Access item banking database for evaluation. Item performance consisted of difficulty, measured by the percent of correct responses, and discrimination, measured by point biserial.

Data from the test-taking questionnaire were analyzed as an assessment of the students' metacognitive skills and test taking ability. Pearson correlation was used to measure association of variables. Assumptions of normality of data were not met according to the Kolmogorov-Smirnov test and the Shapiro-Wilk test, and the non-parametric tests, Kruskal-Wallis and Mann-Whitney, were used to determine differences in group comparisons. For statistical analysis, students were divided into 4 groups (quartiles) based on the range of their summative examination performance and grade point average (GPA). PASW Statistics 18 software package (SPSS Inc., Chicago, IL) was used for data analysis. Data were presented as mean \pm SD unless otherwise noted. Statistical significance was defined as a $p < 0.05$.

For the metacognitive assessment, the number of student-identified incorrect items, the accuracy of incorrect item selection, and weighted accuracy were reviewed. The weighted percentage of accurate selection of incorrect items was calculated as the percent of items accurately identified as incorrect divided by the student's total number of incorrect items on the examination. Weighting was applied as an adjustment since high performing students had less opportunity to select items that they missed than poor performing students. Correlation of overall performance was conducted with the following variables: number of items identified, accuracy, and weighted accuracy. Differences in metacognitive performance were also analyzed based on quartiles of examination performance and GPA.

Students' ability to predict summative examination performance was assessed by correlating differences in examination performance and GPA. The multiple-choice test-taking ratings were assessed using the Kruskal-Wallis test for non-normally distributed data. The Mann-Whitney test was used to determine the location of the difference.

RESULTS

One hundred seven students completed the summative examination. Of those students, 104 selected up to 10 items they felt certain they answered incorrectly. The overall score of the summative examination was correlated

with GPA and weighted accuracy ($r=0.69$ and $r=0.29$, respectively, $p < 0.05$). Table 1 summarizes student identification of incorrect items. The Kruskal-Wallis test revealed a significant difference among the 4 groups of students based on their examination performance quartiles. The Mann-Whitney test showed significant differences between the second and third quartiles and the second and fourth quartiles. The Kruskal-Wallis test did not show a significant difference among the 4 groups of students based on their GPA quartiles.

One hundred students (93.5%) provided a predicted score for the summative examination and 103 (96.3%) rated their multiple-choice test-taking ability. Test-taking assessment data are shown in Table 2.

Student prediction of summative examination performance was moderately correlated with actual performance ($r=0.41$, $p < 0.05$). Most students (86%) underpredicted their performance on the summative examination (actual 73.6 ± 7.1 vs predicted 63.7 ± 10.5 , $p < 0.05$). The difference between the actual and predicted overall score of the 100 students who provided a prediction ranged from -9.60 to 42.80. Students in the lowest performance quartile (quartile 4) were the best predictors of examination performance among the 4 groups of students. The Kruskal-Wallis test was used to assess the differences among the 4 groups based on their examination performance quartiles and GPA quartiles. There was a significant difference in the ability to predict examination performance among the 4 groups based on examination performance quartiles but not GPA quartiles. The Mann-Whitney test revealed a significant difference between

Table 1. Selection of Items Students Felt Certain They Answered Incorrectly

Examination Performance Quartiles (N)	Number of Items Accurately Selected as Incorrect ^a
Quartile 1: ≥ 81.6 (17)	1.86 ± 1.48
Quartile 2: 73.6-81.5 (41)	$1.99 \pm 0.87^{b,c}$
Quartile 3: 65.6-73.5 (35)	1.52 ± 0.59^b
Quartile 4: ≤ 65.5 (11)	1.27 ± 0.49^c
	$p < 0.05$
GPA Quartiles (N)	
Quartile 1: ≥ 3.50 (18)	2.19 ± 0.96
Quartile 2: 3.05-3.49 (27)	1.80 ± 1.20
Quartile 3: 2.60-3.04 (24)	1.62 ± 0.54
Quartile 4: ≤ 2.59 (35)	1.54 ± 0.80
	NS

N=104

^a Weighted by overall performance on the examination

^b Quartile 2 vs 3= $p < 0.05$ according to the Mann-Whitney test

^c Quartile 2 vs 4= $p < 0.05$ according to the Mann-Whitney test

Table 2. Summative examination Actual vs Predicted Performance

Group	N	Actual Score	N	Predicted Score	Difference in Scores	N	Test-Taking Ability Rating ^a
Examination Performance					$p < 0.05^b$		$p < 0.05^c$
Quartile 1: ≥ 81.6	17	85.1 ± 2.9	15	70.5 ± 5.3	14.5 ± 6.1	16	3.8 ± 0.8
Quartile 2: $73.6 - 81.5$	42	76.6 ± 2.3	38	67.9 ± 9.0	8.7 ± 9.2	39	3.7 ± 0.7
Quartile 3: $65.6 - 73.5$	37	69.1 ± 2.4	37	58.0 ± 10.7	11.1 ± 11.2	37	3.4 ± 0.7
Quartile 4: ≤ 65.5	11	61.9 ± 2.7	10	59.1 ± 8.4	2.7 ± 9.3	11	3.4 ± 0.8
Grade Point Average (GPA)					NS		$p < 0.05^d$
Quartile 1: ≥ 3.50	18	81.5 ± 5.1	17	71.5 ± 4.5	9.9 ± 7.1	18	4.0 ± 0.7
Quartile 2: $3.05 - 3.49$	30	76.5 ± 6.0	27	67.0 ± 8.7	9.6 ± 8.2	27	3.7 ± 0.5
Quartile 3: $2.60 - 3.04$	24	73.3 ± 5.1	22	59.3 ± 11.4	13.2 ± 12.1	23	3.5 ± 0.7
Quartile 4: ≤ 2.59	35	68.0 ± 5.4	34	59.9 ± 10.7	7.8 ± 10.8	35	3.3 ± 0.8

^a Likert Scale (1=poor to 5=excellent)

^b Difference in Scores by Examination Performance Quartile: 1 vs 2; 1 vs 4; 3 vs 4 according to the Mann-Whitney test

^c Test Taking Ratings by Examination Performance Quartile: 1 vs 3; 2 vs 3 according to the Kruskal Wallis test

^d Test Taking Ratings by GPA: 1 vs 3; 1 vs 4; 2 vs 4 according to the Kruskal Wallis test

students in performance quartiles 1 and 2, 1 and 4, and 3 and 4 on ability to predict performance.

Test-taking ability was also assessed by evaluating student self-perceived ability to take multiple-choice tests using a 5-point Likert scale (Tables 2 and 3). The scaled ratings ranged from 2-5 (3.56 ± 0.72). Students' estimation of their test-taking ability ($n=103$, 96.3%) was moderately correlated with their actual score on the summative examination, ($r=0.30$, $p < 0.05$), predicted score ($r=0.27$, $p < 0.05$), and GPA, ($r=0.36$, $p < 0.05$). The Kruskal-Wallis test revealed a significant difference in multiple-choice test-taking ability rating based on their summative examination performance and GPA quartiles ($p < 0.05$), depicted in Table 2.

DISCUSSION

This is the first study in the pharmacy literature to evaluate the metacognitive monitoring skills of students in a high-stakes summative examination setting. Since these skills are important in the development of life-long learning, it is critical that we assess such skills in students. Students were somewhat able to predict their examination performance but had limited success in identifying individual incorrect items. Higher performing students showed better metacognitive skills than lower performing

students by providing more accurate identification of incorrect items. This supports the idea that lower performing students are not able to identify what items they answered incorrectly.¹³ Though not significant, a trend was identified showing that students with higher GPAs were better able to identify incorrect items. This study did not reveal a causal relationship between poor performance and poor metacognition, so future studies might want to investigate this relationship as well as methods to help students develop metacognitive skills to perform better on examinations.

There was a moderate correlation between those students who rated their ability to take multiple-choice examinations and examination performance. The majority of students tended to underpredict their overall examination performance. This may be a function of the summative nature of the examination since the breadth of materials covered consisted of 2 professional academic years (approximately 70 semester credit hours). Additionally, a summative examination is concept-based and does not focus on memorization. The lowest performers on this examination were best able to predict their performance. This was also shown in the correlation between self-perceived, multiple-choice test-taking ability and actual and predicted scores, as well as GPA. As confidence in

Table 3. Test Taking Ability with Actual and Predicted Score and GPA

Group by Likert Scale	N	Actual Score	N	Predicted Score	GPA
1=poor	0	—	0	—	—
2	4	68.8 ± 6.4	4	60.0 ± 10.8	2.49 ± 0.31
3	47	71.9 ± 7.5	46	60.8 ± 9.9	2.78 ± 0.42
4	42	75.1 ± 5.9	40	66.5 ± 11.0	3.13 ± 0.48
5=excellent	10	77.4 ± 8.3	10	67.8 ± 8.2	3.14 ± 0.63
Not answered	4	80.2 ± 6.6	7	—	—

test-taking ability increases, predicted and actual scores should also increase. The underprediction of performance in this study could be attributed to the theory of fixed and growth mindsets.¹⁵ Higher achieving students might demonstrate a growth mindset, realizing there is more to learn and thus underestimating their achievement, while lower achieving students might have a more fixed mindset, feeling they are achieving what they need to know and thus, accurately predicting performance.

The students' metacognitive assessment served to validate the evaluation process because 80% of the items students frequently selected failed to meet the criteria for continued inclusion on the examination. The reasons for item difficulty may have included changed curricular content, poor item construction, or items whose focus was too granular for a summative examination. Items frequently selected by students were reviewed in greater detail and 25 that at least 10 students missed were identified as being confusing or too detailed. Of these 25 items, 20 were removed from the pool based on the psychometric criteria for continued use on the summative examination. Thus, results from the questionnaire helped improve the quality of the examination.

As with any project that uses data derived from a real-time assessment, which undergoes continual evaluation, changes and modifications to the process and application result in limitations on reporting and fundamental measurements. This project's constrained sample size likely limited the ability to achieve significance in many comparisons. Scores could have also been affected by students' testwiseness, which was not measured.¹⁶ Additionally, this study reflected early administration of the examination. As more data is gathered on items, the examination will improve and focus on content more conceptual and less granular in nature.

We believe students who improve metacognitive skills will improve performance on examinations because these students will be better able to identify what they do not know and correct the gap in knowledge. Metacognition can be measured in multiple settings by various tools. Repeated use of metacognitive assessment on formative examinations with feedback could result in improved metacognitive skills. In addition to these quantitative data, qualitative assessment measures of metacognition using portfolio and reflective exercises should also be explored. The use of this tool within courses on interim examinations will be explored to determine if metacognitive skills differ on discrete amounts of educational material over a defined period of time. Moreover, further investigation between metacognition and life-long learning should be a focus in pharmacy education as instructors search for methods to increase both in students.

CONCLUSION

Metacognition is important to the development of life-long learning in pharmacy students. Students should be able to monitor what they know and what they don't know in order to improve performance. Using a summative examination, higher performing students evidenced better metacognitive skills by better identifying incorrect items. However, lower performing students were better able to predict their overall performance on the examination. Student responses to the metacognitive questionnaire helped refine the items and resulted in changes to the summative examination. Future studies should investigate methods to increase metacognition in pharmacy students in an effort to improve summative performance.

REFERENCES

1. Flavell JH. Metacognitive aspects of problem solving. In: Resnick LB, ed. *The Nature of Intelligence*. Hillsdale, NJ: Erlbaum. 1976;231-235.
2. Jacobs JE, Paris SG. Children's metacognition about reading: Issues in definition, measurement, and instruction. *Educ Psychol*. 1987;22(3):255-278.
3. Schraw G. Promoting general metacognitive awareness. *Instr Sci*. 1998;26:113-125.
4. Schraw G, Moshman D. Metacognitive theories. *Educ Psychol Rev*. 1995;7(4):351-371.
5. Dutke S, Barenberg J, Leopold C. Learning from text: Knowing the test format enhanced metacognitive monitoring. *Metacogn and Learn*. 2010;5(2):195-206.
6. Accreditation Council for Pharmacy Education: Accreditation standards and guidelines for the professional program in pharmacy leading to the doctor of pharmacy degree. <http://www.acpe-accredit.org/standards/default.asp>. Accessed January 3, 2014.
7. Garrett J, Alman M, Gardner S, Born C. Assessing students' metacognitive skills. *Am J Pharm Educ*. 2007;71(1):Article 14.
8. Austin Z, Gregory PA. Evaluating the accuracy of pharmacy students' self-assessment skills. *Am J Pharm Educ*. 2007;71(5): Article 89.
9. Hacker DJ, Bol L, Horgan DD, Rakow EA. Test prediction and performance in a classroom context. *J Educ Psychol*. 2000; 40(1):160-170.
10. De Carvalho Filho MK. Confidence judgments in real classroom settings: Monitoring performance in different types of tests. *Int J Psychol*. 2009;44(2):93-108.
11. Everson HT, Tobias S. The ability to estimate knowledge and performance in college: A metacognitive analysis. *Instr Sci*. 1998;26:65-79.
12. Turan S, Demirel O, Sayek I. Metacognitive awareness and self-regulated learning skills of medical students in different medical curricula. *Med Teach*. 2009;31(10):e477-483.
13. Sinakovich FJ. Performance and metamemory: do students know what they don't know? *J Instr Psychol*. 1995;22(1):77-87.
14. Eva KW, Cunnington JPW, Reiter HI, Keane DR, Norman GR. How can I know what I don't know? Poor self-assessment in a well-defined domain. *Adv Health Sci Educ Theory Pract*. 2004;9(3):211-224.
15. Dweck, CS. *Mindset: The new psychology of success*. New York, NY: Random House;2006.
16. Rogers W, Yang P. Test-wisness: its nature and application. *Eur J Psychol Assess*. 1996;12(3): 247-259.