RESEARCH

Changes in Pharmacy Students' Metacognition Through Self-Evaluation During Advanced Pharmacy Practice Experiences

Sarah A. Nisly, PharmD, Jamie Sebaaly, PharmD, Amy G. Fillius, PharmD, Wesley R. Haltom, PharmD, Melissa M. Dinkins, PharmD

Wingate University, School of Pharmacy, Wingate, North Carolina

Submitted January 5, 2019; accepted June 10, 2019; published January 2020.

Objective. To explore whether metacognition can be improved in Doctor of Pharmacy (PharmD) students through routine self-assessment over a year-long advanced pharmacy practice experience (APPE) sequence.

Methods. Differences between self-assessment scores and preceptors' scores for three cohorts of pharmacy students between 2015 and 2018 were compared between the first, second, and third trimester to determine whether students more accurately evaluated their performance over time. The primary endpoint was change in the absolute difference between student and preceptor evaluation (rubric and composite scores) between trimesters.

Results. Of 2577 student and preceptor evaluations eligible for inclusion, 1713 were completed, matched, and analyzed. Using the same rubric as preceptors, students overestimated their performance by an average of 16 points during the first trimester, followed by 14 and 12 points during the second and third trimester, respectively. This reflected a significant improvement over time. No significance difference was found between student and preceptor composite scores. Faculty preceptorship, students' pre-APPE grade point average, and type of APPE were not associated with any difference in rubric or composite scores.

Conclusion. This analysis revealed that the difference between student self-evaluation grades and preceptor evaluation grades was greatest during the first trimester and significantly decreased in the second and third trimesters. This could reflect students' development of metacognitive processes over time. Metacognition is a vital skill for pharmacy students to learn, and opportunities to develop this skill should be incorporated throughout the pharmacy curricula.

Keywords: self-evaluations, advanced pharmacy practice experience, metacognition

INTRODUCTION

Metacognition, described as “thinking about thinking,” refers to one’s ability to regulate thinking and learning through three self-assessment skills: planning, monitoring, and evaluating.¹ In health professions, metacognition is an important consideration in the process of becoming a professional and practicing as a clinician. Health professionals must subscribe to lifelong learning, and metacognition guides this process by making the learner or clinician aware of what they do not know and prompting them to bridge the knowledge gaps. Self-awareness of one’s knowledge, skills, abilities, beliefs, motivation, and emotions has been recognized by the Accreditation Council for Pharmacy Education (ACPE)

in Standards 2016 as a key element of personal and professional development (Standard 4).²

Self-awareness is a non-cognitive skill that may be viewed as difficult to foster and assess in health professions learners, but it is a skill that can be developed.³ Self-evaluation assignments have been used in health professions curricula to stimulate reflection on didactic and experiential learning and to assess learners’ self-awareness of knowledge and skills. Change in accuracy of self-evaluations is one measure that can be used to indicate change in the metacognitive process as self-evaluation is an important component of metacognition.¹ In several studies, pharmacy learners have been found to overestimate their performance when compared to the results on evaluations conducted by faculty members or preceptors.⁴⁻⁷ This is consistent with findings in medical and nursing students.^{8,9} Additionally, top performers have been found to more accurately predict or evaluate their

Corresponding Author: Sarah A. Nisly, Wingate University, School of Pharmacy, 515 N Main St., Wingate, NC 28174. Tel: 336-713-1938. Email: s.nisly@wingate.edu.

performance than bottom (third or fourth quartile) performers in studies of pharmacy and medical students, and underperformers tend to highly overestimate their knowledge and skills.^{4,5,8} Findings in some studies suggest that while overall performance scores might be similar between instructor and learner, particular areas of rubrics tend to have larger discrepancies in scores. Examples of rubric items reported with higher score discrepancies include empathy, team relationships, industriousness and enthusiasm, personal attitudes, and global knowledge.^{5,10-11}

Approaches to developing metacognition in health professions students have been described in the literature. Steuber and colleagues evaluated implementation of metacognition processes in student pharmacists taking a semester-long elective course. Strategic feedback was provided by faculty members and peers, which resulted in improved student prediction of performance and self-awareness. Despite the improvement seen, this study found that one semester was not enough time to fully realize the effects of interventions on metacognition.⁴ Mort and Hansen studied the impact of video review on first-year student pharmacists' self-assessment of communication skills following a patient counseling encounter.⁶ Using the same rubric as faculty evaluators, student pharmacists evaluated their encounter prior to and after viewing the recording. Students' self-evaluations were more highly correlated with faculty scores after viewing their video compared to their self-evaluation before viewing the video, suggesting that video review increased student pharmacists' awareness of skill achievement.⁶ However, students' overestimation of their performance remained post-video review, although to a lesser extent. Conversely, use of videoed encounters in a small cohort of physical therapy students did not lead to significant changes in performance between mid-term and final examinations, providing conflicting evidence about the use of video recording and review of performance.¹² Baxter and Norman evaluated senior-year nursing students' ability to self-assess their performance during a simulated emergency situation.¹³ While the simulation did increase their perceived confidence and competence, it had no effect on their perceived ability to collaborate or communicate. The authors went on to state that they believe self-assessments do not provide any additional benefits in development of a student's metacognition and may in fact provide a faux level of confidence.¹³ This conflicting evidence highlights the difficulty of developing metacognition in learners over the course of a semester or during individual events.

This study was designed to explore whether metacognition could be developed over a year-long advanced pharmacy practice experience (APPE) sequence through

repeated use of self-assessment at the end of each APPE. Differences between self-assessment scores and instructor scores for three cohorts of student pharmacists were compared between the first, second, and third trimesters to determine whether students more accurately evaluated their performance near the end of the APPE sequence as compared to the beginning. We hypothesized that repetition of self-assessment, along with preceptor feedback throughout the APPE year, would lead to a smaller difference between student and preceptor evaluation scores, indicating increased self-awareness and metacognitive change throughout the APPE year.

METHODS

Completed student and preceptor matched final evaluations for all APPEs between September 2015 and April 2018 were eligible for analysis. These dates were chosen because of the implementation of identical student and preceptor evaluation rubrics in September 2015. Students who failed to complete a self-evaluation or did not receive an evaluation from their preceptor were excluded. The Wingate University School of Pharmacy APPE curriculum consists of the following required five-week rotations: one adult internal medicine, one ambulatory care, one advanced institutional, one long-term care, and two community experiences. The three remaining APPEs are student selected electives. In total, students completed 1800 hours of APPEs over three trimesters of rotations. Students received a graded evaluation, which sum to 100 points, with the following breakdown: A=89.5-100 points; B=79.5-89.4 points; C=69.5-79.4 points; F=<69.5 points.

Each APPE was curated by the student's preceptor, but usually encompassed supplemental activities beyond patient care, including journal clubs, patient case presentations, and drug information requests. Preceptors and students completed identical evaluations at the midpoint and the conclusion of each experience through an online system. Evaluations consist of a detailed rubric, broken into different sections (Table 1). Supplementary activities (eg, journal clubs, patient case presentations) were evaluated separately using standardized rubrics and manually added to the midpoint and final evaluations as dictated by APPE type or preceptor. Supplementary activities occurred at timepoints dictated by the preceptor. As such, students may have received their grade and verbal feedback for the activity prior to the final evaluation for inclusion in their self-assessment. If none had been received prior to completing the final evaluation, students manually entered in a predicted score for each activity. Student completion of identical supplementary activity grading rubrics is not mandated through the electronic grading system. The combination of the scored rubric and

Table 1. Percentage of Individual Preceptor Evaluation Components by APPE Type

Evaluation Component	Internal Medicine, %	Ambulatory Care, %	Advanced Institutional, %	Community, %	Long-Term Care, %	Patient Care Elective, %	Non-Patient Care Elective, %
Verbal and written communication	21.4	20	16.7	21.4	18.5	20	22.2
Maturity, initiative, responsibility, and dependability	14.3	13.3	16.7	14.3	12.5	13.3	22.2
Understanding of pharmacotherapy	14.3	13.4	16.7	28.6	12.5	13.4	-
Medication use process	-	-	33.2	21.4	-	-	-
Patient-centered pharmaceutical care	35.7	40	-	-	44	40	-
Evidence-based medicine/information retrieval and evaluation	14.3	13.3	16.7	14.3	12.5	13.3	22.2
Problem-solving, planning, and integration and application of information	-	-	-	-	-	-	33.4
Additional evaluation components ^a	30	30	30	30	30	30	30

^a The composite evaluation is weighted with 70% comprising the main components, which are presented in this table, and the remaining 30% consisting of additional evaluation components, as appropriate (eg, case presentation, journal club, site-specific projects)

individual supplementary activities grades received was considered the composite score for the rotation.

Each participant received electronic reminders prior to due dates with final evaluations due one week after the APPE ended. To encourage both student self-reflection and student provision of feedback to the preceptor, students received a 3% grade reduction for failure to complete the preceptor and self-assessments beginning in May 2016. The grading rubric Likert scale (Table 2) used in each evaluation was electronically converted to a numeric grade.

Student self-evaluations were paired with assigned preceptor evaluations, comprising the data for this study. Data was garnered from the electronic evaluation

software and de-identified prior to data analysis. Beyond self-evaluation and preceptor scores for each individual rotation pair, investigators captured the type of rotation, rotation trimester, and grade point average (GPA) prior to the APPE year.

The primary endpoint was change in the absolute difference between student and preceptor scores over trimesters. The absolute difference in rubric scores and composite scores was considered independently over each trimester. Secondary endpoints included the impact of pre-APPE GPA, faculty preceptors, and rotation type. A one-way ANOVA was performed to identify any significant difference in student and preceptor rubric and composite scores between trimesters. A Tukey HSD post-hoc analysis was done to identify where the difference between trimesters was found. A logistic regression analysis was completed to identify factors contributing to the statistical difference found in the primary endpoint. All analyses were performed using SPSS, Version 25 (IBM, Armonk, NY).

RESULTS

A total of 2577 student and preceptor evaluations were eligible for inclusion during the study period. Completed

Table 2. Likert Ranking Terminology within Evaluations

Likert Ranking	Description
1	Unable, Fails to, Never
2	Inconsistently, Sometimes, Usually Unable
3	Often, Usually, Typically
4	Consistently, Appropriately
5	Always, Completely, Exceeds Expectations

student and preceptor evaluations were matched, resulting in 1713 evaluations for analysis. The inclusion of only trimesters two and three resulted in 282 matched evaluations for the graduating class of 2016. There was an increase in completed evaluations by the graduating classes of 2017 and 2018 (696 and 735 evaluations, respectively).

The combined data points by trimester showed a reduction in absolute difference between preceptor and student rubric scores over time (Table 3). This resulted in a significant difference between student and preceptor scores for the third trimester in comparison to each of the first two trimesters. There was no significance noted between trimesters when comparing the absolute difference in composite scores. No significant difference was seen in either rubric or composite scores across APPE types (Table 4).

In a regression analysis, the following factors were considered predictors of a significant absolute difference in student and preceptor rubric scores: long-term care and ambulatory care APPEs and first trimester scores ($R\ 0.32$, $p<.0001$). Faculty preceptors, other APPEs, and graduation year were not significant in the regression analysis. Finally, students were divided into top, middle, and low performers by GPA prior to APPE year. No significant difference in the primary endpoint was found between any GPA ranking.

DISCUSSION

This analysis revealed that the absolute difference between pharmacy students' self-evaluation grades and preceptor evaluation grades was largest during the first trimester. In addition, the difference in grades significantly decreased during the third trimester compared with the first and second trimesters. This could be the result of students developing a metacognitive process from repeated exposure to the grading rubric over time. As part of the program, students must continually monitor and reflect upon their performance in order to thoughtfully complete self-evaluations. With significant overlap of evaluation components across APPE types, the repeat exposure is likely to "sharpen" student's ability to think

critically about their output. Unfortunately, during all trimesters, students overestimated performance on the self-evaluation rubric. Significance was lost when non-patient care activities (eg, journal clubs and presentations) were incorporated into the final composite scores. While the supplementary activities all had rubrics available for individual grading use, the consistent use of these paper forms was not documented electronically. Likewise, this section of the evaluation required manual grade entry, and students likely had received their grade prior to final evaluation for incorporation into their self-assessment. These factors may have contributed to the smaller difference in preceptor and student composite scores by the end of each trimester.

Previous studies have shown that learners frequently have poor self-assessment skills and thus, likely poor metacognition skills.^{4-7,13} This is often demonstrated by a poor correlation between student self-assessments and preceptor evaluations and has been demonstrated in many of the health professions, including medicine, nursing, and pharmacy.^{4,10,13,14} There are limited data assessing the development of metacognition over time, particularly during APPEs. As previously described, Steuber and colleagues assessed the development of metacognition during a third-year pharmacy elective.⁴ In that study, the authors implemented repetitive faculty and peer feedback and creation of action plans for future work into the course to enhance learner cognition. Despite these interventions, students' ability to self-predict performance did not improve over time and there was a trend of overconfidence in self-assessment. The authors advocate for longitudinal continuous feedback in pharmacy curricula to help learners improve self-assessment.⁴ Hill and colleagues compared student self-assessment of knowledge and skills during the last APPE with preceptor evaluations.⁷ At the end of APPEs, students rated themselves significantly higher than preceptors in six areas of drug knowledge. Students also overrated their proficiency in 16 skill areas compared to preceptors' evaluations of the students' proficiency. The authors concluded that as self-assessment is integral to students' professional development, pharmacy educators

Table 3. Absolute Difference in Evaluation Scores by Trimester

Trimester	Matched Evaluations, No.	Rubric Score Difference, ^a	Composite Score Difference, ^a
		Mean (SD)	Mean (SD)
One	464	16 (13) ^b	4 (5)
Two	574	14 (13) ^c	4 (8)
Three	675	12 (13)	4 (8)

^a Absolute difference was calculated by averaging the difference between preceptor and student evaluations

^b $p<.0001$ for absolute difference between first and third trimesters

^c $p=.009$ for absolute difference between second and third trimesters

Table 4. Rubric and Composite Evaluation Score Difference Between Students and Preceptors by Rotation Type

Type of APPE ^a	Rubric Score Difference, ^b	Composite Score Difference, ^b
	Mean (SD)	Mean (SD)
Internal medicine (n=209)	14 (12)	4 (4)
Ambulatory care (n=200)	18 (14)	2 (10)
Community (n=390)	14 (13)	4 (6)
Advanced institutional (n=209)	15 (14)	5 (10)
Long term care (n=197)	3 (2)	4 (4)
Electives (n=508)	15 (14)	4 (8)

Abbreviations: APPE=advanced pharmacy practice experience

^a No significant difference was found between rotation types

^b Absolute difference was calculated by averaging the difference between preceptor and student evaluations

and program curricula should teach students how to improve the accuracy of their self-assessment.⁷ This can be accomplished by providing ample opportunities for students to receive feedback on their self-assessment skills. Our study demonstrated a decrease in grade differences over time, which may indicate improvement in metacognition over time. This supports the idea that longitudinal continuous assessment can help learners' metacognitive skills.

Notably, in our study, the type of preceptor (faculty vs non-faculty) a student was assigned to during an APPE was not associated with a significant difference in either student or preceptor rubric scores. We hypothesized that there may be a smaller difference between student and faculty rubric scores as faculty members may be more likely to provide more frequent, specific, and constructive feedback, allowing students to better self-assess. As this was not seen, preceptor development that encourages preceptors to provide feedback that incorporates learner self-assessment may be useful. A possible explanation for the "assessment gap" not improving while students were on faculty-precepted APPEs could be the interpersonal effects from students working with the same faculty members during their didactic studies or IPPEs in earlier years of the PharmD program. Put simply, a student may not have wanted to admit weaknesses to a faculty member who knew the knowledge and skills the student should have acquired earlier in the PharmD curriculum.

Though students' self-assessment skills improved during the second and third trimesters, there was still an 11-point overestimation on the rubric score when compared with preceptor scores. This difference may have been because of wording of the rubric. By its very nature, evaluation during experiential education is fraught with subjectivity. Rubrics are built to attempt to remove as much subjectivity as possible, but it is, at best, an imperfect approach. Anecdotally, several preceptors have described, through commentary to the experiential department and their use of the evaluation tool, a likelihood

to take Likert scale guidance quite literally; that is, they fail to award scores of 5 to students, based on the premise that "no one is perfect" and that per the school's rubric, a score of 5 indicates perfect performance. As may be expected, students assigned to these preceptors provided feedback to the experiential department on how they were being evaluated. Based on this, and the research findings annotated above, edits to rubrics and anchors were implemented beginning with the 2018-2019 academic year. Edits to Likert scale language were made to address the absolutes in the language and to allow for more open interpretation of when to award higher (or lower) scores. The most significant edits to illustrate come at the top and bottom of the scale. A score of 5 is now described as "almost always" and a score of 1 is now described as "rarely." This comes in tandem with other improvements, such as more explicit mention/assessment of the Pharmacist's Patient Care Process and documentation of experiential interprofessional education (IPE) activities. Moving forward, comparisons of student self-evaluations vs preceptor evaluations will be conducted to assess whether these language changes contribute to better reliability.

This analysis was the first of its kind to review a metacognitive process for final-year APPE students over time. The large sample size and ability to evaluate differences over multiple years strengthen the findings. Conversely, limitations exist. The data was gathered from an electronic database, which limited the amount of granular detail we were able to examine. For example, we were unable to "drill down" to individual rubric components for each APPE. A more detailed analysis of individual rubric sections might yield significant results. Additionally, students had neither any incentive to accurately predict their grade, nor any penalty for failing to do so. Thus, students may have overpredicted their rubric grades because they were not attempting to accurately self-assess, possibly because of a lack of understanding of the importance of self-evaluation in personal and

professional development. Students and preceptors are expected to meet for a final discussion at the APPE conclusion. However, there is no formal process in place to encourage discussion about differences in student and preceptor scoring. Our findings highlight the need for discussions regarding the value of accurate self-evaluation discussions between the preceptor and student. Indeed, periodic structured feedback from the experiential department given to students with regards to their submitted self-evaluations may be a future intervention to study. Another approach to consider is reflective exercises where students are asked to consider feedback from preceptors in past APPEs.

Several strategies beyond routine self-assessment exist to increase metacognition and should be used throughout the pharmacy curriculum.^{1,3} These strategies include providing high-quality, frequent external feedback to guide learners as they develop their own self-assessment skills, reflective-thinking assignments, and documentation of the clinical thought process. Implementation of multiple strategies such as these on a long-term, continuous basis can improve learners' metacognitive skills.

CONCLUSION

Metacognition is a vital skill for pharmacy students to develop because of the ever-changing nature of the pharmacy profession, which requires life-long learning and ongoing self-assessment of knowledge and skills. Requiring pharmacy students to complete the same grading rubric as their preceptor and to reflect on differences in scoring resulted in improvement in students' self-assessment skills over time. This simple strategy may dovetail with larger initiatives to improve metacognition in pharmacy students.

ACKNOWLEDGMENTS

The authors acknowledge Christy Inge and Daniela Rodriguez Reynaga for their assistance with data preparation.

REFERENCES

1. Medina MS, Castleberry AN, Persky AM. Strategies for improving learners' metacognition in health professional education. *Am J Pharm Educ.* 2017;81(4):Article 78.
2. Accreditation Council for Pharmacy Education. Accreditation standards and key elements for the professional program in pharmacy leading to the doctor of pharmacy degree. ("Standards 2016"). <https://www.acpe-accredit.org/pdf/Standards2016FINAL.pdf>. Accessed April 10, 2019.
3. Motycka CA, Rose RL, Ried LD, et al. Self-assessment in pharmacy and health science education and professional practice. *Am J Pharm Educ.* 2010;74(5):Article 85.
4. Steuber TD, Janzen KM, Walton AM, Nisly SA. Assessment of learner metacognition in a professional pharmacy elective course. *Am J Pharm Educ.* 2017;81(10):Article 6034.
5. Austin Z, Gregory PAM. Evaluating the accuracy of pharmacy students' self-assessment skills. *Am J Pharm Educ.* 2007;71(5):Article 89.
6. Mort JR, Hansen DJ. First-year pharmacy students' self-assessment of communication skill and the impact of video review. *Am J Pharm Educ.* 2010;74(5):Article 78.
7. Hill LH, Kirkwood CK. Student and preceptor perception of performance in advanced pharmacy practice experiences. *Am J Pharm Educ* 2005;69(4):Article 63.
8. Edwards RK, Kellner KR, Siström CL, Magyari EJ. Medical student self-assessment of performance on an obstetrics and gynecology clerkship. *Am J Obstet Gynecol.* 2003;188(4):1078-1082.
9. Chen S, Chien L, Kuo M, et al. Exploring discrepancies in perceived nursing competence between postgraduate-year nurses and their preceptors. *J Cont Educ Nursing.* 2017; 48(4):190-196.
10. Weiss PM, Koller CA, Hess LW, Wasser T. How do medical student self-assessment compare with their final clerkship grades? *Med Teach.* 2015;27(5):445-449.
11. Van Rosendaal GMA, Jennett PA. Comparing peer and faculty evaluations in an internal medicine residency. *Academic Medicine.* 1994;69(4):299-303.
12. Kachingwe AF, Phillip B, Beling J. Videotaping practical examinations in physical therapist education: does it foster student performance, self-assessment, professionalism, and improve instructor grading? *J Phys Ther Educ.* 2015;29(1):25-33.
13. Baxter P, Norman G. Self-assessment or self deception? a lack of association between nursing students' self-assessment and performance. *J Adv Nurs.* 2011;67(11):2406-2413.
14. Wagner ML, Suh DC, Cruz S. Peer- and self-grading compared to faculty grading. *Am J Pharm Educ.* 2011;75(7):Article 130.