7-1-2016

Peer Observation and Evaluation Tool (POET): A Formative Peer Review Supporting Scholarly Teaching

Jeffrey L. Crabtree
Indiana University - Purdue University Indianapolis, jlcrabtr@iu.edu

Patricia J. Scott
Indiana University - Purdue University Indianapolis, Scottp@iu.edu

Fengyi Kuo
Indiana University, fkuo@iu.edu

Peer Observation and Evaluation Tool (POET): A Formative Peer Review Supporting Scholarly Teaching

Abstract

The Peer Observation and Evaluation Tool (POET) is a valid and reliable measure developed for formative peer evaluation of pharmacy faculty. The authors replicated a study conducted on the POET in pharmacy in order to report the instrument's validity and reliability in occupational therapy and to explore its potential as a formative teaching evaluation for occupational therapy educators. To verify item importance, seven participants from the faculty in an occupational therapy department rated each item. To establish inter-rater reliability, the participants evaluated one videotaped 55 min lecture. The POET was reliable with ICC at 0.93. There were high levels of agreement with the importance ratings among the participants with all scales. The POET appears to be a valid and reliable formative measure of teaching. At a time of significant change in the level of occupational therapy education, this measure may be an important support for scholarly teaching in two ways: First, this measure offers several opportunities to document the instructor's strengths and, second, it offers the instructor suggestions about ways to improve teaching quality. Finally, the POET may facilitate faculty professional growth and development through systematic, strategic, and constructive peer review feedback.

Credentials Display

Jeffrey L. Crabtree, OTD, MS, OT, FAOT; Patricia Scott, PhD, MPH, OT, FAOTA; Fengyi Kuo, DHS, OTR, CPRP

Background and Literature Review

Scholarly teaching, often associated with the scholarship of teaching and learning (SoTL), is an important concept, and, to varying degrees, is expected of faculty in all occupational therapy education programs (American Occupational Therapy Association [AOTA], 2009). Since Boyer's 1990 report, *Scholarship Reconsidered: Priorities of the Professorate*, a number of researchers have discussed the features of scholarly teaching and of the SoTL (Glanville & Houde, 2004; Glassick, 2000; Grise-Owens, Owens, & Miller, 2016; Spake & Salem, 2005). However, the literature does not always clearly distinguish between the two. Therefore, for the purpose of this paper, scholarly teaching is defined as scholarship that meets the following criteria:

- The work must be made public.
- The work must be available for peer review and critique according to accepted standards.
- The work must be able to be reproduced and built on by other scholars.

(Glassick, 2000, p. 879)

Compared to the SoTL, scholarly teaching has a critical but narrow focus. For example, Potter and Kustra's (2011) definition of scholarly teaching, similar to Boyer's, is "teaching grounded in critical reflection using systematically and strategically-gathered evidence, related and explained by well-reasoned theory and philosophical understanding, with the goal of maximizing learning through effective teaching" (p. 3). The two concepts can be distinguished by the type of faculty position: A tenure-track position with an emphasis on teaching may require faculty to show evidence of the SoTL, which would include scholarly teaching, while a tenure-track position with an emphasis on research may require faculty to show evidence of scholarly teaching but not of the SoTL, as their tenure-track scholarship requirements would likely be met through research.

Evaluation of scholarly teaching is a common practice in virtually all occupational therapy academic settings. For example, the Accreditation Council for Occupational Therapy Education (ACOTE) (2015) requires faculty who are teaching two or more courses to show evidence of teaching effectiveness—often through teaching evaluations. At the authors' educational institution, faculty on tenure or clinical tracks, no matter what area they chose for excellence, must meet at least "satisfactory" standards in teaching. While high standards for teaching performance are ubiquitous among occupational therapy and other health care professions, the practice of evaluating teaching varies, as does the purpose, and there is little agreement in the profession about the most effective approach to evaluating teaching (Papay, 2012).

The 2008-2009 Task Force for the Recognition of Teaching Excellence report from the American Association of Colleges of Pharmacy (Hammer et al., 2010) offers an excellent overview of the various approaches to evaluating teaching. It states that the process

> must be systematic, sound in theory and practice, manageable to implement, and well understood by faculty members. Important elements in the process include classroom observation, observer training, experience

with the forms used, and review of course materials before the observation. (Hammer et al., 2010, p. 5)

The approach to evaluating faculty teaching may be summative (typically assessing the outcome of teaching or the students' learning) and/or formative (typically assessing the process of teaching) (Appling, Naumann, & Berk, 2001; Worrell, Everly, Hamant, & Kiel, 1999). Evaluations may be regularly scheduled or ad hoc (Hubball & Clarke, 2011), and the settings of the evaluations may include classrooms, clinics, or labs (Fernandez & Yu, 2007). In addition, the evaluation instruments vary widely from teaching checklists or rubrics (Wiese et al., 2007) to videotaped lectures (Barber, 1992; Green, Ellis, Frémont, & Batty, 1998) to self-evaluations (Bryan, Krych, Carmichael, Viggiano, & Pawlina, 2005). Whatever form the teaching evaluations take, they are often intended to show teaching effectiveness, and faculty are typically expected to document incremental progress toward improving teaching approaches (ACOTE, 2015; Commission on Institutions of Higher Education, 2011; Papay, 2012).

Peer teaching review is a common element of scholarly teaching evaluation. The reviews are sometimes part of a faculty member's larger teaching portfolio (Kreeber, 2006) or imbedded in a larger evidence-based teaching evaluation program (Hansen et al., 2007). Peers may review "faculty members' facilitation of the learning process for learners and their demonstrated commitment to the educational mission of the department" (Rosenbaum, Ferguson, Kreiter, & Johnson, 2005,

p. 430). Or, peer teaching reviews can be used as a tool for continuous instructional improvement (Papay, 2012). Regardless of the rigor of or approach to the review, at many universities and colleges the outcome of virtually any of these reviews can influence whether a faculty member is promoted, gains tenure or a long-term contract, or receives an award or another form of scholarly acknowledgment (Fincher et al., 2000).

The literature suggests that peer teaching review can be formally or informally done by classroom observation, and that the reviews may contribute to the development of teaching portfolios, comprise part of the mentoring process, or serve as the basis of external reviews from intercampus sources, such as centers for the scholarship of teaching and learning (Worrell et al., 1999). Despite the many forms peer teaching reviews may take, few of the methods appear to be valid or reliable measures of teaching effectiveness (Bernstein, 2008; Brown & Ward-Griffin, 1994).

Any form of peer teaching review can be biased (Lee, Sugimoto, Zhang, & Cronin, 2013). Consequently, there is a need for valid and reliable formative teaching evaluations. This study replicated a prior study in pharmacy of the Peer Observation and Evaluation Tool (POET, Northeastern University Department of Pharmacy Practice © 2007) (Trujillo et al., 2008), which demonstrated the tool to be a valid and reliable measure of lecture-based classroom teaching in a pharmacy program. The aim of the current study was to evaluate the use of the tool in an occupational therapy education setting.

The POET, a formative teaching evaluation, was identified as having promise in what Papay (2012) expressed as "a key purpose to teacher evaluation: to improve instruction by developing teachers' instructional capacity and effectiveness" (p. 133). Furthermore, the POET meets many of the 2008-2009 Task Force for the Recognition of Teaching Excellence recommendations for effective teaching evaluations (Hammer et al., 2010). A task force from the Northeastern University Department of Pharmacy Practice developed the POET. The development process included 10 phases that culminated in the instrument. The POET contains 39 ranked items divided into four sections (preobservation meeting, classroom observation, postobservation meeting, and postobservation assessment). The same 39 items were used in this current study (see Appendix).

Working under the assumption that the content of the POET includes effective teaching standards, the authors identified the following objectives for this study:

1. To establish the content validity as reflected by measuring agreement of POET items
by occupational therapy faculty.

2. To establish the reliability of the POET for use in an occupational therapy classroom as reflected by an acceptable interclass correlation (ICC) following the view of a videotaped lecture.

3. To discuss the faculty professional development benefits of the POET.

4. To explore inter-professional collaboration opportunities for the SoTL.

Method

This study explored the psychometric properties of the POET using occupational therapy faculty. It is a replication of the study published by Trujillo and colleagues (2008) and was approved by the Indiana University IRB # 1304011140.

Participants

Eight occupational therapy core faculty at a research university were eligible to participate in this study. Participation was voluntary, yet motivation was strong, as performance expectations require all faculty to participate in systematic reviews of teaching. Faculty who declare teaching as their area of excellence are specifically challenged to show incremental changes based on regular (usually yearly) evaluations of their teaching. The faculty were given the IRB-approved study information sheet at a regularly scheduled faculty meeting. Explicit agreement to participate was assumed if the faculty member arrived at the stated location ready to complete the study procedures. Seven faculty members participated in the study.

Instrumentation

The POET was the main instrument used in this study.

Demographic Survey

Demographics were collected for informational purposes and were not included in the analysis (see Table 1).

Table 1
Demographics

Gender			
	Male	3	43%
	Female	4	57%
Ethnicity			
	Caucasian	6	85%
	Asian	1	15%
Years in Academia			
	mean	20.1	
	range	8-40	
Areas of Expertise			
	Driving		
	Mental health		
	Management		
	Pediatrics		
	Older adults		
	Hand - UE		
	rehabilitation		
	Cognition		
	Evidence-based practice		

The POET is a formative evaluation designed to provide information about a broad range of topics, from why the lecture is situated in a particular place in the curriculum to how the instructor managed the classroom environment. During the preobservation (Phase 1), the instructor provides the reviewer with the lecture materials and handouts and reviews the teaching pedagogy specific to the lecture to be observed. The reviewer may ask clarifying questions about the instructor's goals for the lecture and how the instructor intends to reach those goals. Then, the reviewer observes the entire lecture (Phase 2). After the observation, the reviewer meets with the instructor two times: once to discuss the instructor's self-reflection about the observed session (Phase 3), and once to offer the instructor possible recommendations (Phase 4) (see Appendix).

The POET authors (Trujillo et al., 2008) reported interclass correlation coefficients for eight observed lectures individually. The overall coefficients ranged from 0.66 to 0.97 for all lectures. During the testing, three raters were present for seven lectures. The authors reported comparisons between rater ICCs ranging from 0.43 to $< +0.98$. All but three were statistically significant with $p < 0.01$. The authors reported comparisons between rater ICCs ranging from 0.43 to 0.98. All but three were statistically significant with $p < 0.01$.

Content Validity Form

All of the three subsections (located in Phase 2)—content, teaching strategies and presentation skills, and classroom climate—of the POET ($n = 27$) were listed along with the request to rank each on a 4-point Likert scale. The scale choices were *not important*, *somewhat important*, *important*, and *essential*.

Procedures for the Current Study

When the participants entered the room, the researchers requested that they spread out so that they were not seated directly next to each other. One of the researchers provided a short introduction to the POET and to the process of a formative evaluation. The participants were allowed to ask questions. The participants then completed the demographic survey and the content validity form. The participants used the content validity form to identify which POET items they agreed were essential aspects of teaching. A researcher collected the forms. The participants were then given the POET Observation Form (see Appendix). After a researcher explained the process of completing the

form, the participants proceeded to complete the form while watching a 55 min videotaped lecture. At the completion of the videotape, the participants were given a few minutes to complete the rating forms. No discussion was allowed during this time. After the researchers collected all of the observation forms, the participants completed the second content validity form. The participants were then thanked for their participation and invited to leave. One week later, when the preliminary results were available, the participants were asked at a regularly scheduled faculty meeting to (a) give input on items in which there was variable rater agreement and (b) discuss the implications for using the POET in the occupational therapy department.

Confidentiality

All of the forms were coded and no names were used. Due to a small sample size, demographic information was used for descriptive purposes and not in the analysis, thus assuring no results could be attributed to a certain individual.

Data Collection and Analysis

To verify item importance, the seven participants rated each item. To establish inter-rater reliability, the participants evaluated one videotaped 55 min lecture. Inter-rater reliability was evaluated using ICC. The consistency type ICC analysis was used to determine average measure reliability (and to replicate Trujillo). Interclass correlation coefficients were computed for the overall classroom observation component of the lecture as well as for the three subsections: content, teaching strategies and presentation skills, and classroom climate. Statistical significance of the proportion of variance that is systemic was set at $p < 0.05$. SPSS version 21 (Chicago, IL) was used to compute all data.

Results

The first objective was to establish the content validity by measuring the participants' agreement of the POET items. Content validity was measured by importance ratings assigned by the participants to the various items of the POET evaluation taken before and after watching the videotaped lecture. The teaching items comprise the majority of the POET observation ratings. Table 2 shows mean ranks of items before watching the videotaped lecture of 2.3 and following the lecture; 2.48 showed no significant difference. Both mean ratings fell at the mean of possible ratings of 1-4, with a range of 3.14 - 1.43 before and 3.14 - 1.29 after. There was a notable difference of up to 26 points in the importance of the item: "Depth of material presented appears appropriate to type of course and student level." Twelve items moved up in importance, five stayed the same, and seven items moved down in importance.

A conservative Friedman's Test found no difference overall for the pre and postratings (see Table 3).

Table 2

Impact of Observed Lecture on Perceived Importance of Ratings Organized by Difference Scores

POET Questions on content of Teaching	Before video		After video		Difference scores
	Mean Rank	Mean Rating	Mean Rank	Mean Rating	
posttest - # 2, Depth of materials appropriate to type of course and student level	28	2.00	2	3.14	26
posttest - # 18, The instructor makes connections with prior learning within curriculum	25	2.14	10	2.71	15
posttest - # 11, Breadth of material appropriate dedicated to this topic	17	2.57	3	3.14	14
posttest - # 24, The instructor encourages critical thinking	20	2.29	9	2.71	11
posttest - # 20, The instructor emphasizes a conceptual grasp of the material	14	2.57	5	3.00	9
posttest - # 29, The instructor effectively uses audio/visual/learning aids	22	2.14	13	2.57	9
posttest - # 31, The instructor creates a classroom atmosphere conducive to learning	13	2.57	4	3.00	9
posttest - # 3, Breadth of materials time dedicated to this topic	9	2.71	1	3.14	8
posttest - # 27, The instructor uses class time efficiently.	23	2.14	15	2.57	7
posttest - # 6, The instructor establishes the relevance of information	18	2.43	12	2.57	6
posttest - # 19, The instructor makes references to the material taught previously	33	1.43	29	2.29	4
posttest - # 1, The instructor appears knowledgeable and up-to-date	4	3.00	6	3.00	2
posttest - # 5, The instructor provides an overview of what is planned for the class	27	2.00	27	2.29	0
posttest - # 8, The instructor is an effective communicator	11	2.57	11	2.57	0
posttest - # 12, Clear distinction between fact and opinion/ practice experience	34	1.43	34	1.71	0
posttest - # 21, Instructor provides periodic summaries and ties things together	32	1.57	32	2.00	0
posttest - # 30, Instructor emphasizes material likely or unlikely to be examined	35	1.43	35	1.29	0
posttest - # 4, Divergent opinions or conflicting views presented when appropriate	29	1.71	30	2.14	-1
posttest - # 16, The instructor explains content clearly, providing examples when appro	5	2.86	7	2.86	-2
posttest - # 17, The instructor is an effective communicator	15	2.57	17	2.57	-2
posttest - # 14, The instructor provides an overview of what is planned for the class	21	2.29	24	2.43	-3
posttest - # 26, The lecture remains focused on its objectives	24	2.14	28	2.29	-4
posttest - # 28, Questions are welcomed and responded to in a professional manner	10	2.71	14	2.57	-4
posttest - # 34, The instructor demonstrates flexibility to student concerns or interests	12	2.57	18	2.43	-6
posttest - # 15, The instructor establishes the relevance of information	16	2.57	23	2.43	-7
posttest - # 22, The learning activities are well organized	2	3.14	16	2.57	-14
posttest - # 33, The instructor reacts to student behavior issues appropriately	1	3.14	19	2.43	-18
posttest - # 23, Breadth of materials appropriate for this topic	3	3.00	22	2.43	-19

Note. Mean rating is based on range of 1-4 for each item with 1 = *not important*, 2 = *somewhat important*, 3 = *important*, to 4 = *essential*. Shaded area includes items not impacted by the viewing of the lecture.

Table 3

ANOVA with Friedman's Test

df	Mean Square	Friedman's Chi-Square	Significance
6	333.476		
1	28.571	0.769	0.380
1	20.249	0.479	0.520

The second objective was to establish the reliability of the POET for use in an occupational therapy classroom as reflected in the consistency of the ratings between the seven participant raters. The inter-class correlation (ICC) for all items is excellent at 0.93. The three sections varied (0.82 - 0.88) with both the content items and the classroom climate items showing a strong ICC of 0.82. The teaching items alone showed strong reliability (0.88) (see Table 4).

Table 4

Reliability as Reflected by Interclass Correlations

Observation Reliability	Intraclass Correlation[a]	95% Confidence Interval		F Test with True Value 0			
		Lower Bound	Upper Bound	Value	df1	df2	Sig
Content items	0.875	0.441	0.991	8.000	3	12	0.003
Teaching items	0.815	0.372	0.987	5.419	3	48	0.003
Classroom climate items	0.815	.404	.964	5.400	6	18	0.002

The final objective was to discuss the faculty professional development benefits of the POET. Specific observation items from the POET that the participants chose to discuss are Item 4, "Divergent opinions or conflicting views presented when appropriate," and Item 30, "Instructor emphasizes which material students are likely or unlikely to be examined [*sic*]."

Item 4. Some raters indicated that this item is essential in lecture content delivery, while several raters identified this item as not important or somewhat important. In judging the importance of this item to classroom instruction, raters may have taken into account the extensiveness of professional standards that need to be covered in certain courses; therefore, faculty may have felt the level of practicality and performance skills in clinical practice might be valued more favorably in entry-level professional curriculum than presenting divergent opinions and conflicting views.

Item 30. Discussion centered on how the professional school curriculum is likely to focus more on graduate-level performance skills and clinical reasoning. The comments from the faculty

related to the use of the POET in this curriculum centered around two themes: One, using the POET was time consuming, but given the high stakes for faculty, the lengthy process appeared to be worth the effort. Two, faculty noted the POET fosters an interactive evaluation of the instructor's teaching skills ranging from lecture content to the instructor's performance to managing the classroom environment.

Discussion

This study sought to identify the degree to which the POET (Trujillo et al., 2008), a valid and reliable formative peer teaching review in a pharmacy program, would meet the needs of a peer teaching review of occupational therapy faculty. It is important that the POET was found to have both concurrent and content validity on "importance" ratings across all seven raters watching the same videotaped lecture. These and other results suggest that the POET may be helpful in identifying instructors' strengths and in providing constructive feedback that can support faculty professional development (Papay, 2012).

The results of this study also provided the opportunity to reflect more deeply on some interesting observed phenomena. For example, Item 2, "Depth of material presented appears appropriate to type of course and student level," was rated 28th, close to the bottom of the scale, before viewing the tape and 2nd after viewing the tape, a difference of 26 points. Item 23, "Breadth of material appropriate for the amount of time dedicated to this topic," was rated 3rd before viewing the tape and 22nd after viewing the tape, a difference of 19 points. This reordering happened independent of the means and ranges of importance assigned to individual items in the POET (see Table 2). These results may suggest the measure is highly sensitive to a videotaped lecture, which begs the question: Would the POET also be effective when observing lectures in person in the classroom?

Also, these results led the researchers to question the reordered ranking of the items. One possible reason for the reordered ranking is the lecturer satisfied the rater by covering the material sufficiently in breadth and insufficiently in depth or visa-versa. Either way, from the researchers' perspectives, that the participants reordered the rankings independently across raters is worth studying more carefully. From the faculty's perspective, these sorts of results could prompt discussion about the importance of the POET's various items to the overall curriculum design and about which POET items are most useful in peer teaching review and in supporting faculty professional development.

As the follow-up study (DiVall et al., 2012) suggests, the POET also fosters a balanced, interactive evaluation of the instructor's teaching skills, ranging from lecture content to the instructor managing the classroom environment. This process facilitates constructive feedback and offers several opportunities to document the instructor's strengths and to offer the instructor suggestions about ways to improve in all aspects of teaching. The results of this study suggest that the POET, as is, can be used in occupational therapy programs as a means of supporting and improving faculty teaching skills as well as documenting teaching skills for merit raises and promotion. Furthermore, given the POET was designed for pharmacy faculty and appears to be useful for occupational therapy faculty, further research is warranted to investigate the POET's applicability to occupational therapy and other professions, such a physical therapy or speech language pathology, for its potential to support faculty professional growth in specific professions and its likelihood for peer teaching review in inter-professional academic settings.

Limitations

In this replication of the study by Trujillo and colleagues (2008), only one videotaped lecture was evaluated. It is not possible, therefore, to make assumptions about the consistency of ratings across multiple videotaped lectures or about in-person observations of lectures. Faculty at this institution demonstrated consistent teaching standards even though the faculty represents divergent practice areas and spans years in ages and in experience in academia. Although the researchers recommend using the POET for faculty evaluation and teaching growth and development, it may be advisable to measure the internal consistency across faculty

raters, as faculty will conduct peer teaching reviews at different times. Finally, the discussion among faculty reflected that some faculty believed items were general enough to apply standards to reflect the higher order thinking needed for clinical reasoning while others were uncertain. As a result, some of the items in the POET may not be a good fit for all occupational therapy education teaching and learning objectives.

Conclusion

The authors of the POET (Trujillo et al., 2008) and of this study acknowledge that this formative system of peer teaching review is time consuming; however, given the high stakes for faculty, the process appears to be worth the effort. This is because evaluation of faculty teaching has essentially two overlapping purposes: to make decisions about the teaching effectiveness for the promotion and tenure process and to promote faculty growth and development in the SoTL over time (Bernstein, 2008; Boehm & Bonnel, 2010).

The POET offers both the instructor and the peer reviewer several opportunities to review the lecture content, the instructor's performance, the review results, and the instructor's reflections. These are valuable aspects of the POET, as the in-depth process facilitates peer teaching review over time and enables constructive feedback, offers several opportunities to document the instructor's strengths, and offers the instructor suggestions about ways to improve (Bernstien, 2008, Boehm & Bonnel 2010).

Implications for further research include reliability and validity testing in in-person classroom observations in occupational therapy and other health care professions, and to explore the reliability and validity of the POET in inter-professional teaching settings.

Implications for Occupational Therapy Practice

The results of this study suggest that the POET is a valid and reliable formative peer teaching review. Several implications for occupational therapy practice include:

- Occupational therapy faculty can use the POET to support and document teaching.
- The POET can support faculty development in teaching and classroom management.
- The POET can stimulate useful and important discussions among faculty about course content and professional development.
- Results of the POET may be used to provide evidence of effective teaching for promotion, tenure, merit raises, and other forms of faculty acknowledgment.

References

American Occupational Therapy Association. (2009). Specialized knowledge and skills of occupational therapy educators of the future. *American Journal of Occupational Therapy*, *63*(6), 804-818. http://dx.doi.org/10.5014/ajot.63.6.804

Accreditation Council for Occupational Therapy Education. (2015). *Accreditation manual*. Bethesda, MD: Author. Retrieved from https://www.aota.org/-/media/Corporate/Files/EducationCareers/Accredit/Policies/ACOTE%20Manual%20Complete.pdf

Appling, S. E., Naumann, P. L., & Berk, R. A. (2001). Using a faculty evaluation triad to achieve evidence-based teaching. *Nursing and Health Care Perspective*, *22*(5), 247-251.

Bernstein, D. J. (2008). Peer review and evaluation of the intellectual work of teaching. *Change: The Magazine of Higher Learning*, *40*(2), 48-51. http://dx.doi.org/10.3200/chng.40.2.48-51

Barber, S. G. (1992). Postgraduate teaching audit by peer review of videotape recordings. *Medical Teacher*,

14(2-3), 149-157.
http://dx.doi.org/10.3109/01421599209079481

Boehm, H., & Bonnel, W. (2010). The use of peer review in nursing education and clinical practice. *Journal for Nurses in Staff Development*, *26*(3), 108-115. http://dx.doi.org/10.1097/nnd.0b013e3181993aa4

Boyer, E. L. (1990). *Scholarship reconsidered: Priorities of the professorate*. New York: The Carnegie Foundation for the Advancement of Teaching.

Brown, B., & Ward-Griffin, C. (1994). The use of peer evaluation in promoting nursing faculty teaching effectiveness: A review of the literature. *Nurse Education Today*, *14*(4), 299-305. http://dx.doi.org/10.1016/0260-6917(94)90141-4

Bryan, R. E., Krych, A. J., Carmichael, S. W., Viggiano, T. R., & Pawlina, W. (2005). Assessing professionalism in early medical education: Experience with peer evaluation and self-evaluation in the gross anatomy course. *Annals of the Academy of Medicine, Singapore*, *34*(8), 486-491.

Commission on Institutions of Higher Education. (2011). *Standards of accreditation*. Burlington, MA: Author.

DiVall, M., Barr, J., Gonyeau, M., Matthews, S. J., Van Amburgh, J., Qualters, D., & Trujillo, J. (2012). Follow-up assessment of a faculty peer observation and evaluation program. *American Journal of Pharmaceutical Education, 76*(4), Article 61. Retrieved from http://www.ajpe.org/doi/pdf/10.5688/ajpe76461

Fernandez, C. E., & Yu, J. (2007). Peer review of teaching. *Journal of Chiropractic Education*, *21*(2), 154-161. http://dx.doi.org/10.7899/1042-5055-21.2.154

Fincher, R.-M. E., Simpson, D. E., Mennin, S. P., Rosenfeld, G. C., Rothman, A., McGrew, M. C., . . . Turnbull, J. M. (2000). Scholarship in teaching: An imperative for the 21st century. *Academic Medicine*, *75*(9), 887-894. Retrieved from http://journals.lww.com/academicmedicine/Fulltext/2000/09000/Scholarship_in_Teaching__An_Imperative_for_the.9.aspx

Glanville, I., & Houde, S. (2004). The scholarship of teaching: Implications for nursing faculty. *Journal of Professional Nursing*, *20*(1), 7-14. http://dx.doi.org/10.1016/j.profnurs.2003.12.002

Glassick, C. E. (2000). Boyer's expanded definitions of scholarship, the standards for assessing scholarship, and the elusiveness of the scholarship of teaching. *Academic Medicine*, *75*(9), 877-880. http://dx.doi.org/10.1097/00001888-200009000-00007

Green, M. E., Ellis, C. L., Frémont, P., & Batty, H. (1998). Faculty evaluation in departments of family medicine: Do our universities measure up? *Medical Education*, *32*(6), 597-606. http://dx.doi.org/10.1046/j.1365-2923.1998.00242.x

Grise-Owens, E., Owens, L. W., & Miller, J. J. (2016). Conceptualizing the scholarship of teaching and learning for social work education. *Journal of Social Work Education*, *52*(1), 6-17.

Hammer, D., Piascik, P., Medina, M., Pittenger, A., Rose, R., Creekmore, F. . . . Scott, S. (2010). Recognition of teaching excellence. *American Journal of Pharmaceutical Education*, *74*(9), Article 164. http://dx.doi.org/10.5688/aj7409164

Hansen, L., McCollum, M., Paulsen, S. M., Cyr, T., Jarvis, C. L., Tate, G., & Altiere, R. (2007). Evaluation of an evidence-based peer teaching assessment program. *American Journal of Pharmaceutical Education*, *71*(3), Article 45. Retrieved from: http://www.ajpe.org/doi/pdf/10.5688/aj710345

Hubball, H., & Clarke, A. (2011). Scholarly approaches to peer-review of teaching: Emergent frameworks and outcomes in a research-intensive university. *Transformative Dialogues: Teaching and Learning Journal*, *4*(3), 1-32. Retrieved from https://www2.viu.ca/integratedplanning/documents/TD437_HubbalClarke_Peer_Review.pdf

Kreeber, C. (2006). Developing the scholarship of teaching through transformative learning. *Journal of Scholarship of Teaching and Learning*, *6*(1), 88-109. Retrieved from http://files.eric.ed.gov/fulltext/EJ854916.pdf

Lee, C. J., Sugimoto, C. R., Zhang, G., & Cronin, B. (2013). Bias in peer review. *Journal of the American Society for Information Science and Technology* [Advances in Information Science article], *64*(1), 2-17. http://dx.doi.org/10.1002/asi.22784

Papay, J. (2012). Refocusing the debate: Assessing the purposes and tools of teacher evaluation. *Harvard Educational Review*, *82*(1), 123-141. http://dx.doi.org/10.17763/haer.82.1.v40p0833345w6384

Potter, M. K., & Kustra, E. D. (2011). The relationship between scholarly teaching and SoTL: Models, distinctions, and clarifications. *International Journal for the Scholarship of Teaching and Learning*, *5*(1), Article 23. http://dx.doi.org/10.20429/ijsotl.2011.050123

Rosenbaum, M. E., Ferguson, K. J., Kreiter, C. D., & Johnson, C. A. (2005). Using a peer evaluation system to assess faculty performance and competence. *Family Medicine*, *37*(6), 429-433. Retrieved from http://www.stfm.org/fmhub/fm2005/June/Marcy429.pdf

Spake, E., & Salem, A. (2005). Setting the stage: The scholarship of teaching and learning. *Journal of*

Physical Therapy Education, *19*(3), 5-13. Retrieved from http://web.a.ebscohost.com/abstract?site=ehost&scope=site&jrnl=08991855&AN=19568457&h=mFscj8tR5x4UELKymwKSclyjbUIODK48tyQqXAQrT97G39DfS%2bFpH0wYlOObNcSD9iutKB5FWaiCe1H5ZywEnw%3d%3d&crl=f&resultLocal=ErrCrlNoResults&resultNs=Ehost&crlhashurl=login.aspx%3fdirect%3dtrue%26profile%3dehost%26scope%3dsite%26authtype%3dcrawler%26jrnl%3d08991855%26AN%3d19568457

Trujillo, J. M., DiVall, M. V., Barr, J., Gonyeau, M., Van Amburgh, J. A., Matthews, S. J., & Qualters, D. (2008). Development of a peer teaching-assessment program and a peer observation and evaluation tool. *American Journal of Pharmaceutical Education*, *72*(6), 1-9. Retrieved from http://www.ajpe.org/doi/pdf/10.5688/aj7206147

Wiese, G. C., Percuoco, R. E., Pickar, J. G., Duray, S. M., Faruqui, S. R., Schmiedel, G. O., & McLean, I. D. (2007). Development of an evidence-based application and rubric for evaluating applicants' qualifications for promotion to professor. *Journal of Manipulative and Physiological Therapeutics*, *30*(7), 527-535. http://dx.doi.org/10.1016/j.jmpt.2007.07.007

Worrell, M. B., Everly, J., Hamant, C., & Kiel, J. (1999). Peer review of teaching for summative and formative evaluation in occupational therapy. In P. Crist (Ed.), *Innovations in Occupational Therapy Education* (pp. 59-67). Bethesda, MD: American Occupational Therapy Association.

Appendix
Peer Observation and Evaluation Tool (POET)

Section 1: Pre-observation visit
*The instructor should provide lecture materials (handouts, resources, etc.) at **least 1 week prior** to this meeting*

	DNO	NSD	ND	A	AW	Comments
1. Lecture objectives are clearly stated in the handout/syllabus						
2. Lecture objectives align with the overall course goals						
3. Handout material appears to be relevant to lecture objectives						
4. Lecture outline and organization are logical						
5. Reading list is provided and relevant to lecture objectives						
6. Planned in-class activities reflect appropriate lecture objectives						
7. Planned assessment strategies are consistent with lecture objectives						
8. Instructor appears well prepared for class						

Questions:

Why did you choose this lecture to be assessed? Is this the first time you are teaching this lecture? If no, what changes have you made to this lecture over the past few times you taught it?

What questions/concerns do you have? What would you particularly like feedback on? Are you interested in having an active learning inventory of your lecture completed?

What is your educational philosophy?

Where is similar content taught in the curriculum? Have you contacted other instructors to determine exactly what they cover? What impact has this had on your lecture and/or student outcomes?

How does this lecture's content fit within the entire course (e.g. one out of several lectures on the same topic)?

Have you planned any in-class learning activities? If yes, what lecture objectives will these activities meet? Share how these activities facilitate student learning.

What is your plan for assessing the content of this lecture?

DNO	Did not observe	either because was not in the class for the entire lecture duration or instructor did not do and it's not applicable
NSD	Needs significant development	the instructor did not do this and should consider adding
ND	Needs development	the instructor attempted to do this but development/ revision is necessary
A	Accomplished	minor improvements can be recommended
AW	Accomplished well	no recommendations for improvement

Northeastern University Department of Pharmacy Practice ©2007

Section 2: Classroom Observation Items

	DNO	NSD	ND	A	AW	Comments
Content						
9. The instructor appears knowledgeable and up-to-date about the content						
10. Depth of material presented appears appropriate to type of course and student level						
Instructor does not spend a lot of time going over material previously taught in other courses; intellectual level of material presented appropriate to the student level						
11. Breadth of material appropriate for the amount of time dedicated to this topic						
Instructor is able to go through majority of the material during the class period. Amount of content appropriate for the time						
12. Clear distinction between fact and opinion/ practice experience						
Instructor differentiates between consensus statements, guidelines, expert opinion and personal views, practice, experiences						
13. Divergent opinions or conflicting views presented when appropriate						
Instructor provides examples of conflicting or different guidelines, clinical trials, practices						
Teaching strategies, presentation skills, organization, and clarity						
14. The instructor provides an overview of what is planned for the class period.						
15. The instructor establishes the relevance of information						
16. The instructor explains content clearly, providing examples when appropriate						
17. The instructor is an effective communicator						
Instructor's command of English is adequate; the instructor effectively holds class attention; the instructor uses eye contact effectively; the instructor speaks clearly and loudly enough to be heard throughout the classroom; the instructor employs an appropriate rate of speech (e.g. for note taking); the instructor emphasizes major points in the delivery of the content by pausing, raising voice, etc.; the instructor is enthusiastic and confident on explaining the subject matter						

DNO	Did not observe	either because was not in the class for the entire lecture duration or instructor did not do and it's not applicable
NSD	Needs significant development	the instructor did not do this and should consider adding
ND	Needs development	the instructor attempted to do this but development/ revision is necessary
A	Accomplished	minor improvements can be recommended
AW	Accomplished well	no recommendations for improvement

Northeastern University Department of Pharmacy Practice ©2007

	DNO	NSD	ND	A	AW	Comments
18. The instructor makes connections with prior learning <u>within curriculum</u>						
19. The instructor makes references to the material taught previously <u>within the course</u>						
20. The instructor emphasizes a conceptual grasp of the material						
Instructor provides clear and comprehensive explanations when required.; instructor points out practical applications of concepts; instructor suggests ways to learn complicated ideas						
21. Instructor provides periodic summaries of the most important ideas and ties things together at the end of the class						
Instructor makes appropriate transitions by summarizing ideas and welcoming questions						
22. The learning activities are well organized						
Appropriate number of activities; spaced out appropriately, students are given appropriate time to complete them, appropriate discussion at the end of each activity takes place.						
23. Instructor's teaching strategies facilitate student learning						
Instructor follows a progressive development of course content and involving active student learning and the application of student involvement building upon Bloom's taxonomy -- knowledge, comprehension, application, analysis, synthesis, and evaluation.						
24. The instructor encourages critical thinking						
Instructor asks stimulating and challenging questions periodically; classroom activities and outside assignments include problem solving; students have chances to discuss or apply concepts during class						
25. The instructor effectively uses in class activities and outside assignments to gauge student progress						
Instructor employs active learning techniques. Activities and assignments supplement lecture content; instructor provides clear directions for each activity; promotes student engagement and is able to involve everyone in the class, not just the most outspoken students.; provides adequate time and resources for completion; instructor facilitates group work well, mediates discussion well, helps students apply theory to solve problems						

DNO	Did not observe	either because was not in the class for the entire lecture duration or instructor did not do and it's not applicable
NSD	Needs significant development	the instructor did not do this and should consider adding
ND	Needs development	the instructor attempted to do this but development/ revision is necessary
A	Accomplished	minor improvements can be recommended
AW	Accomplished well	no recommendations for improvement

Northeastern University Department of Pharmacy Practice ©2007

	DNO	NSD	ND	A	AW	Comments
26. The lecture remains focused on its objectives.						
Instructor stays on the subject; does not spend a considerable amount of time on material not covered by objectives; if questions or discussion lead on a tangent, able to get the class back on the subject						
27. The instructor uses class time efficiently.						
28. Questions are welcomed and responded to in an effective and professional manner.						
Instructor asks students periodically if anyone has questions; repeats student questions and answers so all can hear; responds to questions clearly and thoroughly, and/or tells the class that he/she will research and follow up						
29. The instructor effectively uses audio/visual/learning aids to accompany the verbal presentation						
Handouts and/or PowerPoint slides express content clearly; are legible (appropriate font); contain same or similar content covered during the lecture; at adequate level of detail; shows creativity (if applicable); board work (if used) is legible and organized						
30. Instructor emphasizes which material students are likely or unlikely to be examined						
Classroom climate						
31. The instructor creates a classroom atmosphere conducive to learning						
Instructor appears approachable, comes to class early and stays after the class to talk to students and answer questions						
32. The instructor encourages student participation						
Instructor encourages multiple perspectives.; students seem comfortable asking questions						
33. The instructor reacts to student professional behavior issues appropriately						
34. The instructor demonstrates flexibility in responding to student concerns or interests						
Instructor responds well to student differences; sensitive to individual interests, abilities, and experiences; listens carefully to student questions and comments; actively helpful when students need assistance.						
35. The instructor treats students impartially and respectfully.						

DNO	Did not observe	either because was not in the class for the entire lecture duration or instructor did not do and it's not applicable
NSD	Needs significant development	the instructor did not do this and should consider adding
ND	Needs development	the instructor attempted to do this but development/ revision is necessary
A	Accomplished	minor improvements can be recommended
AW	Accomplished well	no recommendations for improvement

Northeastern University Department of Pharmacy Practice ©2007

Section 3: Post observation meeting
The instructor should be prepared to discuss items below at this meeting

<u>Instructor's self-reflection on the lecture:</u>

How do you think the class went?

Is there anything you wanted to accomplish but were unable to do so? If yes, what was it and was it critical? What would you do differently next time to accomplish it?

In your opinion what went really well? Can you provide evidence that it went well?

In your opinion what did not go well? Can you provide evidence that it did not go well?

For items on the pre-observation and classroom observation forms where you gave yourself a rating of "Needs Development" or "Need Significant Development" what are your plans for improvement?

Did the lecture affect or change your plans for assessment?

Did you collect student evaluations of this lecture? What did students report to be the most effective and least effective aspects of your lecture?

What other constructive feedback did you receive through student evaluations? How do you plan to address it?

DNO	Did not observe	either because was not in the class for the entire lecture duration or instructor did not do and it's not applicable
NSD	Needs significant development	the instructor did not do this and should consider adding
ND	Needs development	the instructor attempted to do this but development/ revision is necessary
A	Accomplished	minor improvements can be recommended
AW	Accomplished well	no recommendations for improvement

Northeastern University Department of Pharmacy Practice ©2007

Section 4: Post assessment meeting
(within 2 weeks of first major assessment)

The instructor should provide the original questions submitted for the exam, final version of exam questions if applicable, and results of any item analysis if available prior to this meeting.

	DNO	NSD	ND	A	AW	Comments
36. The examination content is representative of the lecture content and objectives						
37. The tests used in the course have been well designed and selected						
The examination questions are clearly written.; the examination questions are of appropriate length and level of challenge; assessments include activities to assess higher order thinking						
38. Students are given ample time to complete the assignments and take-home examinations.						
39. The instructor determines the degree of mastery of lecture objectives						
Exam item analysis is performed						

Final Comments and Recommendations (limited to 2-3) will be provided in a letter form
Comments on classroom observation will be provided at the post-observation meeting
Comments on assessment will be provided at the post-assessment meeting

DNO	Did not observe	either because was not in the class for the entire lecture duration or instructor did not do and it's not applicable
NSD	Needs significant development	the instructor did not do this and should consider adding
ND	Needs development	the instructor attempted to do this but development/ revision is necessary
A	Accomplished	minor improvements can be recommended
AW	Accomplished well	no recommendations for improvement

Northeastern University Department of Pharmacy Practice ©2007