JET
The Journal of Effective Teaching
an online journal devoted to teaching excellence

Teaching Dossier Documents: A Comparison of Importance by Major Stakeholders

Charles A. Burnap, Gary F. Kohut[1], and Maria G. Yon
University of North Carolina at Charlotte, Charlotte, NC, 28223

Abstract

Most institutions of higher education require evidence of effective teaching as part of the review process for reappointment and for tenure/promotion. This study examines the relative importance of various documents placed in the teaching section of the dossier as rated by pre-tenured faculty members and review committee members. Results indicate that these documents are viewed quite differently by the two groups. There is a danger that the differences in the value that various documents play may affect the evaluation of teaching effectiveness and the final outcome of reappointment and tenure/promotion decisions. These differences weaken the evaluation process. There must be a common understanding of the importance of particular documents and how those pieces of evidence assist in telling the story of teaching effectiveness.

Keywords: Teaching Dossiers, teaching portfolios, teaching evaluation, teaching effectiveness.

One of the most important and controversial issues during the last 20 years has been the re-evaluation of how faculty performance should be measured and assessed. While traditional methods of faculty assessment can easily quantify the number of student instructional hours offered, activities assigned, and exams given, traditional methods do little to capture evidence of learning effectiveness.

Evaluating teaching effectiveness is a task performed by nearly all faculty in almost every institution of higher education. Yet, the evaluation of teaching is controversial, particularly as it relates to the core elements of teaching effectiveness. Ory (2000) examined the evolution of teaching evaluations from the 1970s (aimed at developmental purposes including faculty improvement) through the 1980s and 1990s (driven by administrative needs related to budget restrictions) through the 2000s (driven by demands for accountability, a renewed interest in teaching improvement, and the search for more valid metrics of teaching effectiveness).

Shao, Anderson, and Newsome (2007) acknowledge the considerable debate concerning the evaluation of teaching effectiveness. They found that there is a significant difference between what respondents feel is currently being used and what should be used to evalu-

[1] Corresponding author's email: gfkohut@uncc.edu

ate teaching effectiveness. With regard to teaching effectiveness, respondents tend to believe that currency in field, peers evaluations, classroom visits, and professor's preparation should be given more weight, while teaching awards and use of technology should not be given as much weight as they currently are. Respondents differed in their opinions depending on whether they were from doctoral, masters, or baccalaureate schools; on whether they were faculty, department chairs or deans; and on years of experience they had in higher education. Some research has found a high degree of similarity between what instructors and students considered important to effective teaching (Miller et al., 2001; Keller, Mattie, Vodanovich, & Piotrowski 1991).

One tool to help evaluate teaching effectiveness is the teaching dossier or portfolio. Broadly speaking, the dossier is a document that offers evidence of teaching/learning effectiveness such as student ratings and classroom observation reports in addition to a teaching philosophy, evidence of instructional innovations, assessment efforts, and teaching-related activities such as mentoring, classroom research, and related work. It is a structured collection of teacher and student work created over time that is framed by reflection and enriched through collaboration (Wolf & Dietz, 1998). Teaching dossiers can have a range of applications useful to teachers in college and university environments (Ouellett, 2007). A dossier is flexible; it might be useful as a developmental tool (formative purpose) to stimulate individual reflection and personal development, or it can be used as an evaluative tool (summative purpose) to offer support for tenure and promotion decisions, to supplement job application information, and to offer evidence of teaching effectiveness for teaching awards.

A potential barrier with evaluating teaching, however, is that evaluations can simultaneously serve formative and summative roles (Casey, Gentile & Bigger, 1997; Mills and Hyle, 1999; Pratt, 1997; Smith and Tillema, 2007). Seldin (2004) and Bernstein (1996) believe that teaching portfolios have the capacity to be used for assessment of teaching quality as well as development of enhanced teaching performance. Cavanagh (1996) argued to keep the two purposes separate and to conflate them, where the peer is also the evaluator, may create situations where faculty may not want to risk having frank discussions of weaknesses necessary to improve teaching. Tigelaar, Dolmans, Wolfhagen and van der Vleuten (2005), in their study of teaching portfolio assessment, suggest that an integration of formative and summative is possible, but acknowledged the considerable effort involved in achieving each end. De Rijdt, Tiquet, Dochy, and Devolder (2006) found fewer than a quarter of academics producing portfolios, partially due to the ambivalence surrounding the multiple functions. Conrad and Bowie (2006) found that the context in which portfolios are read, summative or formative, yield distinct outcomes.

Clearly, there is a tension between formative and summative purposes. Knapper and Wright (2001) claim that the differences between summative and formative portfolios are "not as great as might be expected" (p.25). On the other hand, Chism (2007) notes that the distinction between formative and summative reviews is often difficult to sustain in practice. Pan et al. (2009) note that formative purposes require precise and specific feedback as a guide to improving teaching performance; summative purposes need an overall parsimony that facilitates the evaluation process. Buckridge (2008) questioned whether

portfolios can have mixed formative and summative purposes and still meet the needs of institutions. She notes that while there has been some success to fulfill both purposes, such success can diminish the currency of teaching and replace deeper transformative opportunities with ore expedient concerns of faculty promotion.

Additionally, some experts argue that the definition and measurement of effective teaching are ambiguous and subjective (Dilts, 1980) and that the evaluation of teaching is unsystematic (Seldin, 1984). On the other hand, there are those who believe that teaching can be effectively defined and measured (Cashin, 1988). Chism (1999) found that there is a great deal of consensus on what characterizes effective teaching. Among those factors that are consistently mentioned are subject matter competence, preparation and organization, enthusiasm, and interpersonal rapport. Fink (2008) identified four fundamental tasks of teaching: knowledge of subject matter, designing learning experiences, interacting with students, and course management. He argued that if the teacher does all four well, students will have a good learning experience. To the degree that the teacher does one or more poorly, the quality of the learning experience declines (p. 39). Arreola (2007) identified five broad skill dimensions required for competent teachers: content expertise, instructional design skills, instructional delivery skills, instructional assessment skills, and course management skills. Hatch (2006) described teaching as "a complex intellectual endeavor that demands disciplinary expertise, a deep understanding of students, and sophisticated pedagogical skills."

Many faculty members believe that it is far easier to evaluate the quality of research than of teaching due to its explicit and tangible outcome, generally by number and value of publications. Furthermore, publications can be reviewed by outside experts in a particular field. The effects of teaching are not as easily quantifiable, and rarely is the act of teaching or its products reviewed by experts outside one's institution.

At most institutions, tenure and/or promotion decisions are based on the evaluation of a faculty member's teaching, research, and service. As difficult as some believe it is to evaluate, teaching is an essential part of a faculty member's promotion and tenure dossier. Therefore, faculty members are required to include a section that describes the faculty member's teaching activities and includes indicators of teaching effectiveness.

In order to present a comprehensive picture of teaching, a faculty member includes documents and artifacts that represent products of good teaching. Often included are administrator's annual reviews, course syllabi, teaching materials, teacher-developed tests, journals or diaries, videotapes of teaching, peer observation reports, samples of student work, letters of support, and student course evaluations. Statements of educational philosophy and narratives that help to interpret the artifacts may be included in the dossier and may provide evidence of a faculty member's teaching effectiveness. According to Edgerton, Hutchings, and Quinlan (1991) the ideal dossier should highlight a professor's reflections about a sample of actual work explaining his or her instructional decisions. The dossier should also include documentation of any classroom research that was carried out and an explanation of changes made as a result of the investigation. In addition to the dossier itself, it is not unusual to present a box filled with supporting materials to

the tenure and promotion committee. Cannon (2001) proposed that the three essential components of a teaching portfolio are the personal statement of teaching, an overview of teaching accomplishments and activities, and verification of the success of the activities through feedback from colleagues (classroom observation reports) and students (student evaluations).

Although considerable research exists concerning the relative importance of teaching portfolios, much less work exists on the assessment of portfolios for summative purposes (Burroughs, 2001; Smith & Tillema, 2003, 2007). Reviewers make important decisions about the effectiveness and contributions of faculty members by evaluating the documents the faculty member has placed in the dossier. However, research has supported that those implementing teaching portfolios have different views from those experiencing their direct and indirect impacts (Leggett & Bunker, 2006; Taylor, 2001; Guest & Duhs, 2003; Stephenson, 2004). At issue here is the match between the relative importance the faculty member puts on individual documents placed in the teaching section of the dossier and the importance the review committee and administrators place on these same elements. Naturally, a faculty member would emphasize a particular element in some way in order to present the best case. A mismatch by reviewers and the faculty member involving the perceived importance of particular elements in the teaching dossier could possibly cause a faculty member to misrepresent his or her teaching to the degree that tenure and/or promotion might be jeopardized.

This paper is part of a larger study in which various aspects of teaching across the university were examined. The purpose of this paper is to make a comparison between pretenured faculty members and tenured reviewers regarding the relative importance ascribed to various documents placed in the teaching section of the reappointment or tenure/promotion dossier.

Method

The research was completed at a southeastern urban research university of about 17,000 students, 700 faculty members, and six colleges, with the proposal of a seventh college recently accepted. The first doctoral programs were begun in 1996. At the time of the study, six doctoral programs were in existence and three more programs had been granted permission to plan. (The university has continued to expand in the interim with more students, more faculty, and more doctoral programs.) Since the establishment of doctoral programs, the institution has been moving quickly toward a research agenda. While research expectations for tenure and promotion have increased, the university maintains its focus on the importance of teaching. Several examples illustrate this point. Among them are two university level awards: Each fall, the Bank of America Award for Teaching Excellence is presented at the end of a week of activities that highlight teaching. The Provost's Award for Excellence in Teaching is awarded annually to a department that has demonstrated special emphasis on teaching. In addition, the title of Cone Distinguished Professor for Teaching has been awarded to a number of faculty members who excel in this area. Finally, individual college awards are given each year to faculty members who show evidence of distinguished accomplishments in the area of teaching.

Procedures

Instrumentation. Respondents were asked to consider the usefulness of fifteen documents or artifacts that are usually included in the teaching section of a faculty member's dossier for reappointment or for promotion and tenure (see Table 1 for the specific documents). They were asked to rate each item in terms of its usefulness in evaluating the faculty member's teaching. The following scale was used: very useful, somewhat useful, minimally useful, not useful. We note that a neutral response was not included in order to force the respondent to decide whether they lean more towards one end of the scale or the other. A fifth choice, not applicable (NA), was also included.

Subjects

Two different groups were the focus of this study: pre-tenured faculty members and faculty members who serve as reviewers for reappointment and tenure/promotion cases. Questionnaires were distributed to 343 tenured faculty members, including administrators, in order to capture all who might have been in the position of reviewing faculty for reappointment or tenure and promotion. The questionnaires were delivered to department review committee members and college review committee members in the middle of the fall semester to coincide with the completion of their peer reviews. Questionnaires were distributed to department chairs and college deans at the end of the same semester to coincide with the timeline for their review. Responding to the questionnaires as soon as the round of reviews was completed was considered imperative in giving results that reflected how reviewers used specific documents in the teaching portfolio. Questionnaires were mailed to 127 faculty members who had just been reviewed an all pre-tenured faculty members. New hires were not included due to the fact that they were less familiar with the tenure process at the time of the survey. These numbers represent the entire population involved in the promotion and tenure process at the university.

Results

One hundred fifty tenured faculty members responded to the survey. Although surveys were sent to all tenured faculty members and administrators, only those who served on a reappointment or tenure committee during the semester of the study were asked to comment on the relevance of teaching portfolio documents. Eighty of the one hundred fifty respondents fell into this category, and were able to evaluate the relevance of portfolio documents. Eighty pre-tenured faculty members returned the questionnaire and 53 (those whose dossiers were under review) rated the importance of portfolio documents.

Results were averaged using the following scale: very useful = 3, somewhat useful =2, minimally useful =1, not useful = 0. Missing data or responses that indicated an item was not applicable were not included in the averages. A summary of results is found in table 1. Portfolio items are listed in column one. The data in columns two and three correspond to the two groups (tenured reviewers or pre-tenured faculty) of respondents. Items are ordered in descending importance as judged by the tenured reviewers.

TABLE 1. Comparison of the Relative Importance of Various Components of a Teaching Dossier.

Dossier Item	Reviewers *Item Mean*	Pre-tenured Faculty *Item Mean*
Contributions to teaching within the university	2.43 (rank = 1)	2.16 (rank = 6)
Course and curriculum development	2.42 (rank = 2)	2.11 (rank = 7)
Teaching awards	2.41 (rank = 3)	2.03 (rank = 9-10)
Student evaluations	2.40 (rank = 4)	1.70 (rank = 15)
Contributions to teaching nationally or internationally	2.38 (rank = 5)	2.10 (rank = 8)
Faculty narrative	2.34 (rank = 6)	2.24 (rank = 3)
Department chair's annual review	2.32 (rank = 7)	2.34 (rank =1)
Peer observation reports	2.32 (rank = 8)	1.87 (rank = 12-13)
Description of teaching activities	1.92 (rank = 9)	2.27 (rank = 2)
Letters from outside reviewers	1.85 (rank = 10)	1.89 (rank = 11)
Course syllabi	1.81 (rank = 11)	2.18 (rank = 5)
Letters from faculty within the university	1.59 (rank = 12)	2.03 (rank = 9-10)
Student work/artifacts	1.57 (rank = 13)	2.22 (rank = 4)
Course exams/tests/quizzes	1.44 (rank = 14)	1.86 (rank = 14)
Letters from students	1.36 (rank = 15)	1.87 (rank = 12-13)

Notes:

3 = very useful
2 = somewhat useful
1 = minimally useful
0 = not useful
Rank 1 = highest ranking in terms of importance (most important)
Rank 15 = lowest ranking in terms of importance (least important)

It is apparent that the means in column two (tenured reviewers) fall into two distinct groups: There were eight items with means greater than or equal to 2.32 (group one) and seven items with means less than or equal to 1.92 (group two). Dependent *t*-tests were done comparing review committee responses to pairs of items in group one. At a 95% confidence level, there were no significant statistical differences within this group. However, the same type of tests showed that any pair that included an item from group one and an item from group two, showed statistically significant differences. This two-tiered rating was not apparent among pre-tenured faculty. Indeed, although the overall means of the fifteen items were roughly the same for both groups (overall mean for column 2 is 2.04, while the overall mean for column 3 is 2.06), the pre-tenured faculty members ratings were generally closer to the mean—only ranging from 1.70 to 2.34. Dependent *t*-tests do, however, show that pre-tenured faculty had a statistically significant preference for the chair's annual review, a description of teaching activities, faculty narrative, and

student work over peer observation reports, letters from students, course exams, and student evaluations.

Both tenured reviewers and pre-tenured faculty are in general agreement about the relative importance of several portfolio items: letters from outside reviewers (on teaching), letters from students, and sample course exams are relatively less important, while the faculty narrative and the chair's annual review are relatively important. There are, however, striking differences on some items: Tenured reviewers rate teaching awards, student evaluations, and peer observation reports highly (3^{rd} , 4^{th} and 8^{th} --all in the top tier), while pre-tenured faculty members rate these items 9^{th} , 15^{th}, and 12^{th}. Reversing this difference, tenured reviewers ranked the description of teaching activities, course syllabi, and student work/artifacts as second tier items while pre-tenured faculty ranked them 2^{nd}, 5^{th}, and 4^{th} on their list. Table 2 illustrates these comparisons while taking into account the fact that reviewers and pre-tenured faculty have different levels of variation among items. This table was constructed by standardizing the entries in each column of table 1. A positive score in Table 2, column 2 indicates that tenured reviewers rated this item above their mean response of 2.04 while a negative value indicates that this item rated below the mean. The values measure the number of standard deviations above or below the mean. Column 3 of Table 2 was created for pre-tenured faculty in the same way. The largest differences in column 4 are indications of divergent opinions (in the relative importance of an item) between the two groups (tenured reviewers vs. pre-tenured faculty).

In this article, we generally focus on the relative value ascribed to various dossier items. However, for those who wish to directly compare the average response of the two groups for a given item, please refer to Table 3. This table was generated using a 2-tailed *t*-test for 2 samples with unequal variance. *p*-values that are less than 0.05 indicate statistically significant differences. This analysis does not address the relative importance of a given item, but instead gives a measure of the probability that the responses for reviewers and pre-tenured faculty members come from populations with the same mean (for that dossier item). The smaller the *p*-value, the less likely that the two populations have the same mean. Even items with strong relative agreement such as course exams or letters from students can show significant differences in means.

Discussion and Conclusions

The data reported above indicates that tenured reviewers and pre-tenured faculty members do not rate all portfolio items in a similar manner. It is interesting that tenured reviewers clearly separate portfolio items into two distinct tiers: one with significant value to their review process and the other with much less value. Although pre-tenured faculty had a clear (and statistically significant) preference for chair's annual review, a description of teaching activities, faculty narrative, and student work when compared with peer observation reports, letters from students, course exams, and student evaluations, they were, in general, much more egalitarian in their view of the importance of various dossier items.

TABLE 2. Comparison of Z-scores of Various Components of a Teaching Dossier.

Dossier Item	Reviewers *Z-score*	Pre-tenured Faculty *Z-score*	Difference
Contributions to teaching within the university	0.974	0.550	0.424
Course and curriculum development	0.949	0.280	0.669
Teaching awards	0.925	-0.151	1.076
Student evaluations	0.900	-1.93	2.83
Contributions to teaching nationally or internationally	0.850	0.226	0.624
Faculty narrative	0.751	0.981	-0.230
Department chair's annual review	0.701	1.52	-0.819
Peer observation reports	0.701	-1.01	1.715
Description of teaching activities	-0.291	1.14	-1.434
Letters from outside reviewers	-0.465	-0.906	0.441
Course syllabi	-0.564	0.658	-1.222
Letters from faculty within the university	-1.11	-0.151	0.959
Student work/artifacts	-1.16	0.874	-2.033
Course exams/tests/quizzes	-1.48	-1.07	-0.414
Letters from students	-1.68	-1.01	-0.667

Notes:

Mean for reviewers = 2.04 Mean for pre-tenured faculty = 2.06
Standard deviation = 0.403 Standard deviation = 0.185
Zscore = (item –mean)/(standard deviation)

In view of the high stakes nature of the review process, it is worrisome that some dossier documents are viewed quite differently by reviewers as compared to those faculty members being reviewed. Review committees tend to highly rate the items that have long been regarded as important measures of teaching quality/effectiveness. These items include course and curriculum development, teaching awards, and student evaluations. We note that documentation of curriculum development and teaching awards are more in the nature of the objective evidence that is common when evaluating research. The faculty narrative and peer observation reports are newer requirements and perhaps more subjective, but highly valued. The authors of this article were surprised that student work/artifacts were not more highly regarded by review committees. We speculate that the perceived value may increase as committees become more accustomed to evaluating this material.

TABLE 3. ***P*-values for a 2-tailed *t*-test for 2 samples with unequal variance.**

Dossier Item	*p*-value
Contributions to teaching within the university	0.0639
Course and curriculum development	0.0159
Teaching awards	0.0210
Student evaluations	0.0000104
Contributions to teaching nationally or internationally	0.104
Faculty narrative	0.500
Department chair's annual review	0.914
Peer observation reports	0.00209
Description of teaching activities	0.00670
Letters from outside reviewers	0.959
Course syllabi	0.0097
Letters from faculty within the university	0.0135
Student work/artifacts	0.0000967
Course exams/tests/quizzes	0.0275
Letters from students	0.00576

Notes: For each of the fifteen items, this table compares the responses of review committee members to those of pre-tenured faculty. A *p*-value less than 0.05 means that the difference in responses of these two groups was statistically significant.

Pre-tenured faculty downplayed the importance of student evaluations, teaching awards, and peer observation reports. The chair's annual review, description of teaching activities, faculty narrative, student work/artifacts, and course syllabi were deemed most important. Many of these items give the faculty member a chance to discuss teaching philosophy, goals, and priorities. These items allow pre-tenured faculty a chance to interpret rather than simply report on classroom activities.

While the chair's annual review is a first tier (highly valued) item, review committees do not rate it as highly (relative to other items in the dossier) as pre-tenured faculty members. Presumably, this reflects the review committee's desire to form its own judgment using independent evidence. Obviously, more evidence is generally available at promotion/tenure time than during a typical annual review. In contrast, the chair's review may provide the only significant feedback available to pre-tenured faculty. It is reasonable that their self-assessment is strongly influenced by this evidence.

Peer observation has become an integral part of the evaluation of untenured faculty members at our institution. In 1994, the state assembly required classroom observations of untenured faculty to be part of the review process. Throughout the university, observation guidelines require both pre- and post-observation meetings. The pre-observation meeting allows the untenured faculty member to set the observed class in context and to discuss instructional goals for that class. The post-observation meeting allows the observer to give feedback as well as allowing the untenured faculty member to comment on

any actions that they may have taken during the observed class. Student evaluations are also required in each class taught during the academic year. It is surprising that pre-tenured faculty members have a relatively low regard for both student and peer observations. These are the only direct evidence of classroom technique that is available to the review committee.

One may try to account for the relatively low value that pre-tenured faculty place on peer observation by supposing that the process was particularly stressful, or noting that since very few classes are actually observed, the observed classes may not be representative. Finally, one could conjecture that if peer observation reports tend to include only positive comments, they may not be seen as particularly useful. None of these reasons seem to be valid. In a related article, Kohut, Yon, & Burnap (2007) examined perceptions of observers and those observed during the peer review process. Data was gathered from the same group of pre-tenured faculty currently under discussion. When these pre-tenured faculty were asked whether the observed classes were representative of their teaching, 54.2% strongly agreed, 31.9 % agreed, 11.1% were neutral, 2.8% disagreed, and 0% strongly disagreed. When responding to the statement "Having my classes observed was stressful," 11.3% strongly agreed, 14.1% agreed, 21.1% were neutral, 32.4% disagreed, and 21.1% strongly disagreed. When responding to the statement "Peer observers tend to include only positive comments in their reports," 5.7% strongly agreed, 20% agreed, 34.3% were neutral, 28.6% disagreed, and 11.4% strongly disagreed. Responses to this survey also indicated that pre-tenured faculty rated peer observations as more useful, valid, and reliable than did those doing the observations. None of this seems to indicate that pre-tenured faculty would place such a low value on these reports. Perhaps the fact that peer observations may have formative as well as summative value accounts for this seeming inconsistency. The formative nature is more apparent in the post-observation discussion, while during the promotion/tenure review, peer observation reports are merely summative.

The evaluation of teaching will remain controversial until there is a consensus on what *constitutes* evidence of effective teaching. Agreement on the core issues of teaching rated in a portfolio would be helpful (Cochran-Smith & Fries, 2002; Delandshere & Arens, 2001, 2003) but such an agreement has yet to be adopted in the teaching profession (Babin, Shaffer, & Morgan, 2002; Murphy & MacLaren, 2007; Murphy, MacLaren, & Flynn, 2009; Zuzowsky & Libman, 2002). Work by Sorcinelli (2006, 1993) and Chickering and Gamson (1987) have helped in this regard. Sorcinelli suggested that good teaching has five dimensions: dynamism/enthusiasm, command of the subject, organization/clarity, faculty-group interaction, and faculty-student contact. However, these exists little agreement in this set of characteristics although Chickering and Gamson suggest a set of behaviors that help guide instructors interested in improving their teaching such as giving prompt feedback, communicating high expectations, and respecting diverse talents and ways of learning.

The responses to this survey indicate that there are some significant points of disagreement. This undermines the evaluation process. If pre-tenured faculty and reviewers making high-stakes decisions see the importance of portfolio documents differently, a

faculty member's choices and presentation of this material may adversely affect the promotion/tenure decision. We hope that this article will provide insights into the nature of these differences, and a basis for a collegial discussion of these issues.

Teaching must be recognized within institutions for teaching dossiers to become more widely adopted in higher education. Central to this is having faculty and administrators adopt similar views of the purpose of portfolios and the relative importance of portfolio content. As Arreola (2007) observed, only when the elements of a faculty evaluation program (summative purpose) are carefully integrated into a professional enrichment program (formative purpose) does the institution obtain the greatest benefit.

References

Arreola, R.A. (2007). *Developing a comprehensive faculty evaluation system: A guide to designing, building, and operating large-scale faculty evaluation systems* (3rd ed.). Bolton, MA: Anker.

Babin, L.A., Shaffer, T.R., & Morgan, T.A. (2002). Teaching portfolios: Uses and development. *Journal of Marketing Education* 24, 35.

Bernstein, D.J. (1996). A departmental system for balancing the development and evaluation of college teaching: A reply to Cavanagh. *Innovative Higher Education* 20, 241-248.

Buckridge, M. (2008). Teaching portfolios: Their role in teaching and learning policy. *International Journal for Academic Development* 13, 117-127.

Burroughs, R. (2001). Composing standards and composing teachers: The problem of National Board Certification. *Journal of Teacher Education* 52, 223-232.

Cannon, R. (2001). Broadening the context for teaching evaluation. In Knapper, C. and Cranton, P. Eds. *Fresh approaches to the evaluation of teaching. New directions for teaching and learning.* San Francisco, CA: Jossey-Bass, 87-97.

Casey, J.R., Gentile, P., & Bigger, S.W. (1997). Teaching appraisal in higher education: An Australian perspective. *Higher Education* 34, 459-482.

Cashin, W. E. (1988). *IDEA Paper No. 20. A summary of research.* Kansas State University, Manhatten, Center for Faculty Evaluation and Development.

Cavanagh, R. (1996). Formative and summative evaluation in the faculty peer review of teaching. *Innovative Higher Education* 20, 235-240.

Chickering, A.W., & Gamson, Z.F. (1987). Seven principles for good practice in undergraduate education. *Wingspan Journal* 9, Special insert.

Chism, N.V. (2007). *Peer review of teaching: A sourcebook* (2nd ed.). Bolton, MA: Anker.

Chism, N.V. (1999). *Peer review of teaching: A sourcebook*. Bolton, MA: Anker.

Cochran-Smith, M., & Fries, M.K. (2002). The discourse of reform in teacher education: Extending the dialogue. *Educational Researcher* 31, 26-28.

Conrad, L., & Bowie, C. (2006). The impact of context on readings of teaching portfolios. *International Journal for Academic Development* 11, 31-42.

Delandshere, G., & Arens, S.A. (2001). Representations of teaching and standard-based reform: Are we closing the debate about teacher education? *Teaching and Teacher Education* 17, 547-566.

Delandshere, G., & Arens, S.A. (2003). Examining the quality of the evidence in pre-service teacher portfolios. *Journal of Teacher Education* 54, 57-73.

De Rijdt, C., Tiquet, E., Dochy, F., & Devolder, M. (2006). Teaching portfolios in higher education and their effects: An explorative study. *Teaching and Teacher Education* 22, 1084-1093.

Dilts, D. A. (1980). A statistical interpretation of student evaluation feedback. *Journal of Economic Education* Spring, 10-15.

Edgerton, R., Hutchings, P., & Quinlan, K. (1991). *The teaching portfolio: Capturing the scholarship of teaching.* Washington, DC: American Association of Higher Education.

Fink, D. (2008). Evaluating teaching: A new approach to an old problem. In S. Chadwick-Blossey and D.R. Robertson (eds.), *To improve the academy: Resources for faculty, instructional, and organizational development*, 26, 3-21. San Francisco: Jossey-Bass.

Guest R., & Duhs, A. (2003). Quality assurance and the quality of university teaching. *Australian Journal of Education* 47, 40-57.

Hatch, T. (2006). *Into the classroom: Developing the scholarship of teaching and learning.* San Francisco, CA: Jossey-Bass.

Keller, J.W., Mattie, N., Vodanovich, S.J., & Piotrowski, C. (1991). Teaching effectiveness: Comparisons between traditional and nontraditional college students. *Innovative Higher Education* 15, 177-184.

Knapper, C., & Wright, A. (2001). Using portfolios to document good teaching: Premises, purposes, practices. *New Directions for Teaching and Learning*, 88, 19-29.

Kohut, G., Yon, M., and Burnap, C. (2007, Winter). Peer observation of teaching: perceptions of the observer and observed. *College Teaching* 55, 19-25.

Leggett,M., & Bunker, A. (2006). Teaching portfolios and university culture. *Journal of Further and Higher Education* 30, 269-282.

Miller, J.L., Dzindolet, M.T., Weinstein, L., Xie, Xiaolin, & Stones, C.R. (2001). Faculty and students' views of teaching effectiveness in the United States, China, and South Africa. *Teaching of Psychology* 28, 138-142.

Mills, M., & Hyle, A.E. (1999). Faculty evaluation: A prickly pair. *Higher Education* 38, 351-371.

Murphy, T., MacLaren, I., & Flynn, S. (2009). Toward a summative system for the assessment of Teacher quality in higher education. International Journal of Teaching and Learning in Higher Education 20, 226-236.

Murphy, T., & MacLaren, I. (2007). Theory and practice in the development of teaching portfolios at NUI Galway. In C. O'Farrell (Ed*.), Teaching portfolio practice in Ireland: A handbook* (p. 86-100). Dublin, Ireland: AISHE.

Ory, J.C. (2000). Teaching evaluation: Past, present, and future. *New Directions for Teaching and Learning*, 83, 13-18.

Ouelett, M.L. (2007). Your teaching portfolio: Strategies for initiative and documenting growth and development. *Journal of Management Education* 31, 421-433.

Pan, D., Tan, G.S.H., Ragupathi, K., Booluck, K., Roop, R., & Ip, Y.K. (2009). Profiling teacher/teaching using descriptors derived from qualititative feedback: Formative and summative applications. *Research in Higher Education* 50, 73-100.

Pratt, D.D. (1997). Reconceptualizing the evaluation of teaching in higher education. *Higher Education* 34, 23-44.

Sorcinelli, M.D. (1993). *Dimensions associated with effective teaching*. Amherst: University of Masachusetts, Center for Teaching.

Sorcinelli, M.D. (2006). *Dimensions associated with effective teaching*. Amherst: University of Masachusetts, Center for Teaching.

Seldin, P. (2004). *The teaching portfolio: A practical guide to improved performance and promotional/tenure decisions* (3rd. ed.). Boston, MA: Anker Publishing Co.

Seldin, P. (1984). *Changing practices in faculty evaluation: A critical assessment and recommendations for improvement*. San Francisco: Jossey-Bass.

Shao, L.P., Anderson, L.P., & Newsome, M. (2007). Evaluating teaching effectiveness: Where we are and where we should be. *Assessment & Evaluation in Higher Education* 32, 355-371.

Smith, K., & Tillema, H. (2003). Clarifying different types of portfolio use. *Assessment & Evaluation in Higher Education* 26, 625-648.

Smith, K., and Tillema, H. (2007). Use of criteria in assessing teaching portfolios: Judgemental practices in summative evaluation. *Scandinavian Journal of Educational Research* 51, 103-117.

Stephenson, S.L. (2004). Saving quality from quality assurance. *Perspectives* 8, 62-67.

Taylor, J. (2001). Improving performance indicators in higher education: The academics perspective. *Journal of Further and Higher Education* 25, 379-393.

Tigelaar, D., Dolmans, D., Wolfhagen, I., & van der Vleuten, C. (2005). Quality issues in judging portfolios: Implications for organizing teaching portfolio assessment procedures. *Studies in Higher Education* 30, 595-610.

Wolf, K., & Dietz, M. (1998). Teaching portfolios: Purposes and possibilities. *Teacher Education Quarterly* 25, 9-22.

Zuzowsky, R., & Libman, Z. (2002, August). Standards of teaching performance and teacher tests: Where do they lead us. Paper presented at ATEE conference, Warsaw.