Toward a Summative System for the Assessment of Teaching Quality in Higher Education

Timothy Murphy, Iain MacLaren, and Sharon Flynn
National University of Ireland Galway

This study examines various aspects of an effective teaching evaluation system. In particular, reference is made to the potential of Fink's (2008) four main dimensions of teaching as a summative evaluation model for effective teaching and learning. It is argued that these dimensions can be readily accommodated in a Teaching Portfolio process. The Teaching Portfolio initiative that is in use for the Postgraduate Certificate Programme in Teaching and Learning in Higher Education at the National University of Ireland Galway is a case in point. The challenges encountered when attempting to develop mechanisms for the summative evaluation of quality teaching are explored, as well as some of the possibilities for their resolution.

It is recognized that the task of assessing the nature, quantity and quality of teaching is a highly complex activity. As Chism (1997) pointed out, "any good approach to evaluating teaching will reflect the complexity of teaching itself" (p. 7). It has even been suggested that it is not easy to arrive at a consensus of what constitutes good or effective teaching (Babin, Shaffer, & Morgan, 2002; Casey, Gentile, & Bigger, 1997; Murphy & MacLaren, 2007). The challenge at arriving at such a consensus is compounded because there are real variations in teaching quality "in different courses, between different subject areas, and within subject areas" (Casey et al., 1997, p. 462). While recognizing the complexities involved in any substantive discussions about the value of teaching, Chism (1999) did acknowledge the research that has consistently shown that there is a great deal of consensus on what characterizes effective teaching. Among those factors that are consistently advanced are subject matter competence, preparation and organization, clarity, enthusiasm, and interpersonal rapport (Chism).

Arreola (2007) does appreciate the extent to which the concept of *excellence* has invaded the current lexicon of higher education. With reference to faculty evaluation, he described this as the "Lake Woebegone" model. According to this view, higher education as a profession "is stuck in the silly verbal knot of expecting everyone to be what, by definition, only a few can be" (Arreola, 2007, p. 25). As a consequence, he pointed out that "the pursuit of excellence in higher education has resulted in many faculty receiving neither the time, resources, nor incentives, to develop the skills necessary to become competent teachers" (p. 25). Notwithstanding this challenge, however, he identified five broad skill dimensions required for teaching to emerge. These he specifies as content expertise, instructional design skills, instructional delivery skills, instructional assessment skills, and course management skills.

These dimensions are broadly in line with Fink's (2008) four fundamental tasks of teaching (see Appendix A). It is Fink's (2008) contention that,

> there is a direct relationship between how well a teacher performs these four fundamental tasks and the quality of the student's learning experience. 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)

Both Arreola and Fink then would be in sync with Hatch's depiction of teaching as "a complex intellectual endeavor that demands disciplinary expertise, a deep understanding of students, and sophisticated pedagogical skills" (Hatch, 2006, p. 11).

The summative teaching quality assessment model that is being proposed for implementation at the National University of Ireland Galway is cognizant of Hatch's insight. It aims to support the view that "effective evaluation of teaching requires some combination of evidence from the person whose teaching is being evaluated, from that person's students, and from professional colleagues" (Chism, 1999, p. xi). Such evidence is integral to the existing Teaching Portfolio initiative at NUI Galway. The portfolio is an essential component of the Postgraduate Certificate Programme in Teaching and Learning in Higher Education. In the subsequent sections, core aspects of each of the requirements for effective evaluation of teaching will be developed, specifically those pertaining to student feedback, peer review of teaching, and teaching portfolios.

As a precursor to this whole discussion, however, it is very important to be mindful of the link between teaching and learning. There is little discussion about quality teaching in higher education today that omits some considerations of the link between teaching and learning. Fink (2002) identified the primary purpose of

teaching as generating as much significant learning as possible. Arreola (2007) further elaborated on this link when he depicted teaching "as an interaction between a teacher and a student conducted in such a way that the student is provided with the opportunity to learn" (p. 18). Additionally, he stated that "Faculty must be able to design and deliver a set of experiences to the learner such that, if the learner engages the experiences, there is a high probability that learning will occur" (p. xx).

Quality Assessment of Teaching and Student Feedback

Student Feedback Systems

Feedback from the students was acknowledged by Chism (1999) as being an integral aspect of an effective teaching evaluation system. The extent to which some type of student feedback is used in the assessment of teaching performance was acknowledged by Cashin (1999). He made reference to a US Department of Education survey (1991) of over 40,000 department chairs. The analysis of which revealed that 97 percent of the chairs indicated that they used "student evaluations" to assess teaching performance. He adroitly acknowledged, however, that "there is almost universal agreement that data from a variety of sources, not just student ratings, are required to accurately evaluate teaching" (p. 28). Centra (1993) further underscores this point. He contends that "student evaluations represent only one source of information: student opinion" (p. 89).

Limitations of Student Feedback Systems

Cashin (1999) also put the limitations of student feedback into stark relief as he sketched out a composite overview of the various elements involved in any substantive quality review of teaching. He identified these as subject matter mastery, curriculum development, course design, delivery of instruction, assessment of learning, availability to students, and administrative requirements. He believed that most students know almost nothing about the first three and that, therefore, in order to comprehensively assess the quality of teaching more data is required than what can be obtained solely from student feedback.

Encouraging Faculty Cooperation

In his discussion about student ratings, Centra (1993) made an important contribution to the debate. He contended that such ratings are most likely to have an effect when academics learn something new about their teaching. He prefaced this by stating that

"involving faculty representatives in the decision about which form to use will help ensure the faculty's commitment to its use" (p. 21). There are some noticeable parallels here with Cashin's (1999) claim that

> It goes without saying that the more confidence faculty have in the reliability and validity of a teaching evaluation system, the more likely it is that they will pay attention to the resulting data. (p. 48)

Marincovich (1999) also made a number of helpful contributions to this debate. She was especially concerned about how to use end-of-term student data, as well as other sources of student feedback for the maximum benefit of the faculty member. Specifically, she made reference to Cohen's (1980) contention that faculty members receiving augmented feedback, or more specifically expert consultation, are much more likely to improve. She elaborated on the role that teaching consultation professionals can play here. She stated that when working with faculty,

> the consultant's most important contribution will be in helping clients to pick out those two or three aspects of their teaching in which improvements will have the greatest payoff for their students' learning and in helping to devise improvement strategies. (as cited in Marincovich, 1999, p. 48)

She is very conscious of the view that efforts such as these, however, will be in vain if a college or university's leadership does not "clearly signal the value that it puts on effective teaching and make that value unambiguous through its reward system" (Marincovich, 1999, p. 48).

Benefits of Peer Observation

As mentioned previously, student feedback is an important source for providing information about teaching quality but it is not the only source. In the following section, reference will be made to Fink's (2008) main dimensions of teaching. He demonstrated how student questionnaires provide data for one of these dimensions. He also contended, however, that such feedback is complimented by classroom observations. This requirement of the assessment of quality teaching will be developed in the next section, which also includes a fuller description of Fink's (2008) model and of its significance for summative assessment of quality teaching.

Quality Assessment of Teaching and Peer Review of Teaching

An Inclusive Understanding

While the literature on the peer review of teaching (PRT) certainly recognizes that PRT does involve some observation on actual classroom practice, the literature is also aware that PRT is not necessarily restricted solely to this activity. PRT has also been understood to include reviews of learning materials, assessment and methods of evaluating teaching (see Beaty & McGill, 1995; Gosling, 2005; Keig & Waggoner 1994). Gosling (2005) provided a helpful outline of three PRT models: an evaluative, a developmental, and a collaborative model. His evaluative model aligns closely with the intended outcomes of the summative PRT process that is being proposed for implementation at the National University of Ireland, Galway.   Gosling (2005) specified the intended outcomes of such a process as

- assuring the quality of teaching,
- assisting staff to identify weaknesses in their teaching and put in place an action plan to remedy them,
- helping staff to prepare for internal or external audit processes,
- deciding whether a staff member should successfully complete probationary requirements, or achieve promotion, and
- assuring the quality of teaching and the student learning experience

Applications of Fink's Approach

In terms of designing a summative PRT process for implementation at NUI Galway, it is being proposed that Finks' four fundamental tasks of teaching, as outlined above, can provide a helpful initial template. At the core of his model is the correlation between the four tasks and the quality of the student's learning experience. In order to design a PRT process then, it could be argued that the following points should be carefully considered:

1. the reviewer's knowledge of the subject matter,
2. the effectiveness of the design or plan for the learning experience,
3. the nature of the interaction with the learner, and
4. the handling of course management issues.

The teaching evaluation model that Fink ultimately opts for, however, has a different combination. He contended that for the vast majority of academics, points 1 and 4 above are taken as a given and that therefore it is not necessary to include them. Instead, he proposed inserting gathering data about the learning achieved by students and also about the academic's efforts to improve over time. The following criteria then represent the main dimensions of teaching for him: design of courses, teacher-student interaction, quality of student learning, and getting better over time,

Although he provided us with an important generic rubric for the design and implementation of a summative faculty evaluation system, it is very important that each individual department be given maximum flexibility for designing a system that is most appropriate for their particular needs and contexts (Arreola, 2007).

Sources of Information on Teaching Quality

Regarding the selection and identification of sources of information, Arreola (2007) presented an important "rule-of-thumb." Arreola contended that "the important principle to follow in identifying sources is always to select the source which has the best opportunity to observe first-hand the performance to be evaluated" (p. 47). The application of this principle is clearly evident in Table 1.

Applying the rule-of-thumb for the peer review of teaching, Chism (1999), for example, claimed that

While students are the most appropriate judges of day-to-day teacher behaviors and attitudes in the classroom, they are not the most appropriate judges of the accuracy of course content, use of acceptable teaching strategies in the discipline and the like.

Table 1
Relationship Between Performance Criteria and Assessment Source

Criteria		Primary Source of Information
I.	The Design of Courses	Course design materials
II.	Teacher-student interaction	Student questionnaires, Observations
III.	Quality of student learning	Samples of student learning materials
IV.	Getting better over time	Teacher self-report, Documentation

Table 2
Low and High Standards for the Performance Criteria of Course Design

Low	High
• Is based on only a casual or cursory job of collecting information about the situational factors.	• Is based on a careful and thorough job of obtaining information on and analyzing situational factors.
• Does not have a clear statement of learning goals, only a list of topics to be covered.	• Has a clear statement of learning goals, and the goals go well beyond just learning the content and simple application skills.
• Does not use active learning, only passive learning, i.e., lectures and readings.	• In-class learning activities include active learning, e.g., experiential and reflective activities.
• Does not give frequent and immediate feedback to students on their learning, only 1 or 2 mid-terms and a final.	• Students receive feedback, on their learning, weekly and at times daily, un-graded as well as graded.
• Does not use a dynamic or powerful teaching strategy; most classes are just repetitions of the same learning activities, over and over.	• Teacher uses a teaching strategy with a combination and sequence of learning activities that build on each other and culminate in powerful, integrated learning.

For these kinds of judgments, peers are the most appropriate source of information. (p. 7)

At a later stage, she underscored this point further when she made reference to Cohen and McKeachie's (1980) point that "peer reviewers are in an ideal position to judge course content and design, materials and instruments used to assess student achievement" (as cited in Chism, 1999, p. 30)

Indicators for Quality Teaching

Fink (2008) claimed that it is possible to identify quality indicators for each of the four dimensions of teaching. What is needed, in his view, are specific standards that differentiate higher quality from lower quality teaching (see Appendix B). In Table 2, the standards for high and low quality course design are presented:

Summative Challenge for Evaluation of Teaching Quality

Arreola (2007) engaged in a critically important discussion about objectivity and the summative evaluation of teaching quality. In this connection, he introduced the concept of "controlled subjectivity" (p. xviii). He arrived at this juncture because he contended that "subjectivity in a faculty evaluation system is unavoidable" (p. xviii). In his view, "controlled subjectivity is the consistent application of a pre-determined set of values in the interpretation of measurement data" (p. xix).

His comments have implications for the entire peer review process including the peer observation of classroom teaching. De Zure (1999) offered a number of recommendations to enhance the reliability of such observations. He specified these as

1. training of observers
2. basing conclusions on more than one observation by more than one observer
3. consensus about what constitutes good teaching in the discipline with a focus on shared criteria for teaching effectiveness including the elements colleagues can judge best
4. consistency for all instructors and observers
5. the rules of the game should be known to all (the instructors, the observers, the reviewers or personnel committee)
6. the instructor should have input into the process at several stages (e.g., the selection of observers, selection of class to be observed, interpretation of the classroom experience after the observation, input into the written report).
7. a validated observation instrument should be used (Chism, 1999, 2007, *Peer Review of Teaching : A Sourcebook* contains sample documents that can be adapted for local use).

Links Between Formative and Summative Processes

Arreola (2007) also engaged in a significant conversation about the possible links that can exist between formative and summative processes on the quality of teaching in higher education. He contended that faculty evaluation systems which are implemented "without reference to professional enrichment

opportunities or programs, are inevitably viewed by faculty as being primarily punitive in intent" (p. xxii). Centra (1993) also acknowledged the benefits of using evaluation procedures formatively before adopting them summatively as teachers will become acquainted "with the procedures and the criteria to be used, and have the opportunity to improve their performance before being judged" (p. 5). Ideally, what you want to do then is establish a culture of open discussion and generous critique so that when you move to actually collecting data for decision making purposes, the conversations are more a part of the routine.

Links between Formative and Summative Processes and Developments at NUI Galway

The Partnerships for Learning and Teaching initiative that is being considered for implementation at the National University of Galway is in sync with the previously mentioned comments about the links that can exist between formative and summative processes. It is congruent with developmental processes at other universities where the focus is very much on staff working collegially to enhance the teaching and learning environment and where the outcomes could be used as evidence for quality purposes.

<center>Quality Assessment of Teaching
and Teaching Portfolios</center>

Teaching Portfolios and the Formative/Summative Dilemma: The First Catch-22?

The previously mentioned discussion about formative and summation evaluation processes of teaching quality also has relevance for Teaching Portfolios. Murphy and MacLaren (2007) conducted a research project that examined the potential of teaching portfolios in higher education for staff development and progression purposes. The question as to how it might be possible to introduce them for summative purposes such as tenure and promotion without also losing their potential to stimulate a good deal of reflection about teaching emerged as a significant focus for the study (see also Knapper & Wright, 2001; Seldin, 2004; Way, 2002). The responses to this question in our consultations and questionnaires are divided into those who see real potential in combining both aspects and those who argue that these two purposes are fundamentally at variance and hence should be addressed via different mechanisms.

One of the respondents to the above mentioned study claimed that "summative instruments may be used for formative purposes but not the reverse" (Questionnaire Response [QR] - 14, Ireland [Ire]). However, another challenged this assumption, echoing

Knapper and Wright's (2001) claim that the differences between the summative and the formative portfolio are "not as great as might be expected" (p. 25). In support, reference was made to Snyder, Lippincott, and Bower's (1998) conclusion that "reconciliation is possible when the assessment can be based on a broad archive of portfolio evidence gathered over a longer period of time from which the teacher can select evidence for specific assessment purposes" (Email Discussion [EM] - 20, Europe [Eur]). Nevertheless, she did recognize that "high-stake assessment will sabotage professional development because it obstructs teachers experimenting in their teaching" and that conversely "portfolio evidence for formative assessment purposes might be of insufficient quality to meet minimal acceptable quality requirements for high stake assessment" (EM - 20, Eur).

It is clear that there are particular challenges in using portfolios for judgmental and even comparative assessment of candidates for promotion (Baume & Yorke, 2002; Casey et al., 1997; Tigelaar et al., 2005). Variations abound in what is considered high quality teaching, for example, in different courses, between different subject areas, and even within subject areas (Babin et al., 2002; Casey et al., 1997; Kreber, 2002). Such assessment leaves "the question open to what extent the interpretation of these evaluations would require a conversation about how the assessment categories reflect the standards of the various disciplines" (QR - 9, United States [US]). As a result, this "would require a broad conversation about discipline-based teaching standards in addition to general standards" (QR - 9, United States [US]). Dyrud (1997) goes so far as to suggest that it is comparing apples and oranges.

Teaching Portfolios and the Reliability Question: The Second Catch-22?

There are additional concerns about any grading or ranking schemes that might be employed in terms of reliability, consistency, objectivity, and comparability (Moss, 1994; Murray, 1995; Ross, Bondy, Hartle, Lamme, & Webb, 1995). Given their subjective nature, "creating criteria, ensuring consistency and reviewing, even for a Pass/Fail result, can be challenging and often problematic activities" (QR - 12, Irl).

Concerns such as these underscore the importance of establishing an approach "that enables assessors to interpret meaning in context and that will have a positive effect on the intended assessment consequences" (Tigelaar et al., 2005, p. 602). Baume and Yorke (2002), for example, emphasized that "reliability is enhanced when there are explicit outcome standards against which to judge, and when there are clear and unambiguous performance data upon which to

exercise that judgment" (p. 17). Even Knapper and Wright (2001), who are very conscious about not forcing portfolios into a quantitative paradigm, still recognized the importance of establishing clear criteria for judging them. They also proposed that that it is very helpful to involve the teaching community in the process of determining the appropriate criteria, whether at the institutional, school, or department level (2001) (see also Casey et al., 1997; Felder & Brent, 1996; Ross et al., 1995).

The importance of clearly specifying the criteria by which the teaching portfolios will be assessed was frequently raised by the respondents in the Murphy and McLaren (2007) study. In practice, however, they note that "quite often neither the candidates nor the assessors are clear about the content and performance standards" (QR - 17, Eur). To address this problem, "teachers and their assessors could create criteria together thus making the assessed partners rather than adversaries" (QR - 20, US). It might also be possible, according to one respondent, to ask Heads of Department/Chairs "to outline the standards of their field so that the assessors of tenure committees who come from different fields understand what constitutes excellent teaching in a specific discipline on a national as well as departmental level" (QR - 9, US).

As previously noted, Fink (2008) was primarily concerned with the quality of the student's learning experience. The focus on student learning is also evident in Zubizarreta's (1999) description of the portfolio for "reflective analysis and peer collaboration leading to improvement of teaching and student learning" (p. 64). A number of the items that he identifies as being potentially constitutive of a teaching portfolio readily align with Fink's four main dimensions above. He outlined them as follows:

1. Information from oneself
   - Responsibilities, philosophy, methods, goals
   - Materials
   - Teaching development activities
2. Information from others
   - Student and peer assessments and ratings
   - Year-end evaluations by chair and dean
   - Honors and awards
   - Letters from colleagues, students, alumni
3. Products of student learning
   - Pre/post tests of learning
   - Classroom assessment activities
   - Student exams, projects, presentations, publications, essays in drafts with instructor's formative feedback
   - Alumni assessments

In a manner similar to Fink (2008), the participants enrolled in the Postgraduate Certificate Programme in Teaching and Learning in Higher Education at NUI Galway are very much focused on being excellent practitioners of the teaching and learning process, especially in terms of maximizing the students' learning experiences. It is in evidence that Fink's (2008) four main dimensions of effective teaching are also reflected in the Teaching Portfolio in use at NUI Galway. It also aligns with the review of portfolios intended for ongoing summative purposes as proposed by Chism (2007, see pp. 181-185).

Teaching Portfolios and the Capacity for Reflection: The Third Catch-22?

Critical reflectivity is about opening up knowledge claims "to proper intellectual challenge" (Andresen, 2000). It is suggested that teaching portfolios may be particularly appropriate for promoting this ethos because their construction requires reflection on "what one teaches, how one teaches, why one teaches that way, how effective that is, and, if necessary or desired, effectively communicating that to others" (Babin, 2002, p. 69). As van Manen (1991) points out, such engagement allows us "to make our pedagogical lives conversationally available: debatable, accountable, evaluable" (p. 19). It also encourages practitioners to conduct research on their own sites of practice that will allow them to develop their own contextually sensitive theories of practice (Brookfield, 1995). McLean and Bullard (2000) confirm these sentiments in the following statement where they contend that portfolios:

> which are produced in contexts in which critical reflective practice, authenticity, and serious engagement with ideas about the teaching/learning relationship are promoted may have the potential both to stimulate teachers to articulate and improve their practice and to be a contribution to understanding the nature of the formation of professional university teachers. (p. 94)

It is recognized, however, from Murphy and MacLaren's (2007) consultations, that the reflective nature of teaching portfolios does present some distinct challenges for many academic staff since many people are not "naturally reflective" (QR - 2, English [Eng]) and hence find it difficult, at first, to operate in this mode of writing. There is of course also a lack of agreement in what constitutes "reflection" and "reflective writing" (see Moon, 2000) – "It is difficult to give an appropriate definition of 'reflection', let alone to develop content and performance standards to assess reflection" (QR - 17, Eng). Additionally, there is the added task of trying to distinguish between different

levels of reflection, "that which includes the testing of validity claims and that which is limited to making explicit one's beliefs (which is in a way nothing more but making an assertion)" (QR - 8, Scottish [Scot]).

All this means, according to another respondent, that "many teachers from different disciplines do not really know what reflective practice means, and even if they do, they are not always clear about how to operationalize reflective practice within their own contexts" (QR -11, Irl). This, of course, implies a similar conceptual and practical challenge for the assessors of portfolios and not just for their authors. One of the respondents pointed to research that she had conducted on the status that is accorded to reflection in higher education:

> According to my own research into this matter, assessors think that explicated reflections are subject to multiple modifications and interpretations and as a result will decrease the validity of a portfolio assessment. For this reason in my research assessors tended to give less weight to reflections in the portfolio than other portfolio elements (e.g. artefacts of teacher behaviour as shown on video, which seems to be more objective). (QR - 17, Eur)

Summary and Conclusion

This paper has examined various aspects of an effective teaching evaluation system. In particular, it has explored the potential of Fink's (2008) four main dimensions of teaching as a summative evaluation model for effective teaching and learning. It emerged that these dimensions can be readily accommodated in a Teaching Portfolio process. The Teaching Portfolio initiative that is in use for the Postgraduate Certificate Programme in Teaching and Learning in Higher Education is a case in point. The challenges encountered when attempting to develop mechanisms for the summative evaluation of quality teaching are also addressed, as well as some of the proposed resolutions for same. As Arreola (2007) pointed out, however, what is clearly evident is that only when the elements of a faculty evaluation program are carefully integrated into a professional enrichment program does the institution obtain the greatest benefit from both.

References

Andresen, L. W. (2000). A useable, trans-disciplinary conception of scholarship. *Higher Education Research & Development*, *19*(2), 137-153.

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*(1), 35.

Baume, D., & Yorke, M. (2002). The reliability of assessment by portfolio on a course to develop and accredit teachers in higher education. *Studies in Higher Education*, *27*(1), 7-25.

Brookfield, S. (1995) *Becoming a critically reflective teacher*. San Francisco: Jossey-Bass.

Casey, J. R., Gentile P., & Bigger, S. W. (1997). Teaching appraisal in higher education: An Australian perspective. *Higher Education*, *34*(4), 459–482.

Cashin, E. W. (1999). Student ratings of teaching: Uses and misuses. In P. Seldin and Associates (Ed.), *Changing practices in evaluating teaching: A practical guide to improved faculty performance and promotion/tenure decisions* (pp. 25-44). Bolton, MA: Anker.

Centra, J. A. (1994). The use of the teaching portfolio and student evaluations for summative evaluation. *Journal of Higher Education*, *65*(5), 555-570.

Centra, J. A. (1993). *Reflective faculty evaluation: Enhancing teaching and determining faculty effectiveness*. San Francisco: Jossey-Bass.

Centra, J. A. (1975). Colleagues as raters of classroom instruction. *Journal of Higher Education*, *46*(3), 327-337.

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

Cohen, P. A. (1980). Effectiveness of student rating feedback for improving college instruction: A meta-analysis of findings. *Research in Higher Education*, *13*(4), 321-341.

Cohen, P., & McKeachie, W. (1980). The role of colleagues in the evaluation of college teaching. *Improving College and University Teaching*, *28*(4), 147-154.

Dyrud, M. A. (1997) Portfolio assessment for promotion and tenure: Three assessment challenges. *Proceedings of the Frontiers in Education Conference*, 267-271.

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

Felder, R. M., & Brent, R. (1996). If you've got it, flaunt it: Uses and abuses of teaching portfolios, *Chemical Engineering Education*, *30*(3), 188-189.

Fink, D. (2002). Improving the evaluation of college teaching. In K. Gillespie (Ed.), *A guide to faculty development* (p. 46-58). Bolton, MA: Anker.

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.

Gosling, D. (2005). Peer observation of Teaching. SEDA, Paper 118. London.

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

Jarzabkowski, P., & Bone, Z. (1998). A 'how-to' guide and checklist for peer appraisal of teaching. *Innovations in Education & Training International, 35*(1), 177-182.

Keig, L., & Waggoner, M. (1994). *Collaborative peer review: The role of faculty in improving college teaching.* ASHE-ERIC Higher Education Report No. 2, Washington, D.C.

Kreber, C. (2002) Controversy and consensus on the scholarship of teaching. *Studies in Higher Education, 27*(2), 151-167.

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

Marincovich, M. (1999). Using student feedback to improve teaching. In P. Seldin and Associates (Eds.). *Changing practices in evaluating teaching: A practical guide to improve faculty performance and promotion/tenure decisions* (pp. 45-69). Bolton, MA: Anker.

McLean, M., & Bullard, J. E. (2000). Becoming a university teacher: Evidence from teaching portfolios. *Teacher Development, 4*(1), 79-101.

Moon, J. (2000). Reflection in learning and professional development: Theory and practice. London: Routledge Falmer.

Moss, P.M. (1994) Can there be validity without reliability? *Educational Research*, *23*(2), 5-12.

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.

Murray, J. P. (1995) The teaching portfolio: A tool for department chairpersons to create a climate of teaching excellence, *Innovative Higher Education*, *19*, 163-175.

Ross, D., Bondy, E., Hartle, L., Lamme, L., & Webb, R. (1995) Guidelines for portfolio preparation: Implications from an analysis of teaching portfolios at the University of Florida, *Innovative Higher Education*, *20*(1), 45-62.

Schon, D. A. (1983). *The reflective practitioner.* San Francisco: Jossey-Bass.

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*(5), 595-610.

Van Manen, M. (1991). *The tact of teaching: The meaning of pedagogical thoughtfulness.* New York: State University of New York Press.

Way, D. (2002). Evaluating teaching portfolios: Moving from a competitive to a collaborative culture. *Proceedings of the International Consortium for Educational Development*, Perth, Australia.

Zubizarreta, J. (1999). Evaluating teaching thorough portfolios. In P. Seldin (Ed.), *Changing practices in evaluating teaching: A practical guide to improved faculty performance and promotion/tenure decisions.* Bolton, MA, Anker.

_____

TIMOTHY MURPHY is a Postdoctoral Researcher at the Centre for Excellence in Learning &Teaching at the National University of Ireland Galway, where he has previously worked for the Open Learning Centre and as a lecturer in Education. He holds a Doctorate in Education from Teachers' College (Columbia University) in New York. He has strong interests (and publications) in civic education, as well as academic staff development Correspondence: Dr Timothy Murphy, Postdoctoral Researcher, Centre for Excellence in Learning and Teaching (CELT), National University of Ireland, Galway, Ireland (timothy.murphy@nuigalway.ie).

IAIN MACLAREN (Mac Labhrainn) is the Director of NUI Galway's Centre for Excellence in Learning & Teaching (CELT) and has extensive teaching and research experience in higher education. He also has strong interests in issues of educational technology, reflective practice and higher education policy issues. His undergraduate degree and PhD were in Astrophysics and, over the last 20 years, he has gradually migrated towards a focus on teaching and learning. He has recently co-edited books on enquiry-based learning, academic staff development in e-learning, and civic engagement and higher education.

SHARON FLYNN was a full-time lecturer in the Information Technology department, NUI Galway, for more than 10 years, during which time she developed a strong interest in educational issues and in using IT to support teaching in Higher Education. She also served

as a Plagiarism Advisor in the Faculty of Engineering during this time. Sharon holds a BSc in Computer Science and Mathematics from University College Dublin, where she also completed her MSc in Computer Science in the area of Computer-Based Tutoring Systems. She has a PhD in Computing Science from Glasgow University.

Appendix A

Knowledge of Subject Matter: Whenever we teach, we are trying to help someone learn about something. The "something" is the subject of the teaching and learning, and all good teachers have some advanced level of knowledge about the subject.

Designing Learning Experiences: Teachers also have to make decisions ahead of time about what the learning experience is going to include and how they want it to unfold. For example: What reading material will be used? What kinds of writing activities will they have students do? Will there be field experiences? Will the teacher use small group activities? How will student learning be assessed? Collectively these decisions represent the teacher's design or plan for the learning experience.

Interacting with Students: Throughout a course, the teacher and the students interact in multiple ways. Lecturing, leading whole class or small group discussions, email exchanges, and meeting with students during office hours – these are all different ways of interacting with students.

Course Management: A course is a complex set of events that involves specific activities and materials. One of the responsibilities of the teacher is to keep track of and manage all the information and materials involved. A teacher needs to know: who has enrolled in the course and who has dropped it; who has taken a test and who was absent; who got what grade on their homework and exams. (Fink, 2008)

Appendix B
Criteria for Assessing Excellence in Teaching

1. Course design
   - Situational factors: Course decisions should be based on solid information about multiple situational factors, e.g., the number of students, their prior knowledge, their feelings about this subject, etc.
   - Learning goals: Are focused on higher level learning, more than just content coverage.
   - Learning activities: Are active and not primarily passive.
   - Feedback and assessment: These procedures enhance the learning process (i.e., they constitute educative assessment) and are more than just a basis for assigning grades.
   - Level of integration: The learning goals, teaching/learning activities, and the feedback and assessment procedures reflect and support each other.
2. Interaction with students
   Teacher is perceived by students as,
   - Competent
   - Trustworthy
   - Dynamic (or energetic)
   - Challenging
   - Stimulating
   - Making students feel included
3. Overall quality of the student learning experience
   - During the course: Students are engaged in their learning
   - End of the course: Students have achieved significant kinds of learning
   - After the course: What students learn has the potential to add value to their lives
4. Improvement over time
   - Seeks out new ideas on teaching.
   - Innovates and tries new ideas in one's own teaching.
   - Evaluates own teaching thoroughly.
   - Reflects continuously on "What do I need to learn about and do next, to improve my teaching?"